Social Neuroscience

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/psns20

Neural sensitivity to trustworthiness cues from realistic face images is associated with temperament: An electrophysiological study with 6-month-old infants

Elisa Baccolo, Ermanno Quadrelli & Viola Macchi Cassia

SOCIAL NEUROSCIENCE
https://doi.org/10.1080/17470919.2021.1976271

Neural sensitivity to trustworthiness cues from realistic face images is associated with temperament: An electrophysiological study with 6-month-old infants

Elisa Baccolo[a], Ermanno Quadrelli[a,b] and Viola Macchi Cassia[a,b]

[a]Department of Psychology, Università Degli Studi Di Milano-Bicocca, Milano, Italy; [b]Nero Mi, Milan Center for Neuroscience, Milano, Italy

ABSTRACT

Discriminating facial cues to trustworthiness is a fundamental social skill whose developmental origins are still debated. Prior investigations used computer-generated faces, which might fail to reflect infants' face processing expertise. Here, Event-Related Potentials (ERPs) were recorded in Caucasian adults (N = 20, 7 males, *M* age = 25.25 years) and 6-month-old infants (N = 21, 10 males) in response to variations in trustworthiness intensity expressed by morphed images of realistic female faces associated with explicit trustworthiness judgments (Study 1). Preferential looking behavior in response to the same faces was also investigated in infants (N = 27, 11 males) (Study 2). ERP results showed that both age groups distinguished subtle stimulus differences, and that interindividual variability in neural sensitivity to these differences were associated with infants' temperament. No signs of stimulus differentiation emerged from infants' looking behavior. These findings contribute to the understanding of the developmental origins of human sensitivity to social cues from faces by extending prior evidence to more ecological stimuli and by unraveling the mediating role of temperament.

ARTICLE HISTORY
Received 1 February 2021
Revised 15 July 2021
Published online 11 September 2021

KEYWORDS
Faces; trustworthiness; infancy; event-related potentials; looking preference; temperament

As humans, we are hypersensitive to those facial properties that convey social signals. For example, we spontaneously infer other people's mental states by recognizing affectively relevant facial configurations, like those articulating emotional expressions. However, the universal facial expressions of emotions only represent one type of cues that we extract from faces. Our everyday social interactions are supported by the decoding of a variety of voluntarily and involuntarily produced facial information, including face morphologies that we spontaneously associate with certain social traits. Trustworthiness is one of such traits, defined as the perception of other people's approachability (Willis & Todorov, 2006). Trustworthiness judgments are triggered by specific facial configurations, often defined as trustworthiness Action Units (AUs, see Jack & Schyns, 2015). These include upward/downturned eyebrows, upward/downturned curving mouth, and a wrinkling nose. Recent evidence suggests that the simultaneous movement of the brow raiser muscle, the lip corner, and the nose wrinkler form a unique configural set that vehiculates the perception of approach/avoidance tendencies and differs from any of the sets associated with the six universal emotional expressions (Jack & Schyns, 2015).

Trustworthiness perception in the adult population

Trustworthiness judgments from faces occur very rapidly, automatically, and with high consensus (Bar et al., 2006; Willis & Todorov, 2006). Electrophysiological studies measuring event-related potentials (ERPs) in response to faces varying in the level of expressed trustworthiness show that untrustworthy faces elicit enhanced early visual and perceptual responses at the level of the C1 and the N170 components, as well as enhanced attentional responses, as indexed by modulations of the early posterior negativity (EPN) and the late positive potential (LPP) (e.g., Dzhelyova et al., 2012; Lischke et al., 2018; Marzi et al., 2014; Yang et al., 2011). This evidence supports the adaptive significance of humans' sensitivity to facial cues to trustworthiness as, from an evolutionary perspective, a prompt and heightened attentional response to untrustworthy faces implies higher chances to avoid potential harm (Zebrowitz et al., 2003).

CONTACT Elisa Baccolo e.baccolo@campus.unimib.it Department of Psychology, Università Degli Studi Di Milano-Bicocca, Via Alcide De Gasperi, 24, Inzago, MI 20065, Italy
(for the third author, Macchi Cassia is the double family name) All authors' country of residence is Italy.
Supplemental data for this article can be accessed here.

Developmental origins of sensitivity to trustworthiness cues

If a large part of the literature has focused on disentangling the cognitive and neural processes underlying adults' sensitivity to fine-grained differences in facial cues to trustworthiness, only a small number of studies has investigated the fundamental question of the developmental origins of this sensitivity (see Over & Cook, 2018).

Evidence from children suggests that the ability to use facial information to attribute personality traits might not require extended social experience, as 3-year-old children can judge how "mean" or "nice" a person looks based on its facial appearance. By the age of 6 years these explicit face-trait judgments acquire the same level of consistency as the one produced by adults (Cogsdill et al., 2014; see also Ewing et al., 2015). More recently, Baccolo and Macchi Cassia (2020) explored children's sensitivity to perceptual differences among faces varying in the level of expressed trustworthiness. Results showed that, already at the age of 5 years, children represent perceived differences between faces as a function of the level of trustworthiness they express, just like adults, and this representation becomes more fine-grained with increasing age. Moreover, at 5 years more accurate judgments of trustworthiness were associated with more advanced emotion understanding abilities, suggesting that, in younger children, the ability to use face information to infer trustworthiness traits builds on the ability to consistently use transient facial cues to infer internal emotional states.

Indeed, the ability to discriminate emotional facial expressions (see review by Grossmann, 2010) and regulate behavior accordingly (e.g., Serrano et al., 1995) appears rather early in development, and becomes more fine grained during the first year of life. Even newborns distinguish happy expressions from fearful and disgusted ones (Farroni et al., 2007; Lischke et al., 2018), and by 3 to 4 months infants can reliably discriminate happy expressions from a number of other emotional expressions, while by the age of 7 months they begin to exhibit adult-like attentional preference for fearful expressions (see review by Leppänen et al., 2007). Infants' behavior also tends to vary according to the affective meaning of facial expressions, as 4-to 9-month-old infants prefer to approach individuals posing happy facial expressions and avoid individuals exhibiting negative facial expressions (Serrano et al., 1995).

Preverbal infants have also been shown to modulate their avoidance/approach behavior based on other people's conduct toward others: 5- and 9-month-olds tend to look longer and preferentially reach animated characters that help social targets over those who hinder or cause harm to the same targets (e.g., Hamlin & Wynn, 2011). Gredebäck et al. (2015) found that the early preference for prosocial others is also reflected in infants' electrocortical activity. The P400 Event Related Potential (ERP) component differed when 6-month-infants observed agents that previously hindered others and agents that previously helped others, suggesting that by 6 months of age infants are more readily attuned to process positive compared to negative-valenced social actions.

Sensitivity to facial cues to trustworthiness in infants

In light of this evidence, the question of whether infants might be sensitive to those facial cues that older children and adults perceive as signaling approach or avoidance appears compelling. To our knowledge, only four studies have explored this question by testing infants' behavioral and/or neural discrimination of computer-generated faces varying in the level of expressed trustworthiness. In all studies, infants were presented with face stimuli selected from the extensively validated database of computer-generated faces varying along a trustworthiness continuum created by Todorov (Oosterhof & Todorov, 2008) on the basis of data-driven, computational models. The stimulus set typically included three versions of a single male Caucasian identity showing three different levels of trustworthiness: a high trustworthy and a low trustworthy version taken from the continuum extremes (3 SD above and below the average face) and a neutral version corresponding to the average face from the continuum.

Jessen and Grossmann (2016, 2019) reported that 7-month-old infants showed a linear increase in looking time from untrustworthy to neutral and trustworthy faces when the three faces were presented pairwise. A similar finding was obtained by Sakuta and colleagues (Sakuta et al., 2018) with Japanese infants, who preferred trustworthy over untrustworthy Caucasian faces in a preferential looking task, at least when both faces were also high in dominance. These behavioral results are coherent with evidence showing that infants preferentially approach people displaying happy facial expressions and avoid those who display anger or negative expressions (Serrano et al., 1995). Moreover, in light of the fact that trustworthy-looking faces share at least some subtle features with happy faces (e.g., Oosterhof & Todorov, 2009), it has been proposed (see Jessen & Grossmann, 2019) that infants may view trustworthy faces as more familiar than neutral and untrustworthy faces because, under typical rearing conditions, they are predominantly exposed to social agents who pose positive facial expressions while acting prosocially toward them.

According to the developmental account of the origins of spontaneous face-trait inferences (Trait Inference Model) proposed by Lischke et al. (2018), this repeated experience would lead to the establishment of a predictive contingent relationship between face representations and representation of personality traits in long term memory. Partial support for this hypothesis comes from an ERP study showing that facial trustworthiness impacts infants' object processing in the context of a gaze cueing paradigm in which the frontocentral Nc is larger to objects that are attended to by trustworthy faces compared to untrustworthy faces (Jessen & Grossmann, 2019).

These findings imply that the infants' brain responds to variations in facial trustworthiness. However, a more direct exploration of infants' neural sensitivity to facial cues to trustworthiness comes from two ERP studies conducted by Jessen and Grossmann (2016, 2017) with both supraliminally and subliminally presented stimuli. Under standard stimulus presentation (Jessen & Grossmann, 2016) infants showed an enhanced amplitude for the P400 elicited at occipital sites and for the Nc at frontal and central sites in response to neutral faces as opposed to either highly trustworthy or highly untrustworthy faces. When faces were subliminally presented (Jessen & Grossmann, 2019), infants showed an enhanced NSW at frontal and central sites in response to untrustworthy faces as opposed to neutral faces.

