

How can I trust you? The role of facial trustworthiness in the development of Epistemic and Interpersonal Trust

Alberto Milesi^a, Pietro De Carli^a, Francesca Locati^b, Ilaria M.A. Benzi^c, Chloe Campbell^d, Peter Fonagy^d, Laura Parolin^a

^a Department of Psychology, University of Milano-Bicocca, Milan, Italy

^b Department of Humanities, University of Pavia, Pavia, Italy

^c Department of Brain and Behavioral Sciences, Lab on Attachment and Parenting – LAG, University of Pavia, Pavia, Italy

^d Research Department of Clinical, Educational and Health Psychology, University College London, London, UK

Short Title: Epistemic trust, Interpersonal trust, and Facial Trustworthiness

Corresponding Author:

Laura Parolin

Department of Psychology

University of Milano-Bicocca

Piazza dell'Ateneo Nuovo 1

Milan, MI, 20126, Italy

Tel: [0264483796](tel:0264483796)

E-mail: laura.parolin@unimib.it

Number of Tables: 1

Number of Figures: 1

Word count: 6185

Keywords: Epistemic Trust, Interpersonal Trust, Facial Trustworthiness, Borderline Personality Disorder

Abstract

Recently, researchers from developmental and clinical psychology highlighted epistemic trust (ET) as a key factor for personality disorders. ET is intended as the mental openness to information coming from others during social exchanges. ET develops from signals called ostensive cues, delivered through facial expressions during interactions in a secure attachment context. Similarly interpersonal trust (IT) refers to the perception of others as not harmful, which is also developed through secure attachment relationships.

Our purpose is to suggest a conceptualization of ET as a specific facet of IT. We hypothesize that positive experiences of caregiving promote IT development that includes a specific sense of trust toward others' knowledge. Moreover, we suggest that the early ability to infer a judgment of trustworthiness from facial cues is the starting point for developing both IT and ET. This conceptualization supports the role of considering both IT and ET in the development of borderline pathology.

Introduction

Let's imagine a child, aged around 3-4 years old, that sees for the first time a lit fireplace. In utter astonishment, they do not know what this fascinating show is about. They slowly get closer to it, trying to figure out what is going on. Their parents are paying attention to the scene from a discrete but secure distance, suggesting to the child that this thing seems very curious, but also telling the child to be careful in getting closer, otherwise they will get hurt. The child suddenly stops, starting to think about what they should do. We think that this scene evokes some questions about the ongoing process, in which trust in oneself and others' words and believes is central. In this work, we will try to hypothesize what influence the child's choice and how the co-occurring underlying elements are developed.

Recently, the concept of epistemic trust (ET) has been put under the spotlight of academic research by different authors as a key concept in the understanding of several forms of psychopathology (Fonagy & Allison, 2014; Kamphuis & Finn, 2019; Luyten, Campbell, & Fonagy, 2019). Generally, the term "epistemic" defines the particular active position in acknowledging something, and "trust" is an attitude toward others or ourselves that comes from a positive evaluation of facts, circumstances, and relations; that is, we can rely on others because of trust. Thus, ET can be defined as the individual's openness to the possibility of acquiring new knowledge coming from another individual: this knowledge is perceived as trustworthy and reliable and is generalizable through different life domains (Fonagy & Allison, 2014). ET has been suggested as a key psychological construct in the understanding of personality disorders (Kamphuis & Finn, 2018), especially borderline personality disorder (BPD; Fonagy, Luyten, & Allison, 2015; Luyten et al., 2019). Moreover, the comprehension of how humans accept incoming information allows us to investigate some elements of the psychotherapy process, such as a patient's internalization of information coming from the therapist, and it is possible to hypothesize treatment techniques (Fonagy & Allison 2014).

Indeed, it has been hypothesized that the lack of mentalizing abilities of borderline patients (Fonagy et al., 2015) could originate from a difficulty in the development of ET. Indeed, the traumatic events often experienced by these individuals during their childhood (e.g., verbal, and physical abuse, emotional neglect; Luyten et al., 2019) could be responsible for undermining the development of ET, thus contributing to their mentalization deficits (Fonagy et al., 2015; Luyten et al., 2019). Both Campbell and colleagues (2021) and Kampling

and colleagues (2022) found associations between traumatic experiences and epistemic trust disruption. However, secure attachment experiences, characterized by a sense of security and protection, lead to the development of internalized patterns of trustworthy relationships (Cohn, 1990, Lieberman, 1977). Insecure attachment affects not only the individual's disposition to consider new knowledge from another person as reliable and relevant to the self (ET) but also the more general sense of interpersonal trust (IT) toward others. Rotenberg defines IT as *"defined sets of beliefs about persons which comprises positive expectations of their behavior"* (Rotenberg, 2010, p.11).

Specifically, complex trauma involves experiences of physical and psychological abuse in which infants are exposed to the perception of the other as malevolent and not trustful (Luyten et al., 2019). This repetitive exposure to relational patterns in which the other cannot be trusted results in mental patterns of the relationships with others, named internal working models (IWMs; Main et al., 1985), characterized by suspiciousness and vigilance. According to social information processing theory (Dodge, Pettit, McClaskey, Brown & Gottman, 1986), IWMs are a constitutional part of the mental patterns that allow the processing of social information. The first step of social processing, according to the authors, involves perceptive and sensorial processes in coding social cues, and, subsequently, the second step involves causal attribution to the same cues in terms of intentional attribution. It is possible to affirm that the very first step in the attribution of trustworthiness is at an implicitly perceptive level, which is enriched by a more profound judgment in the second step in which IT and ET are involved.

Some researchers have investigated how humans perceive some social cues, specifically attributing characteristics relevant for social exchanges. According to Willis and Todorov (2006), trustworthiness is the first characteristic processed when meeting a new person. This happens thanks to the face evaluation of trustworthiness at early stages and maintains a central role in humans' social interactions for the entire life (Willis & Todorov, 2006; De Carli et al., 2019).

Given these premises, this contribution aims to present ET within the broader context of IT. Specifically, we focus on how both IT and ET are inferred from facial cues during first impression formation. We hypothesize that early positive experiences of parental care, thanks to ostensive signals (i.e., *"cues designed by a communicator to generate an interpretation of communicative intention in an addressee"*; Szufnarowska et al., 2014, p.1)

driven through facial cues, are significant promoters of the broader construct of IT, which entails ET as well. Additionally, we provide evidence that a disturbance in the early experiences that support this trustworthiness recognition ability is likely to be involved in the development of psychopathological symptoms. More specifically, we suggest a more profound understanding of BPD, in which the impairment of IT and ET significantly influences the quality of life of individuals with this diagnosis.

