Development and Psychopathology (2025), 1–10
doi:10.1017/S0954579425000185

CAMBRIDGE UNIVERSITY PRESS

Regular Article

Reading the mind in infant eyes test: A measure of the recognition of infant emotion

Whitney Barnett[1,2], Julia Garon-Bissonnette[1], Casey Carrow[1], Hannah A. Piersiak[1], Lauren G. Bailes[1,3], Jada Ledgister[1] and Kathryn L. Humphreys[1]

[1]Department of Psychology and Human Development, Vanderbilt University, Nashville, TN, USA, [2]Department of Public Health, Meharry Medical College, Nashville, TN, USA and [3]Department of Psychology, Oklahoma State University, Stillwater, OK, USA

Abstract

Emotion recognition, the ability to interpret others' emotional expressions and infer mental states, is crucial for caregiver–child interactions. The ability to accurately recognize infant emotions may facilitate attuned and responsive caregiving. Across two studies, we validate a novel measure to assess the recognition of infants' emotions (Reading the Mind in Infant Eyes Test [RMIET]) and investigate how this ability relates to observed caregiving. Study 1 examined item-level performance in 55 infant mental health experts and 100 undergraduate students. Study 2 examined RMIET scores in 133 pregnant people and their later caregiving when their children were 18-month-old. In Study 1, agreement was high among both mental health experts (ICC = .82) and undergraduate students (ICC = .93), providing evidence of the content validity of the RMIET. In Study 2, scores assessing the recognition of adult and infant emotions were positively correlated ($r = .22$, $p = .012$). After accounting for covariates, RMIET scores were statistically significantly associated with higher sensitivity and warmth and lower negative regard. Taken together, these studies provide preliminary evidence of content and predictive validity for the RMIET.

Keywords: recognition of infant emotion; emotion perception; caregiver sensitivity; Reading the Mind in Infant Eyes Test

(Received 13 September 2024; revised 28 December 2024; accepted 25 February 2025)

Introduction

Emotion recognition, a core component of social cognition, is central to social interactions, promotes the establishment of healthy relationships, and confers social advantage (Boyatzis & Satyaprasad, 1994; Elfenbein et al., 2002). Emotion recognition represents the *ability to recognize and interpret others' emotional expressions* and enables individuals to infer others' mental states, predict their behavior, and respond appropriately (Adolphs, 2009). Though difficulties are linked to neurodevelopmental and psychiatric disorders (Saghir et al., 2017; Shenoy et al., 2019), emotion recognition abilities vary within the general population of adults (Pavlova & Sokolov, 2022). Given accurately interpreting others' emotional cues is an important factor in establishing positive relationships, caregivers' ability to recognize infant emotions may be uniquely important for the caregiver–child relationship and may foster caregivers' attunement and responsiveness to their infant's experience.

Caregivers' ability to interpret infant emotions may be particularly important during infancy, prior to language acquisition, given that non-verbal cues are the primary way infants are able to communicate needs to caregivers. Caregiver sensitivity is defined as a caregiver's ability to respond appropriately to their children's needs, interests, and cues (Ainsworth et al., 1978; Bornstein et al., 2012). Caregivers' ability to accurately recognize infant emotions may represent a first step to guide behaviors and support appropriate responses to infants' needs, interests, and cues. Meta-analytic results investigating parental reflective functioning and mind-mindedness provide partial support to this idea (Zeegers et al., 2017). Specifically, the degree to which parents consider and coherently reflect upon their infants' internal states, including but not limited to emotions, has a small direct effect on sensitivity as well as a medium direct and small indirect effect (via sensitivity) on attachment security (Zeegers et al., 2017). The ability to consider infant internal states may be, at least in part, driven by emotion recognition abilities.

Several measures have been developed to assess emotion recognition abilities of adult targets. One such measure, the Reading the Mind in the Eyes Test (RMET; Baron-Cohen et al., 2001), assesses one's ability to recognize adult emotions or mental states based on images showing only the eye region of an adult face. Since its creation 20 years ago, the RMET has become one of the most frequently used tools for assessing social capabilities (Pavlova & Sokolov, 2022). Originally developed for evaluation of social cognition in individuals with autism, the RMET has been widely used since in both clinical (Johnson et al., 2022) and neurotypical populations (Pavlova & Sokolov, 2022). Across multiple meta-analytic reviews, the RMET has demonstrated lower ability to recognize emotion in individuals with schizophrenia (Deng et al., 2024) and autism (Peñuelas-Calvo et al., 2019) compared to

Corresponding author: Whitney Barnett; Email: whitney.barnett@mmc.edu

Cite this article: Barnett, W., Garon-Bissonnette, J., Carrow, C., Piersiak, H. A., Bailes, L. G., Ledgister, J., & Humphreys, K. L. (2025). Reading the mind in infant eyes test: A measure of the recognition of infant emotion. *Development and Psychopathology*, 1–10, https://doi.org/10.1017/S0954579425000185

neurotypical controls. Though the ability to accurately assess emotions may be conserved across the age of the target, it is also possible that the recognition of adult compared to developmentally younger targets' emotions may differ. This may be due to several factors. For example, infants tend to express emotions through basic cues (e.g., crying, smiling) whereas adults tend to use a broader range of cues including body language and a wider range of facial expressions (Messinger & Fogel, 2007). Further, adults, specifically compared to young children and infants, have more control over their emotional expressions and are able to mask or modify emotional cues (Martin & Ochsner, 2016). Outside of potential differences in expression complexity, perceptions of the capabilities of others differs by target age. Adults and infants are perceived differently in their agency (see Gray et al., 2007), such that infants are perceived as less competent in their perceptual and cognitive abilities (e.g., telling right from wrong, planning) and in their social and emotional (e.g., finding something funny) abilities (Weisman et al., 2017). These findings suggest the possibility that the ability to interpret emotions accurately may vary as a function of beliefs about the target's emotional capacities.

In this context, Pahnke et al. (2020) developed a measure of child emotion recognition, the Reading the Mind in the Eyes Child Test (RMECT), using emotional state photos of children aged between 8 and 10 years old. In a sample of adults, they found that adult (RMET) and child (RMECT) emotion recognition were positively correlated in small to medium range ($r = .27$). Additionally, they tested whether performance on the RMECT was associated with empathy, a known correlate of one's ability to recognize emotions (Olderbak & Wilhelm, 2017). Accuracy in recognizing child emotional states was found to be related to participants' ability to adopt others' perspective and to be empathetic towards their emotional experience (Pahnke et al., 2020). However, this study did not compare the degree to which the RMECT contributed to variance in relevant behavior with children. Thus, in the current study, we aim not only to assess whether one's ability to recognize infant emotion is distinct from one's ability to recognize adult emotion, but to also investigate whether the ability to recognize infant emotion is, perhaps, more important for caregiving behaviors.

