Journal of Experimental Social Psychology 109 (2023) 104517



Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp





The "Happy Face Killer" in the eyes of the beholder: Relational encoding of facial emotions in context influences trustworthiness attributions☆

Simone Mattavelli *, Gaia Carlotta Fiamberti, Matteo Masi, Marco Brambilla

University of Milano-Bicocca, Italy

ARTICLE INFO

Keywords:
Face perception
Context
Emotions
Trustworthiness
Relational encoding

ABSTRACT

Research on face perception has established that faces surrounded by threatening contexts are perceived as less trustworthy. Moreover, recent studies revealed that such a face-context integration effect is moderated by the nature of the relational qualifier connecting the face and the context: presenting a face as belonging to either the perpetrator or the victim of the threatening context changes its perceived trustworthiness. Here, we asked whether relational qualifiers can also be extracted from subtle facial cues. In two preregistered studies ($N = 225$), we tested whether face-context integration is qualified by facial emotions. In Experiment 1, faces appeared with either happy or fearful emotional expressions in threatening contexts (vs. no context). Facial emotions moderated face-context integration: The negative impact of contextual threat on attributed trustworthiness showed stronger for happy than fearful faces. In Experiment 2, participants judged the emotional stimuli on both trustworthiness and smartness. Emotions altered face-context integration when judging trustworthiness but did not alter smartness judgments. Moreover, participants' tendency to judge happy faces in threatening contexts as less trustworthy correlated with their belief that the target face belonged to the criminal on the scene. The importance of considering relational encoding when studying person perception is discussed.

Between 1990 and 1995, a man later identified as Keith Jesperson strangled eight women and dumped their bodies along the road. In 1994, Jesperson wrote anonymously to a local newspaper bragging about his crimes. In his letter, he described the murders and how he disposed of the bodies. Disconcertingly, his confession was signed with the symbol of a smiley face. Since then, the reporter Phil Stanford coined the moniker "Happy Face Killer". It goes without saying that Jesperson's reputation was already destroyed by his horrible acts. Yet, that smiley face at the end of his letter made him even more evil: in that scribble, people could vividly figure Jesperson's face, smiling at his own crimes. This dark story makes salient the key issues addressed in the present paper, that are, facial emotions (e.g., happiness), threatening scenarios (e.g., a crime scene) and their interplay in social judgments.

1. Face-context integration and relational encoding

The sight of a face is often enough to determine whether a person is friendly, warm, or trustworthy (Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). Morphological facial features can elicit dispositional attributions, in line with the human tendency to overgeneralize facial cues to stable trait inferences (Zebrowitz, Fellous, Mignault, & Andreoletti, 2003). However, in real life faces are rarely processed in isolation. Information conveyed by the surrounding environment can be integrated in the ultimate attribution made on the target face. As a case in point, recent studies showed that such attributions can result from the integration between the face and the relevant context, either visual or auditory (Brambilla, Biella, & Freeman, 2018; Brambilla, Masi, Mattavelli, & Biella, 2021). Face-context integration occurs when the attributions of a dispositional trait (e.g., trustworthiness) to faces are influenced by the type of context in which such faces appear. For instance, Brambilla et al. (2018) exposed participants to trustworthy- and untrustworthy-looking faces presented in contexts selected to be neutral (e.g., a rural landscape), negative (e.g., an abandoned building), or threatening (e.g., a bloody knife) and asked them to categorize each facial stimulus as trustworthy vs untrustworthy. Results showed that untrustworthy faces were more easily categorized as such when surrounded by threatening rather than negative or neutral visual contexts.

Beside corroborating the well-established idea of a privileged bond

☆ This paper has been recommended for acceptance by Prof. Shlomo Hareli.

* Corresponding author at: University of Milano-Bicocca-Bicocca, Department of Psychology, 1, Piazza dell'Ateneo Nuovo, 20126 Milan, Italy.
E-mail address: simone.mattavelli@unimib.it (S. Mattavelli).

https://doi.org/10.1016/j.jesp.2023.104517
Received 23 January 2023; Received in revised form 14 July 2023; Accepted 17 July 2023
Available online 22 July 2023
0022-1031/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

between the two dimensions of trustworthiness and threat (Brambilla & Leach, 2014; Willis & Todorov, 2006), these findings fit with a functional account of face perception: as our judgments of others' trustworthiness are highly related to whether they represent an opportunity or a threat (Ames, Fiske, & Todorov, 2011; Cosmides & Tooby, 1992), contextual information should qualify our judgments by proving some information about the target person being judged. More specifically, a face might be perceived as less trustworthy in a threatening context due to the assumptions on the role played by the target in the specific context.

