Effects of Social Context on Deliberate Facial Expressions: Evidence from a Stroop-like Task

Stephen Katembu1 · Qiang Xu2 · Hadiseh Nowparast Rostami1 · Guillermo Recio3 · Werner Sommer1

Accepted: 3 January 2022 / Published online: 22 April 2022 © The Author(s) 2022

Abstract
Facial expressions contribute to nonverbal communication, social coordination, and interaction. Facial expressions may reflect the emotional state of the expressor, but they may be modulated by the presence of others, for example, by facial mimicry or through social display rules. We examined how deliberate facial expressions of happiness and anger (smiles and frowns), prompted by written commands, are modulated by the congruency with the facial expression of background faces and how this effect depends on the age of the background face (infants vs. adults). Our main interest was whether the quality of the required expression could be influenced by a task-irrelevant background face and its emotional display. Background faces from adults and infants displayed happy, angry, or neutral expressions. To assess the activation pattern of different action units, we used a machine classifier software; the same classifier was used to assess the chronometry of the expression responses. Results indicated slower and less correct performance when an incongruent facial expression was in the background, especially when distractor stimuli showed adult faces. Interestingly, smile responses were more intense in congruent than incongruent conditions. Depending on stimulus age, frown responses were affected in their quality by incongruent (smile) expressions in terms of the additional activation or deactivation of the outer brow raiser (AU2), resulting in a blended expression, somewhat different from the prototypical expression for anger. Together, the present results show qualitative effects on deliberate facial expressions, beyond typical chronometric effects, confirming machine classification of facial expressions as a promising tool for emotion research.

Keywords Deliberate facial expression · Stroop effect · Mimicry · Infant faces · Cognitive control

Guillermo Recio and Werner Sommer joint senior authorship.

Stephen Katembu
katembmu@hu-berlin.de

1 Institut Für Psychologie, Humboldt-Universität Zu Berlin, Rudower Chaussee 18, 12489 Berlin, Germany
2 Ningbo University, Ningbo, China
3 Universitat de Barcelona, Barcelona, Spain
Introduction

Facial expressions constitute an important medium of day-to-day communication, social coordination, and interaction, enabling individuals to relate and participate in collective action. Thus, when we meet people, we usually look at their faces in order to obtain socially relevant information, which informs effective interactions (Blais et al., 2008). We display facial expressions of affective states, which, to varying extent, are deliberate, sending social signals to the communication partners, often in response to the perceived emotional state of the partner, and modulated by the situational context. The present work aims to contribute towards understanding the adaptability and flexibility of deliberate facial expressions, as captured by video analysis.

Various theories attempt to explain the mechanisms controlling facial expressions. Ekman’s emotion theory views facial expressions as basic, at least partially universal, fairly fixed ways in which we communicate our internal emotional states (Ekman, 1999), yet bendable by display rules learned through socialization (Ekman & Friesen, 1969a). In contrast, according to Fridlund’s (1994) behavioral ecology theory, facial expressions signify behavior intentions, and are used as tools for communicating social motives to specific audiences, whether physically present or imagined (e.g., Crivelli & Fridlund, 2018). Further, componential appraisal theorists argue that facial expressions are outcomes of the evaluation or appraisal of events in their context (e.g., Scherer, 2009); sequentially accumulated appraisal outcomes then produce unique, context- and individual-specific response patterns (Scherer, 2009; Scherer et al., 2013). Expressions of emotion have also been conceived as a motor action that may directly induce its imitation (mimicry) by an observer. Mimicry, the overt or covert spontaneous (automatic) and congruent activation of muscles serving facial expressions upon seeing the facial expression of someone else, is the basis of the matched motor hypothesis (Preston & de Waal, 2002). Experimentally, mimicry manifests in fast and unintentional activation of the facial musculature in response to facial displays of emotion in others (e.g., Dimberg & Petterson, 2000) and by mirror neurons, that is, motor neurons that are responsible for certain actions but also activated by observing others performing these actions (Ferrari et al., 2003).

However, facial displays of emotion are not just outcomes of one’s own inner states and situational appraisals or an automatic reflex in response to emotional displays of others. Mimicry of emotional signals has been shown to possess some level of adaptability and flexibility determined by social factors (Hess & Fischer, 2014). In their review, Seibt et al. (2015) showed facial mimicry to be a multifactorial response modulated by personal, emotional, motivational, relational, and social characteristics of both the mimic and the mimickee in a given context (Chartrand & Lakin, 2013; van Baaren et al., 2009). Specifically, stimuli that represent an emotional meaning, such as facial expressions do, engage both emotion regulation mechanisms and cognitive control processes in order to attend to, select, and sustain task-relevant information, which then informs appropriate facial responses (Quaglia et al., 2019; Tottenham et al., 2011).

Emotion regulation is called for when emotional stimuli automatically attract attention, hence interfering with other processes (Tottenham et al., 2011), and lead to cognitive conflict (Song et al., 2017). Such conflicts are operationalized in the emotional Stroop task where affective stimuli interfere with other, non-affective responses (Chiew & Braver, 2010). To resolve such conflicts and achieve the goals of the task (Banich et al., 2009) requires the deployment of attentional resources, such as setting up an
attentional bias in perception, attenuating irrelevant information, while enhancing task-relevant information and maintaining it in working memory (Egner & Hirsch, 2005).