Given the importance of sensitive caregiving in shaping children's development (Cooke et al., 2022; Roger Mills-Koonce et al., 2015), and given important variability in caregivers' interactions with an infant or child versus another adult (i.e., their partner; Brand et al., 2002; van Schaik et al., 2020), examining whether and to what degree infant emotion recognition explains variation in caregiving behavior may be relevant in informing early screening and interventions for caregiver–child relationships. Across two studies, we evaluated a new measure to assess recognition of infant emotion (Study 1), assessed the association between performance on measures of emotion recognition with adult and infant targets, and explored their prospective association with caregiving behavior (i.e., sensitivity, warmth, negative regard, and intrusiveness) (Study 2). In secondary analyses, we investigated whether recognition of infant emotion may explain additional variance in caregiving behavior above the ability to recognize adult emotion.

Study 1

Method

Participants

All procedures and recruitment methods were approved by the Institutional Review Board of Vanderbilt University. All participants provided informed consent prior to participating. There were two groups of participants, comprising expert raters, specifically individuals with substantial work experience with infants or young children ($n = 55$; group 1) and undergraduate students at Vanderbilt University ($n = 100$; group 2). Participants for both groups had to speak and read English fluently and be at least 18 years of age. Additional inclusion criteria for group 1 included having experience in the field(s) of developmental psychology/psychiatry, infant and early childhood mental health, pediatrics, or a related field (Hill et al., 2024) and for group 2 included being enrolled as an undergraduate student. Group 1 participants were recruited via email and participant referral and received a $50 gift certificate for completing the study. Group 2 participants received course credit for completing the study. All participant data was collected online in a Qualtrics database.

Measures

Demographics

Demographic and professional characteristics were collected via self-report, including participant age, gender, race and ethnicity, and highest level of education. For group 1 (i.e., expert raters), field of study and years of practice in their field were also collected. For group 2 (i.e., undergraduate students), number of siblings and years providing care for young children were also collected as an indicator of experience with children.

Reading the mind in infant eyes test (20-item)

Measure development. The Reading the Mind in Infant Eyes Test (RMIET) was created to measure the ability to recognize emotional expressions based on the eye region of an infant's face (Figure 1). This measure was based on the Reading the Mind in the Eyes Test (RMET) (Baron-Cohen et al., 2001). To create this measure, the study team made a public request for images of infants' faces that caregivers would approve for use in research. From this set of 2,157 images, researchers eliminated images that were poor quality, those in which the infant was not facing forward, those in which both of the infant's eyes were not captured or open, and those in which the infant was expressing an ambiguous emotional state. The remaining images ($n = 61$) were cropped to include just the eye region. From a list of 59 words of emotional states, 16 researchers selected words that they thought described what the infant was thinking or feeling for these 61 images. A research team led by the senior author met to reach consensus regarding which words most accurately described the emotional state in each image to determine the target word as well as 3 foil words. The final set of 20 images were selected based on image quality and clarity of the depicted emotional state. In line with open science principles and our commitment to collaborative research, this questionnaire and the scoring approach are available from: https://osf.io/qsn2f/.

Measure administration and rater agreement. For each image, participants indicated whether the depicted expression matched each of the 4 emotional states (target and 3 foil words) on a scale that ranged from 0 (*does not match at all*) to 7 (*perfect match*). Separately within each group, responses for emotional states were averaged across raters.

Analysis

Analyses were conducted using STATA 15.0. Categorical variables (e.g., respondent sex, race, ethnicity, education) were summarized using frequencies and percentages, while continuous variables

https://doi.org/10.1017/S0954579425000185 Published online by Cambridge University Press

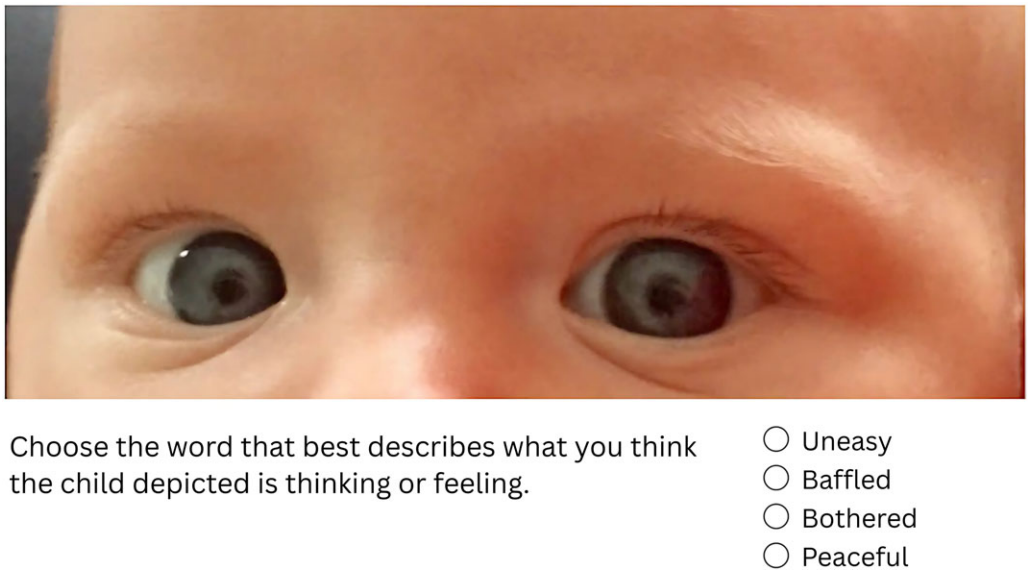

Figure 1. Example test item from the Reading the Mind in Infant Eyes Test (RMIET).

(i.e., image ratings) were summarized using means (standard deviations [*SD*]). For each word ($n = 80$, 4 words per image), an average score was created with higher scores indicating that the word was closer to a "perfect" match for the infant image. Tukey's test for pairwise means (Keselman & Rogan, 1977) were run comparing the mean rating for each target versus foil per image, to assess whether average ratings for target emotional states were statistically significantly different from foils. Intraclass correlation coefficients were run within each respondent group to quantitatively assess the agreement of rater responses. The following criterion were used for qualifying the ICC values: <.50 indicates poor reliability, .50–.75 indicates moderate reliability, .75–.90 indicates good reliability, and > .90 indicates excellent reliability (Koo & Li, 2016).

Results

Participants in group 1 were an average of 43.51 ($SD = 15.39$) years old. The majority were female (98%), White (75%), and had a graduate degree (49% had a master's degree, 24% a Ph.D. or M.D.). On average they had 16.07 ($SD = 10.03$) years of experience in their field and the most common field of training was clinical or counseling psychology (33%) (Supplemental Table 1). Group 2 was comprised of undergraduate students with a mean age of 19.67 ($SD = 0.92$) years. The majority were female (73%) and 45% were White, 27% were Asian. Most respondents had one or more siblings (54%) and had some experience babysitting (6% 1–2 times; 14% 3–5 times; 9% 6–10 times; 50% 11 or more times) (Supplemental Table 2).