Mattavelli, Masi, and Brambilla (2023) offered initial confirmation of the idea that relational encoding can account for the impact of threatening contexts on perceived facial trustworthiness. Indeed, they showed that not all types of threatening contexts lowered perceived trustworthiness in the same fashion. Across three experiments (Experiment 1a-1c), the authors exposed participants to trustworthy- and untrustworthy-looking faces presented either in threatening contexts that were ascribable to a human action (e.g., a bloody knife) or in (more) threatening contexts that were not ascribable to a human action (e.g., a tornado). They found that the perceived trustworthiness of both trustworthy- and untrustworthy-looking faces decreased more strongly when such faces were embedded in threatening context that were ascribable (vs not ascribable) to the human action. Additionally, using neutral faces (Experiment 2), the authors found that face-context integration was altered when the relationship between the face and the context was manipulated via written instructions informing that the face belonged to either the perpetrator or the victim of the surrounding context. Indeed, faces were judged as less trustworthy when they belonged to the perpetrator of the threatening surrounding context. Taken together, these findings unveiled that perceivers can modulate their response when overtly informed about the relationship between the target and the context. However, Mattavelli et al. (2023) did not clarify whether relational qualifiers conveyed more subtly (e.g., by facial expressions) can also moderate face-context integration.

2. Face-context relational cues: the case of emotions

When processing faces in context, perceivers rarely dispose of explicit information that speaks for the role played by a target individual in a specific scene. Most of the time, information of this sort must be derived from subtle cues. Facial emotions are one such cue: they are detected quickly and serve as key regulators of social behavior (De Gelder, 2006; Frijda, 1986). For instance, from emotions, we can infer whether a person feels either comfortable or uncomfortable in a situation. Here, we question whether and how facial emotions interact with the context when perceivers are asked to determine if someone deserves to be trusted.

Previous studies have already addressed the relationship between facial emotions and contextual information. In such studies, the label "context" was used to refer to different types of features that might influence the perception of facial expressions, including body postures (e.g., Aviezer et al., 2008; Aviezer, Bentin, Dudarev, & Hassin, 2011; Meeren, van Heijnsbergen, & de Gelder, 2005), verbal description of the situation (Carroll & Russell, 1996), other emotional faces (Masuda et al., 2008; Russell & Fehr, 1987) and, of relevance for the present contribution, visual scenes (Righart & De Gelder, 2006; Righart & Gelder, 2008). For instance, Righart and Gelder (2008) found that participants categorized facial emotions faster and more accurately when embedded in emotionally congruent (vs. incongruent) scenes. Importantly, the effect of context in easing the recognition of facial emotion is not merely a matter of congruity. Using another face as a context stimulus, Mumenthaler and Sander (2012, 2015) showed that a face expressing fear was judged as more fearful when another face expressing anger gazed at its direction than when it gazed elsewhere. In other words, when participants could establish a relationship between the target and the face in the background, emotion recognition was facilitated. Taken together, these findings speak for the interdependency of emotions and contextual information and suggest that the latter might ease the recognition of the former by signalling specific functional relations.

Whereas prior research has mainly focused on the role of the context in facilitating emotion recognition (see Barrett, 2011 for a review), here we tackle a different question. We ask whether facial emotions and contextual information can be integrated to form inferences that can ultimately impact trustworthiness attributions. Based on Mattavelli et al. (2023, Experiment 2), we anticipate that, while verbal instructions blatantly inform that a target person presented in a threatening context is either the victim or the perpetrator, facial emotion could convey the very same information but in an indirect fashion. For instance, perceivers might be inclined to assume that a fearful face presented in a threatening context belongs to the victim of the scene portrayed in such a context; by contrast, seeing a joyful face in the very same threatening context might lead the perceiver to assume that the target person is the perpetrator. These inferences should impact upon the attribution of different dispositional traits to the target person portrayed in the scene.

In sum, this is the first attempt to face-context integration on attribution of trustworthiness by manipulating dynamic facial cues (i.e., emotional expressions). Previous studies examining this phenomenon focused on the interplay between morphological facial features and contextual scenes (e.g., Mattavelli et al., 2023, Experiments 1a-1c). Differently from morphological features, emotional expressions provide valuable insights into the internal states of the target. For example, when observing a smiling face, people often infer that the target is experiencing happiness in a given situation. Importantly, the appropriateness of an expressed emotion can vary depending on the context in which it is displayed. The appropriateness of an emotion should directly influence the dispositional attributions, such as trustworthiness, made by observers towards the target face. Based on this rationale, we anticipate a significant interaction between facial emotions and the contextual information.

3. The present research

In two experiments, we questioned whether the attribution of trustworthiness to faces in a threatening context is affected by their emotional expressions. Participants viewed a series of real faces expressing either happiness or fear and were surrounded by either threatening scenes potentially ascribable to the human action or a grey background (i.e., no context). In Experiment 1, participants rated facial stimuli on trustworthiness. We hypothesized that facial emotions would moderate face-context integration on trustworthiness judgments (i.e., lower trustworthiness for faces presented in threatening contexts vs. no context). We expected face-context integration to be stronger for happy faces than for fearful faces, such that the former should be perceived as less trustworthy than the latter. This should occur because expressing happiness in threatening situations should be interpreted as a cue for malice or meanness, whereas expressing fear in the same situations might indicate that the target is feeling uncomfortable in that specific context. This assumption was tested directly in Experiment 2. Namely, we explored the specificity of the hypothesized effect. If inferences drawn from emotional expressions (e.g., a happy face in a crime scene likely belongs to a bad individual) explain face-context integration, this should happen when the ultimate judgments are made on a disposition that justifies reliance on such inferences but not on an irrelevant disposition. By asking participants to make attributions on both trustworthiness and smartness, we tested whether judgments' relevance could qualify the interaction hypothesized in Experiment 1.