First, we introduce the crucial role ET plays in understanding interpersonal exchanges. Second, we provide evidence on how ET might develop. Third, we expand the understanding of ET in the broader context of perceived trust (IT). Fourth, we set facial cues as a common ground between ET and IT. Finally, we discuss plausible clinical implications of our contribution to the understanding of borderline pathology.

As many constructs are discussed along this work, we summarize all of them in the subsequent table (Table 1).

Table 1. Definitions of mentioned constructs.

CONSTRUCTS	DEFINITION
<i>Epistemic Trust</i>	The individual's openness to the possibility of acquiring new knowledge coming from another individual: this knowledge is perceived as trustworthy and reliable and is generalizable through different life domains (Fonagy & Allison, 2014)
<i>Interpersonal Trust</i>	Defined sets of beliefs about persons which comprises positive expectations of their behavior (Rotenberg, 2010)
<i>Ostensive Cues</i>	Cues designed by a communicator to generate an interpretation of communicative intention in an addressee; the ostension process, which promotes natural pedagogy, and maternal sensitivity, which promotes attachment security intended in terms of mentalizing or sensitivity to an intentional state, are overlapping constructs (Fonagy et al., 2007; Szufnarowska et al., 2014)

Borderline Personality Disorder	Severe psychiatric condition characterized by emotion dysregulation, impulsiveness, self-harm, suicidal behaviors, and severe interpersonal impairment (APA, 2013)
---------------------------------	--

From Epistemic Vigilance to Epistemic Trust: how does ET develop?

ET is one specific facet of the general trust inherent in any exchange of information and is related to how communication's content is perceived in terms of reliability and relevance. Among the authors who have investigated trust during social exchanges, Sperber and colleagues (2010) pointed out how humans perceive incoming information from others and how they can trust them. They claim that humans are naturally provided with cognitive mechanisms for epistemic vigilance (EV) to evaluate incoming information's trustworthiness. Indeed, being suspicious toward the transmitted information constitutes an evolutionary advantage since it protects from possible deception (Sperber & Wilson, 1995). When EV is activated, which is usually by default, individuals do not trust the information they are receiving *per se*. In contrast, when the EV level is lowered, the information can be perceived as acceptable; the shared knowledge is then internalized as reliable and relevant for future situations. Sperber and colleagues add that informants need two qualities: competence and benevolence. They also emphasize the importance of the informants during a trustful communicative exchange. Our ability to assess the trustworthiness of a speaker probably emerges in biological and cultural evolution, coming from the possibility of understanding who we can avoid and who we can approach for survival. Csibra and Gergely (2009) highlighted specific cues that constitute this "subtle" communication, i.e., ostensive cues, that can emphasize the relevance of the transmitted message. According to Fonagy and Allison (2014), this process of ostension reduces the level of natural EV and makes it possible to experience ET.

The authors point out that in this way, the informant should be trusted about specific topics when they are speaking to a particular audience and under certain circumstances. However, this kind of evaluation of an informant cannot be made because of its high energy and time expenditure. They suppose that humans probably "rely on less costly general impressions of competence, benevolence and overall trustworthiness" (Sperber et al., 2010, p. 369).

The facial cues we use to infer the trustworthiness of individuals in the formation of first impressions have been experimentally studied by Willis and Todorov (2006). They found that people can judge faces very accurately in terms of different characteristics, including trustworthiness, even when the exposure to the facial stimulus is just 100 ms. This kind of judgment even improved the participant's confidence when the exposure was heightened to 500 ms. However, no differences were found when the participants were exposed to 1000 ms, confirming the idea that the judgment of faces in terms of trustworthiness needs is completed in less than a second.

Supported by this model, Sperber and colleagues suggest that when we see a new face, the very first thing we do is assess its trustworthiness, and, relying on this primary evaluation, we subsequently proceed in the communication. This posits the base to understand how facial trustworthiness perception can be strictly linked to ET and, more generally, IT. Indeed, the natural and immediate ability to judge a face as trustworthy or untrustworthy is also modulated during infancy thanks to primary care relationships, during which the infant is exposed to facial cues (Li et al., 2022; Wang et al., 2018). These cues are responsible for attachment development, thanks to which the IWM can be built and introduced among those social patterns' schemata involved in social perception; the very same IWM will be the basis for ET and IT development.

A growing body of research on children's ET and EV shows that in early childhood, individuals do not perceive information from others as always reliable (Corriveau et al., 2009; Heyman, 2008; Koenig & Harris, 2007; Sperber et al., 2010). Fonagy and Allison (2014), in a recent work building on natural pedagogy (Csibra & Gergely, 2009) and epistemic vigilance (Sperber et al., 2010) theories, present the role of ET in children's development through the lens of attachment theory. From an attachment perspective, attachment relationships can be considered the environment in which this trust can be favored. In the attachment context, EV is lowered repeatedly thanks to the natural pedagogy, working with the help ostensive cues that allow the communication to accept exchanged knowledge as reliable. Fonagy and Allison (2014) suggested that the natural pedagogy theory by Csibra and Gergely (2009) can explain how different attachment histories can favor or undermine the development process of ET. Specifically, ET has been recently theorized to be a relevant element that can favor the development of mentalization ability (Fonagy & Allison, 2014). EV is a cognitive mechanism present at very early stages of life and how specific relational

experiences, thanks to natural pedagogy can favor or disrupt ET development. Csibra and Gergely (2009) explain that humans are the only species that can communicate by social learning to transmit cultural knowledge; natural pedagogy is indeed a system that has originated from human evolution to favor collaboration among individuals.

The ostension process must rely on signals that unambiguously specify that the communication is directed toward the infant, discriminable by newborns, and must induce preferential orientation toward their source. Among these signals, there is eye contact, unique tones used with infants, and contingent reactivity to the infant's behavior in a turn-taking manner (Csibra, 2010). The ostension process, which promotes natural pedagogy, and maternal sensitivity, which promotes attachment security intended in terms of mentalizing or sensitivity to an intentional state, are overlapping constructs (Fonagy et al., 2007). Thus, mirroring interactions in which the mother marks with her expressions the baby's emotional states can be considered ostensive cues, which permit the relaxation of EV in favor of ET. Consequently, Fonagy and Allison (2014) hypothesized that infants with secure attachment patterns would perceive their caregivers as a reliable source of information because, more likely, a sensitive caregiver (thanks to which the child is more likely to develop a secure attachment) will have used ostension in his or her communication. Therefore, it is possible to consider the presence of ostensive cues in the caregiver's communications as predictors of secure attachment relationships (Beebe et al., 2010; Fonagy et al., 2007). Gergely (2013) suggests that the caregiver is naturally predisposed to contingently respond to the emotional expressions of the child. This, in turn, permits the infant themselves to be able to acquire further knowledge from the same caregiver in the future. Thanks to these "marked mirroring interactions" (Gergely et al., 2002), intended by Fonagy and Allison (2014) as an overlapping concept to ostensive cues, the caregiver can make the infant aware that the information about his or her emotion at that moment is relevant and generalizable.