Responses for each emotional state were averaged separately for group 1 and 2 across each of the 80 emotional states presented, using the 0 to 7 scale discussed above, with higher scores indicating a better match between the emotional state and infant image. For group 1, target emotional states received an average rating ranging from 3.91–6.40 (Supplemental Table 3). Foils received an average rating ranging from 0.53–5.04. For group 2, target emotional states received an average rating ranging from 4.16–6.43; foils ranged from 0.76–5.11 (Supplemental Table 4). Tukey's pairwise comparison of means results for target emotional state words versus foils are reported in Supplemental Tables 3 & 4. The target emotional state was, on average, rated as a better match for the photo compared to the foils (i.e., received a higher average numerical rating indicating a better match) for all items. In terms of statistical significance in differences between scores, ratings for 49 out of 60 foils were lower than the target in group 1 and 52 out of 60 foils were lower than the target word in group 2 (Supplemental Table 3 & 4). For both groups, ICCs, comparing raters scores (0 to 7) for each foil and target word, were high (group 1 ICC = .82; group 2 ICC = .93) indicating agreement among raters.

Study 1 discussion

Study 1 aimed to provide evidence of the validity a novel measure of emotion recognition in infants, the RMIET. Results from Study 1 provide preliminary evidence of the content validity of the RMIET. Based on the RMET, which has been extensively used to measure ability to recognize adult emotion (Yeung et al., 2024), we developed the RMIET to measure individual ability to recognize infant emotion using photos of infant's eye regions. We followed a similar approach to Pahnke et al. (2020), who developed an adapted measure of the RMET to assess emotion recognition in older children. After selecting photos and cropping them to include only the infant's eye region, we used consensus to generate and select descriptors of emotional states for both target and foil terms.

We evaluated content validity of the RMIET in two samples by asking a group of expert raters, who had extensive training or experience working with children, as well as a group of individuals not selected based on their experience with young children (i.e., undergraduate students), to assess the degree to which each target emotional state matched each photo in the measure. Using both groups enabled us to test the RMIET in both experts and nonexperts. Among expert raters, all target emotional states were rated as the best match, and these target scores were statistically significantly higher than nearly all of the foil words. Among undergraduate students, two of the target emotional states were not rated as the best match. Interrater agreement was high for both undergraduate students (ICC = .93) and expert raters (ICC = .82). As proposed by Boateng et al. (2018), review of items by expert

https://doi.org/10.1017/S0954579425000185 Published online by Cambridge University Press

judges provides an important method to evaluate content validity of a measure. Further, in both groups almost all target words were rated as the best match and rater agreement was high. Thus, results from Study 1 provide preliminary evidence of the content/logical validity of the RMIET.

Study 1 had several strengths, including the development of a novel measure of infant emotion recognition and testing performance in multiple independent samples (e.g., undergraduate students and infant mental health professionals). However, the data collected in Study 1 did not enable us to investigate associations between the ability to recognize infant emotion and correlates of interest for the caregivern–child relationship. In particular, we expect the ability to recognize infant emotion to be associated with caregiving behavior. We address this limitation in Study 2 by investigating the predictive validity of the RMIET and evaluating associations with caregiving behavior (i.e., sensitivity, warmth, negative regard, and intrusiveness).

Study 2

Caregivers' ability to interpret infants' cues, including recognizing infant emotions, may be foundational for guiding caregiving behavior and supporting appropriate responses to infants' needs, which in turn are critical in shaping children's development (Cooke et al., 2022; Roger Mills-Koonce et al., 2015). Therefore, examining whether and to what degree caregivers' ability to recognize infant emotions, as opposed to general emotion recognition, may explain variations in caregiver sensitivity may be useful for informing early screening and, potentially, interventions to support caregiver–child relationships.

Study 2 aimed to: 1) assess the association between emotion recognition performance using adult and infant targets among pregnant individuals, 2) explore the association between emotion recognition in adults and infants and later caregiving behavior (i.e., sensitivity, warmth, negative regard, and intrusiveness), and 3) investigate whether recognition of infant emotion may explain additional variance in caregiving behavior above the ability to recognize adult emotion.

Method

Participants

Participants for Study 2 were drawn from a longitudinal pregnancy cohort study in a large metropolitan city in the central southeastern part of the United States. Pregnant individuals were recruited from local obstetric clinics, print and digital advertisements, listservs, and web-based advertisements (e.g., Facebook, Instagram). Inclusion criteria required that participants be pregnant, at least 18 years of age, fluent in English, and no immediate plans to move from the area. Eligible participants were provided with information about the study and those who agreed to enroll provided informed consent at the first study visit during pregnancy. Participants completed questionnaires on adult and infant emotion recognition and sociodemographic information at a visit between 15 and 38 weeks gestation. Caregiving behavior was then assessed during 11-minutes of "free play" interaction with caregiver–child dyads when children were 18 months old. Vanderbilt University Institutional Review Board approved all recruitment methods and study procedures.

At the time of these analyses, 133 participants had completed the RMIET during pregnancy and engaged in a mother–child interaction when their child was age 18 months. Over half of participants were first-time parents (54%), the majority had completed a Bachelor's degree (80%), and most participants were White (86%). Mean age of caregivers was 32.27 years ($SD = 4.64$) at the baseline assessment. Income-to-needs ratio were calculated by dividing the median point in the reported income bin by the U.S. Department of Housing and Urban Development's low-income threshold (U.S. Department of Housing and Urban Development., 2024) for the number of people in the household for the county in which the University resides. Scores ranged from 0.38 to 3.25 [$M = 1.79$, $SD = 0.70$ (Supplemental Table 5)]; 18% of households had a ratio of <1, which is considered low income based on local thresholds.

Measures

Socio-demographic factors

Socio-demographic variables were collected at the first study visit, during pregnancy. Self-reported socio-demographic data included education level, household income, and whether the study child was the participant's first child.

Reading the mind in the eyes test

The RMET is used to evaluate an individual's capacity to understand the emotional state of others (Baron-Cohen et al., 2001) based on information provided by the eye region of adult faces. The full assessment consists of 36 black and white photos that solely depict the eye portion of someone's face. For this study we administered only the first 20 images to reduce participant burden. For each image, participants are asked to select one of four options (1 target and 3 foils) that best describes the emotion displayed in the image of adult eyes. A final score for adult emotion recognition is calculated by summing the number of total correct emotional states chosen. Possible scores range from 0 to 20, with higher scores indicating a greater ability to recognize adult emotion.

Reading the mind in infant eyes test (17 item)

A subset of items (17 items) from the RMIET used in Study 1 were administered in Study 2. The original version of the RMIET, version 1.1, consisted of 20 items, however 3 items were excluded based on preliminary analyses that indicated poor discrimination between target emotional states and foils (i.e., distractor emotional states) for these 3 items. The RMIET administered in Study 2 was administered prior to completion of expert review for Study 1. Therefore, these 3 items were replaced prior to Study 1 to create a measure with 20 items that showed sufficient discrimination between targets and foils. Using this 17-item RMIET, version 1.2, for each image participants were asked to select which of four emotional states (1 target and 3 foils) best matched the photo. A final score, with a possible range of 0 to 17, was calculated by summing the number of correct responses. Higher scores indicate a greater ability to recognize infant emotion.