All studies received formal approval from the ethics committee of the local university. We preregistered the entire protocols of both experiments on Open Science Framework (Experiment 1: https://osf.io/paebd Experiment 2: https://osf.io/2pbev).

All the analysis codes are also available on Open Science Framework https://osf.io/cdk6p/). We reported all the manipulations and measures

used in each study.

4. Experiment 1

We employed a 2 (context: threatening vs. no context) x 2 (facial emotion: fear vs. happiness) within-subjects design. The dependent variable was the level of trustworthiness attributed to each face.

4.1. Method

4.1.1. Participants and sample size determination

The effect of interest was the interaction between context and facial emotion. We conducted a power analysis in R using the pwr::pwr.f2.test () function. We set the effect size at f = 0.15 (d = 0.30, which qualifies as small to moderate effect size). At alpha = 0.05, with a power = 0.95, the analysis suggested 87 participants.

4.1.2. Stimuli

We employed 16 face identities borrowed from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015) and selected two emotional expressions for each face (i.e., happy vs. fearful). This led to a total of 32 facial stimuli being used in the study.

We used 8 threatening context stimuli, four taken from Mattavelli et al. (2023, Experiment 2) and other four stimuli pretested to be threatening (M = 5.17, SD = 0.54, on a scale from 1 = not at all threatening to 7 = extremely threatening). All the context stimuli portrayed scenes ascribable to the human action (e.g., a gun with bullets, a bloody knife, a burning building, see Supplementary Materials for the full set of stimuli). A grey rectangle was used in the no context condition.

4.1.3. Procedure

93 participants (73 females, M_{age} = 27.94, SD_{age} = 12.43) were recruited via Prolific Academic and asked to participate in a study on face perception. After giving their consent to participate and providing demographics, participants underwent a judgment task. In this task, they saw a series of faces appearing on screen. Each face was embedded in a visual context. In each trial, participants evaluated the extent to which each face appeared as untrustworthy or trustworthy using a 7-points scale (1 = untrustworthy; 7 = trustworthy). The judgment phase consisted of four consecutive blocks of 16 trials, with stimuli administered in random order. Within each class of emotional faces, four faces were presented in threatening backgrounds and four faces in no background scene. Each face identity was associated with one specific emotion and one specific context stimulus. We created eight different combinations to counterbalance face-context assignment across participants. In each block, each face identity appeared once.

4.2. Results

Data were analyzed in a two-level generalized mixed model. The nature of the context (threatening vs. no context), facial emotion (fear vs. happiness), and their interaction were included as fixed factors. The individual intercept and the intercept for the facial identities used across trials were the random factors. As preregistered, we explored the significant interaction by testing the simple effect of facial emotions on the attribution of trustworthiness in each type of context.

We found a main effect of the type of context, b = 0.34, SE = 0.02, t(5701.46) = 18.31, $p < 0.001$, indicating lower trustworthiness ascribed to faces appearing in threatening contexts (vs. no context). The main effect of facial emotion was not significant, b = 0.09, SE = 0.10, t(29.62) = 0.85, p = 0.401. The interaction between the two terms was significant, b = −0.15, SE = 0.02, t(5701.46) = −7.78, $p < 0.001$. Decomposing this interaction revealed no difference in attributed trustworthiness for fearful and happy faces when presented without context scene, b = −0.11, SE = 0.21, t(31.6) = −0.53, p = 0.598. Instead, when faces were presented in threatening contexts, happy faces were judged as significantly less trustworthy than fearful faces, b = 0.47, SE = 0.21, t(31.6) = 2.21, p = 0.034 (see Fig. 1 for bar graphs).

4.3. Discussion

Participants perceived faces in threatening contexts as less trustworthy when displaying happiness rather than fear. This finding is consistent with the idea that face-context integration depends on inferences regarding the role of the target individual in the surrounding context. Whereas fear is an emotional expression justified by contextual threat, expressing happiness in threatening contexts is likely interpreted as a cue for malice or meanness, characteristics that typically apply to the perpetrator. However, an alternative explanation of these findings lies in stimuli conceptual congruency. When faces in contexts are conceived as stimuli-pairs, a conceptual overlap between the face and the context might facilitate the processing of this pair to ultimately increase the evaluation of the face on the ultimate criterion (see Righart & De Gelder, 2006). Stimuli congruency might affect processing fluency, that is, the ease of the mental operations involved in the processing of perceptual stimuli. Past research has shown that higher perceptual fluency increases stimulus evaluation (e.g., Reber, Winkielman, & Schwarz, 1998). In other words, as threatening contexts should overlap more with fearful faces relative to happy faces, the lower trustworthiness attributed to happy faces might reflect the negative experience of processing a disfluent stimuli pair. Experiment 2 was designed to rule out this potential confound.

