Perceiving emotions in neutral faces: expression processing is biased by affective person knowledge

Franziska Suess, Milena Rabovsky, and Rasha Abdel Rahman
Department of Psychology, Humboldt-Universität zu Berlin, 12489 Berlin, Germany

According to a widely held view, basic emotions such as happiness or anger are reflected in facial expressions that are invariant and uniquely defined by specific facial muscle movements. Accordingly, expression perception should not be vulnerable to influences outside the face. Here, we test this assumption by manipulating the emotional valence of biographical knowledge associated with individual persons. Faces of well-known and initially unfamiliar persons displaying neutral expressions were associated with socially relevant negative, positive or comparatively neutral biographical information. The expressions of faces associated with negative information were classified as more negative than faces associated with neutral information. Event-related brain potential modulations in the early posterior negativity, a component taken to reflect early sensory processing of affective stimuli such as emotional facial expressions, suggest that negative affective knowledge can bias the perception of faces with neutral expressions toward subjectively displaying negative emotions.

Keywords: facial expressions; affective knowledge; emotion; ERPs

INTRODUCTION

Person-related knowledge, acquired via direct personal interactions or verbal descriptions of traits and behaviors of personally unfamiliar individuals, is an important factor determining how we evaluate other persons, and even more so when this knowledge contains affective and socially relevant information (Singer et al., 2004; Croft et al., 2010; Todorov, 2011). However, while effects on person evaluation and moral judgments are well-established (Bliss-Moreau et al., 2008), far less is known about the influence of affective knowledge on other aspects of person recognition, in particular those preceding higher level evaluations, such as the perception of facial features and emotional expressions (Abdel Rahman, 2011). The present study was designed to investigate such affective knowledge effects on perceptual aspects of face processing. Specifically, we ask whether long established or recently acquired affective and socially relevant knowledge has the potential to bias the perception of objectively neutral faces toward subjectively conveying emotional expressions that are congruent with the valence of the associated biographical knowledge. Such an influence of knowledge on expression perception, may this knowledge be correct or based for false information or gossip, should have considerable consequences for how we evaluate the mental states of other persons, and should thus strongly influence our social interactions.

According to a widely held view on the expression and perception of facial emotions (Ekman and Friesen, 1976; Ekman, 1993, 2003), basic expressions such as those of anger or happiness are invariant manifestations of specific emotional states that are universally shared and uniquely defined by specific facial muscle movements (Ekman and Friesen, 1976; Ekman, 1993, 2003, but see Barrett, 2006). Thus, expression perception should be stable and unaffected by visually opaque affective information about the person’s biography.

Yet, cumulative evidence on visual context effects suggests that expression perception can be subject to influences from meaningful information outside the face at least when the source of the context information is visual, stemming from, for instance, body postures and gestures (Meeren et al., 2005; De Gelder et al., 2006; Van den Stock et al., 2007; Aviezer et al., 2008). Thus, Aviezer et al. (2008) have reported remarkable shifts in the categorization and perception of basic expressions induced by body contexts, demonstrating a marked vulnerability of expression perception to influences from outside the face (see also Aviezer et al., 2011; for a review, see Hassin et al., 2013).

Comparable evidence for non-visual conceptual effects on expression perception is scant. Some research suggests that general face processing mechanisms such as person evaluations are influenced even when the available person knowledge is based on seemingly sparse behavioral information, e.g. that the person hit a small child (Bliss-Moreau et al., 2008). Furthermore, affective knowledge is reflected in spontaneous activations of neural regions associated with social cognition and emotion (Todorov et al., 2007), and personality information may even influence how we imagine or classify facial features (Hassin and Trope, 2000). Crucially, some evidence suggests that affective knowledge may modulate early sensory processing of the faces of well-known persons (Abdel Rahman, 2011) and, in the case of negative social information associated with initially unfamiliar faces, even alter the potency of faces to gain access to and maintain conscious perception (Anderson et al., 2011). Furthermore, an observer’s knowledge about the situation in which an expression is presented may modulate its classification (Carroll and Russell, 1996, but see Nakamura et al., 1990), and the perception of ambiguous facial expressions is influenced by verbal emotion concepts (e.g. ‘happy’; Halberstadt et al., 2009). In line with theoretical views of person recognition incorporating interactions between perceptual and semantic processing components (Haxby et al., 2000), these findings suggest that not only late evaluative but also early perceptual aspects of face recognition are shaped by affective person knowledge.

