This is the post print version of the article, which has been published in Developmental Science . 2018, 21(6),e12687.
DOI: https://doi.org/10.1111/desc.12687

Infants' attention bias to faces as an early marker of social development

Mikko J. Peltola[a], Santeri Yrttiaho[b], & Jukka M. Leppänen[b]

[a] Human Information Processing Laboratory, Faculty of Social Sciences, University of Tampere

[b] Infant Cognition Laboratory, Tampere Center for Child Health Research, Faculty of Medicine and Life Sciences, University of Tampere

Correspondence: Mikko Peltola, Human Information Processing Laboratory, Faculty of Social Sciences, 33014 University of Tampere, Finland. E-mail: mikko.peltola@staff.uta.fi, tel: +358 50 318 6120

Research Highlights

- Infants showed a robust attention bias to faces at 7 months, particularly when faces displayed a fearful expression.
- Longitudinal analyses tested whether individual variations in infants' attention bias to faces were associated with social developmental outcomes at 24 and 48 months of age.
- Increased attention to faces at 7 months was associated with more frequent spontaneous helping at 24 months and reduced callous-unemotional traits at 48 months of age.
- Infants' attention bias to faces may be specifically linked with the development of affective empathy and responsivity to others' needs.

Abstract

Infants have a strong tendency to look at faces. We examined individual variations in this attentional bias in 7-month-old infants by using a face-distractor competition paradigm and tested in a longitudinal sample whether these variations were associated with outcomes reflecting social behavior at 24 and 48 months of age (i.e., spontaneous helping, emotion understanding, mentalizing, and callous-unemotional traits; *N* = 100-138). The results showed a robust and distinct attention bias to faces at 7 months, particularly when faces were displaying a fearful expression. This bias declined between 7 and 24 months and there were no significant correlations in attention dwell times between 7 and 24 months of age. Variations in attention to faces at 7 months were not associated with emotion understanding or mentalizing abilities at 48 months of age, but increased attention to faces at 7 months (regardless of facial expression) was related to more frequent helping responses at 24 months and reduced callous-unemotional traits at 48 months of age. Thus, while the results fail to associate infants' face bias with later-emerging emotion understanding and mentalizing capacities, they are consistent with a model whereby increased attention to faces in infancy is linked with the development of affective empathy and responsivity to others' needs.

Infants show an early capacity to orient to salient social stimuli, such as faces. In newborns, this bias has been demonstrated as prolonged visual tracking of face-like patterns (Johnson, Dziurawiec, Ellis, & Morton, 1991) and in older, 4-12 month-old infants, as biased orientation to faces amongst multiple competing objects (Amso, Haas, & Markant, 2014; Frank, Vul, & Johnson, 2009; Gluckman & Johnson, 2013; Kwon, Setoodehnia, Baek, Luck, & Oakes, 2014) or relatively longer dwell time to faces compared to other visual objects (Gluckman & Johnson, 2013; Kwon et al., 2014; Leppänen et al., 2011). The bias to faces is regarded as a central and evolutionary conserved component of infant "social engagement", and a prerequisite for the acquisition of more complex capacities, such as the ability to represent others' thoughts and emotional states (e.g., Klin, Shultz, & Jones, 2015).

While the bias to attend to faces is well documented, studies have only recently begun to examine the mechanisms underlying this bias in infants (Frank, Amso, & Johnson, 2014) as well as its hypothesized role in early social development (e.g., Bedford, Pickles, Sharp, Wright, & Hill, 2015). In the current study, we sought to further address these questions in two ways. First, we examined whether individual variations in infants' attention to faces reflect a distinct, domain-specific trait and are not explained by variations in more general attention orienting and holding mechanisms (Cohen, 1972). Second, we examined whether there is continuity in attention to faces between infancy and early childhood (i.e., whether attention dwell times to faces are correlated and show a similar pattern between 7 and 24 months), and whether individual variations in the bias to faces in infancy are associated with multiple aspects of early childhood social development, including measures of empathic responsivity toward others and processes that may be important in mediating individual variations in social behavior (cf. Bedford et al., 2015).

The existence of distinct attentional mechanisms for faces is supported by event-related potential studies showing dissociable patterns of cortical activation to objects vs. faces (de Haan & Nelson, 1999; Halit, Csibra, Volein, & Johnson, 2004; Yrttiaho, Forssman, Kaatiala, &

Leppänen, 2014) and by results showing that variance in young children's visual scanning of dynamic social scenes is best explained by two orthogonal factors: a general social orienting component and a tendency to orient attention to the most informative features of a scene, such as areas of the other person's face (Chawarska, Ye, Shic, & Chen, 2016). There is also evidence showing that infants' tendency to look at faces is consistent across types of face stimuli, such as individual face pictures, faces among multiple competing objects, and faces included in natural scenes (Gillespie-Smith et al., 2016). However, a strictly domain-specific view of infants' visual and attentional processing is challenged by studies showing that attention to faces correlates with measures of general attention capacities, such as visual search (i.e., latency to localize a discrepant target within a matrix of non-face objects; Frank, Amso, & Johnson, 2014) or the average duration of visual fixations during scene exploration (Amso et al., 2014).

To further examine the specificity of infants' attention bias to faces, we examined individual variations in attention to faces and, in particular, the degree of shared variance (i.e., cross-correlation) in infants' dwell times to non-face patterns and neutral, happy, and fearful facial expressions. Following the rationale discussed in Wilmer (2008), a model suggesting partially independent mechanisms underlying attention to faces and objects would predict that dwell time correlations across stimulus categories (non-face patterns vs. faces) are significantly lower than within-category correlations (i.e., associations between two face conditions). Notably, this model can be extended if partially independent mechanisms are assumed for faces displaying different emotional expressions, which would predict correlations between expression categories to be lower than those within expression categories. On the other hand, a model positing infants' attention to faces being driven by domain-general oculomotor functions would predict that correlations between dwell times to non-face patterns and faces are high and also of comparable magnitude to within-category correlations (i.e., correlations in dwell time between different face stimuli).

Regarding the second aim of the current study, evidence for the role of infant attention to faces in the development of social-emotional capacities is limited, but emerging data from recent prospective longitudinal studies are consistent with this notion. Heightened attention to faces signaling negative emotion (i.e., fearful faces) at 7 months of age has been associated with increased odds of secure attachment to the mother at the age of 14 months (Peltola, Forssman, Puura, van IJzendoorn, & Leppänen, 2015). Studies investigating early markers of callous-unemotional (CU) traits (i.e., a marked disregard of others' distress and lack of empathy) have shown that reduced looking at faces vs. non-face objects at 5 weeks of age (Bedford et al., 2015) and reduced looking at the parent's face during moments of disrupted face-to-face interaction in the still-face procedure (Wagner et al., 2016) are associated with increased CU traits later in childhood in unselected populations (but see Bedford et al., 2017, for a lack of direct association).

Complementing prospective longitudinal studies, other studies have shown concurrent associations between relatively quicker orienting to fearful faces with increased altruistic behavior in 4 to 5-year-old children (Rajhans, Altvater-Mackensen, Vaish, & Grossmann, 2016; see also Marsh, Kozak, & Ambady, 2007, for related evidence from adults). Lack of attention to the eyes and reduced accuracy in recognizing fearful facial expressions have been associated with reduced empathy and increased levels of CU traits in school-aged children and adults (Dadds et al., 2006; Dadds, Jambrak, Pasalich, Hawes, & Brennan, 2011; Muñoz, 2009; White et al., 2016; but see Dawel, O'Kearney, McKone, & Palermo, 2012, for evidence of more pervasive emotion processing impairments in individuals with antisocial tendencies). Further, children who may have limited amount and range of exposure to faces during developmentally sensitive periods due to institutionalization (Wismer Fries & Pollak, 2004), maltreatment (Pollak, Cicchetti, Hornung, & Reed, 2000; Pollak & Kistler, 2002), or maternal depression (Székely et al., 2014) tend to show poorer performance in tasks measuring the ability to label emotional facial expressions and understand emotional scenarios.

While attention to faces has been variably linked with various social behaviors, studies in infants are still scarce and covariations among infants' attentional biases to faces and fearful expressions as well as different early-developing social-cognitive capacities, emotion-related behaviors, and behavioural problems have not been systematically examined within the same study. Understanding these covariations is an important first step in identifying mechanisms that mediate developmental outcomes (e.g., emotion understanding may have a central role in mediating CU symptoms; Bedford et al., 2015; White et al., 2016). We addressed these questions by studying how infants' attentional biases to faces in general and fearful faces in particular at 7 months are associated with age-typical social behaviours at 24 and 48 months, spanning from relatively reactive tendencies (spontaneous prosocial behavior and CU traits) to more complex social-cognitive processes (emotion understanding and mentalizing). To this end, we studied spontaneous helping of others in need (Kärtner, Keller, & Chaudhary, 2010; Warneken & Tomasello, 2006), the ability to understand others' emotions from contextual cues (Wismer Fries & Pollak, 2004), the ability to mentalize others' desires and intentions (Wellman & Liu, 2004), and empathic abilities characterized as a lack of disregard of others' distress (i.e., CU traits; Bedford et al., 2015; Wagner et al., 2016). These characteristics have been previously implicated with attention to faces and they develop during the early childhood years, with spontaneous helping appearing before two years of age (Warneken & Tomasello, 2006) and more complex abilities such as mentalizing and emotion understanding showing a more protracted development (e.g., Wellman & Liu, 2004). The onset age of first detectable CU behaviors is currently unclear (Waller et al., 2017), but available evidence indicates that the assessment of behaviors characterizing CU traits becomes increasingly reliable by the age of four (Hyde et al., 2013).

Based on studies showing marked developmental changes in attention to faces and facial expressions at around 4 to 8 months of age (Frank et al., 2009; Jessen & Grossmann, 2015; Kwon et al., 2014; Leppänen & Nelson, 2009; Peltola, Leppänen, Mäki, & Hietanen, 2009; Yrttiaho et al., 2014), and the evidence for distinct mechanisms for face and non-face object processing (Halit et al., 2004; Yrttiaho et al., 2014), we hypothesized that individual variability in attention dwell times to faces is evident at 7 months of age and that dwell times to faces are partially distinct from dwell times to non-face objects. Based on developmental models (Klin et al., 2015) and available empirical data (e.g., Bedford et al., 2015), we predicted that greater attention bias to faces at 7 months is broadly associated with more optimal developmental outcomes later in childhood: greater attention to faces at 24 months of age, more spontaneous helping at 24 months, better emotion understanding and mentalizing abilities at 48 months, and reduced expression of CU traits at 48 months. Given that we measured attention to neutral, happy, and fearful expressions, we were also able to further explore the relative significance of attention bias to faces in general (e.g., Bedford et al., 2015) vs. selective attention to fearful facial expressions (Peltola et al., 2015; Rajhans et al., 2016) in the prediction of early social development.