5. Experiment 2

Experiment 1 offered initial evidence of the idea that facial emotions are key in qualifying the face-context integration. Experiment 2 aimed at shedding light on the underlying mechanism of the observed effect, that is, lower trustworthiness attributed to happy (vs. fearful) faces in a threatening context. To this aim, we tested the specificity of the effect by comparing the interaction between context and emotions on two different types of attribution, that are, trustworthiness vs. smartness.[1] If

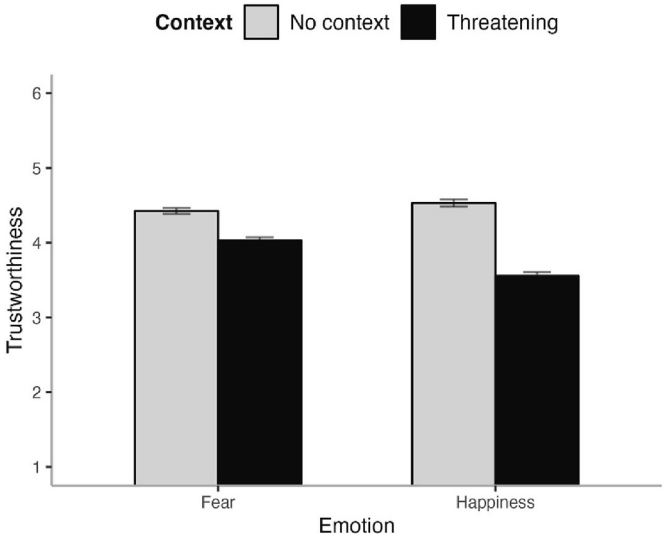


Fig. 1. Results of Experiment 1.

[1] Note that in the preregistered protocol we referred to the second dimension as "slyness". The Italian word used in the study was "*furbo*". Since the translation of this term includes terms such as "*clever*", "*sly*", "*smart*", "*cunning*", we used smartness to capture the underling construct.

the effect observed in Experiment 1 depends on stimuli congruency and its positive impact on the ultimate criterion, it should emerge regardless of whether such a criterion is related to the inferences made upon facial emotions in context. On the contrary, if the effect is driven by relational inferences, then the effect should emerge only in a domain for which such inferences are relevant. Indeed, we expect that the interplay between emotions and contextual threat would influence only trustworthiness attribution. Moreover, we directly tested the nature of such inferences: in a role attribution task, participants judged the likelihood with which the target faces presented in threatening contexts belonged to a criminal.

5.1. Method

We employed a 2 (context: threatening vs. no context) x 2 (facial emotion: fear vs. happiness) x 2 (dispositional attribution: trustworthiness vs smartness) within-subjects design. We also administered a role attribution task to measure the extent to which participants believed that each face presented in a threatening scenario belonged to a criminal.[2]

5.1.1. Sample size determination

The effect of interest was the three-way interaction between context, facial emotion, and dispositional attribution. We conducted a power analysis in R using the pwr::pwr.f2.test () function. We set the effect size at f = 0.10 (d = 0.20, which qualifies as small effect size). At alpha = 0.05, with a power = 0.95 for each contrast, the analysis suggested 131 participants.

5.1.2. Participants and procedure

132 participants (53 females, M_{age} = 29.79, SD_{age} = 11.64) were recruited via Prolific Academic. After giving their consent to participate and providing demographics, participants underwent two consecutive judgment phases. In two blocks of 48 trials each, facial stimuli appeared embedded in a particular visual context. In one block of trials, participants rated facial trustworthiness, using a 7-points scale (1 = not at all trustworthy; 7 = extremely trustworthy). In the other block, participants rated facial smartness (1 = not at all smart; 7 = extremely smart). The order of administration of the two blocks was counterbalanced across participants, and stimuli were administered in random order. In each block, participants were exposed to eight face identities expressing fear and eight identities expressing happiness. Within each class of emotional faces, four faces were presented in a threatening background and four faces in no context. We created four different combinations counterbalanced across participants by assigning each set of face identities to (i) each type of context and (ii) each emotion. In each block, each face identity appeared four times. Next, participants completed a role attribution task. In this task, the same facial stimuli seen in the previous task were presented in threatening contexts only. Upon presentation of each face-context pair, participants indicated, on a 7-points Likert scale, how likely they believed that the target face belonged to a criminal (1 = not at all; 7 = definitely). This phase consisted of one single block of 16 trials, with each face identity appearing once.

5.1.3. Stimuli

To allow for generalizability of Experiment 1's findings, we used emotional faces from a different database. Thus, we employed 16 face identities borrowed from the Karolinska Directed Emotional Faces (KDEF) (Lundqvist, Flykt, & Öhman, 1998) varying on two emotional expressions (i.e., happiness vs. fear). This led to a total of 32 facial stimuli used in the study. Threatening context stimuli were the ones employed in Experiment 1. The same grey rectangle used in Experiment 1 was used in the no context condition.