The present study focuses on investigating emotional knowledge effects specifically on the perception of facial expressions: can objectively neutral faces be perceived in an unbiased way irrespective of what we know about the person’s character and good or evil deeds or are expressions perceived in light of our knowledge associated with the person as more positive or negative, depending on the valence of the associated knowledge? We build on a previous study by Abdel Rahman
that investigated the influence of affective knowledge on the perception of well-known and initially unfamiliar faces. In this study, early event-related brain potential (ERP) modulations induced by emotional knowledge were found only for well-known faces. Here, to assess the effects of newly acquired information about initially unfamiliar faces more directly, we employed a similar learning paradigm but additionally increased the credibility of the newly acquired information with several measures (see below). Furthermore, here, we included a facial expression rating before and after learning to investigate the consequences of this information on the perception of the person’s facial expression.

Based on the evidence discussed above, we expected a bias in expression judgments according to the valence of the affective knowledge associated with the person. Additionally, ERPs were used to localize the effects on early perceptual processing stages. In particular, perceptual effects should be reflected in modulations of the early posterior negativity (EPN), an ERP component associated with early reflexive emotion effects, reflecting attention to and enhanced perceptual processing of affective stimuli such as objects and scenes (Janghofer et al., 2001; Schupp et al., 2003), words (Kissler et al., 2007; Schacht and Sommer, 2009) and even symbolic hand gestures (Flaisch et al., 2011). Crucially, the EPN has been shown to be sensitive to facial expression perception (Schupp, et al., 2004; Schacht and Sommer, 2009), and therefore, any affective knowledge effects influencing expression perception should be reflected in modulations of this component.

METHOD
Participants
Twenty-four participants (17 female; mean age = 23.67 years, range 19–29) with normal or corrected-to-normal vision took part in the experiment for monetary compensation or for course credits. Informed consent was obtained from all participants before the experiment started. The study was approved by the local ethics committee of the Psychology department at Humboldt-Universität zu Berlin.

Materials
Picture stimuli consisted of gray-scale photographs of 18 very well-known, 18 less well-known and 36 initially unfamiliar male and female faces displaying neutral facial expressions (see Supplementary Material I). All photographs were frontal headshots, scaled to 2.7 cm × 3.5 cm at a viewing distance of about 90 cm. To increase believability of the fictitious stories associated with initially unfamiliar faces, we included not only very well-known persons (e.g. Adolf Hitler) but also somewhat less well-known persons (e.g. Sarah Palin) as filler items that were not included in further analyses because, across participants, they could not reliably be assigned to known and initially unfamiliar faces. The selection was based on prior familiarity ratings of face/ name pairs with a different group of participants. Participants were informed that some of the faces might be unfamiliar to them because some persons were not very well-known in general or only well-known in their country. Because participants were able to verify that some faces were less well-known than others, they were presumably more likely to believe that all stories were associated with real persons even in those cases when they could not recognize the person.

Spoken stories were recorded that contained emotionally negative, positive or neutral biographical information (see Supplementary Material II). Each fictitious story was constructed in analogy to one of the real stories to further increase credibility of the information, and the faces of the initially unfamiliar and well-known persons were matched accordingly for age, gender and race. All stories were presented auditorily with a mean duration of ~26s.

Procedure
The experiment consisted of a learning and a test session, separated by 1 or 2 days. Before learning, participants were told that all persons presented really existed but that some would be more familiar than others. The learning session began and ended with a sympathy rating and a rating of facial expressions ranging from negative (−2) to positive (2), on a 5-point scale, analogous to the Self-Assessment Manikin (Bradley and Lang, 1994). Between these ratings, a familiarity classification task (1: unfamiliar, 2: familiar) was administered, which confirmed the status of faces as unfamiliar (initially unfamiliar faces mean: 1.06) and familiar (well-known faces mean: 1.98). Subsequently, participants learned the person’s nationality. This later task-relevant information was acquired before the affective information was presented to avoid any effects of emotion on initial familiarization and learning. Crucially, the task could be completed entirely independent of the affective or neutral biographical information acquired later. Successful learning was tested in a semantic task in which all 72 faces were presented in random order once for 500 ms. Participants were instructed to classify via button press whether the person was European or not.

In addition, participants executed a spatial attention task in which the faces were presented in a cueing paradigm (Posner, 1980) and participants indicated via button press on which side the target was presented. The purpose of the attention task was to further illuminate the underlying mechanisms of the early visual effects using a different task. The results of the attention task would exceed the scope of the present article and are not discussed any further.

Subsequently the biographical knowledge was learned. Each face was presented on the screen while participants listened to a story describing real or fictitious biographical details. For initially unfamiliar faces, the assignment of faces to positive, negative or neutral stories was counterbalanced across participants. All faces and the respective stories were presented twice in random order. The semantic and attention task were administered again after learning.

The test session with electroencephalogram (EEG) recordings consisted of both tasks (semantic and attention task). The semantic task was conducted as described above. All faces were presented eight times in random order. Here, we focus on the sympathy and facial expression rating and ERPs from the semantic task.