Caregiving behaviors

All dyads completed a videorecorded structured interaction in the laboratory at the 18-month time point. The participants were invited to sit on a floor playmat with access to age-appropriate toys and instructed to play with their child as they normally would at home (i.e., a "free play"), which took place in two 5.5-minute episodes at the start and end of a longer behavioral session. Each was split into 3 epochs (two 2-minute segments and one 1.5-minute segment). An average score, weighted for the duration

https://doi.org/10.1017/S0954579425000185 Published online by Cambridge University Press

of the epoch, for caregiving behavior was generated across all segments of the two episodes.

Using the Parent-Child Interaction Rating Scales-Infant Adaptation (PCIRS-IA; Bosquet Enlow et al., 2014), trained independent coders rated caregiving behavior during each free-play segment of the interaction, which were then averaged (weighted for segment duration) to create a composite score. This study coded four domains using the PCIRS-IA scales: 1) Sensitivity (e.g., acknowledging the child's affect; responsiveness to the child's activity and verbalizations; appropriate timing of activities and adapting the pace of play based on child's cues and interests); 2) Intrusiveness (e.g., interrupting child's exploration; insisting on playing with specific toys; rapid/overwhelming presentation of different stimuli); 3) Warmth or Positive regard (e.g., speaking in a warm tone of voice; laughing with the child; smiling); 4) Negative regard (e.g., disapproval; criticism; harshness; sarcasm; name-calling; threats; harsh punishment; roughness; coldness). Possible scores for each variable ranged from 1 (*not at all characteristic*) to 7 (*very characteristic*). Domains are based on intensity (i.e., quality) and frequency (i.e., quantity) of the behaviors. Thus, caregivers may receive higher scores by exhibiting behavior either high in quality, high in quantity, or a combination of both. Based on a randomly selected subset of 46 free-play videos rated by two coders, inter-rater reliability was good to excellent (ICC = .82 for sensitivity, .88 for intrusiveness, .75 for warmth, & .92 for negative regard).

Analysis

All analyses were conducted using STATA 15 (StataCorp Inc., College Station, Texas, USA). Demographic and caregiving behavior data were described using mean (standard deviation) for continuous data or number (%) for categorical data. Pearson correlations were conducted to test associations between key variables. Multiple linear regression models were run to investigate associations between RMIET scores and caregiver sensitivity (primary outcome) as well as caregiver warmth, negative regard, and intrusiveness (secondary outcomes). All models contained household income and number of children as covariates. Regression coefficients with 95% confidence intervals (CIs), standard errors, standardized and unstandardized betas, and *p*-values are reported.

To investigate whether infant emotion recognition explained a greater variance in caregiving behavior than adult emotion recognition, hierarchical linear regression models were conducted. The first block included adult emotion recognition scores, with maternal education and age included as covariates; the second block added infant emotion recognition scores. For all hierarchical linear regressions, 95% CIs, standard errors, standardized and unstandardized betas are reported. To test whether the model including both RMIET and RMET (block 2) explained more variance in caregiving behavior than the model including only RMET (block 1) we report R^2 differences and corresponding *p*-values (Cohen et al., 2013).

Results

Bivariate correlations between all caregiving variables and both the RMIET and RMET are shown in Table 1. RMET and RMIET scores demonstrated a small positive correlation. We conducted linear regression analyses exploring associations between RMIET scores and observed caregiving behaviors, adjusting for household income and number of children in the home (Table 2). Higher RMIET scores were associated with higher sensitivity and warmth, as well as lower levels of negative regard. RMIET scores were not statistically significantly associated with intrusiveness. Further, in hierarchical analyses we calculated the variance in caregiving behavior explained by the RMIET over and above the covariates included (income and number of children). For sensitivity the RMIET explained 3% of the variance above covariates; for warmth, the RMIET explained an additional 10% of variance; and for negative regard, the RMIET explained an additional 3% of variance.

Last, using hierarchical linear regression, we included two blocks of variables to investigate the potential explanatory power of RMIET scores over and above RMET scores.

Block 1 included adult emotion recognition scores, household income, and number of children in the home; block 2 included infant emotion recognition scores (Table 3). RMIET scores explained statistically significant variance over and above RMET scores for warmth and negative regard. Although the RMIET scores explained some additional variance in sensitivity, both RMET and RMIET associations were at the trend level ($p < .10$) when included in the same model.

Study 2 discussion

In a sample of 133 participants followed from pregnancy to child age 18 months, we evaluated the longitudinal association between emotion recognition performance using adult and infant targets and later caregiving behavior. We found a small positive association between performance on the two tests of emotion recognition, suggesting that recognition of adult and infant emotions represent related, but different, abilities. The ability to recognize infant emotion during pregnancy, after accounting for relevant demographic covariates, was prospectively associated with higher levels of sensitivity, higher levels of warmth, and lower levels of negative regard during play interactions between caregivers and their toddlers. Further, the ability to recognize infant emotion was associated with variation in caregiver warmth and negative regard over and above the ability to recognize adult emotion. This is the first study to investigate associations between one's ability to recognize infant emotion and later caregiving behavior. These findings extend those of Study 1, which presented evidence of content/logical validity of the RMIET, by adding evidence of the RMIET's predictive validity for later caregiving behavior.

Results from Study 2, which utilized a prospective, longitudinal design to assess recognition of infant emotion among pregnant people, demonstrated predictive validity for the RMIET on variation in caregiving with one's own child almost two years later, at 18 months. As expected, results suggest that sensitive caregiving, characterized by a caregiver's ability to respond appropriately to infants' needs, interests, and cues, may be preceded by a caregiver's ability to recognize and identify others' emotional states (both infant and adult). We hypothesized that a child-centered measure of emotion recognition (RMIET) would explain additional variance for later caregiving behavior than an adult-centered measure (RMET). When the RMET and RMIET were included separately in models, after adjusting for relevant covariates, each were associated with higher levels of sensitivity. Yet, when both scores were included as predictors of caregiver sensitivity specifically, neither met threshold for statistical significance, suggesting that as the RMET and RMIET relate to caregiver sensitivity, these two abilities are associated with caregiver sensitivity in similar rather than distinct ways. We also

https://doi.org/10.1017/S0954579425000185 Published online by Cambridge University Press

Table 1. Pearson correlations for infant or adult emotion recognition scores and caregiving behavior at 18 Months ($n = 133$)

	M	*SD*	Observed Range	RMIET score	RMET score	Sensitivity	Warmth	Negative regard	Intrusiveness
RMIET score	12.42	1.87	7–17	1					
RMET score	15.07	2.39	10-20	.22*	1				
Sensitivity	4.38	0.71	2.64–5.91	.15	.20*	1			
Warmth	4.81	0.50	3.39–6.00	.29**	.18*	.57**	1		
Negative regard	1.03	0.09	1.00–1.68	−.18*	−.05	−.31**	−.42**	1	
Intrusiveness	1.76	0.66	1.00–4.36	−.02	−.09	−.68**	−.05	.21*	1

Note. RMIET = Reading the Mind in Infant Eyes Test; RMET = Reading the Mind in the Eyes Test.
*$p < .05$. **$p < .001$.