5.2. Results

We followed the same analytical approach adopted in Experiment 1. The nature of the context (threatening vs. no context), facial emotion (fear vs. happiness), dispositional attribution (trustworthiness vs smartness), and the interaction terms were included as fixed factors. The individual intercept and the intercept for the facial identities were the random factors. We also investigated the simple effect of facial emotions on the attributed likelihood of being a criminal.

The main effect of the type of context, $b = 0.20$, $SE = 0.01$, $t(12{,}504.87) = 17.22$, $p < 0.001$, facial emotion, $b = -0.24$, $SE = 0.04$, $t(27.56) = -6.48$, $p < 0.001$, and dispositional attribution, $b = -0.07$, $SE = 0.01$, $t(12{,}504.31) = -6.36$, $p < 0.001$, were all significant. The interaction between attribution and facial emotion was significant, $b = -0.19$, $SE = 0.01$, $t(12{,}504.31) = -16.38$, $p < 0.001$, indicating that happy faces were judged as smarter than fearful faces, $b = -0.85$, $SE = 0.08$, $t(33.1) = -11.09$, $p < 0.001$, whereas no difference emerged on attributed trustworthiness, $b = -0.10$, $SE = 0.08$, $t(33.1) = -1.30$, $p = 0.204$. Also significant was the interaction between facial emotion and context, $b = -0.09$, $SE = 0.01$, $t(12{,}504.87) = -7.71$, $p < 0.001$, and between attribution and context, $b = -0.20$, $SE = 0.01$, $t(12{,}504.31) = -17.89$, $p < 0.001$. Central to our research question, the three-way interaction was significant, $b = -0.07$, $SE = 0.01$, $t(12{,}504.31) = 6.06$, $p < 0.001$. Decomposing this interaction revealed that the interaction between facial emotion and context replicated on trustworthiness, $b = -0.63$, $SE = 0.06$, $t(12505) = -9.74$, $p < 0.001$: participants attributed lower trustworthiness to happy (vs. fearful) faces in threatening context, $b = 0.22$, $SE = 0.08$, $t(45.9) = 2.47$, $p = 0.013$, whereas the opposite emerged in no context, $b = -0.42$, $SE = 0.08$, $t(45.9) = -4.97$, $p < 0.001$. The interaction between facial emotion and context was not significant on smartness, $b = -0.08$, $SE = 0.08$, $t(12505) = -1.17$, $p = 0.241$ (see Fig. 2 for bar graphs).

We tested the effect of facial emotions on the participants' judgment made in the role attribution task. A significant effect, $b = -0.76$, $SE = 0.08$, $t(28.50) = -9.61$, $p < 0.001$, indicated that happy faces ($M = 5.01$, $SD = 1.71$) in threatening contexts were more likely to be judged as criminal than fearful faces ($M = 3.48$, $SD = 1.57$).

Finally, we explored the correlations between the effect observed in the role attribution task, and the face-context integration effects observed on the two domains of trustworthiness and smartness. To this aim, we calculated three separate scores for each participant. One score reflected the difference in face-context integration on trustworthiness (i.e., threatening contexts minus no context) for fearful vs. happy faces; one score reflected the difference in face-context integration on smartness (i.e., threatening contexts minus no context) for fearful vs. happy faces; and one score reflected face-context integration on the attributed probability of being a criminal for fearful vs. happy faces. We found a significant correlation between the difference (i.e., fearful vs. happy faces) in face context integration on attributed trustworthiness and the difference in face context integration on attributed probability of being a criminal in threatening contexts, $r = -0.43$, $p < 0.001$. Instead, the latter score was not correlated with the difference in face context integration on attributed smartness, $r = -0.12$, $p = 0.123$. Neither significant was the correlation between the difference in face context integration on

[2] We conducted another study in which anger was added as a new level in the manipulation of emotional expressions. This was done to test whether face-context relationship (i.e., angry faces in threatening contexts should signal bad intentions, leading to lower trustworthiness) vs. congruency (i.e., processing angry faces in threatening contexts should activate a fluency experience, reducing the negative effect on attributed trustworthiness) could account for the effect on trustworthiness. However, our results showed a baseline imbalance in the level of trustworthiness attributed to facial emotions: in the no context condition, angry faces were judged as less trustworthy than both happy and neutral faces. This implied that angry faces had less room to be evaluated as less trustworthy in threatening context. The entire pre-registered protocol (https://osf.io/2ue3r) and the analyses code (https://osf.io/cdk6p/) are available.