The need for conflict regulation holds true also for facial displays of emotion. Thus, in social contexts, facial expressions may have to be inhibited, amplified, attenuated, or simply masked with the expression of another emotion to fit the social intentions or display rules dictated by social or other situational demands (e.g., Lee & Wagner, 2002). Otte, et al. (2011a) examined how the perception of socially relevant but task-irrelevant facial expressions interacted with the execution of required expressions. A Simon-type task was used, where participants were to respond as fast as possible with a smile or a frown to the gender of a face stimulus. The face stimuli showed smiling or frowning expressions rendering the facial responses of the participants compatible or incompatible with the stimulus expression. Reaction times (RT) of the required facial expressions were measured as EMG onsets of the m. zygomaticus major (for smiles) and the m. corrugator supercili. RTs were shorter for compatible conditions (e.g., smile responses to smiling faces), compared to incompatible stimulus–response conditions (e.g., smile responses to frowning faces). Hence, the perceived facial expressions, even though not relevant to the task, interfered with the production of facial expressions. Similar effects were observed in a dual-task study (Otte et al., 2011b), suggesting the automaticity of the interference.

Whereas the studies of Otte and co-workers indicate that mimicry can interfere with producing emotional expressions, mimicry itself is not invariant. Bourgeois and Hess (2008) found that, irrespective of context, smiles are likely to be mimicked because they are socially acceptable, indicating affiliation intentions, and involving no personal costs. In contrast, mimicry of sad expressions comes with the risk of personal costs, and mimicry of anger may be associated with aggression and perceived as socially maladaptive (Künecke et al., 2017). Hence, facial expressions of negative emotions are less likely to be mimicked than smiles. As compared to faces of acquaintances, faces of strangers have been shown to be mimicked less (Bourgeois & Hess, 2008). Decoding same-age facial expressions has also been shown to be more accurate compared to other-age faces. (e.g., Fölster et al., 2014). Another social factor is the caregiver status of the receiver relative to the sender. For instance, according to Lorenz (1943), infant faces present a special set of cognitive representations (schema) triggered by certain anatomical characteristics, referred to as ‘Kindchenschema,’ representing features such as big eyes and round faces on a large head, inducing motivation for caregiving behavior, especially among adults (Glocker et al., 2009). Although one might expect that caregiver tendencies would inhibit mimicry of negative expressions, Korb et al. (2016) have shown that adults mimicked angry infant faces, especially when oxytocin levels in the observers were increased.

In contexts where different influences converge on an intended facial expression, mixed expressions of several emotions may be produced. For example, Ekman (1985) described ‘miserable’ or ‘masking’ smiles when one tries to conceal a smile by pulling down the lip corners, or fake smiles when the orbicularis is not involved. Further, the concept of evaluative space (Cacioppo & Berntson, 1994), that is, the simultaneous activation of happiness and sadness, has sometimes been reported (e.g., Williams & Aaker, 2002). Facial expressions change (appear and disappear) perpetually and rapidly to suit changing situations and adapt to the goal(s) being pursued. Some contexts may require the inhibition of reflexive (prepotent) nonverbal facial responses in favor of reflective (conscious) facial responses. Complex ‘mixed’ facial expressions have been demonstrated, especially when neurobiological mechanisms (e.g., mirror neurons) and visual signals (e.g., faces of others) interact in a given context (Scherer & Ellgring, 2007).
The Present Study

It was the general aim of the present study to investigate how deliberate facial expressions are influenced in their quality by the congruency with facial expressions of others, depending on whether the person (expressor) is another adult or an infant. To investigate these questions, we used a Stroop-type task that involved the production of one of two facial expressions (i.e., a smile or a frown) in response to a written verbal command presented on different background stimuli. Background stimuli were face images of adults or infants with positive, negative, or neutral expressions and scrambled faces. Unlike most other studies of this kind, the facial responses were video-recorded and analyzed with OpenFace software (Baltrusaitis et al., 2018), a machine classifier that identifies 17 different action units (AUs) in the face. AUs are visually distinguishable actions in the face, often but not exclusively caused by contraction or relaxation of individual facial muscles (Ekman & Friesen, 1976; Rosenberg, 2005). OpenFace software provides activity measures for each identified AU per video frame. Our main interest in the present report was how the produced facial expressions are affected by the congruency with the sender’s facial expression and the caretaker status relative to the sender. We used the frame-based output of OpenFace software to determine on a trial-by-trial basis (1) The error rates and RTs of required responses, and (2) The quality of the expression by assessing all 17 identified AUs. These data allowed us to address two main questions.

Question 1: The effect of congruency of the facial expressions of background faces on deliberate expressions. Firstly, we were interested in congruency effects on the targeted deliberate expression. Specifically, we hypothesized that, (1) For congruent conditions, RTs would be shorter than in the incongruent condition, while expression accuracy and mean AU activation might be higher due to possible facilitation effects of congruence. (2) In incongruent conditions, error rates should be higher, and RT should be longer than in neutral and congruent conditions. Furthermore, the mean AU activity of the expression might be diminished as a result of possible inhibition effects of incongruence. Second, it was of interest, whether AUs that do not ‘belong’ to a required target expression would be co-activated as a function of congruency, possibly ‘contaminating’ the target responses, leading to ‘mixed’ facial expressions. That is, anger-associated AUs might be active during the incongruent smile condition, and smile-associated action units might be activated during incompatible frown trials.

Question 2: The effects of the age of background faces on deliberate facial expressions. Although (Korb et al., 2016) had shown that angry infant faces were also mimicked, caregiving might attenuate the interference from negative relative to positive expressions. Hence, when an infant face shows a distressed expression, mimicry may be overridden or counteracted by empathy. Therefore, one might expect a global facilitation of smiles and attenuation of frown responses when smiling infant faces are seen. Specifically, in smile production trials, facilitation by compatible background faces might be enhanced and inhibition by incompatible stimuli might be diminished by infant as compared to adult background faces. Conversely, for frown trials, facilitation by compatible negative expressions may be diminished for infant background stimuli, whereas incompatible (smiling) infant faces might lead to mixed emotional expressions.
Method

Participants

Twenty-two healthy adults, compensated with either course credits or 8 EUR per hour, participated in the experiment. Prior to participation, written informed consent was obtained. The study was approved by the Ethics Committee of the Department of Psychology of the Humboldt-Universität zu Berlin. We excluded data of three female participants due to high error rates (>40%). During video data processing, we lost one participant’s video data due to technical reasons. Of the remaining 18 participants (11 females and 7 males, mean age: 26.12 years, SD = 5.6, range 21–39), one was left-handed. None reported any history of neurological or mental disorder, and all reported normal or corrected-to-normal visual acuity. Our sample size, N = 18, was not informed by an a priori power analysis.