Methods

Participants and Study Design

The current study uses data from two prospective longitudinal studies. Data collection for the first sample (Cohort 1; Leppänen et al., 2010, 2011) began in October 2007, and comprises of a total of 92 (43 female) participants who were enrolled in the study at 7 months of age, and were subsequently invited for follow-up assessments at 24 and 48 months of age. The second study (Cohort 2; Forssman et al., 2014; Leppänen et al., 2015; Peltola et al., 2015; Peltola, Hietanen, Forssman, & Leppänen, 2013; Yrttiaho et al., 2014) was started in April 2012 and consists of 126 (55 female) infants who participated in laboratory assessments at 5, 7, 14, 24, and 48 months of age. The parents of the participants were contacted through child welfare clinics and birth records maintained by the population information system. In the current study we report data from the 7-, 24-, and 48-month assessments (i.e., overlapping longitudinal time points in the two cohorts). In addition to the data presented in the current report, data on brain activity, genetic variants, parent-child interaction, attachment, pupil dilation, and executive/inhibitory functions were also collected during the longitudinal study, but as these data fell outside the scope of this report or were available for a subset of participants only (brain activity, attachment, parent-child interaction, pupil dilation), they were not included in the current analyses.

Data were available from a total of 215 participants (99% of the original sample) for the 7-month assessment (M_{age} = 7.06 months, *SD* = 0.31, range = 23.93-25.51 weeks), 119 (55%) participants for the 24-month assessment (M_{age} = 24.50 months, *SD* = 0.36, range = 23.57-25.51 months), and 163 (75%) participants for the 48-month assessment (M_{age} = 48.84 months, *SD* = 1.52, range = 46.09-57.04 months). The lower retention rate in the 24-month data is explained by a delay in the commencement of this follow-up visit. Other reasons for loss at the follow-up assessments included withdrawal from the study, decline, relocation, and a failure to contact. Children who participated in the follow-up visits did not differ from those who were invited, but did not participate, in measures of attention at 7 months of age (all $p > .10$), suggesting that the current sample is representative of the original cohorts.

All participants' data were used in the current analyses with the exception of 3 participants whose data were excluded from all analyses due to preterm birth ($N = 2$) or a procedural error in eye-tracking assessment ($N = 1$). The infants included in the analyses had no history of visual or neurological abnormalities based on parent report. In addition, data from a varying number of participants were excluded from the final analyses after applying analysis-specific inclusion criteria for each of the sub-analyses. These criteria were set *a priori* based on previous studies. Thus, the final sample sizes in the main analyses linking the 7- and 24-month, and 7- and 48-month data varied from 100 to 138. These analyses were sufficiently powered (>80%) for detecting small to medium (~.26-.30) bivariate correlations at an alpha of .0125 (corrected for four tests). Associations of this magnitude were expected on the basis of previous research linking infant attention to faces with later outcomes (Bedford et al., 2015; Peltola et al., 2015).

Ethical permission for the study was obtained from the Ethical Committee of Pirkanmaa Hospital District, Finland. An informed consent was given by the parents of the participants before the start of the study.

Measures

7 months: attention to faces. Infants were assessed in a quiet and dimly lit room within an area surrounded by walls and curtains. The infant was seated on his/her parent's lap at a ~60-cm viewing distance in front of a 23-inch computer monitor and the equipment used for recording eye movements, which was a video camera in Cohort 1 and a Tobii TX300 eye-tracker (Tobii Technology, Stockholm, Sweden) in Cohort 2.

A paradigm designed to assess infants' attention to face and non-face stimuli, and competing geometric shapes was programmed on E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA) and presented to the infant after the parent and the infant had been comfortably seated and, in Cohort 2, the eye-tracker had been calibrated. The calibration was performed by using the infant calibration procedure within the Tobii Studio software, which proceeded by showing the infant an audiovisual animation sequentially in five locations on the screen. If the first calibration was not successful (i.e., one or more calibrations were missing or were not properly calibrated), the calibration was repeated at least two times to attain satisfactory calibration for all five locations. If one or more calibration points were missing after at least two recalibration attempts, the final calibration outcome was accepted, and the experiment was started.

The stimulus presentation followed the procedure described in previous studies (Forssman et al., 2014; Leppänen et al., 2011; Peltola et al., 2015, 2013). Each test trial started with a dynamic attention-grabbing stimulus presented on the center of the screen. After the infant fixated on the stimulus, as judged by the experimenter monitoring the infant via a video camera, two test stimuli were presented (Figure 1). The first stimulus measuring 15° and 11° vertically and horizontally, respectively, was presented on the center of the screen for 4000 ms. The first stimulus was a picture of a face-shaped pattern which was phase-scrambled to retain the amplitude and color spectra of the original face stimuli (see Leppänen et al., 2011) or a picture of a face displaying neutral, happy, or fearful emotional expression. The second stimulus (15° x 4°) was presented with a 1000-ms onset asynchrony laterally on the left or right side of the screen with 13.6° eccentricity, and remained on the screen for 3000 ms. The second stimulus was a geometric shape (vertically arranged black and white circles or a checkerboard pattern). Trials were presented until the infants had accumulated at least 5 (Cohort 1) or 12 (Cohort 2) trials per condition. Testing was paused if the infant became fussy or tired and terminated if the experimenter (consulting the parent) determined that continuing the testing would have been too distressing for the infant.

(Figure 1 about here)

Infants were presented with face pictures that were validated to signal the intended emotions by a group of adult raters (see Peltola, Leppänen, Palokangas, & Hietanen, 2008, for further details). In Cohort 1, each participant saw pictures of one of two female models (the model was counterbalanced between participants). In Cohort 2, each participant saw pictures of one model during the first half of the experiment and a second model during the second half of the experiment (the order of the two models was counterbalanced between participants).

Timestamps corresponding to the onset times of each trial were stored in E-Prime log files and video records (Cohort 1) or in Tobii gazedata output files (Cohort 2), along with other information of the trial. Data analyses of saccadic eye movements from the central stimulus to the lateral stimulus were implemented offline by using manual coding of video records (Cohort 1, see Leppänen et al., 2011) or by automatic coding of the x & y coordinates from the eye-tracking data (Cohort 2, see Leppänen et al., 2015), and criteria specified in prior studies (Leppänen et al., 2011, 2015). Briefly, trials with a sufficient length of fixation on the central stimulus (i.e., >70 % of the time) during the first second of the trial (Cohort 1) or during the time preceding gaze disengagement or the end of the analysis period (Cohort 2), sufficient number of valid samples in the gaze data (i.e., no gaps >200 ms), and valid information about the eye movement from the central to the lateral stimulus (i.e., the eye movement did not occur during a period of missing gaze data) were retained for analysis. The duration of attention dwell time on the first stimulus (face or non-face pattern) was determined for the period starting 150 ms from the onset of the lateral stimulus and ending 1000 ms after the lateral stimulus onset. The duration was then converted to a normalized dwell time index score by using the following formula:

$$\text{Dwell time index} = \frac{\sum_{i=1}^{n}\left(1 - \frac{1000 - x_i}{850}\right)}{n},$$

where *x* is the time point of the saccadic eye movement (i.e., the last time point when gaze is in the area of the first stimulus preceding a saccade towards the lateral stimulus) and *n* is the number of scorable trials in a given stimulus condition. In this index, the shortest acceptable saccadic eye movement latency (150 ms), results in a score of 0, and the longest possible latency (or a lack of saccade, which is equal to the last measured time point of the first stimulus at 1000 ms) in a score of 1. The dwell time indices were calculated separately for each of the four stimulus conditions (i.e., non-face, neutral, happy, and fearful) and also averaged across the three different face stimulus conditions to provide an attention dwell time index for faces vs. non-face patterns (see *Statistical Analyses*). To be included in the analyses, an infant was required to have a minimum of 3 valid (i.e., artifact-free) trials in each stimulus condition. Applying this criterion, valid dwell time data at 7 months were available from 190 infants. On average, the infants provided 4.8 (SD = 0.4) and 9.0 (SD = 2.6) valid trials per condition in Cohort 1 and 2, respectively, with no differences across stimulus conditions, $F(3, 567) = 1.62$, $p = .18$.

24 months: attention to faces. For the children participating in Cohort 2, the test assessing attention to faces and non-face patterns was repeated at 24 months of age. The distractor stimuli were changed from geometric shapes to dynamic animations to make the test more appropriate for 24 month-old children. Following the same processing steps as for the 7-month eye-tracking data, valid dwell time data at 24 months were available from 78 infants. On average, the children provided 9.0 (*SD* = 2.3) valid trials per condition. A significant difference in the number of valid trials across conditions was detected, $F(3, 231) = 3.69$, $p = .01$, due to a slightly higher number of valid trials to fearful faces ($M = 9.3$) than to the non-face patterns ($M = 8.6$), $p = .02$ (Bonferroni-corrected).

24 months: spontaneous helping. The assessment was conducted in a quiet room with the child and a female experimenter sitting by a small (100 x 60 cm) table, and the parent ~2 meters away from the child (for detailed description, see Salovaara, 2013). The assessment was recorded by 1-2 cameras. After a 10-minute warm-up phase, three age-appropriate tasks assessing spontaneous helping were administered in the following order. In the *clothespin task* (Warneken & Tomasello, 2006), another female experimenter entered the assessment room with three wet towels and asked the first experimenter to hang the towels to a drying rack. The first experimenter agreed and told the child that she would hang them first and then continue playing. When attaching the third towel to the rack with clothespins, she “accidentally” dropped a clothespin on the floor and said (in Finnish) “Oh, I dropped my clothespin”. The experimenter signalled nonverbally that she needed help by attempting to reach the clothespin three times without success and by expressing frustration on her face. Between every attempt, the experimenter had a short break and straightened herself before reaching again for the clothespin. If the child had not helped the experimenter by the third attempt, the experimenter looked at the child and asked for help. If the child did not help, the experimenter picked up the clothespin by herself. In the *cabinet task* (Warneken & Tomasello, 2006), the experimenter noted a stack of folders on a table. She told the child that she would put the folders into a cabinet and then continue playing. The experimenter lifted the stack of folders in her arms and attempted to put them into the cabinet but was unable to open the doors because her hands were full. She paused and said “Oh, the door is closed”. The experimenter again signalled nonverbally that she needed help by walking slowly towards the door as if she was trying to open it and expressed frustration. This was repeated three times, with short breaks between every attempt. If the child did not help on any of the three cues, the experimenter asked the child for help. In case the child did not respond to the request, the experimenter placed the folders on the table and opened the cabinet by herself. Finally, a *broken tractor task* (modified from Kärtner et al., 2010) was administered. In this task, the experimenter introduced different toys one at a time to the child. The third toy was a plastic tractor, which was designed so that one of its front wheels would be easily detached. The experimenter introduced the tractor by telling a short story of its importance to her, and towards the end of the story, the front wheel of the tractor “accidentally” came off. The experimenter held the tractor and the detached wheel in her hands and said (with a sad voice) “Oh no, now the wheel detached. My tractor is broken!” She placed the tractor and the wheel on the table, and for 20 seconds, gazed at the broken tractor and expressed sadness on her face, posture, and occasional weeping sounds. If the child attempted to fix the tractor, the experimenter stopped expressing sadness and thanked the child. If the child did not attempt to fix the tractor in 20 seconds, the experimenter suggested that they would fix it. Regardless of how the child acted, the tractor was fixed and the experimenter and the child continued playing with the tractor and other toys for a while. A score (1) was given if the child picked up the clothespin for the experimenter, opened the cabinet door, or attempted to fix the broken tractor wheel before the experimenter asked the child for help. As data from each individual helping task was not available from some participants, the average (rather than sum) of the individual task scores was used in the final analyses, with a score of 1 indicating helping in all available tasks and 0 indicating absence of helping in any of the tasks (cf. Warneken & Tomasello, 2006). The independence of the scores from the three helping tasks was assessed by Pearson Chi-Square Tests of Independence. In pairwise tests, the null hypotheses for independence of the task scores were rejected, all χ^2 (1) > 6.5, all p < .05. Thus, the scores in the three spontaneous helping tasks were positively associated.