In contrast, Fonagy and colleagues (2015) point out that traumatic experiences in early childhood obstruct ET development, keeping the individual in a state of epistemic hyper vigilance due to all the adversities experienced in abusive relations. This constant state of strong suspiciousness toward others is the core of borderline psychopathology (Fonagy et al., 2015).

Although there is still little research about ET development, some evidence supports this perspective (Campbell et al., 2021; Kampling et al., 2022). Both Corriveau and Colleagues

(2009) and Venta (2014) empirically investigated the role of attachment on the possibility for the child to accept incoming socially transmitted knowledge. Both studies found that the worse the attachment relation is, the less an individual will experience ET. Moreover, a previous history of adverse childhood experiences could moderate the relation between attachment and ET (Venta, 2014).

These results support the idea that attachment plays a fundamental role in the development of ET. In addition, ostensive cues are involved in the process of forming attachment bonds, suggesting that attachment security could be provided by the ostension process. However, further studies are needed, and longitudinal studies are crucial.

From Epistemic to Interpersonal Trust

ET is a recent concept, and to our knowledge, only a few studies have operationalized the concept beyond clinical reflection (Fonagy & Allison, 2014; Orme et al., 2019; Campbell et al., 2021; Kampling et al., 2022) and explored its influence on the quality of interpersonal exchanges. Also, some authors tried to develop specific measurements for this construct, such as the Epistemic Trust Assessment (Schröder-Pfeifer et al., 2018; Schröder-Pfeifer et al., 2022), and the Epistemic Trust Mistrust Credulity Questionnaire (Campbell et al., 2021). In this context, the broader concept of IT might come in handy.

One of the most successful attempts to operationalize trust is Rotenberg's concept of interpersonal trust (Rotenberg, 2010; 2019). His theory frames trust within a comprehensive model called Bases, Domains, and Targets (BDT; Rotenberg, 2010; Rotenberg, 2019). From this perspective, trust is an ability that relies on three bases (i.e., reliability, emotional trust, and honesty), can be applied to three domains (i.e., cognitive/affective, behavior dependent, and behavior-enacting) and has two targets (T) (i.e., familiarity and specificity) (Rotenberg, 2010; 2019).

First, *reliability* refers to an individual's ability to deliver on their promises and fulfill their word. Second, the *emotional trust* base refers to an individual tendency to avoid emotional harm conduct toward others. Third, *honesty* is defined as an individual's tendency to tell the truth, and their behavior is not driven by malicious intent, avoiding manipulative strategies.

Concerning the domains, the cognitive/affective domain refers to the individual's beliefs and feelings about the perception of the other, showing the three bases of trust. Then, the behavior-dependent domain refers to the ability of the individual to rely on others to act in a trusting way according to the three bases of trust. Finally, the behavior-enacting (trustworthiness) domain refers to individuals showing behavior guided by the three bases of trust.

Looking at the targets, specificity is intended as the trust toward a general category of people or a specific person, and familiarity is linked to how familiar the subject is.

Thus, IT can be considered the main means by which individuals discriminate who to approach and who to avoid, consequently being able to collaborate and develop relationships.

Rotenberg (2010; 2019) built his BDT model on preexisting theories that have investigated from different perspectives the dimensions at the base of social exchanges, including psychosocial theory (Erikson, 1963), attachment theory (Bowlby, 1969; 1973; 1979), social learning theory (Rotter, 1967;1971; 1980) and knowledge acquisition theory (Harris & Koenig 2006; 2012). Attachment theory suggests that IWMs can influence the IT domain of cognitive and emotional beliefs (Rotenberg, 2019). We believe that ET and IT are influenced bidirectionally by one another in a relational process in which both Rotenberg's model and ET theory can participate. We hypothesize that children live in a state of epistemic vigilance and that during their development, they can learn how to self-regulate and live in the world thanks to the cultural knowledge transmitted by caregivers in a secure attachment environment. When a child may experience distress caused by regular needs, their attachment system will be activated, resulting in a search for proximity and holding (Luyten et al., 2020). In a secure attachment relationship, the caregiver will be able to comfort the child thanks to sensitive care, in which ostensive cues are fundamental, resulting in emotion downregulation. This experience allows the child to open his or her epistemic trust channel to acquire this regulatory knowledge, and the continuous repetition of this kind of interaction will favor the development of a flexible and robust epistemic highway (Luyten et al., 2020). Thus, coherent, and repetitive exchanges of this type allow the baby to build IWM on which the IT beliefs will be based. Here, we propose an example to support this perspective.

Let's go back to the abovementioned situation in which a child is with their parents and sees for the first time a lit fireplace. A securely attached child will be able to explore the environment and will try to get closer to the fire to see it better. The caregivers will be activated by the perceived risk in this situation and will tell the child not to get too close; otherwise, they will be burned and hurt. With this recommendation, parents will also use ostensive cues, such as naming the child, eye contacting them, and mirroring the child's surprised emotional state to activate ET. In this relational situation, the child with secure attachment holds high beliefs that the parents are reliable, emotionally trustworthy, and honest. This is the starting point of IT, built on past experiences and IWMs of a securely attached relationship. In this context, the process of ostension, in which the face also plays a fundamental role in eye contact and marked mirroring, allows us to lower the natural implicit suspiciousness of the child and activate the ET channel. Thus, the consequence of this perceived trustworthiness in the information provided by the parents, the child will likely show dependent behavior by relying on the parents' word and reassurance and engage in trust-enacting behavior by staying far from the fire. In this sense, ET is considered something strictly entangled with IT.

Imagining the abovementioned situation within an insecure or disorganized attachment environment, the surprise experience by the baby could be neglected by caregivers, resulting in the child being left free to approach the fire. Subsequently, if the child could hurt themselves, they will experience distress caused by the situation, with their caregivers unable to handle this and to regulate the child's emotions. This may cause the child to be overwhelmed by these intense emotions, interrupting the possibility of communicating with the other in a trustful way. The repetition of such interactions might result in a rigid epistemic mistrust of the individual, in which IWMs could be characterized by avoidance and fear of others, obstructing the possibility of building trustful beliefs toward others.

Rotenberg and colleagues proposed a developmental model divided into several steps (2013) according to which behavior-dependent honesty is the main kind of trust from 0 to 2 years of age, directed both toward parents and strangers. Subsequently, between 2 and 6 years of age, children show trusting behavior toward an increased group of people relying on cognitive bases of honesty and reliability. Later, between 7 and 12 years of age, individuals behave in a trustful way, relying on the bases of honesty, reliability, and emotional bases thanks to which they can understand moral principles and social expectations violations.