Stimuli

Background stimuli were face pictures of adults and infants with different emotional expressions. The face stimuli consisted of pictures of 60 adults (30 females) from the Radboud Faces Database (Langner et al., 2010) and of 60 infants (30 girls) from online media sources. Neutral faces and scrambled faces were used as control conditions (see Fig. 1). Each face stimulus displayed a positive, negative, or neutral expression. The face pictures were matched in size and brightness. The German words for happiness and anger (i.e., Freude, Ärger) written in black were superimposed on the faces in the midline and at the saddle of the nose. From the viewing distance of 70 cm, the visual angle of the pictures was 4.01°×5.35°. The words Freude and Ärger, indicated participants to smile or to frown, respectively. Background stimuli were adult and infant faces, showing negative, positive, or neutral expressions and scrambled faces. Words and types of faces were combined orthogonally, yielding congruent, incongruent, or neutral expression/word combinations or combinations with scrambled faces.

Procedure

Upon arrival at the lab, participants signed an informed consent form, followed by a short questionnaire to capture details of age, handedness, visual acuity, and alcohol consumption. They were then seated in a dimly lit, sound attenuated cabin in front of a computer monitor at a viewing distance of approximately 70 cm. At the bottom of the computer monitor a Sony EVI-D70P video camera was placed, which recorded the participant’s facial expressions at 25 frames per second. To the left and right of the monitor a vertical strip of LEDs was placed to illuminate the participant’s face.

Fig. 1 Examples for face-word compound Stimuli

| Congruent | Incongruent | Neutral | Scrambled |
|-----------|-------------|---------|-----------|
| Arger     | Freude      | Freude  | Arger     |

Springer
Apart from the videos reported here, in the present study we also recorded multichannel EEG and EMG for a different purpose, to be reported elsewhere. The EEG was recorded from electrodes placed inside an elastic cap, covering the top region of the forehead and the EMG was recorded unilaterally on the right side, with two pairs of small electrodes above the *M. zygomaticus major* and *M. corrugator supercilii*, respectively. Although, the four electrodes placed within the face and the electrode cap may have had some impact on the video data a comparison of the chronometric outcomes from EMG and video onsets indicated only small differences (for a systematic comparison of EMG and unoccluded video onsets, see Beringer et al., 2019).

The experimental procedure was controlled by “Presentation” software (version 19.0 build 11.14.16). The experiment consisted of 16 experimental conditions, each containing 60 trials, totalling 960 trials. The images were presented on a light grey background (RGB = 227/227/227), with an oval frame covering hair, ears, and neck of the faces. Trials of all conditions appeared randomly throughout the experiment, with breaks of self-determined duration after every 120 trials. At the beginning of each trial, a fixation cross appeared in the centre of the screen for 500 ms, followed by a face-word compound stimulus for 2 s. After a further 900–1000 ms (randomised) a “stop” signal was presented (for 900–1000 ms, randomised), followed by the next trial (Fig. 2).

Participants were instructed to smile or frown according to the word superimposed on the visual stimulus as fast and accurately as possible (see Fig. 2). When the “stop” signal was presented, participants should relax their facial muscles and return to a neutral facial expression. The required facial response involved the activation of AU4 (*m. corrugator supercilii/brow lowerer*) for frowns and AU12 (*m. zygomaticus major/lip corner puller*) for smiles (Tian et al., 2005). The instruction did not refer to action units or facial muscles but only asked to show a smile or a frown to the corresponding word.

**Video Data Processing**

The video recordings were analysed frame-by-frame with OpenFace (version 2.0.2, Baltrusaitis et al., 2018). The OpenFace algorithm had been trained on faces viewed from several perspectives and shows high accuracy for facial landmark detection and AU recognition even in non-frontal views. Outputs provided measures of activation for 17 AUs, including AU4 and AU12, the AUs taken as representing the target expressions, that is, frowns and smiles, respectively. Values around 0 indicate no activity, and value increases up to 1 and above indicate activation of a given AU. The OpenFace output text files were converted to.xlsx format in Microsoft Excel (v. 2016). Using timestamp information in the output, the data stream
containing stimulus events was merged with the excel files from OpenFace output. Using a similar procedure as reported by Recio and Sommer (2018), the datasets containing stimulus event information were processed with MATLAB (R2016a, The Math Works, 2016).

For each trial a fixed set of 90 consecutive frames was defined, including 5 pre-stimulus frames and 85 post-stimulus frames, covering a 3.6-s time interval. A baseline correction was applied in each trial by subtracting the average AU intensity scores over the 5 pre-stimulus frames (200 ms).

In order to measure participants’ facial expressions and determine response correctness, data for ‘frown’ and ‘smile’ responses were parameterised from AU4 and AU12, respectively. Target versus distractor channels were defined depending on the required response per stimulus condition. That is, in trials where the word prompted a smile, AU12 was the target channel and AU4 was the distractor channel, and vice versa for frowns. Expression onset and offset thresholds were defined for each participant as 50% of the mean activity for all happy congruent trials (for AU12) all angry congruent trials (for AU4). On average, these thresholds were 0.26 (range 0.09–0.68) for AU4 and 0.81 (range 0.43–1.52) for AU12 across participants. Activity onsets occurring within the first three frames (120 ms) after stimulus onset, were considered as fast guesses and excluded from further analyses (Recio & Sommer, 2018). A response was identified as correct when activity in the target AU channel preceded any activity in the distractor AU channel and lasted for at least seven consecutive frames above threshold. All other trials were considered as errors. The threshold latency in a target trials was taken as reaction time (RT) in that trial; RTs were only analysed for correct trials (hits).