48 months: mentalizing. The assessments were conducted in an observation room very similar to that used in the 24-month assessment. To assess children's mentalizing abilities, the following tasks from Wellman and Liu (2004) were translated into Finnish: *Diverse desires*, *Diverse beliefs*, *Knowledge access*, *Contents false belief*, and *Real-apparent emotion*. The task descriptions are provided in Supplementary Table 1. A score (1) was given if the child answered correctly to questions presented by experimenter, with a maximum sum score of 5 for the mentalizing assessment.

48 months: emotion understanding. To assess emotion understanding, a total of 12 short vignettes depicting happy, fearful, and sad emotions were taken from Wismer Fries and Pollak (2004), translated into Finnish, and modified slightly to make them more appropriate for the 48-month-old children in the present study. Descriptions of the vignettes are provided in Supplementary Table 2. The vignettes were presented in the same random order by the experimenter to each child, and while the experimenter told a story, a picture of a neutral face of an elementary school-aged girl or boy was shown. After the story, the child was presented with four pictures of the same person modeling neutral, happy, sad, or fearful facial expression. The child face stimuli were obtained from the Radboud Faces Database (Langner et al., 2010). The child was asked to indicate how the girl or the boy felt in the story by pointing one of the pictures of facial expressions. The child received a score (1) from every accurately recognized emotion, and the emotion understanding composite score was calculated as the sum of all correctly answered items (0-12 in total).

48 months: callous-unemotional traits. After the laboratory assessment, mothers were asked to complete questionnaires at home. Items reflecting children's CU traits were selected from the Child Behavior Checklist (CBCL) for ages 1½ - 5 (Achenbach & Rescorla, 2000), based on previous studies validating the use of these items of the CBCL as a measure of CU traits in children at this age (Willoughby, Mills-Koonce, Gottfredson, & Wagner, 2014; Willoughby, Waschbusch, Moore, & Propper, 2011). The five items used were 27 (“Doesn’t seem to feel guilty after misbehaving”), 58 (“Punishment doesn’t change behavior”), 67 (“Seems unresponsive to affection”), 70 (“Shows little affection toward people”), and 72 (“Shows too little fear of getting hurt”). Responses to each of the problem behavior descriptions were provided with a scale from 0 (“Not true”) to 2 (“Very true or often true”). CU traits were also assessed at 24 months, but due to the paucity of research investigating CU characteristics in children younger than 3 years (but see Bedford et al., 2015) and available evidence suggesting greater reliability of CU trait assessment during the preschool than toddler age (Hyde et al., 2013; Waller et al., 2017), we only report the 48-month ratings of CU traits.

Statistical Analyses

In the first set of analyses, we estimated measurement error in the dwell time variables, examined correlations between dwell times in the non-face condition and the three face conditions, and assessed the stability of dwell times between 7 and 24 months of age. As many of the dwell time variables deviated from univariate normality, the analyses were conducted by using non-parametric tests (Spearman’s rho and Wilcoxon Signed Rank tests). To estimate measurement error in the dwell time variables at 7 months, we calculated Spearman correlations between odd and even trials for a given condition, including all participants with ≥ 2 valid trials per condition (the criterion for minimum number of acceptable trials was lowered from 3 to 2 for this sub-analysis to avoid excessive loss of participants after each stimulus condition was divided into two “sub-conditions” of odd and even trials). Next, we examined correlations in dwell times between the non-face condition and all three face conditions, and between different face conditions by using Spearman’s rho and compared between-category correlations (i.e., non-face vs. faces) to within-category correlations (i.e., those between neutral, happy, and fearful faces) by using Fisher’s *r*-to-*z* transformation and tests of the difference between two correlations with one variable in common (Lee & Preacher, 2013).

Second, we examined changes in attention dwell times between 7 and 24 months of age. For this analysis, we compared dwell times to non-face stimuli and faces at each age, and also performed paired comparisons of dwell times in each stimulus condition between 7 and 24 months of age with Wilcoxon Signed Rank tests (against a Bonferroni-adjusted alpha of .008). Spearman correlations were used to estimate test-retest stability of dwell times between 7 and 24 months of age within each stimulus condition (against a Bonferroni-adjusted alpha of .0125).

In the third set of analyses, we examined whether attention dwell times at 7 months were associated with the outcomes at 24 and 48 months of age (i.e., average score of spontaneous helping, sum of correct items in the mentalizing task, sum of correct responses in the emotion understanding task, and the sum of mothers' ratings of child CU traits). Given that a general attention bias to faces (Bedford et al., 2015) and a specific bias towards fear (Peltola et al., 2015; Rajhans et al., 2016) have both been linked with social development, we performed two separate sets of analyses linking either *dwell time to faces* or *dwell time to fear* with the four outcomes. For the analyses linking dwell time to faces and the outcomes, we used partial Spearman correlation with dwell time to the non-face stimuli as a control variable, adapting an SPSS syntax available at: http://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/partsp. Similar approach was used to examine associations between dwell times to fear and the outcomes, with the exception that the mean dwell times to neutral and happy faces, instead of dwell times to non-face control stimuli, were used as a control variable.

Given the procedural differences between Cohorts 1 and 2, supplementary analyses were conducted to examine differences between the cohorts in dwell times and outcome variables. A detailed description of these analyses is provided in the supplementary online material. The supplementary analyses showed a difference in the mean level of dwell times between the two cohorts (the dwell time index was .10-.12 higher in Cohort 2), but there was no cohort by stimulus interaction effects on dwell times (i.e., the two cohorts showed an identical pattern of differences in dwell times to non-face control stimuli and faces). The cohorts also differed in the mean levels of spontaneous helping and CU traits, but not mentalizing and emotion understanding. To control for the cohort difference in mean dwell times, the dwell time indices were mean-centered for Cohorts 1 and 2 separately before the original outcome analyses. Additional analyses were also conducted to confirm that adding cohort as a control variable in the partial Spearman correlation analyses did not change the original results concerning the association between dwell times and the outcomes (see supplementary online material).

Results

Attention Dwell Times at 7 Months

Cross-correlations of attention dwell times at 7 months in each stimulus condition, divided to odd and even trials, are shown in Supplementary Table 3. Although the minimum number of trials was set at 2, in the data the number of trials available in the odd and even conditions ranged from 2 to 10 trials, with an average of 4.5 trials included in both types of conditions. Odd-even split-half correlations were .61 in the non-face condition and between .50 and .67 (M = .59) in the face conditions. These values are within the range of those reported in prior infant studies using look-based measures (Gillespie-Smith et al., 2016; Rose, Feldman, & Jankowski, 2012), indicating comparable reliability of the infant attention bias assessment. Dwell times in the non-face condition (M = .37) were positively correlated with dwell times to neutral faces (M = .51, Spearman's rho = .43, *95% CI* = 0.31, 0.54), happy faces (M = .53, Spearman's rho = .51, *95% CI* = 0.40, 0.61), and fearful faces (M = .63, Spearman's rho = .41, *95% CI* = 0.29, 0.52), all p < .001. However, these cross-category correlations were lower than the within-category correlations across the three different face conditions (range = .61-.73, all p < .001). Direct comparisons of the correlation coefficients using the Lee and Preacher (2013) method indicated that apart from one comparison (non-face vs. happy correlation while controlling for dwell time to neutral faces, $p = .10$), all potential cross-category correlations (i.e., those between non-face and face conditions while controlling for the correlation of the unshared variables) were significantly lower than the correlations between different face conditions, all $z > 1.96$, all $p < .05$, suggesting partial independence of dwell times to the face stimuli. This result is illustrated in Figure 2, showing that short dwell time to the non-face stimulus was not uniformly associated with short dwell times to faces. The high correlations of dwell times across the three face conditions indicate a high degree of shared variance in dwell times to different face stimuli.

(Figure 2 about here)

Stability of Dwell Times

At 7 months, dwell times were shortest in the non-face condition, intermediate in the neutral and happy face conditions, and longest in the fearful face condition (Figure 3). The difference between the non-face condition and the combined face condition was significant, $Z = 10.50$, $p < .001$, Cohen's $d = .99$, *95% CI* = 0.16, 0.21. The difference between neutral and happy faces was not significant, $Z = 1.05$, $p = .30$, $d = .10$, *95% CI* = -0.01, 0.05, but dwell times to neutral and happy faces were shorter than dwell times to fearful faces, all $Z > 7.83$, all $p < .001$, all $d > .65$. At 24 months, a significant difference between the non-face condition and the combined face condition was found, $Z = 5.92$, $p < .001$, $d = .82$, *95% CI* = 0.05, 0.09. Comparisons of dwell times to the three face conditions at 24 months showed shorter dwell times to neutral as compared to happy, $Z = 2.74$, $p = .006$, $d = .33$, *95% CI* = 0.01, 0.05, and fearful expressions, $Z = 3.93$, $p < .001$, $d = .44$, *95% CI* = 0.02, 0.06, but there were no differences in dwell times to happy vs. fearful expressions, $Z = 1.38$, $p = .17$, $d = .15$, *95% CI* = -0.01, 0.04. Dwell times shortened between 7 and 24 months of age across all stimulus conditions, although the magnitude of this change was smallest in the non-face condition ($M = 0.11$, $SD = 0.20$, $Z = 3.86$, $p < .001$, $d = .56$, *95% CI* = 0.06, 0.17), intermediate in neutral ($M = 0.21$, $SD = 0.23$, $Z = 5.44$, $p < .001$, $d = .91$, *95% CI* = 0.15, 0.28) and happy face (*M* = 0.19, *SD* = 0.25, *Z* = 4.88, *p* < .001, *d* = .75, *95% CI* = 0.13, 0.27) conditions, and largest in the fearful face condition (*M* = 0.30, *SD* = 0.24, *Z* = 6.05, *p* < .001, *d* = 1.22, *95% CI* = 0.25, 0.37), indicating a marked reduction in the attention bias to fearful faces from 7 to 24 months. The test-retest correlations (Spearman's rho) of dwell times to non-face patterns, neutral faces, happy faces, and fearful faces between 7 and 24 months of age (against a Bonferroni-adjusted alpha of .0125) were .13 (*p* = .31, *95% CI* = -0.12, 0.37), -.22 (*p* = .09, *95% CI* = -0.45, 0.03), -.30 (*p* = .02, *95% CI* = -0.51, -0.06), and -.11 (*p* = .40, *95% CI* = -0.35, 0.14), respectively, thus indicating generally low stability of dwell times from 7 to 24 months.