Table 2. Adjusted linear regression investigating associations between infant emotion recognition in pregnancy and caregiving behavior at 18 Months ($n = 133$)

Sensitivity						
			95% CI for B			
Variable	B	*SE* B	*LL*	*UL*	β	*p*
Number of children	0.11	0.08	−0.06	0.27	0.10	.214
Household income	0.21	0.06	0.09	0.32	0.30	< .001
RMIET score	**0.07**	**0.03**	**0.005**	**0.13**	**0.18**	**.033**
Warmth						
			95% CI for B			
Variable	B	*SE* B	*LL*	*UL*	β	*p*
Number of children	0.11	0.06	0.002	0.23	0.16	.055
Household income	0.14	0.04	0.06	0.21	0.28	.001
RMIET score	**0.09**	**0.02**	**0.04**	**0.13**	**0.32**	**< .001**
Negative Regard						
			95% CI for B			
Variable	B	*SE* B	*LL*	*UL*	β	*p*
Number of children	0.02	0.01	−0.001	0.04	0.16	.056
Household income	−0.02	0.01	−0.04	−0.01	−0.25	.003
RMIET score	**−0.01**	**0.004**	**−0.02**	**−0.001**	**−0.18**	**.031**
Intrusiveness						
			95% CI for B			
Variable	B	*SE* B	*LL*	*UL*	β	*p*
Number of children	0.04	0.08	−0.13	0.20	0.04	.657
Household income	−0.09	0.06	−0.20	0.02	−0.14	.124
RMIET score	**−0.01**	**0.03**	**−0.07**	**0.05**	**−0.02**	**.802**

Note. A separate model was run for each of the caregiving behaviors; all models included number of children and household income as covariates. RMIET = Reading the Mind in Infant Eyes Test. Rows documenting the RMIET score effects are bolded.

found that recognition of infant emotion was associated with more warmth and less negative regard. Further, when both the RMIET and RMET were included as predictors of warmth and negative regard, the ability to recognize infant emotion did explain statistically significantly more variance in these caregiving behaviors than the ability to recognize adult emotion. Consistent with our hypothesis that a child-centered measure may be uniquely relevant for caregiving behavior, our findings suggest that the recognition of infant emotion appears to be a distinct predictor of both warmth and negative regard, above the ability to recognize adult emotion.

Literature on other types of caregiver cognitions and representations, including parental reflective functioning, may offer insights into caregiver social cognitive processes and abilities that elucidate why caregiver identification of *infant* emotion may be especially important for caregiving behavior. Indeed, parental reflective functioning–caregivers' ability to reflect upon their child's mental states and to give meaning to their child's behavior, their own experiences as caregivers, and their relationship with their child (Slade, 2005)–likely requires, as a first step, caregivers to attempt to accurately identify their infant's internal states, including emotions. Parental reflective functioning has been linked to children's socioemotional development (Borelli et al., 2021; Madsen et al., 2023), partly through parental responsiveness and sensitivity (e.g., Kelly et al., 2005; Stacks et al., 2014a; Zeegers et al., 2017). Prenatal parental reflective functioning is associated with more positive engagement at 6 months post-partum (Smaling, Huijbregts, Suurland, et al., 2016). Furthermore, higher postnatal parental reflective functioning has been associated with less negative regard towards the child and controlling caregiving behaviors (Huth-Bocks et al., 2014; Stacks et al., 2014b) and more positive caregiving defined as a composite variable including warmth and sensitivity (Huth-Bocks et al., 2014). Though not seen with caregiver sensitivity and intrusiveness, importantly, infant emotion recognition explained substantially more variance in caregiver warmth and negative regard than adult emotion recognition. This suggests that a more granular approach, including narrowing the focus from general abilities to infant- and child-focused cognitive abilities (e.g., caregiver empathy towards the child specifically rather than caregiver dispositional empathy; Salo et al., 2020), is important to understand how caregiver social cognitive abilities translate to behaviors when interacting with their child. Thus, parental reflective functioning, and its association with caregiving behaviors, may be partly dependent of caregivers' accuracy in identifying and recognizing infant emotions. Despite theoretical associations, the relation between infant emotion recognition and parental reflective functioning and their respective or interactive association with caregiving behaviors, including sensitivity, remains unknown.

Strengths of Study 2 include the use of a longitudinal dataset, which enabled us to test prospective associations between performance on the RMIET and caregiving with one's own child. Additionally, Study 2 evaluated the relative explanatory power of the RMIET above that of the RMET, enabling us to test whether the ability to recognize infant emotion explains caregiving behavior above the ability to recognize adult emotion. There are several

https://doi.org/10.1017/S0954579425000185 Published online by Cambridge University Press

Table 3. Hierarchical linear regression investigating recognition of infant and adult emotions and caregiving behavior ($n = 133$)

Sensitivity								
			95% CI for *B*					
Variable	*B*	*SE B*	*LL*	*UL*	β	*p*	R^2	ΔR^2
Block 1							.12**	
Number of children	0.09	0.08	−0.08	0.26	0.09	.281		
Household income	0.19	0.06	0.08	0.31	0.28	.001		
RMET score	0.06	0.02	0.01	0.11	0.19	.021		
Block 2							.14**	.02
Number of children	0.10	0.08	−0.06	0.27	0.10	.218		
Household income	0.20	0.06	0.09	0.32	0.29	.001		
RMET score	0.05	0.03	−0.002	0.10	0.16	.058		
RMIET score	**0.05**	**0.03**	**−0.01**	**0.12**	**0.14**	**.092**		
Warmth								
			95% CI for *B*					
Variable	*B*	*SE B*	*LL*	*UL*	β	*p*	R^2	ΔR^2
Block 1							.11*	
Number of children	0.09	0.06	−0.03	0.21	0.13	.125		
Household income	0.12	0.04	0.04	0.20	0.25	.004		
RMET score	0.04	0.02	0.001	0.07	0.17	.042		
Block 2							.19**	.08*
Number of children	0.11	0.06	−0.003	0.22	0.15	.057		
Household income	0.13	0.04	0.06	0.21	0.27	.001		
RMET score	0.02	0.02	−0.01	0.06	0.11	.195		
RMIET score	**0.08**	**0.02**	**0.04**	**0.12**	**0.29**	**.001**		
Negative Regard								
			95% CI for *B*					
Variable	*B*	*SE B*	*LL*	*UL*	β	*p*	R^2	ΔR^2
Block 1							.09*	
Number of children	0.02	0.01	0.001	0.05	0.18	.038		
Household income	−0.02	0.01	−0.04	−0.01	−0.24	.006		
RMET score	−0.002	0.003	−0.01	0.005	−0.04	.631		
Block 2							.12*	.03*
Number of children	0.02	0.01	−0.01	0.04	−0.16	.057		
Household income	−0.02	0.01	−0.04	−0.01	−0.25	.003		
RMET score	−0.001	0.003	−0.01	0.01	0.004	.996		
RMIET score	**−0.01**	**0.004**	**−0.02**	**−0.001**	**−0.18**	**.036**		
Intrusiveness								
			95% CI for *B*					
Variable	*B*	*SE B*	*LL*	*UL*	β	*p*	R^2	ΔR^2
Block 1							.03	
Number of children	0.04	0.08	−0.12	0.20	0.04	.644		
Household income	−0.08	0.06	−0.19	0.03	−0.13	.137		
RMET score	−0.02	0.02	−0.07	0.02	−0.09	.325		