Activities for each of the 17 AUs were calculated during a period of 8-frames (320 ms) following the frame at which the threshold was reached in the target AU (=RT) in a given trial. These activities were averaged condition- and participant-wise but only when the trial was correct.

Data Analysis

The combination of required response (smile, frown) and stimulus type (happy, angry, neutral, scrambled) resulted in compatible trials (smile response plus happy face; frown response plus angry face), incompatible trials (smile response plus angry face; frown response plus happy face), neutral (smile or frown response plus neutral face), and scrambled trials (smile or frown response plus scrambled face). Data from accuracy, RTs (of correct responses), and AU mean activity were submitted to analyses of variance with repeated measures (rmANOVA) on factors congruency (congruent, incongruent, neutral, scrambled), stimulus face age (infant, adult), and the type of response (smile, frown). In an additional step, the global congruency effect was separated into an inhibition component (incongruent vs. neutral) and a facilitation component (congruent vs. neutral). Post-hoc tests were corrected for multiple testing with the Bonferroni method.

Results

Error Rates

Figure 3 shows the mean error rates for smile and frown responses in the different conditions. The present study had used both neutral and scrambled faces as reference conditions in order to assess the effects of presenting a face stimulus as compared to an unstructured
visual noise pattern. We determined in a first step whether the error rates in the two neutral conditions were significantly different from each other, using an rmANOVA including the factors reference condition (neutral, scrambled), stimulus face age (infant, adult), and response (smile, frown). The reference conditions did not significantly differ from each other, $F(1, 17) = 0.44$, $p = 0.52$, $\eta^2 = 0.03$, and did not differentially affect error rates (see Table 1). Hence, the following rmANOVA on response and stimulus face age excluded scrambled faces from factor congruency (congruent, incongruent, neutral).

As shown in Table 2, smile responses were fairly accurate ($M=90.2\%$ correct), while frown responses were significantly less accurate ($M=75.5\%$), $F(1, 17) = 16.35$, $p = 0.001$, $\eta^2 = 0.49$. Congruency as a main effect was significant ($F(2, 34) = 3.30$, $p = 0.049$, $\eta^2 = 0.16$) and interacted with response ($F(2, 34) = 7.64$, $p = 0.002$, $\eta^2 = 0.31$). Post-hoc tests for smile responses confirmed a main effect of congruency ($F(2, 34) = 5.38$, $p < 0.01$, $\eta^2 = 0.24$) but without significant facilitation or inhibition components relative to the neutral condition. In a post-hoc test just for frown responses, the factor congruency failed significance.

**Reaction Times**

As was the case for error rates, we started the analysis of RTs, measured in AU12 (smile responses) and AU4 (frown responses) (Fig. 4) by comparing the two reference conditions. Again, two reference conditions did not significantly differ ($F(1, 17) = 0.08$, $p = 0.79$, $\eta^2 = 0.05$).
consistent across all conditions, congruent and incongruent trials elicited the shortest and longest RTs, respectively, with neutral trials in between (see Table 3). This was confirmed by a main effect of congruency ($F(2, 34) = 16.48, p < 0.001, \eta^2 = 0.49$), indicating longer RTs in incongruent than in congruent conditions (Stroop effect) by $M_{\text{diff}} = 67.0$ ms. However, this pattern appeared to be modulated by both stimulus age and type of response, with smallest congruency effects for frown responses with infant background faces (Fig. 5), as reflected in a 3-way interaction of congruency, stimulus age, and response ($F(2, 34) = 6.80, p = 0.003, \eta^2 = 0.29$).

Post-hoc analyses of RTs of smile responses alone yielded a significant main effect of congruency ($F(2, 34) = 24.48, p < 0.001; \eta^2 = 0.59$) but no effect of stimulus age ($F < 1$). The factor congruency involved significant inhibition ($F(1, 17) = 14.76, p = 0.002; \eta^2 = 0.47$) ($M_{\text{diff}} = 38.0$ ms) and facilitation components ($F(2, 34) = 13.20, p = 0.004; \eta^2 = 0.44$) ($M_{\text{diff}} = 32.13$ ms) but no interactions with stimulus age.

In RTs of frown responses, the main effect of congruency ($F(2, 34) = 6.24, p = 0.01, \eta^2 = 0.27$), was modulated by stimulus age ($F(2, 34) = 5.65, p = 0.02, \eta^2 = 0.25$). When congruency was tested separately for adult and infant faces, adult faces yielded a significant main effect ($F(2, 34) = 14.17, p < 0.001, \eta^2 = 0.46$), whereas infant faces did not ($F < 1$). In addition, the factor congruency (for adult faces only) involved a significant facilitation component ($F(1, 18) = 16.47, p = 0.003, \eta^2 = 0.45$) ($M_{\text{diff}} = 77.7$ ms) but no inhibition component ($M_{\text{diff}} = 23.5$ ms). Therefore, the three-way interaction of congruency, stimulus age, and response seems to be due to the absence of a congruency effect for infant stimuli when frowns were the required facial response. Probably as a spin-off of the three-way interaction, there was no congruency effect for infant faces when smiles were the required facial response. A significant facilitation component ($F(1, 18) = 16.47, p = 0.003, \eta^2 = 0.45$) ($M_{\text{diff}} = 73.7$ ms) was found for adult faces, indicating that congruent stimuli elicited shorter RTs than incongruent stimuli.