Albeit confirming that the infants' brain differentiates between different levels of facial trustworthiness, these findings are partially at odds with behavioral evidence of infants' longer looking times to trustworthy faces (Jessen & Grossmann, 2016, 2017; Sakuta et al., 2018). Indeed, infant Nc component is linked to the allocation of attention to the presented stimuli (Reynolds & Richards, 2005), and indexes familiarity in face processing in the first year of life, in that it is larger for familiar faces as compared to unfamiliar faces (e.g., De Haan & Nelson, 1997). Although less consistent, face familiarity effects have been reported also for the P400 (e.g., Balas et al., 2010). Therefore, although the interpretation that infants are responding to familiarity applies smoothly to the finding of infants' longer looking times to trustworthy over untrustworthy faces, it fails to explain ERP data indicating lack of differentiation between high-trustworthy and high-untrustworthy faces at the level of the P400, Nc and NSW.

The current study

In light of the limited number of studies and the uneven results emerging from behavioral and ERP investigations, our goal in the current study was to replicate and extend earlier demonstrations of infants' sensitivity to variations in facial trustworthiness using a different, more ecological, set of stimuli. As already mentioned, all prior studies with infants (and children, but see Crookes et al., 2015) used artificial, computer-generated Caucasian male faces obtained from data-driven modeling (Oosterhof & Todorov, 2008).

Although artificial faces allow for a strictly controlled manipulation of the features of interest, they may not fully reflect infants' expertise at face processing, including perceptual discrimination (see Crookes et al., 2015). Moreover, it is known that face processing abilities during the first year of life tune in response to salient people in the infant's environment (e.g., Stets et al., 2012), resulting in increased sensitivity to faces that match the characteristics of their primary caregiver (e.g., Ramsey-Rennels & Langlois, 2006), which is typically female (Rennels & Davis, 2008; Sugden et al., 2014). Therefore, using computer-generated male faces as stimuli may result in underestimation of infants' sensitivity to subtle variations in physical cues to trustworthiness.

With the aim to overcome these limitations, in the current study we used as stimulus material three variations of one real female face identity selected from a previously validated set of seven parametrically manipulated variations, differing in the level of perceived trustworthiness (see Baccolo & Macchi Cassia, 2019). As in Jessen and Grossmann (2016), we presented our participants with the most trustworthy face of the continuum (High Trustworthy, HT), the least trustworthy face (Low Trustworthy, LT), and the face that lies at the center of the continuum (Neutral, N). Infants' sensitivity to perceptual differences among these three stimuli was investigated by comparing ERP responses evoked by each stimulus (Study 1) and by contrasting looking times to each face while presented pairwise in a standard preferential looking task (Study 2).

By measuring infants' ERPs we aimed to test whether, with the use of more ecological stimuli, we would observe modulations of ERP responses to HT versus LT faces that were absent in Jessen and Grossmann (2016) study. By measuring infants' preferential looking responses, we aimed to test whether the attentional preference for artificial, computer-generated HT male faces generalizes to realistic female face images.

In order to verify the suitability of the realistic face stimuli used in the current study to replicate prior electrophysiological evidence of enhanced perceptual and attentional processing of Low Trustworthy faces compared to High Trustworthy faces (Dzhelyova et al., 2012; Lischke et al., 2018; Marzi et al., 2014; Yang et al., 2011), Study 1 included also a group of adult participants, whose Event-Related Potentials (ERPs) in response to the N, LT and the HT faces were measured using the same exact procedure adopted with infants. A final aim of the current study was to explore the presence of individual differences in infants' neural sensitivity to facial cues to trustworthiness by exploring its association with temperamental traits. Such an association exists in the case of emotion discrimination, in that research has found significant relations between infants' temperament and their neural attention toward faces and bodies varying in emotional expression. Studies are not always consistent in the direction of the reported associations. For example, higher scores on Negative Affect have been reported to be associated with larger Nc responses to happy faces in 3- to 13-month-old infants (Martinos et al., 2012), and with larger P400 responses to angry faces in 7-month-olds (Quadrelli et al., 2019), and higher scores on the Approach/Surgency dimension are associated with larger Nc responses to fearful body expressions (Rajhans et al., 2015). Nonetheless, data converge in showing that individual differences in temperament affect infants' processing of socio-emotional cues in their environment. In light of this, in the current study we explored whether the impact of infants' temperament on ERP measures of emotion discrimination generalizes to neural discrimination of facial cues to trustworthiness.

Indeed, associations between individual differences in personality and social behavior and trustworthiness perception from faces have been reported in adults. Individuals with high anxiety traits showed enhanced working memory neural processing of untrustworthy faces (Meconi et al., 2014), and face discrimination based on physical cues to trustworthiness in a perceptual similarity behavioral task was slower in introverted compared to extraverted individuals (Baccolo & Macchi Cassia, 2019). In light of this evidence, here we explored the association between infants' neural response to HT, N and LT faces and their scores on the broad temperament dimensions of Negative Affect (i.e., propensity to experience negative feelings) and Surgency (i.e., the likelihood to experience and display high levels of activity), measured using the very short form of the revised Infant Behavior Questionnaire (IBQ-r VSF; Putnam et al., 2014). We focused on these two temperament dimensions as they are considered to be precursors to the adults' Neuroticism/Extraversion personality traits, which have been shown to modulate trustworthiness perception (Baccolo & Macchi Cassia, 2019). We hypothesized that infants' ERP responses to HT faces with respect to LT faces would vary according to their score on one or both temperament dimensions.

Study 1 – ERP task

Materials and methods

Participants

The final adult sample comprised 20 Caucasian young adult participants (7 males, *M* age = 25.25 years; *SD* = 3.06; range = 20–32), who were either undergraduate university students receiving course credits for their participation or were recruited from the community by word of mouth. Sample size was based on previous studies investigating neural sensitivity to facial cues to trustworthiness in adults (e.g., Marzi et al., 2014; Yang et al., 2011). Eight additional participants were tested but excluded from the final sample due to excessive EEG artifacts. All participants had no history of neurological or psychological disorders and had normal or corrected-to-normal vision; they all signed an informed consent before testing.

The final infant sample consisted of 21 six-month-old Caucasian participants (10 males, *M* age = 200 days; *SD* = 7.43; range = 184–211), who were born full-term with a normal weight at birth (> 2400 g). Eighteen additional infants were tested but excluded from the analyses due to fussiness (*N* = 1) or failure to contribute at least 10 good trials per condition (*N* = 17). Sample size was determined based on prior comparable studies investigating neural sensitivity to social cues from faces in 6-month-old infants (e.g., Jessen & Grossmann, 2019; Quadrelli et al., 2019). Attrition rate was comparable to that from other infant face processing ERP studies with infants over 6 months of age using purely visual stimulation (e.g., Righi et al., 2014; see also Stets et al., 2012). Infants were recruited via a written invitation that was sent to parents based on birth records provided by neighboring cities, and parents gave their written informed consent. The protocol followed the ethical standards of the Declaration of Helsinki and was approved by the Ethics Committee of the University of Milano-Bicocca.

Stimuli

Stimuli were selected from a previously validated set (see Baccolo & Macchi Cassia, 2020) including seven variations of one Caucasian female facial identity reflecting a continuum of trustworthiness that ranged from 1 (very untrustworthy) to 7 (very trustworthy), interleaved by a neutral face (Figure 1). The continuum resulted from morphing an averaged neutral face toward an averaged untrustworthy and an averaged trustworthy face using an online program for image transformation (WebMorph, DeBruine, 2017), and all the averaged faces were created by averaging three face identities taken from the Chicago Face Database (Ravicz et al., 2015). Details of the morphing and stimulus validation procedures are described in full in Baccolo and Macchi Cassia (2020). The face at center of the continuum (face 4: Neutral, N) and those at the two extremes were selected to be used as stimuli in the current study (face 1: Low Trustworthy, LT and face 7: High Trustworthy, HT). To ensure that infants focused their attention on the internal facial features, the three faces were cropped into an oval shape to remove the hair and the ears (Figure 1).

One of our goals in the current study was to test whether earlier demonstrations of infants' neural discrimination between computer-generated faces on the basis of their perceived trustworthiness (Jessen & Grossmann, 2016) would generalize to more ecological, realistic face images. To provide a direct comparison between the intensity of perceived trustworthiness elicited by these two stimulus sets we ran a further validation study in which we asked an independent sample of 37 young Caucasian adults (4 males, *M* age = 26.84 years; *SD* = 6.01; range = : 21–46) to rate each of the nine computer-generated faces used by Jessen and Grossmann (2016; the neutral, the +3 SD, and the −3 SD versions of faces 005, 010, and 016 from the database by Oosterhof & Todorov, 2008) and the N, LT, and HT faces used in the current study on a 9-point scale ranging from 1 ("I wouldn't trust this person at all) to 9 ("I would definitely trust this person"). The same adult participants were also asked to rate the perceived similarity of all possible pairwise combinations of the faces in each set on a 9-point scale ranging from 1 ("the two faces look identical") to 9 ("the two faces look completely different"). Details of the procedure and data analyses are reported in the Supplementary Information S1. Overall, results confirmed that a similar linear increase in participants' ratings from the Low trustworthy face to the High trustworthy face occurred for both stimulus sets. However, the faces from the realistic set were perceived as more similar to one another than those from the Todorov set, suggesting that the physical cues that differentiate the faces used as stimuli in the current study are more subtle than those differentiating the stimuli used by earlier investigations.