(*Continued*)

https://doi.org/10.1017/S0954579425000185 Published online by Cambridge University Press

Table 3. (*Continued*)

Sensitivity								
Block 2							.03	.00
Number of children	0.04	0.08	−0.13	0.20	0.04	.649		
Household income	−0.08	0.06	−0.20	0.03	−0.13	.140		
RMET score	−0.02	0.02	−0.07	0.03	−0.09	.342		
RMIET score	**−0.001**	**0.03**	**−0.06**	**0.06**	**−0.003**	**.974**		

Note. RMIET = Reading the Mind in Infant Eyes Test; RMET = Reading the Mind in the Eyes Test. Rows documenting the RMIET score effects are bolded.

study limitations. First, although we used longitudinal data to assess prospective associations between the RMIET and caregiving behavior, we cannot assert causality. Though we believe the association is likely causal, without being able to manipulate one's recognition of infant emotions (e.g., through interventions or experiments) we are unable to test this. Second, we used the short form of the RMET in order to reduce participant burden, though a consequence is that it is more difficult to compare our findings to studies in which the full measure was used. Third, Study 2 included caregivers drawn from a U.S. sample, though differences in facial expressions and interpretations across cultures have been documented (Elfenbein et al., 2002; Matsumoto & Ekman, 1989). This may limit generalizability to non-Western (and non-US) samples. Further, the sample in Study 2 was relatively well-educated and high-income (compared to national averages); factors linked to more positive caregiving behaviors (e.g., sensitivity, warmth; Azad et al., 2014). This may limit generalizability to families with fewer resources.

General discussion

Our goal across two studies was to develop and test a novel measure of infant emotion recognition, the RMIET, and to begin the process of establishing its validity. We explored the relation between recognizing infant emotions and recognizing adult emotions. We investigated whether the ability to recognize infant emotions explained variance in caregiving behaviors above and beyond the ability to recognize adult emotions. The RMIET was developed to measure individual ability to recognize infant emotion, using images of the eye region (as with the original RMET). We hypothesized that the ability to recognize infant emotion may represent a different, if related, ability compared to recognizing adult emotions. Age of the target may be relevant to the ability to recognize infant emotion given that pre-verbal infants may rely on more basic emotional cues than adults (Messinger & Fogel, 2007), have less control over their emotional expressions (Martin & Ochsner, 2016), and infants are perceived as less competent in their emotional abilities (Weisman et al., 2017). We further hypothesized that caregivers' ability to interpret infant emotions would relate to variations in caregiving behavior. Evidence from Study 1 and Study 2 suggests that the RMIET exhibits content and preliminary predictive validity as a measure of the ability to recognize infant emotion. Scores documenting a small positive association between tasks focused on the recognition of infant vs. adult emotions suggests the possibility that the ability to recognize infant and recognize adult emotions represent somewhat unique constructs. However, although the RMIET was associated with higher sensitivity, higher warmth, and lower negative regard, the RMIET explained variance above the RMET only for warmth and negative regard. While our findings demonstrate associations between the RMIET and caregiving behaviors, the correlations observed were small. As discussed in more detail below, further work is needed to establish the reliability and validity of the RMIET as a measure of recognition of infant emotion.

Evidence from intervention studies has shown that higher quality caregiving contributes to children's social and emotional development and, importantly, that caregiving quality in early life is modifiable (Almas et al., 2015; Bernard et al., 2012). Given this, understanding factors that contribute to caregiving behavior is critical. The ability to recognize infant emotions may represent an *early* indicator of likely caregiving behavior, assessable prior to a child's birth. Other studies have documented prospective associations between assessments during pregnancy and later caregiving behavior. For example, using the Prenatal Attachment Inventory (Muller, 1993), one study found that mothers' thoughts, feelings, and relationship to their unborn child was associated with later maternal involvement and stimulation in video-taped interactions with their child (Siddiqui & Hägglöf, 2000). Findings align with recommendations that improved understanding of infant- or child-focused cognitive abilities may be uniquely relevant for understanding caregiving behavior (Humphreys et al., 2024; Smaling et al., 2016). Moreover, child-focused cognitive abilities during pregnancy may be an early modifiable factor for later caregiving behavior.

This study has several notable strengths, including the development of a novel measure of infant emotion recognition, the use of longitudinal data, and the inclusion of several independent samples (e.g., college students, infant mental health professionals, and pregnant people). Despite the strengths of this study, there are limitations that need to be considered. First, although the study used longitudinal data to explore associations, causality cannot be determined. Second, a recent systematic review of evidence for the validity of the RMET (Higgins et al., 2024) found that a substantial portion of literature (63%) did not provide evidence of validity for the RMET and relatively low alpha levels for items was pointed to as a limitation of this measure. For the purposes of this study, we evaluated content validity of the RMIET as proposed by Boateng et al. (2018) through review of the RMIET by expert judges (Study 1). Future efforts to evaluate the RMIET and to establish reliability and validity should include assessments of convergent and discriminant validity, test-retest reliability, and examining whether there is an underlying factor structure of the RMIET. Additional research is also needed to explore the extent to which the RMIET is predictive of caregiving behaviors across diverse samples and settings, including those with greater socioeconomic diversity and in non-Western cultural contexts. Third, our study examined pregnant people and caregivers drawn from a U.S. sample that was relatively well-educated and high-income (compared

https://doi.org/10.1017/S0954579425000185 Published online by Cambridge University Press

to national averages); factors linked to more positive caregiving behaviors (e.g., sensitivity, warmth; Azad et al., 2014). This may limit generalizability to families with fewer resources. Further, there are documented differences in emotion expression and perception for Western versus non-Western samples (Lim, 2016; Stern et al., 2022). Several studies have documented cultural differences in emotional expression (Immordino-Yang et al., 2016), facial expression and recognition of emotions (Elfenbein et al., 2002; Matsumoto & Ekman, 1989), and affect valuation (Tsai et al., 2006). For example, Westerners experience high arousal emotions more than low arousal emotions, whereas, in Eastern or collectivist culture, low arousal emotions are valued more than high arousal emotions (Lim, 2016). Additionally, research from one group found differences in how Americans compared to Japanese rely on facial cues, finding that Japanese participants relied more on signals from the eye area whereas Americans rely more on signals from the mouth region (Yuki et al., 2007). These cultural differences in emotion valuation and reliance on facial cues likely mean that the findings presented here may not be generalizable to non-US or non-Western samples.

Conclusions

Results from these two studies provide evidence that recognition of infant emotion may be an important factor for understanding variation in caregiving behavior. If causal, it would suggest a path forward for fostering positive and inhibiting negative caregiving behaviors, with potential implications for child attachment and functioning. Enhancing the ability to recognize infant emotions in caregivers may be an important focus for interventions aiming to improve caregiving quality and the caregiver–child relationship. Our results suggest that the ability to recognize, and respond appropriately to infant emotion, may be an important and potentially modifiable intervention target with implications for the caregiver–child relationship.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0954579425000185.