(Figure 3 about here)

Attention Bias to Faces and Developmental Outcomes

Descriptive data for the outcome measures at 24 and 48 months of age are provided in Table 1. The mean levels of spontaneous helping at 24 months, with approximately 50% frequency of helping behavior, are slightly lower (Warneken & Tomasello, 2006) or similar (Kärtner et al., 2010) to those reported in previous studies. The mean levels of performance in the 48-month behavioral assessments (emotion understanding and mentalizing) correspond to those in previous studies with children at the same age (Wellman & Liu, 2004; Wismer Fries & Pollak, 2004). The levels of CU traits are also consistent with previous studies and with the low base rate of these symptoms in low-risk samples (Willoughby et al., 2014, 2011).

(Table 1 about here)

The correlations between the study variables are depicted in Table 2. The outcome variables at 24 and 48 months were independent with the exception of a positive association between mentalizing and CU traits. The associations between dwell times to faces at 7 months and the outcomes were analyzed with partial Spearman correlations using dwell times to non-face stimuli as a control variable. No associations were observed between dwell times to faces and emotion understanding (Spearman's rho = .05, p = .58) or mentalizing (Spearman's rho = -.14, p = .11) at 48 months. However, dwell times to faces at 7 months were positively correlated with spontaneous helping at 24 months (Spearman's rho = .21, p = .039) and negatively correlated with CU traits at 48 months (Spearman's rho = -.25, p = .006). The correlation with CU traits remained significant at a corrected alpha of .0125. In the second set of analysis examining the predictive significance of attention to fearful expressions (controlling for dwell times to neutral and happy faces), no significant associations with any of the four outcomes were found, Spearman's rho -.11-.08, all p > .22.

Reflecting the low base rate of CU symptoms in the population (e.g., Wagner et al., 2016; Willoughby et al., 2011), the distribution of the CU scores tends to be "zero-inflated" in low-risk samples such as the current sample so that low scores predominate the data. Although correlation coefficients should remain reliable even when used with zero-inflated data (e.g., Huson, 2007), we conducted additional analyses to examine whether the choice of the analysis method affected the pattern of associations between dwell times and CU traits. In these analyses, we fitted a poisson regression model to the data, using the Generalized Linear Models in SPSS. A model with dwell time to faces as an independent variable, dwell time to the non-face pattern and cohort as covariates, and CU score as the response variable showed a significant effect of dwell time to faces on CU scores, Wald Chi-Square = 8.6, df = 1, p = .003. The similarity of this result with the correlational results clearly indicates that the observed association between dwell times and CU scores was not dependent on the choice of the analysis method.

(Table 2 about here)

Discussion

Three main results emerge from this study. First, analyses of the 7-month dwell time data suggest partial independence of attention to non-face patterns and faces. Second, the results suggest that the attention bias to faces may be transiently pronounced at 7 months as i) this bias declines markedly between 7 and 24 months of age and ii) the correlations in dwell times between 7 and 24 months of age were low. Finally, the attention bias to faces in infancy may be importantly associated with the development of responsivity to others' needs and emotional distress: while variability in the attention bias was not related to emotion understanding or mentalizing abilities, relatively increased levels of attention bias to faces at 7 months were associated with a tendency for more frequent helping responses at 24 months and reduced callous-unemotional traits at 48 months of age.

The present results showed partial independence of dwell times to non-face patterns and faces, which may indicate that a single domain-general mechanism contributing to looking times does not sufficiently explain variation in attention biases to faces. One limitation of postulating independence of the mechanisms contributing to dwell time to faces on the basis of the current results is that our non-face condition consisted of a single scrambled image with a face shape and contour. This limitation may affect our results in at least two ways. First, if infants perceived this degraded stimulus as face-like, given its contour, this might have reduced the contrast between the non-face and face conditions in the current study as well as the hypothesized independence of underlying attentional mechanisms in these two conditions. Second, it is not known whether the results of this study extend to different categories of objects, and identifiable non-face objects in particular. Other studies indicate, however, that infants direct their attention with equal likelihood to phase-scrambled non-face patterns (similar to those used in the present study) and various other object categories (Gliga, Elsabbagh, Andravizou, & Johnson, 2009), and that the differences in infants' dwell times across non-face object categories are relatively small compared to the difference between non-face objects and faces (Gluckman & Johnson, 2013). These data suggest that the demonstrated independence of dwell times for non-faces and faces in the current study is a generalizable result, although it will be important to further corroborate this result by applying the present correlational approach to a more extensive array of object stimuli.

The high correlations of dwell times to different facial expressions suggest a common component contributing to attention to faces irrespective of variations in facial expressions. This result could be further interpreted to suggest that aside from a general attention bias towards faces, the current paradigm does not capture independent variations that are specific to particular facial expressions (e.g., fear). A broadly tuned "face bias" may be sufficient to explain the current pattern of results, including differences in dwell times to neutral/happy vs. fearful faces, if assumed that the face bias is more consistently activated by faces that resemble a prototypical face stimulus (e.g., fearful faces with open eyes and mouth) as compared to faces in which these elements are not as salient (Johnson, 2005). While this account is parsimonious and potentially sufficient, we are hesitant to interpret the current results as strong evidence against the possibility of specificity in responses to fearful facial expressions. First, our design did not allow for robust comparison of correlations between vs. within emotion categories as we did not have two subsets of stimuli in each category. Second, previous studies have shown that attention to neutral/happy vs. fearful expressions can be differentially associated with variables describing infants' rearing environment (e.g., parental stress, depression, or sensitivity; Forssman et al., 2014; Taylor-Colls & Fearon, 2015) and that a difference in dwell times to non-fearful and fearful expressions can have independent predictive value in terms of later development (Peltola et al., 2015; see also Rajhans et al, 2016).

The magnitude of the attentional bias to faces reduced markedly between 7 and 24 months of age, although the bias was still evident at 24 months. Comparisons of distinct stimulus categories further showed that the reduction in dwell times was evident for all stimulus conditions but largest for fearful faces. Consequently, the often replicated attention bias to fearful faces during the first year (Ahtola et al., 2014; Forssman et al., 2014; Heck, Hock, White, Jubran, & Bhatt, 2016; Leppänen et al., 2010; Nakagawa & Sukigara, 2012; Peltola, Leppänen, & Hietanen, 2011; Peltola et al., 2008) was absent at 24 months in this sample, with no significant difference in dwell times between happy and fearful faces. The correlations of dwell times between 7 and 24 months of age were also low. It is unlikely that the decline in the attention biases is simply explained by procedural differences between the 7- and 24-month assessments (i.e., change in distractor type) or problems in administering the face-distractor paradigm with 24-month-old children. The biases to faces and fearful expressions in the current paradigm have been observed with various types of distractor stimuli in infants, including dynamic stimuli that are similar to those used with 24-month-old children in the current study (Forssman et al., 2017). Also, among infants who provided longitudinal dwell time data, there were no differences in the number of valid trials between 7 and 24 months, suggesting that the quality of the data did not differ for the two age groups, and the paradigm was not less engaging for the 24-month-olds.

It is possible that the robust attention bias to faces at 7 months reflects processes that are pronounced at this age but decline thereafter. Such transient processes are not uncommon in early development, with well-known examples being infants' "sticky fixation" (Hood, Willen, & Driver, 1998) and broadly tuned phoneme and face discrimination abilities during the first months of life (see Watson, Robbins, & Best, 2014, for a review). Infants also exhibit a pronounced focus on the eye region during the first year (e.g., Peltola, Leppänen, Vogel-Farley, Hietanen, & Nelson, 2009), but there are indications that this tendency may decline and shift toward a more adult-like (e.g., Green, Williams, & Davidson, 2003) distributed scanning pattern over time. Further research is needed to examine whether some of these processes contribute to infants' attention holding on faces (and, particularly, fearful faces) and the potential subsequent reduction of these biases in early childhood. Currently, research on the longitudinal development of attention to faces is very limited. While Nakagawa and Sukigara (2012) observed in a small longitudinal sample that attention to neutral, happy, and fearful faces declined from 24 to 36 months of age, cross-sectional studies investigating attention to neutral, happy, and angry faces in 4- to 24-month-old (Morales et al.,

2017) and 9- to 48-month-old (Burris, Barry-Anwar, & Rivera, 2017) children pointed to a more stable pattern by showing that the patterns of attention biases toward angry and happy faces were not affected by age. An important task for future research is to examine whether attention to different negative emotions (e.g., anger and fear) show different developmental trajectories.

The associations of the 7-month attention bias to faces with 24-month spontaneous helping and 48-month CU traits are consistent with a model suggesting that infants who are more responsive to faces develop to be more responsive to others and less likely to exhibit behaviors that involve disregard of others (cf. Bedford et al., 2015; Dadds et al., 2006; Rajhans et al., 2016). The current results further showed that after controlling for the general attention bias towards faces (i.e., neutral and happy faces), attention to fearful faces did not have additional predictive power in terms of later helping and CU behaviors. This result extends previous studies by providing direct support for the hypothesis that variations in attention to faces (Bedford et al., 2015), but not variations in attention to fearful facial expressions (e.g., Peltola et al., 2015; Rajhans et al., 2016) in infants are primarily related to later CU traits. It remains open whether this result replicates across studies using different paradigms and more diverse samples. Continued investigation of the possible overlap in mechanisms underlying attentional biases to faces and fearful expressions in infants and their prospective associations with CU symptoms is likely to be informative as research on the associations between emotion processing and CU traits in older children and adults have resulted in partially inconsistent results. Indeed, meta-analytic evidence suggests that a general deficit in processing multiple facial expressions (Dawel et al., 2012), instead of a specific deficit in responding to fear (Dadds et al., 2006; Muñoz, 2009) is associated with empathic responsivity and CU traits.