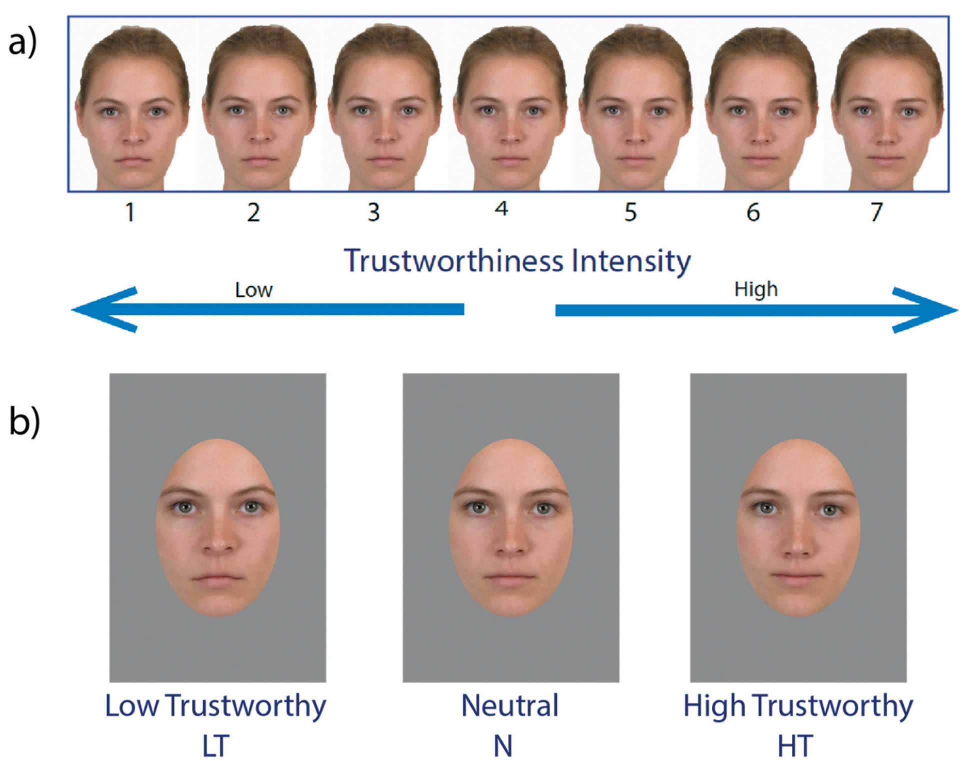

Figure 1. The seven variations of the female face identity included in the trustworthiness continuum created by Baccolo and Macchi Cassia (2020) where trustworthiness intensity ranged from 1 (very untrustworthy) to 7 (very trustworthy) (a). Face 1 (Low Trustworthy, LT), face 4 (Neutral, N) and face 7 (High Trustworthy, HT) were cropped and used as experimental stimuli (b).

Procedure

Data acquisition took place in a darkened audiometric and electrically shielded cabin. Stimuli were shown on a 24-inch monitor (1600 x 1200 pixels). Adult participants seated approximately 60 cm away from the monitor and were instructed to sit as still as possible during the entire procedure. Before testing, they completed the Italian version of the Big Five Questionnaire (BFQ; Caprara et al., 1993), a self-report questionnaire designed to measure the Big Five dimensions of personality. Only the items contributing to the Extraversion and the Neuroticism scales were scored, and the obtained scores were standardized by converting them into z-scores.

Infants seated on their parent's lap, approximately 60 cm away from the monitor; parents were instructed to sit as still as possible during the entire procedure and asked to avoid interfering with the experimental session by talking to the child or pointing to the screen. The whole experimental session was recorded through a digital video-camera hidden above the stimulus presentation monitor and connected to the data acquisition computer and a digital video recorder, both located outside the testing cabin. The live image of the infant's face and body allowed the experimenter to pause or terminate the session as soon as the infant became fussy. Each subject was presented with the N, LT, and HT stimuli, which appeared one at a time on the screen in a pseudo-randomized order, with the only constraint that the same stimulus could not appear more than four times in a row. Stimulus presentation was controlled through E-Prime software v2.0 (Psychology Software Tools Inc., Pittsburgh, PA). A trial consisted of a 1000 ms stimulus presentation followed by an interstimulus interval varying randomly between 900 and 1100 ms. Whenever necessary the experimenter presented a looming fixation point between trials to reorient the infant's attention to the monitor. Stimuli presentation continued until the infant became too fussy or bored to attend, with a maximum of 270 trials. Before testing, the infant's mother or the primary caregiver completed the Infant Behavior Questionnaire-Revised in its very short form (IBQ-R VSF; Putnam et al., 2014). The questionnaire included queries aimed to assess the frequency of specific temperament-related behaviors observed within the last week. We focused on the Negative Affect (NA, tendency to experience negative feelings and difficulty being soothed) and Surgency (SU, tendency to show high levels of activity and positive emotions and to act impulsively) subscales, which are considered as the precursors of the personality dimensions of Neuroticism and Extraversion. The obtained scores were standardized by converting them into z-scores.

EEG recording and processing

Electroencephalographic (EEG) data were recorded continuously using a 128-electrode HydroCel Geodesic Sensor Net (Electrical Geodesic Inc., Eugene, OR), referenced to the vertex electrode (Cz), and amplified using an EGI NetAmps 300 amplifier with a sampling rate of 500 Hz and an online band-pass filter of 0.1–100 Hz. Impedances were checked online before the session started and considered adequate if lower than 50 KΩ. EEG data were further processed offline using NetStation v4.6.4 (Eugene, OR). A band-pass filter of .3–30 Hz was applied to the continuous EEG signal, which was then segmented into epochs centered on the stimulus onset from 100 ms pre-stimulus onset to 1000 ms post-stimulus onset. Data were corrected to the baseline using the average voltage of the 100 ms prior to stimulus onset, and re-referenced to the algebraic mean of all channels. An automatic artifact rejection was applied to the segmented data so that, whenever the signal exceeded ± 200 μV at any electrode in a sliding window of 80 ms, channels were automatically rejected. Any remaining artifacts were hand-edited. Trials for which more than 15% of the channels (N ≥ 18) were marked as bad were excluded from further analyses (e.g., Halit et al., 2003). Of the remaining trials, individually bad channels were replaced using spherical spline interpolation. Individual subject averages for each of the three trustworthiness conditions (N, LT, and HT) were computed separately for each channel across all trials. For what concerns the adult sample, a repeated-measures ANOVA with trustworthiness level (N, LT, HT) as the within-subjects factor confirmed that a similar number of trials contributed to the average ERP for the HT condition (M = 44.4, SD = 11.87), the LT condition (M: 43.45, SD = 12.49) and the N condition (M = 42.75, SD = 13.73), all ps > .39. Similar to other infant visual ERP studies (e.g., Quadrelli et al., 2019), an inclusion criterion of 10 good trials for each stimulus category was adopted to include infant participants in the final sample (see also Stets et al., 2012). A repeated-measures ANOVA with trustworthiness level (N, LT, HT) as the within-subjects factor revealed that a similar number of trials contributed to the average ERP for the HT condition (M = 13.90, SD = 4.17), the LT condition (M = 13.62, SD = 3.47), and the N condition (M = 12.86, SD = 3.41), all ps > .39.

As regards the adult sample, we focused our analyses on the face sensitive N170 (120–180 ms) and the attentional LPP (520–720 ms) components, which were both reported to be modulated by variations in trustworthiness intensity perceived from faces (Lischke et al., 2018; Marzi et al., 2014). After Yang et al. (2011), we also analyzed the C1 (70–100 ms) component. Data were analyzed by averaging electrodes within the occipital-temporal region of each hemisphere where the N170 was more clearly visible (left: 65, 66, 69, 70; right: 83, 84, 89, 90), and by averaging electrodes over the fronto-central region where the C1 and the LPP components were observed (31, 53, 54, 55, 61, 62, 78, 79, 80, 86) (Figure 2). The time windows components were chosen based on previous ERP reports of the three components (e.g., Lischke et al., 2018; Yang et al., 2011), and visual examination of the components' peak across participants.

As for the infant sample, after Jessen and Grossmann (2016), we analyzed the face-sensitive N290 (200–260 ms) and P400 (330–430 ms) ERP components, and the attentional Nc (300–600 ms). A prominent P1 (120–170 ms) was also visible in our data, and was therefore included in the analyses. Data were analyzed by averaging one cluster of electrodes over the occipital-temporal regions of each hemisphere where the P1, the N290 and the P400 were more prominent (left: 65, 66, 69, 70; right: 83, 84, 89, 90) (Figure 3). The Nc was analyzed at fronto-central electrode sites over each of the two hemispheres (left: 36, 30, 37, 42; right: 87, 93, 104, 105). These electrode sites were chosen based on visual inspection of the component topography and correspond to electrode clusters in which the components of interest have been recorded in previous studies (e.g., Quadrelli et al., 2019; Peykarjou et al., 2014). Time windows were chosen based on previous infant ERP reports of the four components, and on examination of the peak of each component across participants (e.g., Leppänen et al., 2007; Quadrelli et al., 2019) and visual examination of the components' peak for each participant. For both adults and infants, peak latency (ms) and mean amplitude (μV) values were extracted for each of the considered components, and entered in the statistical analyses.

EEG Data analysis

Analyses were conducted with SPSS 25.0 (IBM Corporation, Armonk, NY, USA) and RStudio.1.0.136 (Ravicz et al., 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/). As for the adults, peak latency and mean amplitude of the N170 were analyzed through a 3 × 2 repeated-measures Analysis of Covariance (ANCOVA) with trustworthiness level (N, LT, HT) and hemisphere (left, right) as within-subjects factors, and Neuroticism and Extraversion scores derived from the BFQ entered as covariates. For the C1 and the LPP, the same repeated-measures ANCOVA was computed without including hemisphere as a factor.