### Table 2 Mean accuracy rates for smile and frown responses to infant and adult face stimuli per condition

| Condition | Smiles | | Frowns | |
|-----------|--------|--------|--------|
|           | Infants | Adults | Infants | Adults |
| % Correct | Congruent | 0.92 (0.02) | 0.91 (0.10) | 0.77 (0.16) | 0.76 (0.16) |
|           | Incongruent | 0.89 (0.10) | 0.88 (0.10) | 0.76 (0.14) | 0.76 (0.16) |
|           | Neutral | 0.91 (0.08) | 0.90 (0.10) | 0.75 (0.14) | 0.73 (0.15) |
|           | Scrambled | 0.89 (0.12) | 0.90 (0.09) | 0.76 (0.14) | 0.75 (0.15) |
| Effect    | Stroop (IC-C) | 0.03 | 0.003 | 0.01 | 0.00 |
|           | Facilitation (N-C) | 0.01 | 0.01 | 0.02 | 0.03 |
|           | Inhibition (IC-N) | 0.02 | 0.02 | 0.01 | 0.03 |

*Fig. 4 Overall mean RTs (ms) and SDs for AU4 and AU12*
interaction, there was also an interaction of stimulus age and congruency \((F(2, 34) = 4.12, p = 0.03, \eta^2 = 0.20)\).

**AU Activities**

For the sake of simplicity, analyses of the activity of facial expressions was performed separately for smile and frown responses. Within each of these response conditions we analysed in a first step the activity of all AUs that are involved in the corresponding expression. Hence, we analysed AU12 (*Lip Corner Puller; m. zygomaticus major*) and AU6 (*Cheek Raiser; m. orbicularis oculi, pars orbitalis*) as part of the smile expression and analysed AU4 (*Brow Lowerer; m. depressor glabellae, m. depressor supercili; m. corrugator*), AU5 (*upper lid raiser, m. lavator palpebrae sup.*), AU7 (*Lid Tightener, m. orbicularis oculi, pars palpbralis*), and AU23 (*Lip Tightener, m. orbicularis oris*) as constituting the frown response (Rosenberg, 2005). For the activity of each of these AUs we assessed effects of *stimulus face age* (adult, infant) and congruency (congruent, incongruent, neutral).

In a second step, we checked whether activation of non-target AUs might leak into or mix with the target expression, especially in the incongruent conditions (for smiles: AU4, AU5, AU7, AU23; for frowns: AU6, AU12).

Finally, we assessed—for each required facial response—any experimental effects on other AUs (AU1 to AU3, AU5, AU9 to AU11, AU14, AU15, AU17, AU20, AU25, AU26, and AU45).
AU Activities in Smile Responses

Activities of the smile-associated AU12 and AU6 are shown in Fig. 6 and Table 4. The ANOVA of AU12 activity yielded a significant main effect of congruency ($F(2, 34) = 4.77, p = 0.03, \eta^2 = 0.22$). Post-hoc tests of the inhibition and facilitation components showed stronger activity of AU12 in congruent than in neutral conditions ($F(1, 17) = 9.53; p = 0.014; \eta^2 = 0.36$), while there was no significant inhibition component (neutral vs. incongruent). ANOVA of AU6 activity did not yield significant effects.

Leakage of frown activity into required smile responses were assessed by analysing experimental effects on AU4, AU5, AU7, and AU23 within the smile condition. After Bonferroni correction, none of these AUs yielded statistically significant effects (Table 4).

During smile responses there was also no statistically significant effect in AU activity (all $p > 0.05$) in any of the AUs that are neither related to smile or frown expressions (i.e., AU1 to AU3, AU5, AU9 to AU11, AU14, AU15, AU17, AU20, AU25, AU26, and AU45).

In sum, in deliberate smile expressions the activity of AU12 (lip raiser; m. zygomaticus major) increased in compatible relative to neutral trials; no other effects were observed.

Table 4 RTs ANOVA for all four conditions (congruent, incongruent, neutral, scrambled), and ANOVA for RTs without scrambled face.

| Conditions            | Four levels of congruency | Three levels of congruency |
|-----------------------|---------------------------|---------------------------|
|                       | $df$ | $F$  | $p$  | $df$ | $F$  | $p$  |
| Stimulus face age (A) | 1.17 | 2.03 | 0.17 | 1.17 | 1.84 | 0.19 |
| Congruence (C)        | 3.51 | 11.68| 0.00 | 2.34 | 16.48| 0.00 |
| Response (R)          | 1.17 | 0.23 | 0.64 | 1.17 | 0.30 | 0.59 |
| $A \times R$          | 1.17 | 0.16 | 0.70 | 1.17 | 0.04 | 0.84 |
| $A \times C$          | 3.51 | 3.33 | 0.027| 2.34 | 4.12 | 0.025|
| $R \times C$          | 3.51 | 0.23 | 0.88 | 2.34 | 0.26 | 0.78 |
| $A \times R \times C$ | 3.51 | 4.18 | 0.01 | 2.34 | 6.80 | 0.003|

Fig. 6 Activity in AU12 and AU6 for Smile Responses. Activity was a measure of AUs 12 and AU 6 activation under the different stimulus conditions.
AU Activities in Frown Responses

Activities in the frown response-associated AU4, AU5, AU7, AU23 are shown in Fig. 7 and Table 5. ANOVAs of the measures of each of these AUs did not yield any statistically significant effects of the experimental factors congruency or stimulus age ($p > 0.05$).

Leakage of smile activity into the required frown responses were assessed by analysing experimental effects on AU12 and AU6 within the frown condition. After Bonferroni correction, none of these AUs yielded statistically significant effects (Table 5).

When mean activities in AUs not belonging to the frown or smile expression (i.e., AU1 to AU3, AU5, AU9, AU10, AU11, AU14, AU15, AU17, AU20, AU25, AU26, and AU45) were analysed for the frown condition, we found the AU2 (outer brow raiser, m. frontalis, pars lateralis) to be modulated by an interaction of congruency and stimulus age ($F(2,34) = 7.32, p = 0.02, \eta^2 = 0.30$). Post-hoc analysis indicated a significant interaction of congruence by stimulus age for the facilitation component ($F(1, 17) = 26.64, p < 0.001, \eta^2 = 0.61$) (Table 6). That is, relative to neutral background faces, angry infant faces elicited increased AU2 activity ($M_{\text{diff}} = 0.019$), whereas angry adult background faces decreased AU2 activity ($M_{\text{diff}} = −0.020$).