Although the present results provided corroborating evidence for the role of early limited attention to faces in the etiology of deficits in prosociality and empathic responding, it is not known whether limited attentiveness to faces – that may be present from very early on in development (Bedford et al., 2015) – is a factor that *causally* affects the development of unresponsive behavior. One could speculate, for example, that reduced attention to faces may limit infants' opportunities to learn about others and to detect instances when others are in need, ultimately leading to reductions in empathic abilities. To directly address this question, it would be important to examine whether infants' attentiveness to faces can be modified, especially in individuals who show naturally reduced attention to faces, and whether increased attentiveness to faces leads to changes in age-typical social behaviors (cf. Dadds, Cauchi, Wimalaweera, Hawes, & Brennan, 2012; Hyde, Waller, & Burt, 2014). This line of work could further examine whether attentiveness to faces is particularly important during certain "sensitive periods" of development. An alternative possibility for the causal interpretations is that a yet unknown third factor (e.g., genetic or environmentally caused variations in social motivation) modulates the expression of social behaviors in children, including attentiveness to faces early in life and empathy-related traits, thus explaining the observed co-variations among measures of these constructs. From this viewpoint, variations in attention to faces at 7 months may provide a potentially accessible marker of infants' early social development, but they may not necessarily have continuity over time or any causal relation to the development of empathy and its disorders.

Contrary to our hypothesis, variations in the strength of attentional dwell time bias to faces at 7 months were not associated with emotion understanding and mentalizing abilities at 48 months of age. Assuming that the observed variations in attention to faces at 7 months generalize to real-life contexts, our results may suggest that within populations experiencing ample exposure to faces during development (i.e., with a clear majority of infants expressing an attention bias towards faces in the current study), the range of variation in normative attention biases to faces is not sufficient to be associated with the development of the ability to understand emotional situations and match them with facial expressions, or understand others' mental states. It remains possible that these abilities are associated with attention to faces in populations experiencing more profound atypicalities in the amount of exposure to faces due to environmental (e.g., Wismer Fries & Pollak, 2004) or genetic factors (Klin et al., 2015). The lack of significant associations between infant dwell time to faces and emotion understanding, and between emotion understanding and CU traits also suggests that reductions in the ability to recognize others' emotions, as measured here, may not be critical in mediating CU behaviors (cf. Bedford et al., 2015; White et al., 2016). Again, this result should be confirmed in further studies incorporating different measures of emotion recognition (e.g., tasks assessing physiological responsiveness or labeling of less intense facial expressions) and potentially populations with higher risk for CU traits (White et al., 2016).

The divergent associations of infants' face bias with more complex and partially language-dependent abilities (emotion understanding and mentalizing) vs. more reactive tendencies related to responsivity to others (spontaneous helping and CU traits) also suggest that efforts to understand infants' attention bias to faces may benefit from a distinction between cognitive and affective empathy (cf. Dadds et al., 2011). Related to this distinction, previous research has shown intact cognitive perspective-taking abilities (theory of mind) but impairments in affective empathy (responsivity to others' distress) in children with CU symptoms (e.g., Jones, Happé, Gilbert, Burnett, & Viding, 2010). Possibly reflecting a similar dissociation between cognitive and affective empathy, the current results showed a positive correlation between mentalizing and CU traits, although the interpretation of this unexpected result (cf. Song, Waller, Hyde, & Olson, 2016) remains unclear.

Taken together, the current results point to partial independence of the mechanisms subserving attention holding for non-face patterns and faces, and indicate the latter half of the first year as a potentially transient period of increased sensitivity to faces and facial emotions. Importantly, regardless of the possibility that some processes involved in the attention bias to faces may be transient and diminish by two years of age, these processes appear to be meaningfully related to early social development. In particular, our results suggest a nuanced picture where the early attention bias may be specifically related to later-emerging affective empathy, but not more complex emotional and perspective-taking abilities. The results of this project highlight the possibility of uncovering the foundations of core social abilities and their impairments already during infancy, and substantiate the role of infant social information processing as a factor that is correlated with and potentially affects the development of children's ability to successfully interact with others. Limitations to the generalizability of the current results include the focus on a low-risk sample, low base rate of some of the measured variables, and pooling of data from two separate cohorts of children that differed in some of the outcome variables. Although controlling for these factors did not affect the pattern of results in the current study, it will be important to replicate these results in a more heterogeneous sample. Future studies with larger sample sizes should also incorporate factors related to the rearing environment (e.g., parental depression and sensitivity) and infant-related factors (e.g., temperamental reactivity, genetic predispositions) in the analyses to better understand how the variations in infant face and emotion processing emerge and impact later development.

Acknowledgements

We gratefully acknowledge the efforts of the families who participated in the studies and the help of Linda Forssman, Henna Salovaara, Mari Fadjukoff, Mia Huolman, Taru Vihinen, Johanna Sorsa, and Jonna Jokinen in the data collection. This work was supported by grants from the Academy of Finland (#275519 and #307657 to MJP; #218284 to JML) and the European Research Council (#283763 to JML).

References

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA Preschool Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Ahtola, E., Stjerna, S., Yrttiaho, S., Nelson, C. A., Leppänen, J. M., & Vanhatalo, S. (2014). Dynamic eye tracking based metrics for infant gaze patterns in the face-distractor competition paradigm. *PLoS ONE*, 9(5):e97299. https://doi.org/10.1371/journal.pone.0097299

Amso, D., Haas, S., & Markant, J. (2014). An eye tracking investigation of developmental change in bottom-up attention orienting to faces in cluttered natural scenes. *PLoS ONE*, 9(1):e85701. https://doi.org/10.1371/journal.pone.0085701

Bedford, R., Pickles, A., Sharp, H., Wright, N., & Hill, J. (2015). Reduced face preference in infancy: A developmental precursor to callous-unemotional traits? *Biological Psychiatry*, *78*(2), 144–150. https://doi.org/10.1016/j.biopsych.2014.09.022

Bedford, R., Wagner, N. J., Rehder, P. D., Propper, C., Willoughby, M. T., & Mills-Koonce, W. R. (2017). The role of infants’ mother‐directed gaze, maternal sensitivity, and emotion recognition in childhood callous unemotional behaviours. *European Child & Adolescent Psychiatry*, *26*(8), 947–956. https://doi.org/10.1007/s00787-017-0967-1

Burris, J. L., Barry-Anwar, R. A., & Rivera, S. M. (2017). An eye tracking investigation of attentional biases towards affect in young children. *Developmental Psychology, 53*(8), 1418–1427. https://doi.org/10.1037/dev0000345

Chawarska, K., Ye, S., Shic, F., & Chen, L. (2016). Multilevel differences in spontaneous social attention in toddlers with autism spectrum disorder. *Child Development*, *87*(2), 543–557. https://doi.org/10.1111/cdev.12473

Cohen, L. B. (1972). Attention-getting and attention-holding processes of infant visual preferences.

Child Development, *43*(3), 869–79.

Dadds, M. R., Cauchi, A. J., Wimalaweera, S., Hawes, D. J., & Brennan, J. (2012). Outcomes, moderators, and mediators of empathic-emotion recognition training for complex conduct problems in childhood. *Psychiatry Research*, *199*(3), 201–207. https://doi.org/10.1016/j.psychres.2012.04.033

Dadds, M. R., Jambrak, J., Pasalich, D., Hawes, D. J., & Brennan, J. (2011). Impaired attention to the eyes of attachment figures and the developmental origins of psychopathy. *Journal of Child Psychology and Psychiatry*, *52*(3), 238–245. https://doi.org/10.1111/j.1469-7610.2010.02323.x

Dadds, M. R., Perry, Y., Hawes, D. J., Merz, S., Riddell, A. C., Haines, D. J., … Abeygunawardane, A. I. (2006). Attention to the eyes and fear-recognition deficits in child psychopathy. *The British Journal of Psychiatry*, *189*(3), 280–281. https://doi.org/10.1192/bjp.bp.105.018150

Dawel, A., O’Kearney, R., McKone, E., & Palermo, R. (2012). Not just fear and sadness: Meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neuroscience and Biobehavioral Reviews*, *36*(10), 2288–2304. https://doi.org/10.1016/j.neubiorev.2012.08.006

de Haan, M., & Nelson, C. A. (1999). Brain activity differentiates face and object processing in 6-month-old infants. *Developmental Psychology*, *35*(4), 1113–1121. https://doi.org/10.1037/0012-1649.35.4.1113

Forssman, L., Ashorn, P., Ashorn, U., Maleta, K., Matchado, A., Kortekangas, E., & Leppänen, J. M. (2017). Eye-tracking-based assessment of cognitive function in low-resource settings. *Archives of Disease in Childhood*, *102*(4), 311–315. https://doi.org/10.1136/archdischild-2016-311693

Forssman, L., Peltola, M. J., Yrttiaho, S., Puura, K., Mononen, N., Lehtimäki, T., & Leppänen, J. M. (2014). Regulatory variant of the TPH2 gene and early life stress are associated with heightened attention to social signals of fear in infants. *Journal of Child Psychology and Psychiatry*, *55*(7), 793–801. https://doi.org/10.1111/jcpp.12181

Frank, M. C., Amso, D., & Johnson, S. P. (2014). Visual search and attention to faces during early infancy. *Journal of Experimental Child Psychology*, *118*, 13–26. https://doi.org/10.1016/j.jecp.2013.08.012

Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, *110*(2), 160–170. https://doi.org/10.1016/j.cognition.2008.11.010

Gillespie-Smith, K., Boardman, J. P., Murray, I. C., Norman, J. E., O'Hare, A., & Fletcher-Watson, S. (2016). Multiple measures of fixation on social content in infancy: Evidence for a single social cognitive construct? *Infancy*, *21*(2), 241–257. https://doi.org/10.1111/infa.12103

Gliga, T., Elsabbagh, M., Andravizou, A., & Johnson, M. H. (2009). Faces attract infants' attention in complex displays. *Infancy*, *14*(5), 550–562. https://doi.org/10.1080/15250000903144199

Gluckman, M., & Johnson, S. P. (2013). Attentional capture by social stimuli in young infants. *Frontiers in Psychology*, 4:527. https://doi.org/10.3389/fpsyg.2013.00527

Green, M. J., Williams, L. M., & Davidson, D. (2003). In the face of danger: Specific viewing strategies for facial expressions of threat? *Cognition & Emotion*, *17*(5), 779–786. https://doi.org/10.1080/02699930302282

Halit, H., Csibra, G., Volein, A., & Johnson, M. H. (2004). Face-sensitive cortical processing in early infancy. *Journal of Child Psychology and Psychiatry*, *45*(7), 1228–1334. https://doi.org/10.1111/j.1469-7610.2004.00321.x

Heck, A., Hock, A., White, H., Jubran, R., & Bhatt, R. S. (2016). The development of attention to dynamic facial emotions. *Journal of Experimental Child Psychology*, *147*, 100–110. https://doi.org/10.1016/j.jecp.2016.03.005

Hood, B. M., Willen, J., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*(2), 131–134. https://doi.org/10.1111/1467-9280.00024

Huson, L. W. (2007). Performance of some correlation coefficients when applied to zero-clustered data. *Journal of Modern Applied Statistical Methods*, *6*(2), 530–536. https://doi.org/10.22237/jmasm/1193890560

Hyde, L. W., Shaw, D. S., Gardner, F., Cheong, J., Dishion, T. J., & Wilson, M. (2013). Dimensions of callousness in early childhood: Links to problem behavior and family intervention effectiveness. *Development and Psychopathology*, *25*(2), 347–63. https://doi.org/10.1017/S0954579412001101