As for the infants, peak latency and mean amplitude of the P1, N290, P400 and Nc were analyzed through a 3 × 2 repeated-measures ANCOVA with trustworthiness level

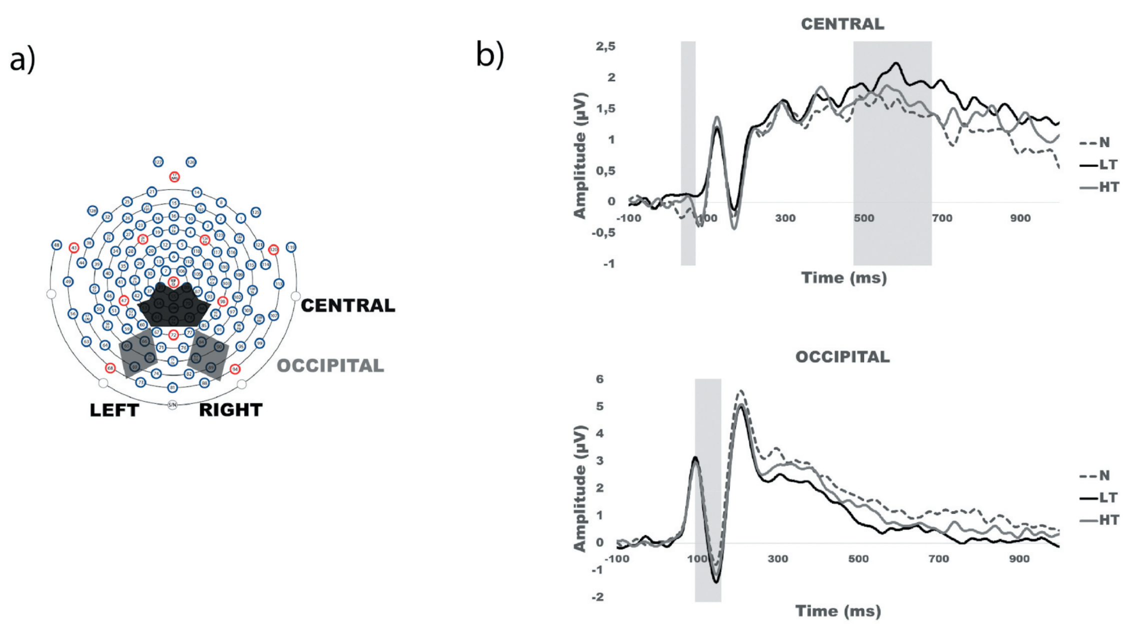

Figure 2. The topographic map shows the electrodes included in the fronto-central cluster (31, 53, 54, 55, 61, 62, 78, 79, 80, 86) and those included in the left (65, 66, 69, 70) and right (83, 84, 89, 90) occipital-temporal clusters used to obtain peak latency and mean amplitude values for each of the three analyzed ERP components (N170, C1, and LPP) (a). Waveform plots depict grand-average ERPs for the N170, LPP and C1 components recorded in adult participants in response to the Low Trustworthy face (solid dark line), the Neutral face (dashed line), and the High Trustworthy face (solid gray line) (b).

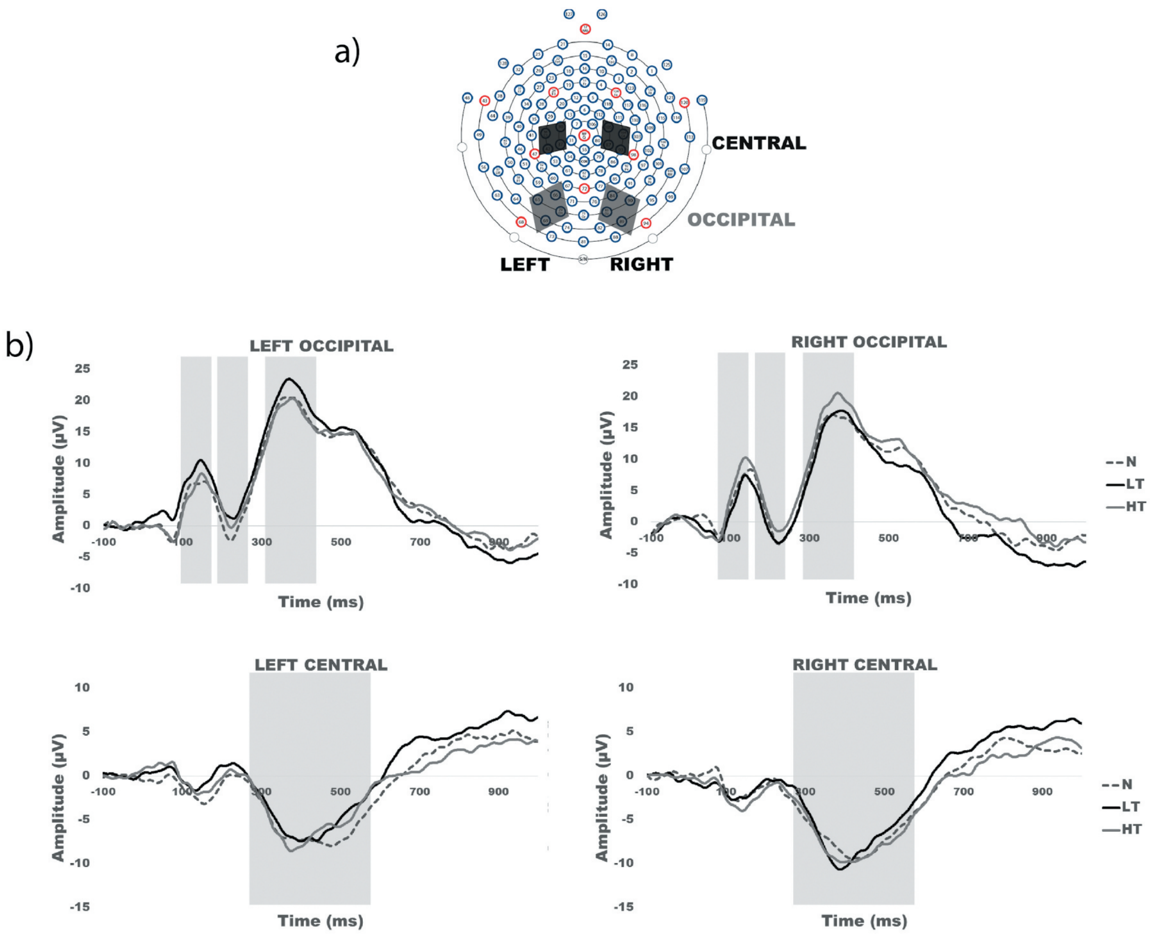

Figure 3. The topographic map shows the electrodes included in the left and right fronto-central (left: 36, 30, 37, 42; right: 87, 93, 104, 105) and occipital-temporal (left: 65, 66, 69, 70; right: 3, 84, 89, 90) clusters used to obtain peak latency and mean amplitude values for each of the four analyzed ERP components (P1, N290, P400, and Nc) (a). Waveform plots depicting grand-average ERPs for the P1, N290, P400 and Nc components recorded in infants in response to the Low Trustworthy face (solid dark line), the Neutral face (dashed line), and the High Trustworthy face (solid gray line) (b).

(N, LT, HT) and hemisphere (left, right) as within-subjects factors, and Negative Affect and Surgency scores derived from the IBQ-R entered as covariates. Post-hoc tests with Bonferroni correction were computed to further analyze interaction effects. Interaction effects involving personality (BFQ) or temperamental (IBQ-R) scores were followed up through correlational analyses (Pearson) including BFQ or IBQ-R subscale scores and ERP difference scores computed by subtracting mean amplitude or peak latency values for the N face from those recorded for the HT face (i.e., HT-N) and for the LT face (i.e., LT-N), and values for the LT face from those recorded for the HT face (i.e., HT-LT). Only significant correlations are reported. Effect sizes are reported as partial eta-square ($\eta 2$) for ANCOVAs and Cohen's *d* for Pearson correlations.

Results

Adults

N170 latency. We did not observe any significant effects of the considered variables on the latency of the N170 (all ps > .24).

N170 Amplitude. The ANCOVA revealed a significant main effect of trustworthiness level, $F(2,34) = 8.05$, $p = .001$, $\eta_p^2 = .32$. Post-hoc tests revealed a significant difference in the amplitude of the N170 in response to the LT face compared to the N face, $p < .001$. More interestingly, a test of within-subjects contrasts showed a significant linear trend, $F(1,17) = 26.99$, $p < .001$, $\eta_p^2 = .614$, with the LT face evoking a larger response

(M = −.30, SD = .50) compared to the HT (M = −.06, SD = .40) and the N (M = .37, SD = .46) faces (Figure 2). No other main effect or interaction attained significance (all ps > .16).

C1 latency. No significant effects were observed (all ps > .63).

C1 Amplitude. The ANCOVA showed a significant main effect of trustworthiness level, $F(2,34) = 4.82$, $p = .019$, $\eta_p^2 = .22$. A test of within-subjects contrasts revealed a significant linear trend, $F(1,17) = 5.506$, $p = .031$, $\eta_p^2 = .25$, with a larger C1 for N (M = −.28, SD = .20) compared to HT (M = −.22, SD = .18) and LT (M = .19, SD = .25) faces (Figure 2). No other main effect or interaction was significant (all ps > .416).

LPP Latency. No significant effects were observed (all ps > .327).

LPP Amplitude. We found a significant main effect of trustworthiness level, $F(2,34) = 3.47$, $p = .042 \eta_p^2 = .17$, and a significant linear trend, $F(1,17) = 8.02$, $p = .012$, $\eta_p^2 = .32$, due to LPP amplitude decreasing linearly across the LT (M = 1.97, SD = .26), the HT (M = 1.60, SD = .25), and the N (M = 1.46, SD = .20) conditions (Figure 2). No other effects attained significance (all ps > .767).