From adolescence, all bases and domains are extended to less proximal targets and more general themes such as politics and others (Rotenberg, 2019).

Zhou and colleagues (2018) and Sakai (2010) investigated the role of genetic and environmental factors in twin studies and found that shared and nonshared environments contribute to trust belief development. Rotenberg (2019) points out that attachment seems to be a relevant environmental factor that influences the formation of interpersonal trust. Indeed, repetitive relational experiences of a trustworthy caregiver permit the development of mental representations of others seen as not harmful. Moreover, interpreting Sakai's findings about the role of environmental factors in trust development (2010), attachment relation seems to contribute to children's sense of trust both within and outside the family. Rotenberg then adds that attachment theory can explain only the contribution of caregivers to the development of trust and claims that this is formed in an interdependent exchange between the child and his or her parents, with a relevant role played by the child himself (Rotenberg, 2019). From this perspective, there is a bidirectional relationship between parents and child trust beliefs and behaviors that result in a complex structure built on past experiences that influences mental representations of relationships concerning the trust dimension. As the abovementioned traumatic experiences can result in insecure or, more often, disorganized attachments, these life events are responsible for trust disruption.

Indeed, the three bases individuated by Rotenberg (2010; 2019) are involved in the expression of trust in the three domains built on the IWM. Here, we hypothesize that in this structural model, ET represents the relational process that allows going from a representational level (cognitive and emotional beliefs domain) to a behavioral level (behavior-dependent and behavior-enacting domains). In summary, from mental representations of the other as more or less trustworthy, thanks to special signals during communicative exchange, the individual can answer behaviorally, relying on the knowledge exchanged.

Disentangling perceived trust: let me see your face.

Todorov and colleagues (2015) underlined the paramount role of facial expressions in interpersonal relationships, highlighting specific traits such as trustworthiness, i.e., the perception of trust and reliability in others. Additionally, Fiske, Cuddy, and Glick (2007)

explain how trustworthiness can be used as a cue to approach or avoid someone or to avoid. Moreover, this process of face trustworthiness evaluation is very rapid and intuitive (Todorov, Pakrashi, & Oosterhof, 2008). Due to its relevance in human social exchanges, both in developmental and adult age, researchers have studied whether this ability is learned through socialization or is innate.

Since very early stages, even with very little social experience, children demonstrate a systematic preference toward human faces (Valenza, Simion, Cassia, & Umiltà, 1996), and they already show a preferential tendency toward trustworthy faces at the age of 6–8 months (Sakuta, Kanazawa, & Yamaguchi, 2018). Sakuta and colleagues (2018) used a preferential-looking paradigm to show that facial stimuli can be judged in terms of trustworthiness and dominance by babies. Their results support the idea that trustworthiness can be judged at very early stages of life, starting from perceptual cues such as face trait perception, allowing us to hypothesize that this ability to detect who to trust does not need high social competence (Sakuta et al., 2018).

In an ERP study, Jessen and Grossman (2019) found that the capacity to detect face trustworthiness is also present at the age of 7-month infants subliminally presented with face stimuli.

Ewing and colleagues (2015) have demonstrated that the capacity of face evaluation according to specific characteristics in children is comparable in accuracy to the same ability in adults. Additionally, Cogsdill and colleagues (2014) found that three-year-old children tend to judge trustworthy faces as friendly. In addition, the same authors (Cogsdill et al., 2014) found that from 5 years of age, children can explicitly attribute trustworthiness to faces. Therefore, according to these authors, it seems that explicit face evaluation of trustworthiness is already present since toddlerhood.

Indeed, facial stimuli during attachment interactions in infancy are central to the development of ET. Building on the works of Tronick's Still Face paradigm, where the exchange of facial expressions between the caregiver and the baby is the core of emotional regulation and conceptualizing marked mirroring interactions intended as ostensive cues by Fonagy and Allison (2014), Cohn and colleagues (1991) found that some negative or positive responses to Still face at the age of the baby of 6 months could predict attachment security

or insecurity, highlighting the crucial role of facial exchanges in attachment development and consequentially in IT and ET construction (Beebe et al., 2010).

Additionally, some authors have investigated the ability to evaluate trust from facial stimuli in adults and children using Rotenberg's IT theorization (Ewing, Caufield, Read, & Rhodes, 2014). Specifically, the authors asked participants to judge how trustworthy the face stimuli were and to explain to the children what trust is, they used the definition of Rotenberg's BDT. Once explained how trust was defined, they tested the comprehension of the construct in children using some items of the Early Childhood Generalized Trust Belief Scale (ECGTBS; Betts, Rotenberg, & Trueman, 2009).

Here, we propose the idea that there could be a developmental trajectory in which a baby at very early stages of life is constantly exposed to caregivers' faces to which they are able to attribute a certain degree of trustworthiness in a rudimentarily implicit way. As mentioned above, in moments of distress, the baby seeks proximity when the attachment system is activated to downregulate their emotions or satisfy their needs. When caregivers respond in a sensitive way, they use ostensive cues in the context of marked mirroring interactions, in which the baby can experience the feeling of being recognized and helped in the downregulation of emotions. This cycle helps the baby to open his or her channel of epistemic trust, thanks to which interpersonal trust can also be developed. Consequently, the primitive ability to attribute trustworthiness to a certain face is improved with the development of a deeper sense of trust that involves more elements of the relationship that are also included in the immediate perception of facial stimuli. Thus, the implicit and perceptual level of trustworthiness attribution can influence and be influenced by ET and IT. To better clarify the model proposed in this work about the development of ET and IT, the influence of facial judgements, and the disruption of this, we here propose two images that allows to visualize the hypothesized processes (Figure 1 and 2).

Figure 1. Development of ET and IT in a secure attachment context.