In sum, frown expressions did not clearly differ in quality by smile intrusions. However, the outer brow raiser (AU2) was modulated by compatible frowning background faces in different directions, depending on whether the background face showed adults or infants (Fig. 8).

Discussion

We examined how deliberate facial expressions of happiness (smiles) and anger (frowns), prompted by written commands, are modulated by the congruency with the facial expression of background faces and how this effect depends on the age of the background face (adults vs. infants). A machine classifier for facial expressions enabled us to analyse multiple aspects of these effects. In addition to response accuracy and the chronometrics of the deliberate expressions, we considered the mean AU activity of the target expressions and leakage by non-target emotions into the target expressions.

Before discussing the main questions of this study, we should shortly mention two side issues. Firstly, we had used two kinds of reference conditions, faces with neutral expressions, and scrambled faces because we were concerned whether presenting a background face with neutral expression might interfere per se with deliberate expressions. However,
Table 5  Mean activity of AUs in different conditions during ‘Smile’ responses and ANOVA results

| AU    | Adult face | Infant face | ANOVA Factors |
|-------|------------|-------------|---------------|
|       | IC         | N           | C             | C   | A   | C x A |
| 01    | -0.155     | -0.146      | -0.124        | -0.145 | -0.154 | -0.146 | 2.0  | 0.78 | 2.8  |
| 02    | -0.072     | -0.073      | -0.060        | -0.067 | -0.082 | -0.075 | 1.72 | 1.50 | 1.84 |
| 04    | -0.206     | -0.212      | -0.220        | -0.211 | -0.211 | -0.213 | 1.00 | 0.04 | 0.56 |
| 05    | -0.044     | -0.048      | -0.059        | -0.050 | -0.054 | -0.059 | 3.50 | 1.42 | 0.30 |
| 06    | 1.206      | 1.200       | 1.223         | 1.222 | 1.207 | 1.212 | 1.44 | 0.35 | 1.55 |
| 07    | 0.227      | 0.213       | 0.242         | 0.221 | 0.225 | 0.221 | 0.31 | 0.11 | 0.74 |
| 09    | 0.114      | 0.103       | 0.103         | 0.108 | 0.965 | 0.113 | 1.10 | 0.05 | 1.49 |
| 10    | 0.498      | 0.503       | 0.520         | 0.525 | 0.513 | 0.508 | 0.15 | 0.77 | 1.42 |
| 12    | 1.505      | 1.492       | 1.527         | 1.519 | 1.505 | 1.515 | 4.77*| 0.45 | 1.92 |
| Facil | –          | 1.492       | 1.527         | –     | 1.505 | 1.515 | 9.53*| 0.00 | 3.09 |
| Inhib | 1.504      | 1.505       | –             | 1.519 | 1.505 | –     | 3.53 | 1.91 | 0.01 |
| 14    | 0.834      | 0.811       | 0.821         | 0.807 | 0.819 | 0.803 | 0.20 | 3.42 | 1.14 |
| 15    | -0.156     | -0.163      | -0.166        | -0.144 | -0.198 | -0.143 | 1.36 | 0.00 | 1.77 |
| 17    | -0.250     | -0.253      | -0.270        | -0.244 | -0.261 | -0.263 | 0.60 | 0.26 | 0.43 |
| 20    | 0.137      | 0.156       | 0.150         | 0.150 | 0.103 | 0.131 | 2.31 | 5.18 | 2.53 |
| 23    | 0.331      | 0.302       | 0.298         | 0.310 | 0.308 | 0.319 | 1.21 | 0.04 | 1.67 |
| 25    | 0.682      | 0.700       | 0.755         | 0.729 | 0.672 | 0.715 | 3.80 | 0.46 | 1.98 |
| 26    | -0.163     | -0.166      | -0.148        | -0.165 | -0.134 | -0.191 | 1.16 | 0.20 | 4.97 |
| 45    | 0.246      | 0.233       | 0.278         | 0.261 | 0.238 | 0.277 | 5.19 | 0.39 | 0.24 |

C congruence; A stimulus age. Shaded areas indicate the results for AU12, the main AU for smile responses

*p < 0.05; **p < 0.01; ***p < 0.001; †p < 0.1
| AU  | Adult face | Infant face | ANOVA |
|-----|------------|-------------|-------|
|     | IC         | N           | C     | C     | A     | C x A |
|     |            |             |       |       |       |       |
| 01  | 0.080      | 0.098       | 0.080 | 0.096 | 0.063 | 0.083 |
| 02  | -0.061     | -0.051      | -0.071| -0.070| -0.084| -0.065|
| Facil | -         | -0.051      | -0.071| -     | -0.084| -0.065|
| Inhib| -0.061     | -0.051      | -     | -0.070| -0.084| -     |
| 04  | 0.556      | 0.558       | 0.567 | 0.554 | 0.562 | 0.566 |
| 05  | -0.045     | -0.049      | -0.045| -0.053| -0.053| -0.048|
| 06  | 0.130      | 0.133       | 0.141 | 0.148 | 0.145 | 0.137 |
| 07  | 0.363      | 0.364       | 0.362 | 0.397 | 0.355 | 0.386 |
| 09  | 0.055      | 0.047       | 0.058 | 0.058 | 0.059 | 0.060 |
| 10  | 0.030      | 0.038       | 0.038 | 0.039 | 0.04345| 0.045 |
| 12  | -0.063     | -0.053      | -0.059| -0.055| -0.051| -0.066|
| 14  | -0.061     | -0.059      | -0.057| -0.057| -0.061| -0.073|
| 15  | 0.065      | 0.066       | 0.071 | 0.073 | 0.059 | 0.081 |
| 17  | 0.149      | 0.172       | 0.204 | 0.177 | 0.195 | 0.199 |
| 20  | 0.011      | -0.001      | 0.001 | 0.016 | -0.011| -0.005|
| 23  | 0.024      | 0.026       | 0.021 | 0.018 | 0.023 | 0.027 |
| 25  | -0.097     | -0.075      | -0.065| -0.063| -0.092| -0.060|
| 26  | 0.019      | 0.000       | 0.015 | -0.004| 0.0192| 0.027 |
| 45  | 0.143      | 0.109       | 0.134 | 0.148 | 0.123 | 0.137 |