Infants

P1 Latency. The ANCOVA revealed a significant Trustworthiness level x Hemisphere interaction, $F(2,36) = 4.02$, $p = .027$, $\eta_p^2 = .183$, which however proved spurious, as no post-hoc comparisons attained significance (all ps > .177). No other main effect or interaction attained significance (all ps > .461).

P1 Amplitude. There were no significant main effects or interaction (all ps > .050).

N290 Latency. The ANCOVA revealed no significant main effects or interaction (all ps > .388).

N290 Amplitude. We found a significant interaction between hemisphere and Negative Affect, $F(1,18) = 20.30$, $p < .001$, $\eta_p^2 = .530$, which was explored through correlational analyses performed separately for each hemisphere. The analyses showed a negative correlation between N290 amplitudes over the left hemisphere and Negative Affect scores, $r = -0.33$, $p = .031$, $d = -.70$ (Figure 3). No other main effects or interactions attained significance (all ps > .112).

P400 Latency. The ANCOVA revealed a significant interaction between trustworthiness level and Surgency scores, $F(2,36) = 3.96$, $p = .028$, $\eta_p^2 = .18$, which was followed up through correlational analyses. Surgency scores were positively correlated with ERP differential scores obtained from subtracting latency values for the LT face from those recorded for the HT face (i.e., HT-LT), $r = 0.49$, $p = .024$, $d = 1.12$, indicating that infants who scored higher on Surgency showed faster P400 responses to the LT relative to the HT face. We also found a significant Trustworthiness level x Hemisphere interaction, $F(2,36) = 3.386$, $p = .045$, $\eta_p^2 = .158$, which was due to the P400 peaking earlier for the HT face (M = 367.17, SD = 7.15) than to the N face (M = 378.76, SD = 6.32) in the left hemisphere only, $p = .038$ (Figure 3).

P400 Amplitude. The ANCOVA revealed a significant Trustworthiness level x Hemisphere interaction, $F(2,36) = 3.36$, $p = .046, \eta_p^2 = .16$. Post-hoc comparisons showed larger amplitude in response to the LT face over the left (M = 21.45, SD = 2.36) than the right (M = 16.55, SD = 2.31), hemisphere, $p = .011$ (Figure 3). No other main effect or interaction attained significance (all ps > .085).

Nc Latency. The ANCOVA revealed a significant interaction between trustworthiness level and Negative Affect, $F(2,36) = 4.37$, $p = .024$, $\eta_p^2 = .20$. Correlational analysis showed a significant negative association between the latency of the Nc evoked by the N face with respect to the HT face (i.e., HT-N) and Negative Affect scores, $r = -0.54$, $p = .012$, $d = -1.28$ (Figure 4), indicating that infants who scored higher on Negative Affect showed faster Nc responses to the HT face relative to the N face. There was also a significant Trustworthiness level x Hemisphere interaction, $F(2,36) = 4.25$, $p = .022$, $\eta_p^2 = .19$. Post-hoc tests showed that the Nc peaked earlier in response to the HT face (M = 404.93, SD = 9.73) than to the N face (M = 445.12, SD = 11.52), $p = .004$, and the LT face (M = 439.29, SD = 12.06), $p = .027$, over the left hemisphere, and that Nc latency to the HT face was faster over the left (M = 404.93, SD =9.73) than over the right (M =425.74, SD = 9.42) hemisphere, $p = 0.023$ (Figure 3). No other main effects or interactions attained significance (all ps > .14).

Nc Amplitude. The ANCOVA revealed no significant effects (all ps > .08).

Summary of the ERP results

Adults ERP data provided evidence for a maximal amplitude of the N170 and the LPP in response to the LT face compared to the HT and N faces. Also, C1 amplitude for the LT face was smaller than for the HT and N face.

Infants ERP data provided evidence for a left-lateralized latency advantage for the HT face at the level of the P400 and the attentional Nc, which both peaked earlier for the HT than for the N and/or the LT face over the left hemisphere. Also, interindividual variations in the latency of both components in response to variations in trustworthiness intensity were associated

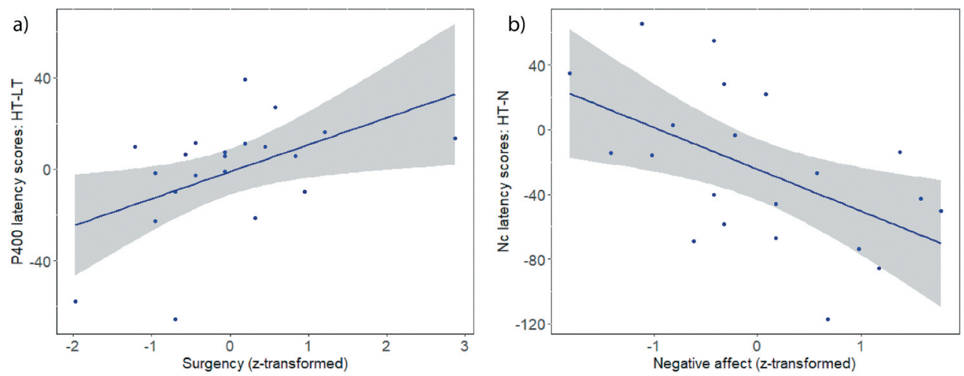

Figure 4. Correlation plots depicting the positive relationship between Surgency scores and the P400 latency evoked by the LT face expressed through the HT minus LT differential scores (i.e., faster P400 latencies to the LT face were associated with higher Surgency scores) (a) and the negative association between Negative Affective scores and Nc latency responses to the HT face expressed through the HT minus N differential scores (i.e., faster Nc latencies to the HT face associated with higher Negative Affect scores) (b).

with temperamental traits in the form of faster P400 responses to LT faces in more extraverted infants, and faster Nc responses to the HT face in infants who score higher on negative affectivity.

Study 2 – Preferential Looking Task

Materials and methods

Participants

Twenty-eight 6-month-old infants were invited to participate in the study. One infant was excluded from the analyses due to fussiness, leading to a final sample of N = 27 (11 males, *M* age = 217.89 days; *SD* = 9.45 days; range: 200–231 days). All infants were born full-term with a normal weight at birth (> 2.250 g), and Caucasian. Sample size was based on prior research on infants' visual preference based on facial cues to trustworthiness (e.g., Jessen & Grossmann, 2016, 2019; Sakuta et al., 2018). Participants' recruitment followed the same method described for Study 1; parents gave their written informed consent before the start of the testing session. The protocol was carried out in accordance with the ethical standards of the Declaration of Helsinki and was approved by the ethics committee of the University of Milano-Bicocca.

Stimuli

Stimuli consisted of the uncropped version of the three N, LT and HT colored faces used in Study 1 (Figure 5). Uncropped stimuli were used to maximize the ecological validity of the face images so as to stimulate infants' preferential looking responses.

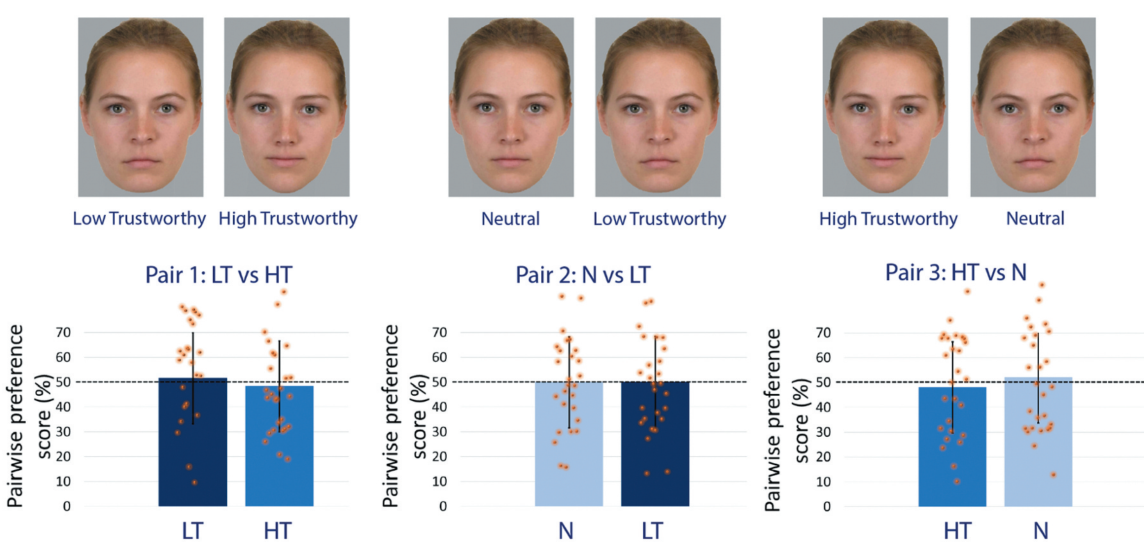

Figure 5. The three pairs of stimuli presented to the infants in the Visual preference task of Study 2, and the percentage of time spent looking at each stimulus in the pairs (i.e., Pairwise preference score).

Procedure

Infants were tested using a standard preferential-looking paradigm (Fantz, 1958) within the same setting used in Study 1. The parent was blind to the hypothesis of the study and was instructed to remain silent and keep the baby aligned with the monitor' midline. All infants viewed one 10-s bilateral presentation of three stimulus pairs, for a total of three trials. The three pairs were obtained by all the possible pairwise combinations of the three faces: Pair 1 consisted of the HT and LT stimuli, Pair 2 consisted of the LT and the N stimuli, and Pair 3 consisted of the HT and the N stimuli (Figure 4). The three pairs were shown in a random order, with the constraint that, because each face image appeared twice on the screen, its left/right position on the screen was reversed between the first and the second presentation. Trial presentation order, as well as the initial left/right position of the stimuli, was counterbalanced across participants. Each trial started with an animation that attracted the infant's attention to the center of the screen; as soon as the infant looked at the screen, the experimenter started the trial. Videotapes of eye movements were recorded and subsequently analyzed frame by frame to the nearest 40 ms by a coder who was blind to the specific position of the stimuli on each trial. The coder recorded, separately for each stimulus and each trial, the total fixation time (i.e., the sum of all fixations). As a measure of inter-observer reliability (Pearson correlation), total fixation duration on each of the three stimuli was recorded by a second coder for all of the infants in the sample; the level of agreement was high, $r = .988$, $p < .001$, $d = 12.79$.