EEG data recording and analysis
The EEG was recorded with Ag/AgCl electrodes from 62 sites according to the extended 10–20 system, referenced to the left mastoid, at a sampling rate of 500 Hz. The horizontal and vertical electrooculograms were measured with external electrodes attached to the left and right canthi of both eyes and beneath and above the left eye. Electrode impedance was kept below 5 kΩ. A short calibration procedure was used to obtain prototypical eye movements to correct for eye movement artifacts. Offline, the continuous EEG was transformed to average reference and low-pass filtered at 30 Hz. Eye movement artifacts were removed with a spatiotemporal dipole modeling procedure using BESA (Berg & Scherg, 1991); remaining artifacts were eliminated with a semiautomatic artifact rejection procedure (amplitudes over ±200 μV, changing >50 μV between samples or >200 μV within single epochs, or containing baseline drifts). Error- and artifact-free EEG data were segmented into epochs of 2 s with a 100 ms pre-stimulus baseline.

Amplitude differences were assessed with repeated-measures analyses of variance (ANOVAs) with the factor Emotion (negative, positive, neutral) separately for very well-known and newly learned faces. Huynh–Feldt corrections were applied when appropriate. Analyses focused on a region of interest (PO7, PO8, PO9, PO10, TP9, TP10),...
chosen based on previous findings of early emotion effects in the EPN (Abdel Rahman, 2011).

RESULTS

Sympathy ratings

The results of the sympathy rating are presented in Figure 1. A 2 (Learning: pre- vs post-learning) x 2 (Familiarity: well-known vs newly learned faces) x 3 (Emotion: negative, positive and neutral) repeated-measures ANOVA yielded main effects of Learning, $F(1,23) = 11.3, P < 0.01, \eta^2 = 0.330$, Familiarity, $F(1,23) = 74.1, P < 0.001, \eta^2 = 0.763$, and Emotion, $F(2,46) = 252.8, P < 0.001, \eta^2 = 0.917$, and interactions of Learning and Emotion, $F(2,46) = 159.3, P < 0.001, \eta^2 = 0.874$, Familiarity and Emotion, $F(2,46) = 158.9, P < 0.001, \eta^2 = 0.874$, and a three-way interaction of Learning, Familiarity and Emotion, $F(2,46) = 47.3, P < 0.001, \eta^2 = 0.673$.

A separate analysis for well-known faces revealed main effects of Learning, $F(1,23) = 19.5, P < 0.001, \eta^2 = 0.459$, and Emotion, $F(2,46) = 257.8, P < 0.001, \eta^2 = 0.918$, and an interaction of Learning and Emotion, $F(2,46) = 11.7, P < 0.001, \eta^2 = 0.338$. Faces associated with negative biographical information were rated as more dislikeable than those associated with relatively neutral knowledge, before learning, $F(1,23) = 158.1, P < 0.001, \eta^2 = 0.873$, as well as after learning, $F(1,23) = 366.6, P < 0.001, \eta^2 = 0.941$. Likewise, faces associated with positive information were rated as more likeable both before, $F(1,23) = 28.9, P < 0.001, \eta^2 = 0.557$, and after learning, $F(1,23) = 54.8, P < 0.001, \eta^2 = 0.704$.

For newly learned faces, there was a main effect of Emotion, $F(2,46) = 141.4, P < 0.001, \eta^2 = 0.860$, and an interaction of Learning and Emotion, $F(2,46) = 141.3, P < 0.001, \eta^2 = 0.860$. Before learning, sympathy ratings did not differ between emotion conditions, $F(1,23) = 3.9, P > 0.05, \eta^2 = 0.146$ (negative vs neutral knowledge condition), and $F(1,23) = 0.66, P > 0.1, \eta^2 = 0.028$ (positive vs neutral knowledge condition), whereas after learning faces associated with negative information were rated as more dislikeable, $F(1,23) = 158.0, P < 0.001, \eta^2 = 0.873$, and faces associated with positive information were rated as more likeable, $F(1,23) = 57.1, P < 0.001, \eta^2 = 0.713$, than those associated with relatively neutral biographical information. As expected, these findings confirm that affective knowledge has a strong influence on person evaluation. Furthermore, they demonstrate successful learning of the affective information associated with initially unfamiliar faces.

Facial expression ratings

The ANOVA on mean expression ratings (see Figure 2) yielded main effects of Familiarity, $F(1,23) = 147.8, P < 0.001, \eta^2 = 0.865$, and Emotion, $F(2,46) = 119.9, P < 0.001, \eta^2 = 0.839$, and interactions of Learning and Emotion, $F(2,46) = 5.0, P < 0.01, \eta^2 = 0.180$, and Familiarity and Emotion, $F(2,46) = 86.1, P < 0.001, \eta^2 = 0.789$.