* C congruence; A stimulus age. The shaded area indicates the results for AU4, the main AU for frown responses
* * p < .05; ** p < .01; *** p < .001; † p < .1
the comparison of error rates and RTs between neutral and scrambled faces did not yield notable differences, with one exception discussed below, indicating no distracting effect by a neutral background face in particular. Therefore, we simplified the analyses of congruency effects by dropping the scrambled face condition. Second, the type of deliberate response had a marked effect on error which were considerably higher for frowns than for smiles. We hold this to be partly due to the fact that in parallel to the videos, we had recorded facial EMG, requiring to place electrodes above the corrugator and zygomaticus muscles; hence the corrugator electrodes may have occluded AU4 to some extent. This conclusion is supported by the fact that the analysis of the EMG (reported elsewhere) yielded more accurate performance for frown responses than the present video analysis. Importantly, the RTs of correct responses were not significantly different for the required expression. Therefore, we believe that the video analysis yielded interpretable results and that smiles and frowns per se are similar in their demands on the participants.

The Effect of Congruency between Response and Background Expressions

The most consistent finding in the present data set are the effects of congruency. In error rates, congruency effects were seen only for smiles. However, in RTs and with adult background faces, the congruency effect was strong and indistinguishable for frowns and smiles. There was a mean increase of RTs from congruent to incongruent conditions (Stroop effect) of 67.0 ms, when averaged over both frown and smile responses. Relative to the neutral condition, there was a clear facilitation due to congruence, and inhibition due to incongruence.

The congruency effects may be explained by facial mimicry eliciting automatic matched response tendencies. Facial mimicry is considered to be rooted in perception-behaviour links where facial expressions of others are spontaneously copied by the observers (e.g., Chartrand & Bargh, 1999; Dimberg & Petterson, 2000; Otte et al., 2011a). Such mimicry involves reflexive responses with little or no cognitive conflicts, thus facilitates performance in compatible trials. In contrast, incompatible trials such as smiling at an angry face, may involuntarily activate facial expressions of the frowning distractor face stimulus as opposed to the required smile. Hence, in incompatible trials, mimicry may activate incorrect responses, or may call for a more reflective response, causing conflicts that need to be solved. Increased RTs in incompatible conditions are presumably due to the necessary and time consuming recruitment and deployment of cognitive processes required to detect and resolve conflicts in incongruent trials (Egner et al., 2008; Etkin et al., 2006).
The present results that congruency effects in RTs do not depend on the type of response appears to be at variance with the idea that mimicry is stronger in conditions with affiliative intent (e.g., Bourgeois & Hess, 2008). At the least it demonstrates that requiring frowns versus smiles does not constitute a manipulation of affiliative intent or else the congruency effect should have been smaller for frowns than for smile responses. However, this was not the case for our adult background faces. Also, the more specific facilitation and inhibition components were not modulated by the type of response.

Of main interest for the present analysis were the AU activity measures provided by the OpenFace software. Hence, for each required deliberate expression, we measured the effects of congruency on the intensity of the expression itself, whether there was any leakage or intrusion from the alternative expression, and whether there was any effect on AUs not related to any of the target expressions. Activity at the AU12 for smile responses yielded a main effect of congruency, and—specifically—a significant facilitation effect (congruent vs. neutral), but no inhibition effect (neutral vs. incongruent). This indicates that the intensity of smile responses benefited from congruent conditions.

We assessed possible intrusions or leakage into required smile responses by analysing experimental effects on frown-related AUs (AU4, AU5, AU7, and AU23) within the smile condition, none of which yielded any statistically significant effects. There was also no congruency effect in AUs that were not related to any target expression. Together, these findings indicate that smile responses were robust against the intrusion of frown activity due to background distractor faces.

Activities at the frown-related AU4, AU5, AU7, AU23 were stable across all stimulus conditions, with no congruency, facilitation, or inhibition effects, and no intrusions of smile activity into the required frown response as assessed by analysing experimental effects on AU12 and AU6 within the frown condition.

In sum, the speed of smile and frown responses are remarkably and similarly sensitive to the congruency of task-irrelevant background faces. It seems that deliberate smile expressions become more intense when the background face is also smiling. In contrast, the mean AU activities of frown responses were unaffected by background faces.

The increased sensitivity to background faces during smile as compared to frown responses in terms of intensity may be complementary to findings that negative perceived emotional expressions, such as anger, are mimicked less than positive expressions (smiles). It is held that mimicry of smiles is more acceptable and bears no personal costs (Bourgeois & Hess, 2008; Chartrand & Bargh, 1999; Dimberg & Petterson, 2000; Künecke et al., 2017). Further, socialisation processes instil display rules that nurture friendliness (e.g., smiling) instead of aggression (e.g., frowning), in favour of facial expressions are socially adaptive (Ekman & Friesen, 1969a). Whereas these previous findings relate to the effects of mimicry as a function of the perceived emotion, the present findings indicate a differential role of the deliberate expression. We suggest that requiring smiles versus frowns may influence the communicative intent, facilitating mimicry during emotionally positive actions (smiles), but less so during emotionally negative actions (frowns).