Acknowledgments. We are grateful to the study participants for their time, particularly during a busy stage of life. We thank the research team involved in data collection and coding, especially Tiffany Tran for her help constructing the Qualtrics database for this study and preliminary data analysis to inform this work.

Funding statement. This work was supported by the National Science Foundation (2042285 to KLH); the Brain and Behavior Research Foundation (29,593 to KLH); a Vanderbilt Institute for Clinical and Translational Research Grant (VR53419); a Vanderbilt Strong Grant; a Vanderbilt Kennedy Center Grant; and National Institutes of Health (T32MH18921 supported LGB & HAP; K01HD112541 supported WB).

Competing interests. None.

References

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology*, *60*(1), 693–716. https://doi.org/10.1146/annurev.psych.60.110707.163514

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

Almas, A. N., Degnan, K. A., Walker, O. L., Radulescu, A., Nelson, C. A., Zeanah, C. H., & Fox, N. A. (2015). The effects of early institutionalization and foster care intervention on children's social behaviors at age 8. *Social Development (Oxford, England)*, *24*(2), 225–239. https://doi.org/10.1111/sode.12089

Azad, G., Blacher, J., & Marcoulides, G. (2014). Longitudinal models of socio-economic status: Impact on positive parenting behaviors. *International Journal of Behavioral Development*, *38*(6), 509–517. https://doi.org/10.1177/0165025414532172

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the mind in the eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(2), 241–251.

Bernard, K., Dozier, M., Bick, J., Lewis-Morrarty, E., Lindhiem, O., & Carlson, E. (2012). Enhancing attachment organization among maltreated children: Results of a randomized clinical trial. *Child Development*, *83*(2), 623–636. https://doi.org/10.1111/j.1467-8624.2011.01712.x

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

Borelli, J. L., Lai, J., Smiley, P. A., Kerr, M. L., Buttitta, K., Hecht, H. K., & Rasmussen, H. F. (2021). Higher maternal reflective functioning is associated with toddlers' adaptive emotion regulation. *Infant Mental Health Journal*, *42*(4), 473–487. https://doi.org/10.1002/imhj.21904

Bornstein, M. H., Suwalsky, J. T. D., & Breakstone, D. A. (2012). Emotional relationships between mothers and infants: Knowns, unknowns, and unknown unknowns. *Development and Psychopathology*, *24*(1), 113–123. https://doi.org/10.1017/S0954579411000708

Bosquet Enlow, M., Carter, A., Hails, K., King, L., & Cabrera, I. (2014). Parent-Child Interaction Rating Scales (PCIRS)–Infant Adaptation. Unpublished manual, Boston Children's Hospital.

Boyatzis, C. J., & Satyaprasad, C. (1994). Children's facial and gestural decoding and encoding: Relations between skills and with popularity. *Journal of Nonverbal Behavior*, *18*(1), 37–55. https://doi.org/10.1007/BF02169078

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for "motionese": Modifications in mothers' infant-directed action. *Developmental Science*, *5*(1), 72–83. https://doi.org/10.1111/1467-7687.00211

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/Correlation analysis for the behavioral sciences*. Routledge.

Cooke, J. E., Deneault, A.-A., Devereux, C., Eirich, R., Fearon, R. M. P., & Madigan, S. (2022). Parental sensitivity and child behavioral problems: A meta-analytic review. *Child Development*, *93*(5), 1231–1248. https://doi.org/10.1111/cdev.13764

Deng, F., Bueber, M. A., Cao, Y., Tang, J., Bai, X., Cho, Y., Lee, J., Lin, Z., Yang, Q., Keshavan, M. S., Stone, W. S., Qian, M., Yang, L. H., & Phillips, M. R. (2024). Assessing social cognition in patients with schizophrenia and healthy controls using the reading the mind in the eyes test (RMET): A systematic review and meta-regression. *Psychological Medicine*, *54*(5), 847–873. https://doi.org/10.1017/S0033291723003501

Elfenbein, H. A., Mandal, M. K., Ambady, N., Harizuka, S., & Kumar, S. (2002). Cross-cultural patterns in emotion recognition: Highlighting design and analytical techniques. *Emotion*, *2*(1), 75–84. https://doi.org/10.1037/1528-3542.2.1.75

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619–619. https://doi.org/10.1126/science.1134475

Higgins, W. C., Kaplan, D. M., Deschrijver, E., & Ross, R. M. (2024). Construct validity evidence reporting practices for the reading the mind in the eyes test: A systematic scoping review. *Clinical Psychology Review*, *108*, 102378. https://doi.org/10.1016/j.cpr.2023.102378

Hill, K. E., Guyon-Harris, K. L., & Humphreys, K. L. (2024). Data representing expert ratings of emotional tone in adjectives used to describe infants and young children. *Data in Brief*, *55*, 110745. https://doi.org/10.1016/j.dib.2024.110745

Humphreys, K. L., Garon-Bissonnette, J., Hill, K. E., Bailes, L. G., Barnett, W., & Hare, M. M. (2024). Caregiving relationships are a cornerstone of developmental psychopathology. *Development and Psychopathology*, *36*(5), 1–14. https://doi.org/10.1017/S0954579424000300

Huth-Bocks, A. C., Muzik, M., Beeghly, M., Earls, L., & Stacks, A. M. (2014). Secure base scripts are associated with maternal parenting behavior across contexts and reflective functioning among trauma-exposed mothers.

Attachment & Human Development, *16*(6), 535–556. https://doi.org/10.1080/14616734.2014.967787

Immordino-Yang, M. H., Yang, X.-F., & Damasio, H. (2016). Cultural modes of expressing emotions influence how emotions are experienced. *Emotion (Washington)*, *16*(7), 1033–1039. https://doi.org/10.1037/emo0000201

Johnson, B. N., Kivity, Y., Rosenstein, L. K., LeBreton, J. M., & Levy, K. N. (2022). The association between mentalizing and psychopathology: A meta-analysis of the reading the mind in the eyes task across psychiatric disorders. *Clinical Psychology: Science and Practice*, *29*(4), 423–439. https://doi.org/10.1037/cps0000105

Kelly, K., Slade, A., & Grienenberger, J. F. (2005). Maternal reflective functioning, mother-infant affective communication, and infant attachment: Exploring the link between mental states and observed caregiving behavior in the intergenerational transmission of attachment. *Attachment & Human Development*, *7*(3), 299–311. https://doi.org/10.1080/14616730500245963

Keselman, H. J., & Rogan, J. C. (1977). The Tukey multiple comparison test: 1953–1976. *Psychological Bulletin*, *84*(5), 1050–1056. https://doi.org/10.1037/0033-2909.84.5.1050

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lim, N. (2016). Cultural differences in emotion: Differences in emotional arousal level between the East and the West. *Integrative Medicine Research*, *5*(2), 105–109. https://doi.org/10.1016/j.imr.2016.03.004