Hyde, L. W., Waller, R., & Burt, S. A. (2014). Commentary: Improving treatment for youth with callous-unemotional traits through the intersection of basic and applied science - reflections on Dadds et al. (2014). *Journal of Child Psychology and Psychiatry*, *55*(7), 781–783. https://doi.org/10.1111/jcpp.12274

Jessen, S., & Grossmann, T. (2015). The developmental emergence of unconscious fear processing from eyes during infancy. *Journal of Experimental Child Psychology*, *142*, 334–343. https://doi.org/10.1016/j.jecp.2015.09.009

Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience*, *6*(10), 766–774. https://doi.org/10.1038/nrn1766

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6

Jones, A. P., Happé, F. G. E., Gilbert, F., Burnett, S., & Viding, E. (2010). Feeling, caring, knowing: Different types of empathy deficit in boys with psychopathic tendencies and autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *51*(11), 1188–1197. https://doi.org/10.1111/j.1469-7610.2010.02280.x

Klin, A., Shultz, S., & Jones, W. (2015). Social visual engagement in infants and toddlers with autism: Early developmental transitions and a model of pathogenesis. *Neuroscience and Biobehavioral Reviews*, *50*, 189–203. https://doi.org/10.1016/j.neubiorev.2014.10.006

Kwon, M.-K., Setoodehnia, M., Baek, J., Luck, S. J., & Oakes, L. M. (2014). The development of visual search in infancy: Attention to faces versus salience. *Developmental Psychology*, *52*(4), 537–555. https://doi.org/10.1037/dev0000080

Kärtner, J., Keller, H., & Chaudhary, N. (2010). Cognitive and social influences on early prosocial behavior in two sociocultural contexts. *Developmental Psychology*, *46*(4), 905–914. https://doi.org/10.1037/a0019718

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Lee, I. A., & Preacher, K. J. (2013). Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. Available from http://quantpsy.org.

Leppänen, J. M., Forssman, L., Kaatiala, J., Yrttiaho, S., & Wass, S. V. (2015). Widely applicable MATLAB routines for automated analysis of saccadic reaction times. *Behavior Research Methods*, *47*(2), 538–548. https://doi.org/10.3758/s13428-014-0473-z

Leppänen, J. M., & Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions.

Nature Reviews Neuroscience, *10*(1), 37–47. https://doi.org/10.1038/nrn2554

Leppänen, J. M., Peltola, M. J., Mäntymaa, M., Koivuluoma, M., Salminen, A., & Puura, K. (2010). Cardiac and behavioral evidence for emotional influences on attention in 7-month-old infants. *International Journal of Behavioral Development*, *34*(6), 547–553. https://doi.org/10.1177/0165025410365804

Leppänen, J. M., Peltola, M. J., Puura, K., Mäntymaa, M., Mononen, N., & Lehtimäki, T. (2011). Serotonin and early cognitive development: Variation in the tryptophan hydroxylase 2 gene is associated with visual attention in 7-month-old infants. *Journal of Child Psychology and Psychiatry*, *52*(11), 1144–1152. https://doi.org/10.1111/j.1469-7610.2011.02391.x

Marsh, A. A., Kozak, M. N., & Ambady, N. (2007). Accurate identification of fear facial expressions predicts prosocial behavior. *Emotion*, *7*(2), 239–251. https://doi.org/0.1037/1528-3542.7.2.239

Morales, S., Brown, K. M., Taber-Thomas, B. C., LoBue, V., Buss, K. A., & Pérez-Edgar, K. E. (2017). Maternal anxiety predicts attentional bias towards threat in infancy. *Emotion*, *17*(5), 874–883. https://doi.org/10.1037/emo0000275

Muñoz, L. C. (2009). Callous-unemotional traits are related to combined deficits in recognizing afraid faces and body poses. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(5), 554–562. https://doi.org/10.1097/CHI.0b013e31819c2419

Nakagawa, A., & Sukigara, M. (2012). Difficulty in disengaging from threat and temperamental negative affectivity in early life: A longitudinal study of infants aged 12 to 36 months. *Behavioral and Brain Functions*, 8:40. https://doi.org/10.1186/1744-9081-8-40

Peltola, M. J., Forssman, L., Puura, K., van IJzendoorn, M. H., & Leppänen, J. M. (2015). Attention to faces expressing negative emotion at 7 months predicts attachment security at 14 months.

Child Development, *86*(5), 1321–1332. https://doi.org/10.1111/cdev.12380

Peltola, M. J., Hietanen, J. K., Forssman, L., & Leppänen, J. M. (2013). The emergence and stability of the attentional bias to fearful faces in infancy. *Infancy*, *18*(6), 905–926. https://doi.org/10.1111/infa.12013

Peltola, M. J., Leppänen, J. M., & Hietanen, J. K. (2011). Enhanced cardiac and attentional responding to fearful faces in 7-month-old infants. *Psychophysiology*, *48*(9), 1291–1298. https://doi.org/10.1111/j.1469-8986.2011.01188.x

Peltola, M. J., Leppänen, J. M., Mäki, S., & Hietanen, J. K. (2009). Emergence of enhanced attention to fearful faces between 5 and 7 months of age. *Social Cognitive and Affective Neuroscience*, *4*(2), 134–42. https://doi.org/10.1093/scan/nsn046

Peltola, M. J., Leppänen, J. M., Palokangas, T., & Hietanen, J. K. (2008). Fearful faces modulate looking duration and attention disengagement in 7-month-old infants. *Developmental Science*, *11*(1), 60–68. https://doi.org/10.1111/j.1467-7687.2007.00659.x

Peltola, M. J., Leppänen, J. M., Vogel-Farley, V. K., Hietanen, J. K., & Nelson, C. A. (2009). Fearful faces but not fearful eyes alone delay attention disengagement in 7-month-old infants. *Emotion*, *9*(4), 560–565. https://doi.org/10.1037/a0015806

Pollak, S. D., Cicchetti, D., Hornung, K., & Reed, A. (2000). Recognizing emotion in faces: Developmental effects of child abuse and neglect. *Developmental Psychology*, *36*(5), 679–688. https://doi.org/10.1037//0012-1649.36.5.679

Pollak, S. D., & Kistler, D. J. (2002). Early experience is associated with the development of categorical representations for facial expressions of emotion. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(13), 9072–9076. https://doi.org/10.1073/pnas.142165999

Rajhans, P., Altvater-Mackensen, N., Vaish, A., & Grossmann, T. (2016). Children's altruistic behavior in context: The role of emotional responsiveness and culture. *Scientific Reports*, 6:24089. https://doi.org/10.1038/srep24089

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2012). Implications of infant cognition for executive functions at age 11. *Psychological Science*, *23*(11), 1345–1355. https://doi.org/10.1177/0956797612444902

Salovaara, H. (2013). *Allelic variation in the oxytocin receptor gene and early-emerging social behaviors in boys and girls*. Unpublished Master's thesis, University of Tampere. Retrieved from http://urn.fi/URN:NBN:fi:uta-201401161030

Song, J. H., Waller, R., Hyde, L. W., & Olson, S. L. (2016). Early callous-unemotional behavior, Theory-of-Mind, and a fearful/inhibited temperament predict externalizing problems in middle and late childhood. *Journal of Abnormal Child Psychology*, *44*(6), 1205–1215. https://doi.org/10.1007/s10802-015-0099-3

Székely, E., Lucassen, N., Tiemeier, H., Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., Kok, R., … Herba, C. M. (2014). Maternal depressive symptoms and sensitivity are related to young children's facial expression recognition: The Generation R Study. *Development and Psychopathology*, *26*(2), 333–345. https://doi.org/10.1017/S0954579413001028

Wagner, N. J., Mills-Koonce, W. R., Propper, C. B., Willoughby, M. T., Rehder, P. D., Moore, G. A., & Cox, M. J. (2016). Associations between infant behaviors during the face-to-face still-face paradigm and oppositional defiant and callous-unemotional behaviors in early childhood. *Journal of Abnormal Child Psychology*, *44*(8), 1439–1453. https://doi.org/10.1007/s10802-016-0141-0

Waller, R., Shaw, D. S., Neiderhiser, J. M., Ganiban, J. M., Natsuaki, M. N., Reiss, D., … Hyde, L. W. (2017). Toward an understanding of the role of the environment in the development of early callous behavior. *Journal of Personality*, *85*(1), 90–103. https://doi.org/10.1111/jopy.12221

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301–1303. https://doi.org/10.1126/science.1121448

Watson, T. L., Robbins, R. A., & Best, C. T. (2014). Infant perceptual development for faces and spoken words: An integrated approach. *Developmental Psychobiology*, *56*(7), 1454–1481. https://doi.org/10.1002/dev.21243

Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind tasks. *Child Development*, *75*(2), 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

White, S. F., Briggs-gowan, M. J., Voss, J. L., Petitclerc, A., Mccarthy, K., Blair, R. J. R., … Wakschlag, L. S. (2016). Can the fear recognition deficits associated with callous-unemotional traits be identified in early childhood ? *Journal of Clinical and Experimental Neuropsychology*, *38*(6), 672–684. https://doi.org/10.1080/13803395.2016.1149154

Willoughby, M. T., Mills-Koonce, W. R., Gottfredson, N. C., & Wagner, N. J. (2014). Measuring callous unemotional behaviors in early childhood: Factor structure and the prediction of stable aggression in middle childhood. *Journal of Psychopathology and Behavioral Assessment*, *36*(1), 30–42. https://doi.org/10.1007/s10862-013-9379-9

Willoughby, M. T., Waschbusch, D. A., Moore, G. A., & Propper, C. B. (2011). Using the ASEBA to screen for callous unemotional traits in early childhood: Factor structure, temporal stability, and utility. *Journal of Psychopathology and Behavioral Assessment*, *33*(1), 19–30. https://doi.org/10.1007/s10862-010-9195-4

Wilmer, J. B. (2008). How to use individual differences to isolate functional organization, biology, and utility of visual functions; with illustrative proposals for stereopsis. *Spatial Vision*, *21*(6), 561–579. https://doi.org/10.1163/156856808786451408

Wismer Fries, A. B., & Pollak, S. D. (2004). Emotion understanding in postinstitutionalized Eastern European children. *Development and Psychopathology*, *16*(2), 355–369. https://doi.org/10.1017/S0954579404044554

Yrttiaho, S., Forssman, L., Kaatiala, J., & Leppänen, J. M. (2014). Developmental precursors of social brain networks: The emergence of attentional and cortical sensitivity to facial expressions in 5 to 7 months old infants. *PLoS ONE*, 9(6):e100811. https://doi.org/10.1371/journal.pone.0100811

Table 1. Descriptive statistics for the outcome measures at 24 and 48 months.