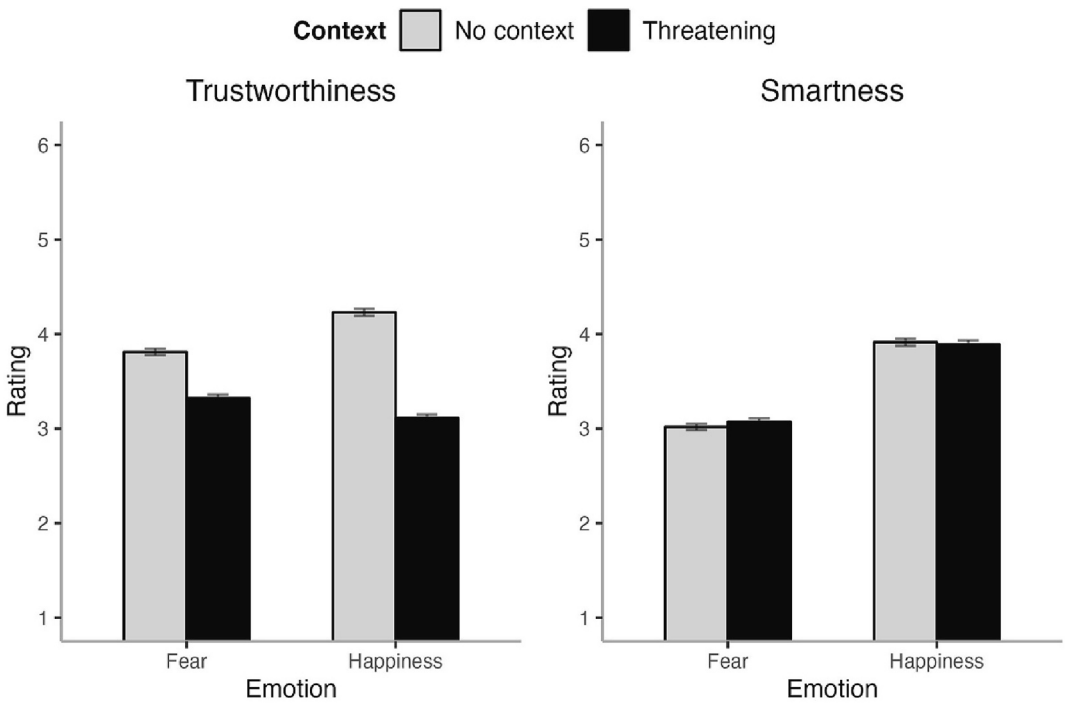


Fig. 2. Results of Experiment 2.

attributed trustworthiness that on attributed smartness, $r = 0.10$, $p = 0.259$.

5.3. Discussion

Experiment 2 proved the specificity of the interaction between facial emotion and contextual threat. Happy faces were judged more negatively than fearful faces when presented in threatening contexts (vs. no context), but this effect was unique for attribution of trustworthiness. In fact, no significant interaction between facial emotion and context emerged in smartness. This latter result rules-out a fluency interpretation of the interaction between context and emotional expressions on judgments of trustworthiness. Moreover, testing directly the inferences made by participants on the role played by the target faces in the threatening context corroborated the idea that this effect might be explained by relational encoding. First, happy faces in threatening contexts were judged as more likely to be the criminal of the surrounding scene. Second, such attributions correlated significantly with participants' tendency to attribute less trustworthiness to happy (vs. fearful) faces in threatening contexts, whereas no correlation emerged when focusing on smartness.

6. General discussion

Across two studies, we tested whether facial emotions could moderate face-context integration on the attribution of trustworthiness to target faces. We manipulated faces' emotional expressions to induce different inferences about how target stimuli felt in the specific context situation and, ultimately, on the role played by target stimuli in the relevant context. Namely, expressing fear in a threatening situation should signal discomfort, whereas expressing happiness in the very same situation should signal the target's malice. In Experiment 1, a significant interaction between context and facial emotion indicated that when faces were presented in threatening scenarios, faces expressing happiness were judged as less trustworthy than those expressing fear. Experiment 2 clarified the underlying mechanism of this effect in two ways. First, it showed that the same interaction could not replicate on an irrelevant dispositional attribution, therefore ruling out the idea of a general amplification effect due to face-context congruency. Second, inferences about the role played by the target faces in the threatening context were measured, rather than merely assumed. Not only we found that happy (vs. fearful) faces in threatening contexts were more likely judged as the persons responsible for the portrayed threatening scene, but this difference was also correlated with the differential face-context integration effect (i.e., the tendency to attribute less trustworthiness to faces in threatening contexts vs. no context) observed in the two types of emotional expressions. In other words, the higher the gap between happy and fearful faces in the attributed likelihood of being a criminal, the higher the gap in attributed trustworthiness to the two classes of facial emotions in context.

A good deal of work demonstrated the interconnection between facial emotions and context stimuli (e.g., Righart & Gelder, 2008). Context stimuli, presented in the form of visual scenes (e.g., de Gelder & Van den Stock, 2011), vignettes, (e.g., Carroll & Russell, 1996) or linguistic labels (e.g., Lindquist & Gendron, 2013) ease emotion recognition. Although prior research offered initial evidence to support the idea that contextual information affects emotions in a relational fashion rather than via mere congruency (Mumenthaler & Sander, 2012, 2015), this effect was largely confined to emotion recognition tasks. In fact, Barrett and Kensinger (2010) focused on contextual encoding by measuring participants' recall of neutral contextual scenes: they found better memory performances when participants categorized the emotion seen in a face compared to when they made an affective judgment about the face. They concluded that whereas emotion perception inherently involves contextualization, its integration with the context remains marginal when it comes to other attributions. Our work complements previous literature by showing that context stimuli are integrated into the processing of facial emotions also when participants are asked to make dispositional attributions. On attributed trustworthiness, face-context integration showed stronger when the facial emotional expression was context-incongruent (i.e., happiness expressed in a threatening context). This finding, replicated across two experiments, did not generalize on another disposition that perceivers could not infer from their relational encoding of emotions and context (e.g., smartness). The fact that inferences made on the role of the target in the surrounding context (i) depended on the nature of the expressed emotions and (ii)