[Insert Figure 1 here]

Ostensive cues, trustworthiness, and trust

Facial expressions seem to play a role in forming and sustaining trust in children, suggesting a potential role of ostensive cues in supporting children development of ET. In regard to first impressions formations, children's perception of face trustworthiness depends not only on individual face characteristics, but also on facial expressions (Caulfield et al., 2014; Tang et al., 2019). In fact, it has been hypothesized that trustworthiness detection relies on an overextension of the ability to sensitively respond to facial expressions (Engell et al., 2010; Jessen & Grossmann, 2016; Said et al., 2009), to the extent that 5 years old's preference for more trustworthy face is associated with the ability to recognize emotions (Baccolo & Macchi Cassia, 2020). In addition, adult studies show that facial gaze (Bayliss et al., 2009; Bayliss & Tipper, 2006; Sutherland et al., 2017) and the gaze of another person (Kaisler & Leder, 2016) modulate the trustworthiness of the face, while no studies addressed these issues in children. To our knowledge, no studies provided information on the role of facial expressions in the development specifically of epistemic and interpersonal trust, especially in the very first months of life. However, evidence shows that children rely on previous experiences (Siddique et al., 2022) in order to modulate the perception of trustworthiness of the face. Notably, it has been suggested a parental role in forming impressions, since a recent study provide evidence of a parental role in reinforcing the first impression formation in 5-6 year old children (Eggleston et al., 2021). A much longer tradition of research investigated facial expressions in social referencing processes, that are directly associated with trust, since children are assumed to select the referring sources on the basis of their trustworthiness and expertise (Feinman, 1982; Feinman et al., 1992). Evidence shows a powerful role of caregiver's facial emotional expressions in regulating infants' behavior in context of uncertainty or ambiguity (Sorce et al., 1985; Striano et al., 2006; Vaish & Striano, 2004). We could expect that the contingency between emotional expression and the behavioral response of the caregiver reinforces the association between trust and social reference. Parental expressed anxiety seems to be associated to infant behavioral avoidance in a social referencing task such as the visual cliff task (Möller et al., 2014), suggesting a role of ostensive cues in the process of trusting parental reference. In fact, infants' neural processing of facial cues seems highly dependent on parental psychological (Bowman et al., 2022; Sandre et al., 2022) and behavioral characteristics (Boomen et al., 2021; Rayson et al., 2017), to the extent that infants with insecure attachment do not show age typical neural discrimination between fearful and non-fearful

faces (Peltola et al., 2020), and in middle childhood avoidant children fail in discriminating between stranger's and caregiver's face (Kungl et al., 2022).

Model clinical application: the case of borderline conditions

We focus here on BPD as it has been theorized as one of the possible psychopathological outcomes of ET and IT development disruption. Moreover, empirical findings concerning ET, IT and facial trustworthiness judgement impairments in BPD may support this psychopathology developmental model.

BPD is a severe psychiatric condition characterized by emotion dysregulation, impulsiveness, self-harm, suicidal behaviors, and severe interpersonal impairment (APA, 2013; Benzi et al., 2020). Specifically, interpersonal impairment has been investigated as a core feature of borderline personality pathology (Gunderson, 2007; Gunderson et al., 2018; Section III, APA, 2013). One of the main elements of borderline impaired interpersonal functioning is the perception of the other as malevolent and untrustworthy. This trust impairment has been identified in the literature through different perspectives, such as facial evaluation (Nicol et al., 2013), hypermentalization (Sharp & Vanwoerden, 2015), and oxytocin levels (Servan et al., 2018). Individuals diagnosed with BPD seem to be biased in face evaluation of trustworthiness, perceiving faces as more untrustworthy than the general population (Nicol et al., 2013). Additionally, complex trauma, identified as attachment trauma that often characterizes these patients' histories, occurs in primary relationships where they are exposed to highly salient stimuli. Here, we hypothesize that starting from early life experiences of maladaptive attachment patterns, during which maternal sensitivity is not delivered through the correct ostension process, the caregiver's face represents a repetitive malevolent stimulus that is then generalized to others.

Fonagy and colleagues (2015) theorized that individuals diagnosed with BPD live in an epistemic hypervigilance state that causes unstable relationships. They suggested a developmental model of BPD in which early life adverse experiences, misuse of ostensive cues, and insecure attachment relationships are the core of this clinical condition. According to this model, ET is triggered by ostensive cues such as eye contact between the caregiver and the children and turn-taking in interactions. Thus, in maladaptive early interactions, when the normal process of ostension is not permitted, infants maintain a continuous state

of suspiciousness toward others that might foster borderline-like epistemic hypervigilance. These considerations about the connection between borderline pathology and ET are supported by preliminary findings by Orme and colleagues (2019), according to which the expression of borderline symptomatology is correlated with ET. Specifically, a clinical adolescent sample was administered with self-reports assessing ET, BPD symptomatology and BPD categorical diagnosis, and significant correlations between these elements were found. Also, preliminary findings by Campbell and colleagues (2021) and Kampling and colleagues (2022) support the association between epistemic trust disruption, traumatic experiences, and psychopathology.

Currently, there is no structured theorization on IT for borderline pathology. However, it is possible to track several studies investigating trusting behavior difficulties in this population (Fertuck et al., 2013; King-Casas et al., 2008; Richetin et al., 2019). Indeed, the abovementioned studies investigated how individuals diagnosed with BPD tend to be more suspicious and less trustful toward others in trust games tasks. Moreover, considering the connections between IT bases of reliability, emotional trust and honesty and ET might widen our understanding of maladaptive interpersonal functioning, providing a broader framework that might shape clinical understanding and intervention. Indeed, the possibility of comprehending whether trust difficulties are the core of some personality pathologies may lead to technical issues in psychotherapy treatment. For example, if the patient cannot structurally trust others, the initial focus of the intervention should be on restoring the ability to perceive others in a trustful way, especially perceiving what others are communicating as trustworthy and not deceiving.

This brief overview highlights the crucial role of ET in understanding interpersonal exchanges and suggests its inclusion in the broader context of perceived trust (IT). Additionally, it emphasizes the contribution of early experiences of facial cues evaluation in developing trust as a common ground between ET and IT. Future studies might investigate a specific network of associations between these constructs. In the end, these considerations about IT and ET also allow us to further comprehend the interpersonal issues of borderline pathology in terms of trust impairments to emphasize the need to focus on these dimensions in psychotherapy treatment. Specifically, with this clinical population, starting from the restoration of the possibility to trust what the other (in this case the therapist) says could be the first step toward the construction of a generalized sense of the other as trustworthy.

Conclusions

This study presents an innovative conceptualization in which ET, IT and facial trustworthiness evaluation are linked together theoretically, building on the existing literature. After the presentation of ET and IT development process, we explain how faces are judged in terms of trustworthiness in early stages of life and eventually how all these elements are interwoven in interpersonal impairments in borderline personality disorder.

The aim of this study is proposing a theoretical framework in which ET, IT and facial trustworthiness can be seen as elements of a more complex developmental model of a wider sense of trust. This could contribute to our knowledge about these processes and be further supported by empirical works investigating these constructs.

Indeed, some authors developed self-report measures for the assessment both of IT and ET, but to our knowledge no study investigated the association between these two constructs. Moreover, these tools could be used to test convergent and divergent validity mutually. The same point can be addressed for facial trustworthiness evaluation, which has never been investigated in association with ET or IT. In addition, longitudinal study could profoundly contribute to test these theoretical hypotheses, in which researcher could investigate the developmental path here proposed. We think that empirical studies concerning the association among these concepts could create a more robust ground onto which clinicians could rely on in using these assessment tools. Also, the application of this developmental model, that include both a normative and a pathological explanation of the process, to the borderline pathology could contribute to the comprehension of such disorder, which is characterized by several interpersonal impairments often related to trust issues.