For well-known faces, there was a main effect of Emotion, $F(2,46) = 149.3, P < 0.001, \eta^2 = 0.867$, whereas neither the effect of Learning, $F(1,23) = 3.9, P = 0.06, \eta^2 = 0.145$, nor the interaction of Learning and Emotion, $F(1,23) = 1.3, P > 0.1, \eta^2 = 0.053$, reached significance. As expected, independent of learning participants perceived the expressions of well-known faces related to negative biographical information as more negative than faces associated with relatively neutral information, $F(1,23) = 209.1, P < 0.001, \eta^2 = 0.901$. No difference was found between the positive and neutral condition, $F(1,23) = 2.2, P > 0.1, \eta^2 = 0.087$.

For newly learned faces, a main effect of Emotion, $F(2,46) = 4.5, P < 0.05, \eta^2 = 0.163$, and an interaction of Learning and Emotion, $F(2,46) = 5.3, P < 0.01, \eta^2 = 0.219$, was found. Before learning, there were no differences between the negative, $F(1,23) = 0.18, P > 0.1, \eta^2 = 0.008$, or positive, $F(1,23) = 0.00, P > 0.1, \eta^2 = 0.000$, relative to the neutral knowledge condition. In contrast, after learning participants rated the expressions of persons newly associated with negative information as more negative than faces paired with neutral information, $F(1,23) = 7.8, P < 0.01, \eta^2 = 0.253$. No difference was found between the positive and neutral condition, $F(1,23) = 0.1, P > 0.1, \eta^2 = 0.005$.

Semantic task

Behavioral results

ANOVAs on mean reaction times (RTs) yielded main effects of Familiarity, $F(1,23) = 30.2, P < 0.001, \eta^2 = 0.568$, and Emotion,
expressions or other physical attributes, rather than induced by the biographical information associated with both well-known and initially unfamiliar faces. As discussed in the introduction, comparable EPN modulations have been interpreted to reflect sensory processing of facial expressions in the visual cortex (Schupp et al., 2004). Thus, our findings suggest that abstract affective person knowledge has an influence on how we perceive a person’s facial expression. Mirroring the pattern of results found in the expression ratings, the EPN was most pronounced for negative information, with robust effects for both well-known and initially unfamiliar faces. Stronger effects of stimuli with negative compared to positive valence have often been reported (Abdel Rahman, 2011; Anderson et al., 2011), and may relate to the enhanced processing of potentially threatening stimuli. A person with a negative biography, e.g., a murderer, could pose a potential threat, and to be aware of such danger might secure survival.

Although the EPN modulations were present in both familiarity conditions, affective knowledge effects for well-known and initially unfamiliar faces differed slightly in their time course and topographical distributions. These differences may be attributed to different levels of associative strength of the person-related information with the face and/or a different quality of the emotional responses. While for well-known faces of persons that were familiar already prior to the experiment, biographical knowledge is well established in long-term memory, the traces of the newly acquired information associated with these faces may be more readily available for processing.
with initially unfamiliar faces may be much weaker. As a consequence, the sensory responses reflected in the EPN may not yet be as strong as for well-known faces. Moreover, the factors familiarity and emotion in person recognition and face processing might be closely entangled, and repeated encounters with a person, each associated with affective responses, may form the basis for affective knowledge effects as the ones observed here (cf. Barrett and Bar, 2009; Abdel Rahman, 2011). From this perspective, prior experiences with a person, including verbal descriptions of the person’s past behavior, can shape how we perceive emotional expressions of that person’s face in the future.

To summarize, we present direct evidence from facial expression ratings and more indirect evidence from ERP responses to faces associated with socially relevant affective biographies suggesting that, literally, facial expressions cannot be viewed independent of the affective knowledge associated with the person. This holds in particular for negative information. Such effects of negatively valenced knowledge on how we perceive a person’s facial expression, may the available information be based on proper facts, rumors or gossip, should have a massive influence on social interactions. Specifically, the tendency to perceive the facial expression of another person as more negative may (negatively) influence our interpretations of the mental state and intentions of that person, which may in turn modulate our own behavior toward that person, potentially resulting in a self-fulfilling prophecy.

To conclude, the observed influences of knowledge on expression perception add to the cumulating evidence on context effects in this field and suggest that not only visual contexts (Aviezer et al., 2008) but also semantic contexts can shape expression perception, in contrast to the widely held view (Ekman, 1993) that basic emotions are immune to such influences.

**SUPPLEMENTARY DATA**

Supplementary data are available at SCAN online.
Conflict of Interest
None declared.

REFERENCES