correlated with the magnitude of face-context integration on trustworthiness attribution, corroborates the idea that emotions are functional cues that perceivers extract to better understand whether someone is worthy of trust. Although we acknowledge that the latter finding is correlational, we highlight the consistency between these findings and Mattavelli et al.'s (2023), in which experimentally manipulating the role played by the target in threatening contexts led to moderated face-context integration.

Beyond emotions, our findings are of theoretical relevance to understanding person-perception in context. When we encounter other individuals, perception serves the primary need to avoid threats (Fiske, 1992; Zebrowitz & Collins, 1997). In fact, our perception of others' faces is essentially explained by the binary decision to either approach or avoid the target individual (Jones & Kramer, 2021). When asked to make decisions of this sort, humans gather information from all the available sources that can be informative about a person. Whereas it is true that facial features might suffice to make dispositional attribution about others (Todorov et al., 2015), our research highlights that contextual information plays a key role in determining our dispositional attributions on the perceived target. Not only is the context integrated into the processing of facial stimuli: perceivers tend to actively make sense out of all available information via relational reasoning, even when such information is subtle and not explicitly translated into relational qualifiers. Thus, the negative impact of a threatening context on attributed trustworthiness is reduced when contingent cues suggest that the person is experiencing negative feelings, rather than positive ones, within such a context.

Our findings fit well with a motivated approach to face perception (Kenrick, Neuberg, Griskevicius, Becker, & Schaller, 2010; Maner, Miller, Moss, Leo, & Plant, 2012; Young, Slepian, & Sacco, 2015), whereby trustworthiness attribution to facial stimuli is influenced by self-protection motives made salient by the visual cues available in the experimental scenario. Namely, we propose that processing a facial stimulus in a threatening context might enhance one's need for self-protection, which ultimately influence attributional processes towards the target. For instance, for a perceiver experiencing danger (activated by the threatening context), the importance of determining whether a target face is either trustworthy or untrustworthy would be motivationally increased by the need to identify the potential source of danger or, alternatively, a source of relief. With that said, we acknowledge that this interpretation rests on the assumption that participants experienced an actual sense of threat or danger when exposed to threatening scenes. Future studies should better explore this possibility. For instance, one could think of a correlational study that explores how face-context integration relates to physiological indicators of fear. Alternatively, one could experimentally manipulate the target of contextual threat: according to a motivated approach, one should expect stronger face-context integration effects when contextual threat is addressed to the perceiver than someone else.

Furthermore, we highlighted the malleable nature of trustworthiness by showing that its perception is readily pushed around by scene context. The findings complement prior research on impression formation showing that prior knowledge regarding a target person may affect the evaluation of facial trustworthiness (Mende-Siedlecki, Cai, & Todorov, 2013). Prior findings reveal that extraneous information from the face (i.e., person knowledge) may affect evaluations of the face. Our findings extend these prior insights by revealing that other forms of extraneous information of the face in in the form of a visual context may alter the evaluations of facial cues.

What remains partly unanswered in our studies is why face-context integration was only reduced, but not eliminated (or reversed) when facial stimuli expressed fear. In fact, fearful faces were judged as less trustworthy when presented in threatening contexts than in no context. This finding is also in line with Mattavelli et al.'s (2023, Experiment 2), who found that faces presented as victims in threatening scenes were perceived as less trustworthy than faces presented in no scene. We advance two potential explanations for such an effect. On the one hand, it might be the case that at least a partial component of face-context integration is accounted for by the overall valence conveyed by processed stimuli (i.e., the negativity carried by the threatening context is transferred to the fearful faces). On the other hand, this effect might also have relational roots: people might come up with negative relational inferences even when the person presented in a threatening context expresses fear (e.g., murderers might express fear in a crime scene because they are afraid of being caught by the police). In line with the latter hypothesis is the fact that, in Experiment 2, the average score in the role attribution task for fearful faces was just below the midpoint. This means that, overall, fearful faces in threatening contexts were judged as (at least) "possibly criminals". Future studies should better clarify this issue, for instance by using stronger manipulations where emotional expressions are combined with additional cues (e.g., the voice of the target stimulus asking for help) that could clearly define the role played by the target in the context.

In conclusion, across two studies we showed that attributions of trustworthiness to target individuals are moderated in concert by facial expressions and contextual information. We confirm that face-context relational encoding is key to determining whether a target person is worthy of trust. Importantly, relational encoding could change by simply altering the type of emotional expressions that target stimuli exhibited in threatening contexts. These findings speak for the active role played by perceivers in person perception and for their tendency to use and integrate available cues to construe meaningful relationships between facial and context stimuli.