	N	Min	Max	Mean	*SD*
24-Month Spontaneous Helping	100	0.00	1.00	0.45	0.36
48-Month Emotion Understanding	138	1.00	11.00	6.75	2.17
48-Month Mentalizing	134	0.00	5.00	2.80	1.10
48-Month Callous-Unemotional Traits	118	0.00	5.00	0.85	1.11

Table 2. Spearman correlations between dwell times to faces (all faces or only fearful faces) at 7 months and the outcome measures at 24 and 48 months. In the partial Spearman correlations between dwell times and the outcomes, dwell time to the non-face stimulus (when using the combined dwell time to faces variable as the predictor) or mean dwell time to neutral and happy faces (when using dwell time to fearful faces as the predictor) is included as the control variable. The numbers in parentheses below the correlation coefficients display the 95% confidence interval.

	1.	2.	3.	4.	5.
1. Dwell Time (Faces)					
2. Dwell Time (Fear)	.87***				
	[0.83, 0.90]				
3. Helping	.21*	.08			
	[0.02, 0.39]	[-0.12, 0.27]			
4. Emotion Understanding	.05	.03	.03		
	[-0.12, 0.22]	[-0.14, 0.20]	[-0.18, 0.23]		
5. Mentalizing	-.14	.02	-.21	.11	
	[-0.30, 0.03]	[-0.15, 0.19]	[-0.40, -0.002]	[-0.06, 0.27]	
6. CU Traits	-.25**	-.11	-.19	-.14	.29**
	[-0.41, -0.07]	[-0.29, 0.07]	[-0.40, 0.04]	[-0.32, 0.05]	[0.11, 0.45]

Note: *p < .05, **p < .01, ***p < .001

Figure Captions

Figure 1. Dwell times were measured to non-face control stimuli and faces displaying neutral, happy, or fearful expressions while a competing stimulus (“distractor”) was presented to the left or right visual field.

Figure 2. A scatterplot showing individual participants’ dwell time to the non-face control stimuli (x-axis) and faces (y-axis). Most participants exhibited relatively longer dwell times to faces (i.e., values above the red diagonal line).

Figure 3. Dwell times to the non-face patterns and faces at 7 and 24 months. Error bars represent the standard error of mean.

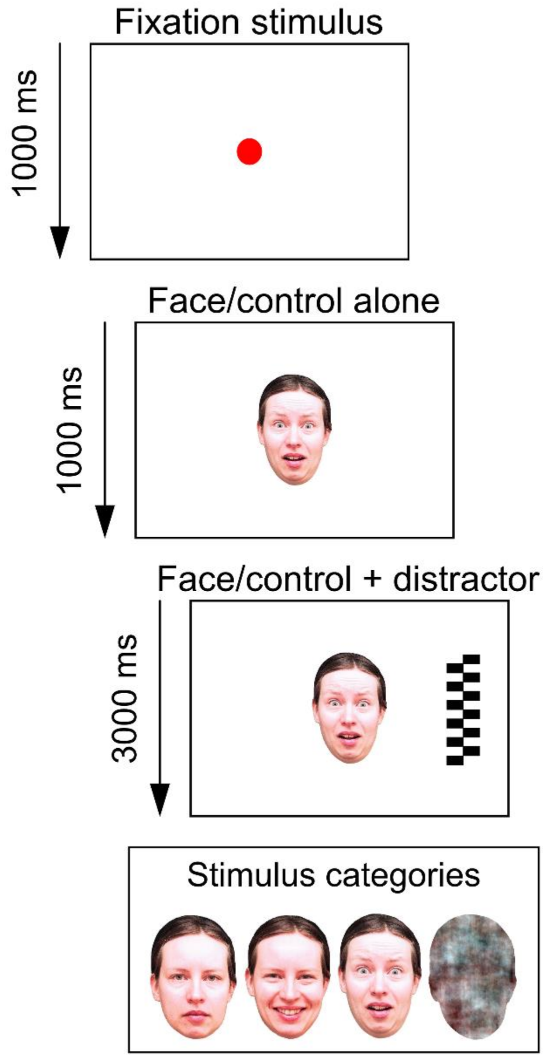

Figure 1. Dwell times were measured to non-face control stimuli and faces displaying neutral, happy, or fearful expressions while a competing stimulus ("distractor") was presented to the left or right visual field.

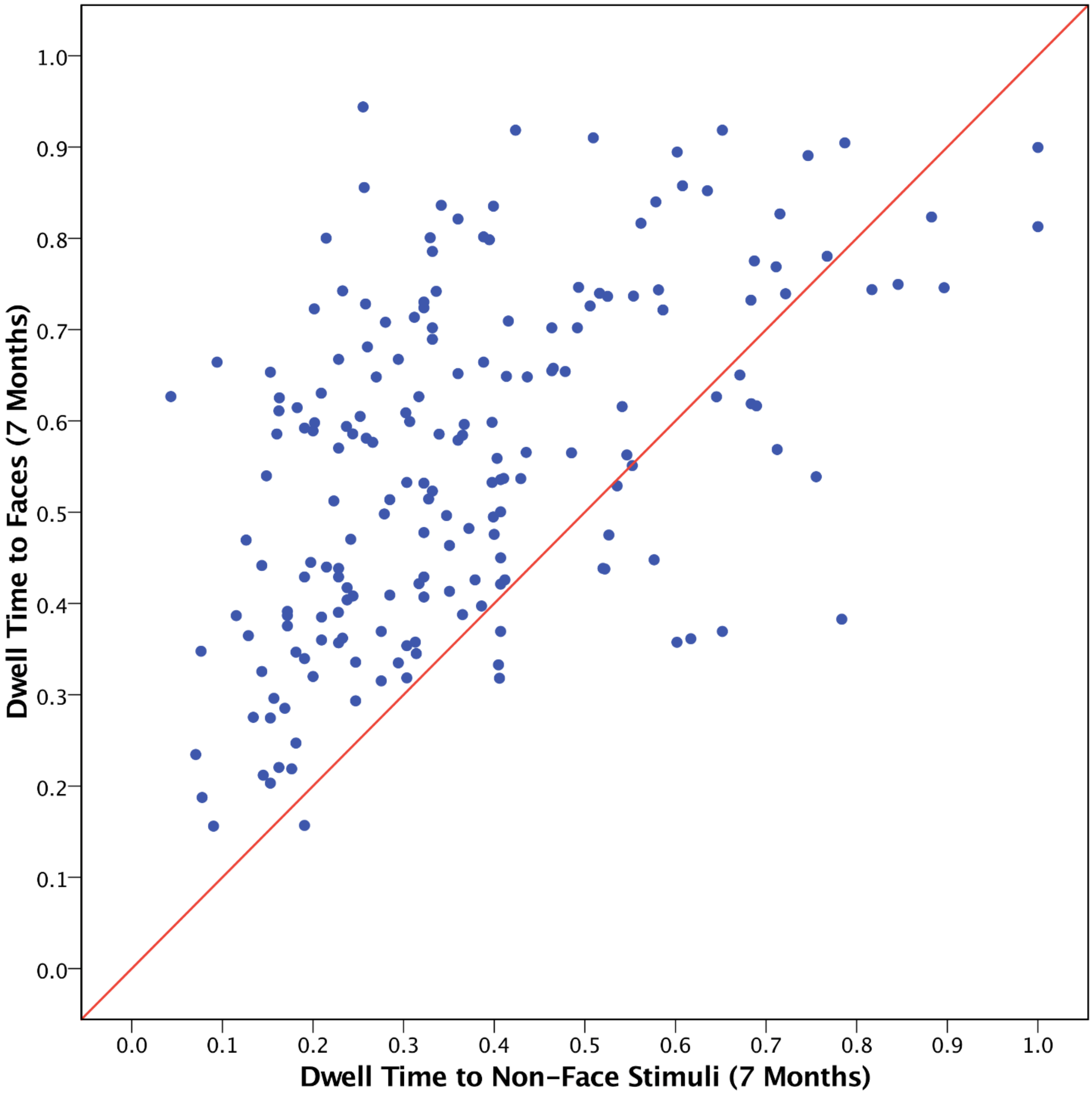

Figure 2. A scatterplot showing individual participants' dwell time to the non-face control stimuli (x-axis) and faces (y-axis). Most participants exhibited relatively longer dwell times to faces (i.e., values above the red diagonal line).

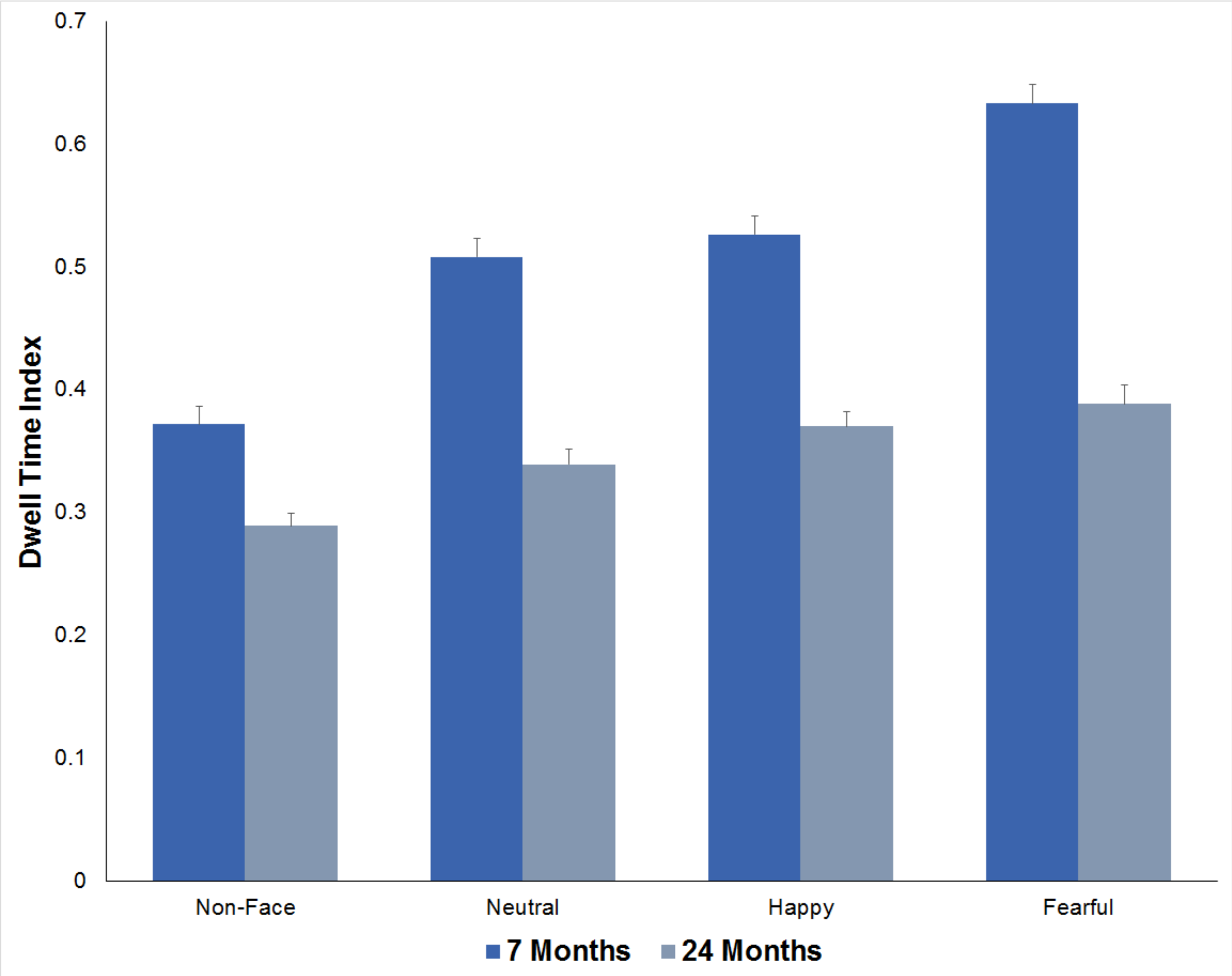

Figure 3. Dwell times to the non-face patterns and faces at 7 and 24 months. Error bars represent the standard error of mean.