Moreover, from a wider perspective, this contribution could favor the empirical investigation involving how attachment and consequentially IT and ET are associated through the using of ostensive cues. Specifically, future studies could investigate how facial trustworthiness is bidirectionally linked to the presence of ostension during early communications, and how this represent and constitute a solid ground onto which attachment and trust can be built. On the contrary, longitudinal studies investigating such relationships could contribute to our knowledge both on typical and atypical development of such variables.

In conclusion, we believe that this work could pique the interest under two perspectives: a developmental one, and psychopathological one. On one hand, developmental researchers might appreciate this theoretical contribution that allows to enrich our comprehension of trust development that integrates theories coming from different perspective. On the other hand, clinical researchers and clinicians might appreciate this attempt to explain how the disruption of some processes in developmental phases could contribute to borderline pathology and consequentially to design interventions based on these assumptions.

Statement of Ethics

Not applicable

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

Not applicable.

Author Contributions

Conceptualization, A.M., P.D.C., F.L., I.M.A.B., C.C., P.F. and L.P.; writing—original draft preparation, A.M., P.D.C., F.L., I.M.A.B.; writing—review and editing C.C., P.F., L.P.; supervision, C.C., P.F., L.P. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

No data are presented.

References (Alphabetical)

- Allen, J. G. (2013). Treating attachment trauma with plain old therapy.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.
- 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>
- Bayliss, A. P., Griffiths, D., & Tipper, S. P. (2009). Predictive gaze cues affect face evaluations: The effect of facial emotion. *European Journal of Cognitive Psychology, 21*(7), 1072–1084. <https://doi.org/10.1080/09541440802553490>
- Bayliss, A. P., & Tipper, S. P. (2006). Predictive Gaze Cues and Personality Judgments. *Psychological Science, 17*(6), 514–520. <https://doi.org/10.1111/j.1467-9280.2006.01737.x>
- Beebe, B., Jaffe, J., Markese, S., Buck, K., Chen, H., Cohen, P., ... & Feldstein, S. (2010). The origins of 12-month attachment: A microanalysis of 4-month mother–infant interaction. *Attachment & human development, 12*(1-2), 3-141.
- Benzi, I., Di Pierro, R., De Carli, P., Cristea, I. A., & Cipresso, P. (2020). All the faces of research on borderline personality pathology: drawing future trajectories through a network and cluster analysis of the literature. *Journal of Evidence-Based Psychotherapies, 20*(2).
- Benzi, I. M. A., Milesi, A., & Parolin, L. A. L. (2021). Unconscious in Psychology. *The Palgrave Encyclopedia of the Possible. Palgrave Macmillan, Cham.* https://doi.org/10.1007/978-3-319-98390-5_162-1.
- Betts, L. R., Rotenberg, K. J., & Trueman, M. (2009). The early childhood generalized trust belief scale. *Early Childhood Research Quarterly, 24*(2), 175-185.
- Boomen, C., Munsters, N. M., Deković, M., & Kemner, C. (2021). Exploring emotional face processing in 5-month-olds: The relation with quality of parent–child interaction and spatial frequencies. *Infancy, 26*(6), 811–830.
<https://doi.org/10.1111/infa.12420>
- Bowlby, J. (1969). Attachment and loss v. 3 (Vol. 1).
- Bowlby, J. (1973). Attachment and loss: Volume II: Separation, anxiety, and anger. In *Attachment and loss: Volume II: Separation, anxiety, and anger* (pp. 1-429). London: The Hogarth press and the institute of psychoanalysis.
- Bowlby, J. (1979). The making and breaking of affectional bonds. London. *Tavistock Publications.* Boyd, J. H., & Weissman, MM (1981). *Epidemiology of affective disorders: A re-examination of future directions. Archives of General Psychiatry, 38*, 1039-1046.

Bowman, L. C., McCormick, S. A., Kane-Grade, F., Xie, W., Bosquet Enlow, M., & Nelson, C. A. (2022). Infants' neural responses to emotional faces are related to maternal anxiety. *Journal of Child Psychology and Psychiatry*, *63*(2), 152–164. <https://doi.org/10.1111/jcpp.13429>

Campbell, C., Tanzer, M., Saunders, R., Booker, T., Allison, E., Li, E., ... & Fonagy, P. (2021). Development and validation of a self-report measure of epistemic trust. *PloS one*, *16*(4), e0250264.

Caulfield, F., Ewing, L., Burton, N., Avard, E., & Rhodes, G. (2014). Facial trustworthiness judgments in children with ASD are modulated by happy and angry emotional cues. *PloS one*, *9*(5), e97644.

Caulfield, F., Ewing, L., Bank, S., & Rhodes, G. (2016). Judging trustworthiness from faces: Emotion cues modulate trustworthiness judgments in young children. *British Journal of Psychology*, *107*(3), 503-518.

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.

Cohn, D. A. (1990). Child-mother attachment of six-year-olds and social competence at school. *Child development*, *61*(1), 152-162.

Cohn, J. F., Campbell, S. B., & Ross, S. (1991). Infant response in the still-face paradigm at 6 months predicts avoidant and secure attachment at 12 months. *Development and Psychopathology*, *3*(4), 367-376.

Corriveau, K. H., Harris, P. L., Meins, E., Fernyhough, C., Arnott, B., Elliott, L., Liddle, B., Hearn, A., Vittorini, L., & de Rosnay, M. (2009). Young Children's Trust in Their Mother's Claims: Longitudinal Links with Attachment Security in Infancy. *Child Development*, *80*(3), 750–761.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, *25*(2), 141-168.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in cognitive sciences*, *13*(4), 148-153.

De Carli, P., Costantini, I., Sessa, P., Visentin, S., Pearson, R. M., & Simonelli, A. (2019). The expectant social mind: A systematic review of face processing during pregnancy and the effect of depression and anxiety. *Neuroscience & Biobehavioral Reviews*. <https://doi.org/10.1016/j.neubiorev.2019.04.013>

Dodge, K. A., Pettit, G. S., McClaskey, C. L., Brown, M. M., & Gottman, J. M. (1986). Social competence in children. *Monographs of the society for research in child development*, i-85.

Eggleston, A., McCall, C., Cook, R., & Over, H. (2021). Parents reinforce the formation of first impressions in conversation with their children. *PLOS ONE*, *16*(8), e0256118. <https://doi.org/10.1371/journal.pone.0256118>

Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: automatic coding of face properties in the human amygdala. *Journal of cognitive neuroscience*, *19*(9), 1508-1519.