Declaration of Competing Interest

The authors have no conflict of interest to declare

Data availability

Data and analysis code are available on OSF via the links reported in the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2023.104517.

References

Ames, D. L., Fiske, S. T., & Todorov, A. (2011). Impression formation: A focus on others' intents. In J. Decety, & J. T. Cacioppo (Eds.), *The Oxford handbook of social neuroscience* (pp. 419–433). Oxford University Press.

Aviezer, H., Bentin, S., Dudarev, V., & Hassin, R. R. (2011). The automaticity of emotional face-context integration. *Emotion, 11*(6), 1406–1414.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., … Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science, 19*(7), 724–732.

Barrett, L. F. (2011). Was Darwin wrong about emotional expressions? *Current Directions in Psychological Science, 20*(6), 400–406.

Barrett, L. F., & Kensinger, E. A. (2010). Context is routinely encoded during emotion perception. *Psychological Science, 21*(4), 595–599.

Brambilla, M., Biella, M., & Freeman, J. B. (2018). The influence of visual context on the evaluation of facial trustworthiness. *Journal of Experimental Social Psychology, 78*, 34–42.

Brambilla, M., & Leach, C. W. (2014). On the importance of being moral: The distinctive role of morality in social judgment. *Social Cognition, 32*(4), 397–408.

Brambilla, M., Masi, M., Mattavelli, S., & Biella, M. (2021). Faces and sounds becoming one: Cross-modal integration of facial and auditory cues in judging trustworthiness. *Social Cognition, 39*(3), 315–327.

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70*(2), 205–218.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the genera- tion of culture* (pp. 163–228). Oxford University Press.

De Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience, 7*(3), 242–249.

de Gelder, B., & Van den Stock, J. (2011). The Bodily Expressive Action Stimulus Test (BEAST). Construction and validation of a stimulus basis for measuring perception of whole body expression of emotions. *Frontiers in Psychology, 2*, 181.
Fiske, S. T. (1992). Thinking is for doing: Portraits of social cognition from daguerreotype to laserphoto. *Journal of Personality and Social Psychology, 63*(6), 877–889.
Frijda, N. H. (1986). *The emotions*. Cambridge University Press.
Jones, A. L., & Kramer, R. S. (2021). Facial first impressions form two clusters representing approach-avoidance. *Cognitive Psychology, 126*, Article 101387.
Kenrick, D. T., Neuberg, S. L., Griskevicius, V., Becker, D. V., & Schaller, M. (2010). Goal-driven cognition and functional behavior: The fundamental-motives framework. *Current Directions in Psychological Science, 19*(1), 63–67.
Lindquist, K. A., & Gendron, M. (2013). What's in a word? Language constructs emotion perception. *Emotion Review, 5*(1), 66–71.
Lundqvist, D., Flykt, A., & Öhman, A. (1998). *Karolinska directed emotional faces (KDEF)* [Database record]. APA PsycTests
Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*(4), 1122–1135.
Maner, J. K., Miller, S. L., Moss, J., Leo, J., & Plant, E. A. (2012). Motivated social categorization: Fundamental motives enhance people's sensitivity to basic social categories. *Journal of Personality and Social Psychology, 103*, 70–83.
Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., & Van de Veerdonk, E. (2008). Placing the face in context: Cultural differences in the perception of facial emotion. *Journal of Personality and Social Psychology, 94*(3), 365–381.
Mattavelli, S., Masi, M., & Brambilla, M. (2023). Not just about faces in context: Face–context relation moderates the impact of contextual threat on facial trustworthiness. *Personality and Social Psychology Bulletin, 49*(3), 376–390.
Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences, 102*(45), 16518–16523.
Mende-Siedlecki, P., Cai, Y., & Todorov, A. (2013). The neural dynamics of updating person impressions. *Social Cognitive and Affective Neuroscience, 8*(6), 623–631.
Mumenthaler, C., & Sander, D. (2012). Social appraisal influences recognition of emotions. *Journal of Personality and Social Psychology, 102*(6), 1118–1135.
Mumenthaler, C., & Sander, D. (2015). Automatic integration of social information in emotion recognition. *Journal of Experimental Psychology: General, 144*(2), 392–399.
Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science, 9*(1), 45–48.
Righart, R., & De Gelder, B. (2006). Context influences early perceptual analysis of faces—An electrophysiological study. *Cerebral Cortex, 16*(9), 1249–1257.
Righart, R., & Gelder, B. D. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective, & Behavioral Neuroscience, 8*(3), 264–272.
Russell, J. A., & Fehr, B. (1987). Relativity in the perception of emotion in facial expressions. *Journal of Experimental Psychology: General, 116*(3), 223–237.
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.
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.
Young, S. G., Slepian, M. L., & Sacco, D. F. (2015). Sensitivity to perceived facial trustworthiness is increased by activating self-protection motives. *Social Psychological and Personality Science, 6*(6), 607–613.
Zebrowitz, L. A., & Collins, M. A. (1997). Accurate social perception at zero acquaintance: The affordances of a Gibsonian approach. *Personality and Social Psychology Review, 1*(3), 204–223.
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, 7*(3), 194–215.