Results

In order to test whether infants showed a preference for one of the stimuli presented, three different preference scores were computed using three different procedures adopted in previous studies: a pairwise preference score (after Macchi Cassia, Kuefner, Westerlund & Nelson, 2006), a total preference score (after Jessen & Grossmann, 2016), and a delta preference score (after Montoya et al., 2017). Statistical analyses are reported for each of the three preference scores.

Pairwise preference score. The pairwise preference score provides a measure of infants' preference for one stimulus within each pair. Following the procedure adopted by Macchi Cassia et al. (2006), each infant's looking time at the HT stimulus for Pairs 1 and 3 and at the N stimulus for Pair 2 was divided by the total time spent looking at either stimuli within the pair and then converted into a percentage score. Hence, only scores significantly above 50% indicated a preference for the considered stimuli. Three preliminary one-way Analyses of Variance (ANOVAs), one for each of the considered stimuli (HT in Pair 1, HT in Pair 3, and N in Pair 2), performed on the pairwise preference scores manifested by the three groups of infants who saw Pair 1, 2 or 3 as the first stimulus pair within the testing session, revealed that order of pair presentation did not affect infants' visual preferences (all $ps > .56$). To determine whether the pairwise preference scores differed from chance (50%) for each of the three stimulus pairs, three separate one-sample *t*-tests were applied, one for each pair. All tests failed to reach significance (all $ps > .61$), suggesting that infants' looking times were equally distributed across the stimuli within each pair (Figure 5).

Total preference score. The total preference score provides a measure of infants' overall preference for the N, the HT, and the LT stimuli across the two trials in which they are presented. Following the procedure adopted by Jessen and Grossmann (2016), total looking time for each of the three stimuli was obtained by summing the time spent looking at each stimulus across the two trials in which it was presented (Pairs 1 and 3 for the HT stimulus, Pairs 1 and 2 for the LT stimulus, and Pairs 2 and 3 for the N stimulus), and subsequently divided by the overall total time spent looking at all stimuli across the three trials and converted into a percentage score. A repeated-measure ANOVA with trustworthiness level (N, LT, HT) as the within-subjects factor proved non-significant, $F(2,52) = .041$, $p = .96$, $\eta_p^2 = .002$, indicating that none of the three stimuli preferentially attracted infants' attention across trials (N: $M = .33$, $SD = .09$; LT: $M = .34$, $SD = .09$; HT: $M = .33$, $SD = .08$).

Delta preference score. The delta preference score measures the magnitude of infants' preference for one of the stimuli within each pair. Following the procedure adopted in Montoya et al. (2017), for each infant and for each pair we computed the difference between the time spent looking at one stimulus and the time spent looking at the other, and divided the obtained delta by the total fixation time on both stimuli within the pair. Specifically, the delta was computed by subtracting looking time on the LT stimulus from the time spent looking at the HT stimulus for Pair 1, by subtracting looking time on the LT stimulus from looking time on the N stimulus for Pair 2, and by subtracting looking time on the N stimulus from looking time on the HT stimulus for Pair 3. A repeated-measure ANOVA with Pair (Pair 1, Pair2, Pair 3) as the within-subjects factor proved non-significant, $F(2,52) = .07$, $p = .89$, $\eta_p^2 = .003$, indicating once again that none of the three stimuli elicited a preferential visual response, irrespective of the stimulus with which it was paired (Pair 1: *M* = −.033, *SD* = .37; LT: *M* = −.001, *SD* = .38; HT: *M* = −.04, *SD* = .41).

General Discussion

The aim of the current study was to investigate whether infants are sensitive to those physical cues embedded in real-life face images that drive adults' trustworthiness judgments at an age, i.e., 7 months, when they show sensitivity to other social cues from real face images, like those signaling emotional states (see Hoehl, 2013). Infants' sensitivity was investigated both electrophysiologically (Study 1) and behaviorally (Study 2) by measuring the extent to which the ERP and the visual preference responses elicited by realistic face images perceived by adults as laying at the center or the extremes of a trustworthiness continuum differ. In order to verify the suitability of the realistic face stimuli used in the current study to replicate prior electrophysiological evidence, Study 1 included also a group of adult participants, whose Event-Related Potentials (ERPs) in response to the N, LT and the HT faces were measured using the same exact procedure adopted with infants.

Adults' data replicated previous reports of enhanced electrophysiological responses to untrustworthy computer-generated faces taken from the Todorov set. In line with earlier evidence (Dzhelyova et al., 2012; Lischke et al., 2018; Marzi et al., 2014; Yang et al., 2011), the amplitude of the N170 and the LPP components was maximal in response to the LT face compared to the HT and N faces. Also, ERP data replicated earlier reports of a smaller C1 for the LT face than for the HT and N faces (Yang et al., 2011). These findings indicate that the subtle physical cues to trustworthiness embedded in our realistic face stimuli were extracted already at the structural encoding stage of face processing and triggered allocation of differential attentional resources in the same way as the cues available in the computer-generated faces, which our validation studies proved to be more easily distinguishable. In contrast, we failed to replicate earlier reports of individual differences related to personality traits in behavioral (Baccolo & Macchi Cassia, 2019) and neural (Meconi et al., 2014) sensitivity to trustworthiness intensity in computer-generated faces.

The infant ERP data replicated and extended earlier demonstrations by Jessen and Grossmann (2016) that the infants' brain responds to variations in trustworthiness intensity expressed by computer-generated faces. The latency of the perceptual P400 and the attentional Nc over the left hemisphere differentiated between the HT face and the N face, and in case of the Nc, also between the HT and LT face, while no effects were observed at the level of the N290. In the infant face-processing literature, both the P400 and Nc components are more frequently reported to show amplitude, as compared to latency, modulations in response to facial familiarity (e.g., De Haan & Nelson, 1997) or emotional expressions (e.g., Van den Boomen et al., 2019), and this is also true for facial trustworthiness, which modulated the amplitude of the P400 and the Nc in Jessen and Grossmann (2016) study. This is potentially due to inter-individual variability in peak amplitude localization, which is known to be high in infants (Luck, 2005). Nonetheless, there are reports of shorter P400 latency for familiar (i.e., human) than unfamiliar (i.e., monkey) face categories in 12-month-old infants (Halit et al., 2003). Similarly, shorter P400 and Nc latencies were found for positively valenced (i.e., happy) expressions compared to angry and fearful faces within the first year of life (Quadrelli et al., 2019; Rigato et al., 2010). In light of this evidence, the observed latency advantage for the HT face in the current study might reflect an overall processing advantage for the most positively valenced face in our trustworthiness continuum.

Indeed, the direction of the modulation of the P400 and Nc responses observed in the current study differs from that reported by Jessen and Grossmann (2016), who found amplitude advantages in favor of the Neutral faces with respect to the High Trustworthy faces at the level of the P400, and with respect to both the High and the Low Trustworthy faces at the level of the Nc. The authors noted that their findings were incongruent with previous reports of a larger P400 for prosocial agents (Gredebäck et al., 2015), which, instead, is in line with our finding of shorter P400 latencies for the HT face. Indeed, sensitivity to variance in facial trustworthiness is associated with approach behavior in older children (Ewing et al., 2015), and impacts object processing in infants (Jessen & Grossmann, 2019). It is thus possible that the processing advantage enjoyed by the realistic HT face in our data was triggered by the positively valenced and friendly approach signaled by this stimulus, which was possibly less evident in the computer-generated stimuli used in Jessen and Grossmann (2016) study.

Another aspect of the current ERP findings that differs from those reported by Jessen and Grossmann (2016) relates to the lateralization of the modulatory effect of trustworthiness on the P400 and Nc parameters. The P400 and Nc latency advantage for the HT face in the current study was left-lateralized, while the P400 amplitude enhancement for the Neutral faces in Jessen and

Grossmann (2016) study was right-lateralized. The finding that the processing advantage for the HT was evident in the left hemisphere only in our data is partially congruent with the Approach-Withdrawal model of emotion-related Prefrontal Cortex (PFC) asymmetries (Davidson, 1983). This model proposes that stimuli that trigger approach and withdrawal behaviors are lateralized, respectively, to the left and the right hemispheres, as they selectively engage the left or the right PFC, whose activation is likely reflected in the infant attentional Nc (e.g., Ravicz et al., 2015). Within this view, infants' reiterate experience with positively valenced faces in their social environment (Vaish et al., 2008) might lead them to perceive the HT face as appealing, thus recruiting the left PFC, which is part of a motivational system that facilitates approach behaviors to engaging stimuli. Again, it is possible that the positive valence of the High Trustworthy face was more evident in the realistic stimuli used in the current study than in the computer-generated faces used in Jessen and Grossmann (2016) study.