Supporting online information for:

Infants' attention bias to faces as an early marker of social development

Mikko J. Peltola, Santeri Yrttiaho, & Jukka M. Leppänen

Supplementary Table 1. Descriptions of the tasks used in the 48-month mentalizing assessment. The tasks were adapted from the Theory of Mind scale by Wellman and Liu (2004) and translated into Finnish (Salovaara, 2013). In tasks 3 and 4, the child was required to answer correctly to the control question in order to receive a score from that task.

Task	Description of the task (equipment used)	Questions to the child
1. Diverse desires	The doll wants to have a different snack than what the child wants (a doll, a colorful picture of a cookie, and a carrot).	"Which snack will Jaakko (the doll) choose? A carrot or a cookie?"
2. Diverse beliefs	The doll and the child have different beliefs about where the doll's cat is hiding (a doll, a black-and-white picture of a bush, and a garage).	"So where will Linda (the doll) look for her cat? In the bushes or in the garage?"
3. Knowledge access	The child knows what is in a box but the doll has never seen inside that box (a doll and a small toy dog inside a wooden box).	"Does Maija (the doll) know what is in the box?" Control question: "Has Maija seen what is in the box?"
4. Contents false belief	The child knows that inside a bag of chewing gum is a toy horse instead of chewing gum, but the doll does not know that (a doll, a small toy horse inside a bag of chewing gum).	"What does Pekka (the doll) think is in the bag? Chewing gum or a horse?" Control question: "Has Pekka seen what is in the bag?"
5. Real-apparent emotion	The child is told a story in which a girl is trying to hide her real emotion by expressing another emotion (a paper doll and three black-and-white schematic face pictures).	"How did the girl feel? How did she try to look on her face?" (answer to the former must be more negative, e.g., sad and happy)

Supplementary Table 2. Vignettes used in the emotion understanding task and the target emotion in each vignette. The vignettes were modified from Wismer Fries & Pollak (2004) and translated into Finnish (Salovaara, 2013) .

Vignettes about a girl (emotion)	Vignettes about a boy (emotion)
1. "Once this girl participated in a running competition. She won the competition, and her friends were cheering for her at the finish line" (happy)	1. "This boy was playing outside with his friends. The boy fell down on the sidewalk and hurt his knee" (sad)
2. "Once this girl and her mom planned a trip to their favorite park on Saturday. But when Saturday came, it was raining so they could not go to the park" (sad)	2. "This boy woke up in the middle of the night and noticed a big thunder and lightning storm outside" (fear)
3. "Once this girl had a bad dream about a monster" (fear)	3. "This boy had a pet bird. One day he got home from school and saw that the bird was not in its cage. The boy thought that his bird might be gone forever" (sad)
4. "This girl's best friend, who she really likes to play with, moved away. Now the girl cannot play with her friend anymore" (sad)	4. "Once this boy drew a picture and showed it to her mom. Mom said that the boy did a good job and that the picture was fantastic" (happy)
5. "This girl and her friend were walking through a forest. They heard rustle coming from the bushes and thought it might be a bear" (fear)	5. "This boy went shopping with his dad. There were lots of people in the store, and the boy got lost and could not find his dad anywhere" (fear)
6. "This girl loves dogs. On her birthday her dad gave her a dog" (happy)	6. "Once this boy and his mom had a picnic together at the boy's favorite place" (happy)

Supplementary Table 3. Spearman correlations of dwell times between odd and even trials within stimulus conditions and across stimulus conditions. The critical within-stimulus odd-even correlations are shown in bold. All correlations except the one marked as “ns” are statistically significant at $p \leq .01$. The numbers in parentheses below the correlation coefficients display the 95% confidence interval.

	C_{Odd}	C_{Even}	N_{Odd}	N_{Even}	H_{Odd}	H_{Even}	F_{Odd}
C_{Even}	**.61** [0.46, 0.73]						
N_{Odd}	.29 [0.09, 0.47]	.40 [0.21, 0.56]					
N_{Even}	$.19^{ns}$ [-0.02, 0.38]	.47 [0.29, 0.62]	**.50** [0.33, 0.64]				
H_{Odd}	.43 [0.24, 0.59]	.51 [0.34, 0.65]	.60 [0.45, 0.72]	.50 [0.33, 0.64]			
H_{Even}	.38 [0.19, 0.55]	.45 [0.27, 0.60]	.56 [0.40, 0.69]	.49 [0.31, 0.63]	**.67** [0.54, 0.77]		
F_{Odd}	.27 [0.07, 0.45]	.40 [0.21, 0.56]	.46 [0.28, 0.61]	.56 [0.40, 0.69]	.59 [0.44, 0.71]	.62 [0.47, 0.73]	
F_{Even}	.37 [0.18, 0.54]	.45 [0.27, 0.60]	.50 [0.33, 0.64]	.45 [0.27, 0.60]	.61 [0.46, 0.73]	.66 [0.52, 0.76]	**.59** [0.44, 0.71]

Note: C = non-face control, N = neutral, H = happy, F = fear

Supplementary Analyses.

Differences between Cohorts 1 and 2

Cohorts 1 and 2 differed in the number of trials, number of models (infants saw either 1 of 2 models in Cohort 1 and both of the two models in Cohort 2), and data coding method (manual/video-based vs. automated/eye-tracking based). In addition, Cohort 1 participants were enrolled in the study at 7 months, whereas Cohort 2 participants were enrolled at 5 months and tested at 5 and 7 months as we explained in the manuscript. The criteria used for extracting dwell times were very similar for the two cohorts, but some criteria for classifying trials as valid were different for the manual and eye-tracking based coding given the differences in these techniques (previous analyses have, however, shown near-perfect (>97%) concordance between manually and automatically processed data; Leppänen et al., 2015). The formula for calculating normalized dwell times was identical for the two cohorts.

A 2 (Cohort) x 4 (Stimulus Condition) analysis of variance (ANOVA) on dwell times showed a significant effect of Cohort, $F(1, 188) = 22.4$, $p < .001$, and Stimulus Condition, $F(3, 564) = 111.7$, $p < .001$, on dwell times, but there was no Cohort x Stimulus Condition interaction, $F(3, 564) = 0.38$, $p = 768$. There is no clear explanation for the difference in the mean level of dwell times between the two cohorts (the dwell time index was .10-.12 higher in Cohort 2). Our further analyses showed, however, that this difference was not caused by the differences in the number of trials, models, and coding methods. An analysis using all original trials from Cohort 1 and only the first 24 manually coded trials from Cohort 2 (when infant saw only one of the two models)[1] showed a similar .10-.12 mean level difference in dwell times across all conditions:

Cohort 1: non-face ($M = 0.32$), neutral ($M = 0.45$), happy ($M = 0.47$), fearful ($M = 0.57$). All differences significant ($p < .001$) except neutral vs. happy.
Cohort 2: non-face ($M = 0.42$), neutral ($M = 0.56$), happy ($M = 0.58$), fearful ($M = 0.69$). All differences significant ($p < .001$) except neutral vs. happy.

Critically, neither of these two analyses show evidence for a Cohort x Stimulus Condition interaction, suggesting that the difference in dwell time between non-face and face stimuli (i.e., the bias for faces) was similar across the two cohorts. Given that the cohort difference was a shift in the location of dwell times that did not have an effect on the pattern of results across conditions (i.e., bias for faces) or distribution of the scores[2], dwell times were mean-centered to remove any effect of cohort on association analyses.

Regarding the outcome variables, the cohorts did not differ on mentalizing and emotion understanding, $p > .10$, but did differ on mean levels of spontaneous helping, $Z = 5.2$, $p < .001$, and

---

[1] Manually coded data for the first 24 trials were available for Cohort 2 from an earlier study (Leppänen et al., 2015) that was conducted to compare video- and eye-tracking based extraction of dwell times, and validate the automated eye-tracking based coding method.
[2] A two-sample Kolmogorov-Smirnov test showed no differences between cohorts after mean-centering in dwell times for non-face stimuli or faces, indicating that, while the location of the dwell-time distributions differed for the two cohort, the distribution of individual dwell times within the cohort was similar.

CU traits, $Z = 2.8$, $p < .01$, with Cohort 1 on average showing less helping and higher CU traits than Cohort 2.

Cohort as a control variable

Given the cohort difference in dwell times (Cohort 1 < Cohort 2, across stimulus conditions), helping (Cohort 1 < Cohort 2), and CU traits (Cohort 1 > Cohort 2), the possibility raises that the cohort effect confounds the association analyses by artificially inflating correlations. Adding cohort as a control variable in the partial Spearman correlation analyses did not, however, change the original results concerning the association between dwell times and the outcomes. Dwell time to faces was significantly correlated with spontaneous helping, Spearman's rho = .22, $p = .03$. There were no correlations between dwell times and mentalizing or emotion understanding, $p > .10$, but dwell times to faces were negatively correlated with callous-unemotional traits, Spearman's rho = .25, $p = .007$.

References

Leppänen, J. M., Forssman, L., Kaatiala, J., Yrttiaho, S., & Wass, S. (2015). Widely applicable MATLAB routines for automated analysis of saccadic reaction times. *Behavior Research Methods, 47*, 538-548. doi:10.3758/s13428-014-0473-z

Salovaara, H. (2013). *Allelic variation in the oxytocin receptor gene and early-emerging social behaviors in boys and girls*. Unpublished Master's thesis, University of Tampere. Retrieved from http://urn.fi/URN:NBN:fi:uta-201401161030

Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind tasks. *Child Development*, *75*(2), 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Wismer Fries, A. B., & Pollak, S. D. (2004). Emotion understanding in postinstitutionalized Eastern European children. *Development and Psychopathology*, *16*(2), 355–369. https://doi.org/10.1017/S0954579404044554