Madsen, E. B., Væver, M. S., Egmose, I., Krogh, M. T., Haase, T. W., de Moor, M. H. M., & Karstoft, K.-I. (2023). Parental reflective functioning in first-time parents and associations with infant socioemotional development. *Journal of Child and Family Studies*, *32*(7), 2140–2152. https://doi.org/10.1007/s10826-023-02565-5

Martin, R. E., & Ochsner, K. N. (2016). The neuroscience of emotion regulation development: Implications for education. *Current Opinion in Behavioral Sciences*, *10*, 142–148. https://doi.org/10.1016/j.cobeha.2016.06.006

Matsumoto, D., & Ekman, P. (1989). American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion*, *13*(2), 143–157. https://doi.org/10.1007/BF00992959

Messinger, D., & Fogel, A. (2007). The interactive development of social smiling. In *Advances in child development and behavior*. (Vol. *35*, pp. 327–366). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-009735-7.50014-1

Muller, M. E. (1993). Development of the prenatal attachment inventory. *Western Journal of Nursing Research*, *15*(2), 199–215. https://doi.org/10.1177/019394599301500205

Olderbak, S., & Wilhelm, O. (2017). Emotion perception and empathy: An individual differences test of relations. *Emotion (Washington)*, *17*(7), 1092–1106. https://doi.org/10.1037/emo0000308

Pahnke, R., Mau-Moeller, A., Hamm, A. O., & Lischke, A. (2020). Reading the mind in the eyes of children test (RME-C-T): Development and validation of a complex emotion recognition test. *Frontiers in Psychiatry*, *11*, 376. https://www.frontiersin.org/articles/10.3389/fpsyt.2020.00376

Pavlova, M. A., & Sokolov, A. A. (2022). Reading language of the eyes. *Neuroscience & Biobehavioral Reviews*, *140*, 104755. https://doi.org/10.1016/j.neubiorev.2022.104755

Peñuelas-Calvo, I., Sareen, A., Sevilla-Llewellyn-Jones, J., & Fernández-Berrocal, P. (2019). The "Reading the mind in the eyes" test in Autism-spectrum disorders comparison with healthy controls: A systematic review and meta-analysis. *Journal of Autism and Developmental Disorders*, *49*(3), 1048–1061. https://doi.org/10.1007/s10803-018-3814-4

Roger Mills-Koonce, W., Willoughby, M. T., Zvara, B., Barnett, M., Gustafsson, H., & Cox, M. J. (2015). Mothers' and fathers' sensitivity and children's cognitive development in low-income, rural families. *Journal of Applied Developmental Psychology*, *38*, 1–10. https://doi.org/10.1016/j.appdev.2015.01.001

Saghir, H., Dupuis, A., Chau, T., & Kushki, A. (2017). Atypical autonomic nervous system complexity accompanies social cognition task performance in ASD. *Research in Autism Spectrum Disorders*, *39*, 54–62. https://doi.org/10.1016/j.rasd.2017.04.004

Salo, V. C., Schunck, S. J., & Humphreys, K. L. (2020). Depressive symptoms in parents are associated with reduced empathy toward their young children. *PLOS ONE*, *15*(3), e0230636. https://doi.org/10.1371/journal.pone.0230636

Shenoy, S., Desai, G., Venkatasubramanian, G., & Chandra, P. S. (2019). Parenting in mothers with schizophrenia and its relation to facial emotion recognition deficits—A case control study. *Asian Journal of Psychiatry*, *40*, 55–59. https://doi.org/10.1016/j.ajp.2019.01.022

Siddiqui, A., & Hägglöf, B. (2000). Does maternal prenatal attachment predict postnatal mother-infant interaction? *Early Human Development*, *59*(1), 13–25. https://doi.org/10.1016/s0378-3782(00)00076-1

Slade, A. (2005). Parental reflective functioning: An introduction. *Attachment & Human Development*, *7*(3), 269–281. https://doi.org/10.1080/14616730500245906

Smaling, H. J. A., Huijbregts, S. C. J., Suurland, J., van der Heijden, K. B., Mesman, J., van Goozen, S. H. M., & Swaab, H. (2016). Prenatal reflective functioning and accumulated risk as predictors of maternal interactive behavior during free play, the still-face paradigm, and two teaching tasks. *Infancy*, *21*(6), 766–784. https://doi.org/10.1111/infa.12137

Smaling, H. J. A., Huijbregts, S. C. J., van der Heijden, K. B., van Goozen, S. H. M., & Swaab, H. (2016). Maternal reflective functioning as a multidimensional construct: Differential associations with children's temperament and externalizing behavior. *Infant Behavior & Development*, *44*, 263–274. https://doi.org/10.1016/j.infbeh.2016.06.007

Stacks, A. M., Muzik, M., Wong, K., Beeghly, M., Huth-Bocks, A., Irwin, J. L., & Rosenblum, K. L. (2014a). Maternal reflective functioning among mothers with childhood maltreatment histories: Links to sensitive parenting and infant attachment security. *Attachment & Human Development*, *16*(5), 515–533. https://doi.org/10.1080/14616734.2014.935452

Stacks, A. M., Muzik, M., Wong, K., Beeghly, M., Huth-Bocks, A., Irwin, J. L., & Rosenblum, K. L. (2014b). Maternal reflective functioning among mothers with childhood maltreatment histories: Links to sensitive parenting and infant attachment security. *Attachment & Human Development*, *16*(5), 515–533. https://doi.org/10.1080/14616734.2014.935452

Stern, J. A., Barbarin, O., & Cassidy, J. (2022). Working toward anti-racist perspectives in attachment theory, research, and practice. *Attachment & Human Development*, *24*(3), 392–422. https://doi.org/10.1080/14616734.2021.1976933

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of Personality and Social Psychology*, *90*(2), 288–307. https://doi.org/10.1037/0022-3514.90.2.288

U.S. Department of Housing and Urban Development. (2024). FY 2024 Income limits summary. https://www.huduser.gov/portal/datasets/il/il2020/2020summary.odn.

van Schaik, J. E., Meyer, M., van Ham, C. R., & Hunnius, S. (2020). Motion tracking of parents' infant- versus adult-directed actions reveals general and action-specific modulations. *Developmental Science*, *23*(1), e12869. https://doi.org/10.1111/desc.12869

Weisman, K., Dweck, C. S., & Markman, E. M. (2017). Rethinking people's conceptions of mental life. *Proceedings of The National Academy of Sciences of The United States of America*, *114*(43), 11374–11379. https://doi.org/10.1073/pnas.1704347114

Yeung, E. K. L., Apperly, I. A., & Devine, R. T. (2024). Measures of individual differences in adult theory of mind: A systematic review. *Neuroscience & Biobehavioral Reviews*, *157*, 105481. https://doi.org/10.1016/j.neubiorev.2023.105481

Yuki, M., Maddux, W. W., & Masuda, T. (2007). Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. *Journal of Experimental Social Psychology*, *43*(2), 303–311. https://doi.org/10.1016/j.jesp.2006.02.004

Zeegers, M. A. J., Colonnesi, 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. *Psychological Bulletin*, *143*(12), 1245–1272. https://doi.org/10.1037/bul0000114