Indeed, prior work with adults suggests that artificial face stimuli may not fully tap face processing expertise (e.g., Crookes et al., 2015). Our validation study showed that variance in trustworthiness intensity was perceived similarly in the computer-generated faces from the Todorov set used in Jessen and Grossmann (2016) and in the realistic set used in the current study. However, the overall distinguishability of the faces was higher for the Todorov set than for the realistic set, and the LT face was perceived as more trustworthy than the Low Trustworthy faces from the Todorov set. Assuming that adults' data can be generalized to infants, these differences might have contributed to the dissimilarities in the direction of the effects reported in the current and the Jessen and Grossmann (2016) study. Interestingly, these same stimulus differences did not affect adults' ERP data, which replicated earlier reports of neural sensitivity to face trustworthiness using Todorov stimuli (e.g., Marzi et al., 2014). This might result from adults' stronger expertise with faces, which allowed them to easily generalize neural responses to variance in trustworthiness across different face-types. It should also be considered that, as reported in the Supplementary information S1, faces from our realistic set were perceived as more similar to one another than those from the Todorov set. Future studies using realistic faces that display more intense cues to trustworthiness than those displayed by our stimuli might disentangle whether the present results diverge from those obtained by Jessen and Grossmann (2016) with computer-generated faces due to the intensity of the trustworthiness cues displayed by the stimuli in the two studies. The fact that our stimuli were more difficult to discriminate compared to the artificial stimuli from the Todorov set may also explain why, in Study 2, we failed to replicate Jessen and Grossmann (2016) finding of a looking time preference for the High Trustworthy face over the Neutral and the Untrustworthy faces. Indeed, we did not observe any clear pattern in preferential looking behavior, with infants looking equally longer to the HT, the N, and the LT faces. Discrepancies between ERP measures and looking times in a preferential-looking task are not uncommon in the infant literature (e.g, De Haan & Nelson, 1997; Macchi Cassia et al., 2006), and were also present in Jessen and Grossmann (2016) study, where infants' looking behavior, but not ERP responses, differentiated between faces at the two extremes of the trustworthiness continuum (i.e., Low Trustworthy vs. High Trustworthy). Such discrepancies are prone to different possible interpretations, including methodological differences in stimulus duration (i.e., longer simultaneous presentations in the behavioral task vs. several individual short presentations of each stimulus in the ERP task). The pattern of results obtained in the current study, where ERP measures were more sensitive to the stimulus manipulation as compared to behavioral measures, may likely indicate that, at 7 months, the attunement of neural circuitries to the subtle physical cues that support trustworthiness perception in our stimuli does not translate yet into overt attention responses. Nonetheless, it should also be noted that the absence of a significant preferential response in our preferential looking task does not imply that infants were incapable of discriminating among the stimuli. This could be rather explored in future studies by testing infants in a visual habituation task.

Another aspect of the current findings that extends earlier demonstration of neural sensitivity to facial cues to trustworthiness, in addition to the use of realistic face images, is the observed association between ERP measures of such sensitivity and infants' temperamental traits, which highlights the presence of individual differences in the processing of trustworthiness cues. We observed faster P400 responses to the LT face with respect to the HT face in more extraverted infants (i.e., those who scored higher on Surgency), and faster Nc responses to the HT face with respect to the N face in infants who have more frequently experience negative feelings and difficulty being soothed (i.e., those who scored higher on Negative Affect).

Although at first sight not overlapping, these findings seem to depict a coherent picture consistent with the literature. The P400 processing advantage for the LT face in extraverted infants resonates well with prior longitudinal reports of associations between an attention bias to thread-related facial expressions at 7 months and the development of responsivity to others' needs and emotional distress at 24 and 48 months (Peltola et al., 2018). At the same time, this result also matches with the finding that, at the level of the Nc, attention allocation to the HT face was faster in infants showing higher levels of withdrawal and distress to limitation. Indeed, both findings point to a pattern in which low levels of activity and positive emotions and high levels of withdrawal and emotional distress are associated with enhanced neural processing of the more trustworthy face. High levels of early life stress are known to be associated with attentional avoidance of negative and threat-related stimuli like fearful facial expressions (Humphreys et al., 2016) or anger (e.g, Nelson et al., 2013). Although we did not find direct evidence of avoidance-like responses to the LT face in association to high Negative Affect scores, our findings may possibly suggest that infants who experience higher levels of distress are faster to allocate attention to the HT face as a coping strategy to reduce the distress that negatively valenced stimuli evoke in them (see In-Albon et al., 2010).

To sum up, current findings show that the infants' brain distinguishes subtle differences between realistic face images that generate explicit trustworthiness judgments in adults, and that interindividual variations in neural sensitivity to these differences are associated with infants' temperamental traits. These findings extend previous evidence obtained with computer-generated stimuli (Jessen & Grossmann, 2016) to more ecological, realistic stimuli. In light of data from our stimulus validation showing that perceptual differences associated with variance in trustworthiness perception were less easily detectable (by adults) in the realistic stimuli used in the present study than in the artificial stimuli used in prior work, the present results indicate a very fine tuning of the perceptual mechanisms supporting recognition of social signals from faces at 7 months.

An important goal for future studies will be to test whether the present results would be replicated using male, as opposed to female, realistic faces. Indeed, the artificial faces from the Todorov set used by Jessen and Grossmann (2016) depicted male identities, and even if the number of studies investigating how face gender affects perceived trustworthiness are very limited, adults tend to evaluate female faces more positively than male faces in terms of first impressions (e.g., Sutherland et al., 2015). It is therefore possible that face gender might have contributed as well to the dissimilarities between the current results and those reported by Jessen and Grossmann (2016). Future studies shall also test whether these results would hold under even more ecological conditions, for example, by using uncropped faces as stimuli, in which visible external features are preserved. Moreover, although our trustworthiness continuum was created by averaging different face identities precisely to reduce the influence of idiosyncratic morphologic features, the use of one single identity continuum in the current study might have still impacted the results, as facial morphology is known to influence perceptual judgments of emotions (e.g., El Zein & Grèzes, 2018) and personal traits (e.g., Todorov et al., 2015). Therefore, the use of a more variable set of stimuli including different averaged identities would allow testing for the strength and generalizability of the observed effects.

Acknowledgments

The authors thank Carlo Toneatto for programming the experiments and Veronica Gaetani, Michelle Giraud and Melissa Savoldi for help in testing participants.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Disclosure of interest

The authors report no conflict of interest.

Data availability statement.

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Funding

This work was supported by a PhD Scholarship from the Italian Minister of University and Research to the first author, and a grant from the Università di Milano-Bicocca to the last author [2017-ATE-0566].

References

Baccolo, E., & Macchi Cassia, V. (2019). Individual differences in perceptual sensitivity and representation of facial signals of trustworthiness. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(2), 224–236. https://doi.org/10.1037/xhp0000601

Baccolo, E., & Macchi Cassia, V. (2020). Age-related differences in sensitivity to facial trustworthiness: Perceptual representation and the role of emotional development. *Child Development*, *91*(5), 1529–1547. https://doi.org/10.1111/cdev.13340

Balas, B. J., Nelson, C. A., Westerlund, A., Vogel-Farley, V., Riggins, T., & Kuefner, D. (2010). Personal familiarity influences the processing of upright and inverted faces in infants. *Frontiers in Human Neuroscience*, *4*, 1. https://doi.org/10.3389/neuro.09.001.2010

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, *6*(2), 269–278. https://doi.org/10.1037/1528-3542.6.2.269

Caprara, G. V., Barbaranelli, C., Borgogni, L., & Perugini, M. (1993). The "big five questionnaire": A new questionnaire to assess the five factor model. *Personality and Individual Differences*, *15*(3), 281–288. https://doi.org/10.1016/0191-8869(93)90218-R

Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring character from faces: A developmental study. *Psychological Science*, *25*(5), 1132–1139. https://doi.org/10.1177/0956797614523297

Crookes, K., Ewing, L., Gildenhuys, J., Kloth, N., G., Oxner, M., W. H., Rhodes, G., Pond, S., & Rhodes, G. (2015). How well do computer-generated faces tap face expertise? *Plos One*, *10*(11), e0141353. https://doi.org/10.1371/journal.pone.0141353

Davidson, R. (1983). Hemispheric specialization for cognition and affect. In A. Gale & J. Edwards (Eds.), *Physiological correlates of human behaviour* (pp. 203–216). Academic Press.

de Haan, M., & Nelson, C. A. (1997). Recognition of the mother's face by six-month old infants: A neurobehavioral study. *Child Development*, *68*(2), 187–210. https://doi.org/10.2307/1131845

DeBruine, L. (2017). Webmorph (Version v0.0.0.9001). *Zenodo*. https://doi.org/10.5281/zenodo.1073696

Dzhelyova, M., Perrett, D. I., & Jentzsch, I. (2012). Temporal dynamics of trustworthiness perception. *Brain Research*, *1435*, 81–90. https://doi.org/10.1016/j.brainres.2011.11.043

El Zein, M., Wyart, V., & Grèzes, J. (2018). Pervasive influence of idiosyncratic associative biases during facial emotion recognition. Scientific reports, 8(1), 8804. https://doi.org/10.1038/s41598-018-27102-z

Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2015). Perceived trustworthiness of faces drives trust behaviour in children. *Developmental Science*, *18*(2), 327–334. https://doi.org/10.1111/desc.12218

Fantz, R. L. (1958). Pattern vision in young infants. *The Psychological Record*, *8*(2), 43–47. https://doi.org/10.1007/BF03393306

Farroni, T., Menon, E., Rigato, S., & Johnson, M. (2007). The perception of facial expressions in newborns. *European Journal of Developmental Psychology*, *4*(1), 2–13. https://doi.org/10.1080/17405620601046832

Gredebäck, G., Kaduk, K., Bakker, M., Gottwald, J., Ekberg, T., Elsner, C., Reid, V., & Kenward, B. (2015). The neuropsychology of infants' pro-social preferences. *Developmental Cognitive Neuroscience*, *12*, 106–113. https://doi.org/10.1016/j.dcn.2015.01.006

Grossmann, T., Oberecker, R., Koch, S. P., & Friederici, A. D. (2010). The developmental origins of voice processing in the human brain. Neuron, 65(6), 852–858. https://doi.org/10.1016/j.neuron.2010.03.001