Engell, A. D., Todorov, A., & Haxby, J. v. (2010). Common Neural Mechanisms for the Evaluation of Facial Trustworthiness and Emotional Expressions as Revealed by Behavioral Adaptation. *Perception*, *39*(7), 931–941. <https://doi.org/10.1068/p6633>

Erikson, E. H. (1963). *Childhood and society* 2nd ed. *Erikson-New York: Norton*.

Ewing, L., Caulfield, F., Read, A., & Rhodes, G. (2015). Appearance-based trust behaviour is reduced in children with autism spectrum disorder. *Autism*, *19*(8), 1002-1009.

Feinman, S. (1982). Social referencing in infancy. *Merrill-Palmer Quarterly*, *28*(4), 445–470.

Feinman, S., Roberts, D., Hsieh, K.-F., Sawyer, D., & Swanson, D. (1992). A Critical Review of Social Referencing in Infancy. In *Social Referencing and the Social Construction of Reality in Infancy* (pp. 15–54). Springer US. https://doi.org/10.1007/978-1-4899-2462-9_2

Fertuck, E. A., Grinband, J., & Stanley, B. (2013). Facial trust appraisal negatively biased in borderline personality disorder. *Psychiatry research*, *207*(3), 195-202.

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in cognitive sciences*, *11*(2), 77-83.

Fonagy, P., & Allison, E. (2014). The role of mentalizing and epistemic trust in the therapeutic relationship. *Psychotherapy*, *51*(3), 372.

Fonagy, P., Gergely, G., & Target, M. (2007). The parent–infant dyad and the construction of the subjective self. *Journal of Child Psychology and Psychiatry*, *48*(3-4), 288–328.

Fonagy, P., Luyten, P., & Allison, E. (2015). Epistemic petrification and the restoration of epistemic trust: A new conceptualization of borderline personality disorder and its psychosocial treatment. *Journal of personality disorders*, *29*(5), 575-609.

Gergely, G. (2013). Ostensive communication and cultural learning: The natural pedagogy hypothesis. In J. Metcalfe & H. S. Terrace (Eds.), *Agency and joint attention* (p. 139). Oxford, UK: Oxford University Press.

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*(6873), 755-755.

Gunderson, J. G. (2007). Disturbed Relationships as a Phenotype for Borderline Personality Disorder. *American Journal of Psychiatry*, *164*(11), 1637–1640.

Gunderson, J. G., Fruzzetti, A., Unruh, B., & Choi-Kain, L. (2018). Competing Theories of Borderline Personality Disorder. *Journal of Personality Disorders, 32*(2), 148–167.

Harris, P. L., & Koenig, M. A. (2006). Trust in testimony: How children learn about science and religion. *Child development, 77*(3), 505-524.

Harris, P. L., & Koenig, M. (2012). Imagination and testimony in cognitive development: The cautious disciple? In *Imaginative minds*. Oxford University Press.

Heyman, G. D. (2008). Children's Critical Thinking When Learning from Others. *Current Directions in Psychological Science, 17*(5), 344–347.

Imber, S. (1973). Relationship of trust to academic performance. *Journal of Personality and Social Psychology, 28*(1), 145.

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.

Jessen, S., & Grossmann, T. (2019). Neural evidence for the subliminal processing of facial trustworthiness in infancy. *Neuropsychologia, 126*, 46-53.

Kaisler, R. E., & Leder, H. (2016). Trusting the Looks of Others: Gaze Effects of Faces in Social Settings. *Perception, 45*(8), 875–892.
<https://doi.org/10.1177/0301006616643678>

Kamphuis, J. H., & Finn, S. E. (2018). Therapeutic assessment in personality disorders: Toward the restoration of epistemic trust. *Journal of personality assessment*.

Kampling, H., Hettich, N., Braehler, E., Sachser, C., Kruse, J., Lampe, A., ... & Riedl, D. (2022). Epistemic trust and personality functioning as mediators for the association of adverse childhood experiences and posttraumatic as well as complex posttraumatic stress symptoms in adulthood. *JOURNAL OF PSYCHOSOMATIC RESEARCH, 157*.

King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., & Montague, P. R. (2008). The rupture and repair of cooperation in borderline personality disorder. *science, 321*(5890), 806-810.

Koenig, M. A., & Harris, P. L. (2005). Preschoolers Mistrust Ignorant and Inaccurate Speakers. *Child Development, 76*(6), 1261–1277.
<https://doi.org/10.1111/j.1467-8624.2005.00849.x>

Koenig, M. A., & Harris, P. L. (2007). The Basis of Epistemic Trust: Reliable Testimony or Reliable Sources? *Episteme, 4*(3), 264–284.

Kungl, M., Vrticka, P., Heinisch, C., Beckmann, M. W., Fasching, P. A., Ziegler, C., & Spangler, G. (2022). Deactivating attachment strategies associate with early processing of facial emotion and familiarity in middle childhood: an ERP study. *Attachment & Human Development, 1–19*.
<https://doi.org/10.1080/14616734.2022.2132050>

Li, X., Jorgensen, N. A., McElwain, N. L., & Telzer, E. H. (2022). Toddler–mother attachment moderates adolescents' behavioral and neural evaluation of trustworthiness. *Social Cognitive and Affective Neuroscience*.

Lieberman, A. F. (1977). Preschoolers' competence with a peer: Relations with attachment and peer experience. *Child Development*, 1277-1287.

Luyten, P., Campbell, C., & Fonagy, P. (2020). Borderline personality disorder, complex trauma, and problems with self and identity: A social-communicative approach. *Journal of personality*, 88(1), 88-105.

Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition*, 112(3), 367-380.

Meltzer, H., Vostanis, P., Goodman, R., & Ford, T. (2007). Children's perceptions of neighbourhood trustworthiness and safety and their mental health. *Journal of child psychology and psychiatry*, 48(12), 1208-1213.

Möller, E. L., Majdandžić, M., & Bögels, S. M. (2014). Fathers' versus mothers' social referencing signals in relation to infant anxiety and avoidance: a visual cliff experiment. *Developmental Science*, 17(6), 1012–1028.
<https://doi.org/10.1111/desc.12194>

Nicol, K., Pope, M., Sprengelmeyer, R., Young, A. W., & Hall, J. (2013). Social judgement in borderline personality disorder. *PLoS One*, 8(11), e73440.

Orme, W., Bowersox, L., Vanwoerden, S., Fonagy, P., & Sharp, C. (2019). The relation between epistemic trust and borderline pathology in an adolescent inpatient sample. *Borderline personality disorder and emotion dysregulation*, 6(1), 1-9.

Peltola, M. J., van IJzendoorn, M. H., & Yrttiaho, S. (2020). Attachment security and cortical responses to fearful faces in infants. *Attachment & Human Development*, 22(2), 174–188. <https://doi.org/10.1080/14616734.2018.1530684>

Randall, B. A., Rotenberg, K. J., Totenhagen, C. J., Rock, M., & Harmon, C. (2010). 12 A new scale for the assessment of adolescents' trust beliefs. *Interpersonal trust during childhood and adolescence*, 247.

Rayson, H., Bonaiuto, J. J., Ferrari, P. F., & Murray, L. (2017). Early maternal mirroring predicts infant motor system activation during facial expression observation. *Scientific Reports*, 7(1), 11738. <https://doi.org/10.1038/s41598-017-12097-w>

Richetin, J., Poggi, A., Ricciardelli, P., Fertuck, E. A., & Preti, E. (2018). The Emotional Components of Rejection Sensitivity as A Mediator Between Borderline Personality Disorder and Biased Appraisal of Trust in Faces. *Clinical Neuropsychiatry*, 15(4).

Rotenberg, K. J. (Ed.). (2010). *Interpersonal trust during childhood and adolescence*. Cambridge University Press.

Rotenberg, K. J. (2019). *The Psychology of Interpersonal Trust: Theory and Research*. Routledge.

Rotenberg, K. J., Petrocchi, S., Lecciso, F., & Marchetti, A. (2013). Research Article Children's Trust Beliefs in Others and Trusting Behavior in Peer Interaction.

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of personality*.

Rotter, J. B. (1971). Generalized expectancies for interpersonal trust. *American psychologist*, 26(5), 443.

Rotter, J. B. (1980). Interpersonal trust, trustworthiness, and gullibility. *American psychologist*, 35(1), 1.

Sabbagh, M. A., & Baldwin, D. A. (2001). Learning Words from Knowledgeable versus Ignorant Speakers: Links Between Preschoolers' Theory of Mind and Semantic Development. *Child Development*, 72(4), 1054–1070. <https://doi.org/10.1111/1467-8624.00334>

Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion*, 9(2), 260–264. <https://doi.org/10.1037/a0014681>

Sakai, A. (2010). Children's sense of trust in significant others: Genetic versus environmental contributions and buffer to life stressors. *Interpersonal trust during childhood and adolescence*, 56-84.

Sakuta, Y., Kanazawa, S., & Yamaguchi, M. K. (2018). Infants prefer a trustworthy person: An early sign of social cognition in infants. *PloS one*, 13(9), e0203541.

Sandre, A., Freeman, C., Renault, H., Humphreys, K. L., & Weinberg, A. (2022). Maternal symptoms of depression and anxiety during the postpartum period moderate infants' neural response to emotional faces of their mother and of female strangers. *Cognitive, Affective, & Behavioral Neuroscience*, 22(6), 1370–1389. <https://doi.org/10.3758/s13415-022-01022-y>

Schröder-Pfeifer, P., Georg, A. K., Talia, A., Volkert, J., Ditzen, B., & Taubner, S. (2022). The Epistemic Trust Assessment—An experimental measure of epistemic trust. *Psychoanalytic Psychology*, 39(1), 50.

Schroder-Pfeifer, P., Talia, A., Volkert, J., & Taubner, S. (2018). Developing an assessment of epistemic trust: a research protocol. *Research in Psychotherapy: Psychopathology, Process, and Outcome*, 21(3).

Servan, A., Brunelin, J., & Poulet, E. (2018). The effects of oxytocin on social cognition in borderline personality disorder. *L'encephale*, 44(1), 46-51.

Sharp, C., & Vanwoerden, S. (2015). Hypermentalizing in borderline personality disorder: A model and data. *Journal of Infant, Child, and Adolescent Psychotherapy*, 14(1), 33-45.

Siddique, S., Jeffery, L., Palermo, R., Collova, J. R., & Sutherland, C. A. M. (2022). Children's dynamic use of face- and behavior-based cues in an economic trust game. *Developmental Psychology*. <https://doi.org/10.1037/dev0001438>

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, 25(4), 359-393.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Malden, MA: Blackwell.

Sorce, J. F., Emde, R. N., Campos, J. J., & Klinnert, M. D. (1985). Maternal emotional signaling: Its effect on the visual cliff behavior of 1-year-olds. *Developmental Psychology*, 21(1), 195–200. <https://doi.org/10.1037/0012-1649.21.1.195>

Striano, T., Vaish, A., & Benigno, J. P. (2006). The meaning of infants' looks: Information seeking and comfort seeking? *British Journal of Developmental Psychology*, 24(3), 615–630. <https://doi.org/10.1348/026151005X67566>

Sutherland, C. A. M., Young, A. W., & Rhodes, G. (2017). Facial first impressions from another angle: How social judgements are influenced by changeable and invariant facial properties. *British Journal of Psychology*, 108(2), 397–415. <https://doi.org/10.1111/bjop.12206>

Szufnarowska, J., Rohlfing, K. J., Fawcett, C., & Gredebäck, G. (2014). Is ostension any more than attention? *Scientific Reports*, 4(1), 1-4.

Tang, Y., Harris, P. L., Zou, H., & Xu, Q. (2019). The impact of emotional expressions on children's trust judgments. *Cognition and Emotion*, 33(2), 318–331. <https://doi.org/10.1080/02699931.2018.1449735>

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, 519-545.

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, 27(6), 813-833.

Vaish, A., & Striano, T. (2004). Is visual reference necessary? Contributions of facial versus vocal cues in 12-month-olds' social referencing behavior. *Developmental Science*, 7(3), 261–269. <https://doi.org/10.1111/j.1467-7687.2004.00344.x>

Valenza, E., Simion, F., Cassia, V. M., & Umiltà, C. (1996). Face preference at birth. *Journal of experimental psychology: Human Perception and Performance*, 22(4), 892.

Venta, A. (2020). Attachment Facilitates Acculturative Learning and Adversity Moderates: Validating the Theory of Epistemic Trust in a Natural Experiment. *Child Psychiatry & Human Development*, 51(3), 471–477.

Wang, X., Wang, N., Han, S., Liu, S., & Zhang, L. (2018). The influence of facial trustworthiness on helping behavior: The role of attachment type. *Acta Psychologica Sinica*.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological science*, 17(7), 592-598.

Zhou, H. Y., Wong, K. K. Y., Shi, L. J., Cui, X. L., Qian, Y., Jiang, W. Q., ... & Chan, R. C. (2018). Suspiciousness in young minds: Convergent evidence from non-clinical, clinical and community twin samples. *Schizophrenia research*, 199, 135-141.