Halit, H., de Haan, M., & Johnson, M. H. (2003). Cortical specialisation for face processing: Face-sensitive event-related potential components in 3- and 12-month-old infants. *Neuroimage*, *19*(3), 1180–1193. https://doi.org/10.1016/S1053-8119(03)00076-4

Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development*, *26*(1), 30–39. https://doi.org/10.0.3.248/j.cogdev.2010.09.001

Hoehl, S. (2013). Emotion processing in infancy. *Contributions to Human Development*, *26*, 1–12. https://doi.org/10.1159/000354346

Humphreys, K., Kircanski, K., Colich, N., & Gotlib, I. (2016). Attentional avoidance of fearful facial expressions following early life stress is associated with impaired social functioning. *Journal Of Child Psychology And Psychiatry, And Allied Disciplines, 57, 1174-82*, *57*(10), 1174–1182. https://doi.org/10.1111/jcpp.12607

In-Albon, T., Kossowsky, J., & Schneider, S. (2010). Vigilance and avoidance of threat in the eye movements of children with separation anxiety disorder. *Journal of Abnormal Child Psychology*, *38*(2), 225–235. https://doi.org/10.1007/s10802-009-9359-4

Jack, R. E., & Schyns, P. G. (2015). Review: The human face as a dynamic tool for social communication. *Current Biology*, *25*(14), R621. https://doi.org/10.1016/j.cub.2015.05.052

Jessen, S., & Grossmann, T. (2016). Neural and behavioral evidence for infants' sensitivity to the trustworthiness of faces. *Journal Of Cognitive Neuroscience*, *28*(11), 1728–1736. https://doi.org/10.1162/jocn_a_00999

Jessen, S., & Grossmann, T. (2017). Neural evidence for the subliminal processing of facial trustworthiness in infancy. *Neuropsychologia*, *126*, 46–53. https://doi.org/10.1016/j.neuropsychologia.2017.04.025

Jessen, S., & Grossmann, T. (2019). Neural evidence for the impact of facial trustworthiness on object processing in a gaze-cueing task in 7-month-old infants. *Social Neuroscience*, *9*, 1–9. https://doi.org/10.1080/17470919.2019.1651764

Leppänen, J., Moulson, M. C., Vogel-Farley, V. K., & Nelson, C. A. (2007). An ERP study of emotional face processing in the adult and infant brain. *Child Development*, *78*(1), 232–245. https://doi.org/10.1111/j.1467-8624.2007.00994.x

Lischke, A., Junge, M., Hamm, A. O., & Weymar, M. (2018). Enhanced processing of untrustworthiness in natural faces with neutral expressions. *Emotion*, *18*(2), 181–189. https://doi.org/10.1037/emo0000318

Luck, S. J. (2005). *An introduction to the event-related potential technique*. MIT Press.

Macchi Cassia, V., Kuefner, D., Westerlund, A., & Nelson, C. A. (2006). A behavioural and ERP investigation of 3-month-olds' face preferences. *Neuropsychologia*, *44*(11), 2113–2125. https://doi.org/10.1016/j.neuropsychologia.2005.11.014

Martinos, M., Matheson, A., & de Haan, M. (2012). Links between infant temperament and neurophysiological measures of attention to happy and fearful faces. *Journal of Child Psychology and Psychiatry*, *53*(11), 1118–1127. https://doi.org/10.1111/j.1469-7610.2012.02599.x

Marzi, T., Righi, S., Ottonello, S., Cincotta, M., & Viggiano, M. P. (2014). Trust at first sight: Evidence from ERPs. *Social Cognitive And Affective Neuroscience*, *9*(1), 63–72. https://doi.org/10.1093/scan/nss102

Meconi, F., Luria, R., & Sessa, P. (2014). Individual differences in anxiety predict neural measures of visual working memory for untrustworthy faces. *Social Cognitive And Affective Neuroscience*, *9*(12), 1872–1879. https://doi.org/10.1093/scan/nst189

Montoya, L., Westerlund, A., Troller-Renfree, S., Righi, G., & Nelson, C. A. (2017). The effect of heterogeneous race exposure during infancy. *Cognitive Development*, *42*, 74–83. https://doi.org/10.1016/j.cogdev.2017.02.014

Nelson, C. A., Westerlund, A., McDermott, J., Zeanah, C., & Fox, N. (2013). Emotion recognition following early psychosocial deprivation. *Development And Psychopathology*, *25*(2), 517–525. https://doi.org/10.1017/S0954579412001216

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 11087. https://doi.org/10.1073/pnas.0805664105

Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion*, *9*(1), 128. https://doi.org/10.1037/a0014520

Peltola, M. J., Yrttiaho, S., & Leppänen, J. M. (2018). Infants' attention bias to faces as an early marker of social development. *Developmental Science*, *21*(6), e12687. https://doi.org/10.1111/desc.12687

Putnam, S. P., Helbig, A. L., Gartstein, M. A., Rothbart, M. K., & Leerkes, E. (2014). Development and assessment of short and very short forms of the infant behavior questionnaire–revised. *Journal of Personality Assessment*, *96*(4), 445–458. https://doi.org/10.1080/00223891.2013.841171

Quadrelli, E., Conte, S., Macchi Cassia, V., & Turati, C. (2019). Emotion in motion: Facial dynamics affect infants' neural processing of emotions. *Developmental Psychobiology*, *61*(6), 843–858. https://doi.org/10.1002/dev.21860

Rajhans, P., Missana, M., Krol, K., & Grossmann, T. (2015). The association of temperament and maternal empathy with individual differences in infants' neural responses to emotional body expressions. *Development and Psychopathology*, *27*(4pt1), 1205–1216. https://doi.org/10.1017/S0954579415000772

Ramsey-Rennels, J. L., & Langlois, J. H. (2006). Infants' differential processing of female and male faces. *Current Directions in Psychological Science*, *15*(2), 59–62. https://doi.org/10.0.4.87/j.0963-7214.2006.00407.x

Ravicz, M. M., Perdue, K., Westerlund, A., Vanderwert, R. E., & Nelson, C. A. (2015). Infants' neural responses to facial emotion in the prefrontal cortex are correlated with temperament: A functional near-infrared spectroscopy study. *Frontiers in Psychology*, *6*, 922. https://doi.org/10.3389/fpsyg.2015.00922

Rennels, J. L., & Davis, R. E. (2008). Facial experience during the first year. *Infant Behavior & Development*, *31*(4), 665678. https://doi.org/10.0.3.248/j.infbeh.2008.04.009

Reynolds, G. D., & Richards, J. E. (2005). Familiarization, attention,and recognition memory in infancy: An event-related potential and cortical source localization study. *Developmental Psychology*, *41*(4), 598–615. https://doi.org/10.1037/0012-1649.41.4.598

Rigato, S., Farroni, T., & Johnson, M. H. (2010). The shared signal hypothesis and neural responses to expressions and gaze in infants and adults. *Social Cognitive and Affective Neuroscience*, *5*(1), 88–97. https://doi.org/10.1093/scan/nsp037

Righi, G., Westerlund, A., Congdon, E. L., Troller-Renfree, S., & Nelson, C. A. (2014). Infants' experience-dependent processing of male and female faces: Insights from eye tracking and event-related potentials. *Developmental Cognitive Neuroscience*, *8*, 144–152. https://doi.org/10.1016/j.dcn.2013.09.005

Sakuta, Y., Kanazawa, S., & Yamaguchi, M. (2018). Infants prefer a trustworthy person: An early sign of social cognition in infants. *PlosOne*, *13*(e0203541), e0203541. https://doi.org/10.1371/journal.pone.0203541

Serrano, J. M., Iglesias, J., & Loeches, A. (1995). Brief report: Infants' responses to adult static facial expressions. *Infant Behavior and Development*, *18*(4), 477–482. https://doi.org/10.1016/0163-6383(95)90036-5

Stets, M., Stahl, D., & Reid, V. M. (2012). A meta-analysis investigating factors underlying attrition rates in infant ERP studies. *Developmental Neuropsychology*, *37*(3), 226–252. https://doi.org/10.1080/87565641.2012.654867

Sugden, N. A., Mohamed-Ali, M. I., & Moulson, M. C. (2014). I spy with my little eye: Typical, daily exposure to faces documented from a first-person infant perspective. *Developmental Psychobiology*, *56*(2), 249–261. https://doi.org/10.1002/dev.21183

Sutherland, C. A. M., Young, A. W., Mootz, C. A., & Oldmeadow, J. A. (2015). Face gender and stereotypicality influence facial trait evaluation: Counter-stereotypical female faces are negatively evaluated. *British Journal of Psychology*, *106*(2), 186208. https://doi.org/10.1111/bjop.12085

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, *66*(1), 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Vaish, A., Grossman, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383–403. https://doi.org/10.1037/0033-2909.134.3.383

van den Boomen, C., Munsters, N. M., & Kemner, C. (2019). Emotion processing in the infant brain: The importance of local information. *Neuropsychologia*, *126*, 62–68. https://doi.org/10.1016/j.neuropsychologia.2017.09.006

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592. https://doi.org/10.1111/j.1467-9280.2006.01750.x

Yang, D., Qi, S., Ding, C., & Song, Y. (2011). An ERP study on the time course of facial trustworthiness appraisal. *Neuroscience Letters*, *496*(3), 147–151. https://doi.org/10.1016/j.neulet.2011.03.066

Zebrowitz, L. A., Fellous, J. M., Mignault, A., & Andreoletti, C. (2003). Trait impressions as overgeneralized responses to adaptively significant facial qualities: Evidence from connectionist modeling. *Personality And Social Psychology Review: An Official Journal Of The Society For Personality And Social Psychology*, *7*(3), 194–215. https://doi.org/10.1207/S15327957PSPR0703_01