The Effect of the Age of Background Faces on Deliberate Facial Expressions

The second main question of the present analysis was a comparison of adult and infant background faces. In particular, we expected that mimicry of negative infant expressions would be attenuated or even overridden by empathy or caregiving responses, possibly leading to a tendency to smile, providing comfort to the infant. This tendency should interfere
with deliberate frowns directed at infants and block the facilitation of frown responses by images of infants in distress (displaying anger).

Accuracy and RTs of deliberate expressions were not affected by background stimulus age as main effects. However, RTs were significantly modulated by the interaction of face stimulus age, congruency, and type of response. Importantly, the smallest Stroop effect in this study was recorded when the target expression was frowning with infant faces in the background. Chronometric analysis found significant congruency effects with a facilitation component for frown responses to adult distractor faces but not for infant distractor faces. Considering that the congruency effect for smiles was not modulated by face age the strong modulation shows clearly that stimulus age affected the control over deliberate frown expressions. This is broadly in line with Fridlund’s (1994) behavioural ecology theory that facial expressions are audience-specific tools of communicating behaviour intentions. The different congruency effects in RTs for frowns to infant and adult face stimuli may therefore represent a difference in audience. Hence, addressing an infant with a frown may be experienced as inappropriate, counteracting any tendency of mimicking both negative and positive expressions. The observed face age effect in frowns may also be explained by findings that mimicry is influenced by personal, emotional, motivational, relational and social characteristics of both the mimicker and the mimickee (Chartrand & Lakin, 2013; van Baaren et al., 2009). Specifically, cognitive representations (schemata) of infant and adult faces seem to differ and the anatomical characteristics of infant faces (Kindchen-schema), such as big eyes and round faces on a large head, induce caregiving tendencies (Glocker et al., 2009). Obviously, infant faces with negative expressions are more likely to be appraised as needing care rather than as a threat.

Hence, frowning at infant faces may trigger very different mechanisms than frowning at adult faces. In contrast, there seems to be no principled difference between smiling to adult and infant faces. However, can we derive evidence from the present data on the nature of the mechanisms driving the differences between frowning at adult and infant face? One clue may come from the analysis of activities of AUs not belonging to the smile or frown expression (i.e., AU1, AU2, AU9, AU10, AU14, AU15, AU17, AU20, AU25, AU26, AU45). During the deliberate smile condition, no statistically significant activity was found, indicating no expression distortions or leakage in the form of irrelevant facial expressions. However, when activities of the AUs not belonging to the frown or smile expression were analysed during the frown condition, AU2 (outer brow raiser) activity was modulated by the congruent condition (i.e., angry background expressions) into opposite directions for adult and infant faces. Specifically, relative to the neutral condition AU2 activity increased for angry infant faces and decreased for angry faces of adults. This may indicate that seeing the face of an infant in distress triggers an expression that is somewhat different from the prototypical expression for anger, resulting in a blended expression (Ekman, 1985).

Generally speaking, faces of infants in distress may have initiated a different kind of reappraisal (Scherer & Ellgring, 2007), tapping into other response actions than involuntary mimicry. According to Scherer and Ellgring (2007), appraisal sequence and eventual appropriate response outcomes seem to cumulatively recruit AU2 as appraisals shift from inferences of novelty and importance (AU1 + AU2), to possible obstruction of goals (AU1 + AU2 + AU4), to situations appraised as unfair and unethical (AU1 + AU2 + AU4 + 17 + 23 + 10 + 14). Hence, the present observation indicates that during the deliberate frown condition, infant faces, especially distressed infant faces, may trigger a reappraisal of the situation differing from angry adult faces. Activation of AU2 may therefore be an accumulative effect of an appraisal of showing anger rather than empathy to distressed infants.
as ‘unfair’ and against caregiving tendencies (Glocker et al., 2009). This may not be the case for encountering an angry adult, in which case detection of anger and corresponding response is appropriate as it could signify danger.

Infant faces represent a unique set of conditions (Glocker et al., 2009). When used as stimuli, infant faces may present a context leading to complex ‘multimodal’ facial expressions as neurobiological bases (mirror neurons, Kindchenschema, caretaker reflex), visual signals (frowning at infant faces), and reappraisals interact. We presume that the activation of AU2 may be part of a complex multi-layered facial expression due to possible reappraisal of the faces of infants in distress under unique conditions. The “inhibition hypothesis” (Darwin, 1872) posits that highly intense facial expressions may not be subject to voluntarily control. In some situations, neuroanatomical and cultural differences may interact in a way that produces facial expressions that escape efforts to control or mask, thus emerging as leaked expressions (Ekman & Friesen, 1969b). Such leakage of expressions has been shown to be most likely under incongruent conditions (Jens, 2017). Seeing the face of an infant in distress may have elicited reactions that cannot be inhibited, and therefore will affect AU2.

Limitations and Perspectives

The present study has demonstrated that rich information can be derived from machine classifier-based video analysis about how cognitive control can modulate the quality of deliberate facial expressions beyond effects in RTs. However, the present study has some limitations that might be overcome in future work. While infant smiles are unambiguous, it is challenging to differentiate negative emotions on infant faces (Camras & Shutter, 2010), which makes it difficult to define “compatible” response conditions. As mentioned, the accuracy of recording frown responses with video may have been compromised to some extent by the EMG electrodes. Although we believe that the electrodes did not invalidate the present results, especially those from unoccluded AUs, future studies should avoid such occlusion. This seems to be all the more feasible, the accumulating evidence that machine classifiers provide rich information about both chronometry and quality of facial expressions (see also Beringer et al., 2019). Another limitation is the camera perspective; the camera recorded faces from below rather than horizontally. However, the OpenFace software algorithm was trained for different viewpoints and is able to deal with this situations, which was also our impression, when we compared video outputs from different angles. Further, the results are limited by a rather low number of participants, which requires a minimum effect size of 0.65 in order to detect a significant effect with 95% probability.

Conclusions

Previous studies have shown that in facial expression production tasks, incongruent background faces can modulate the speed of facial expressions, whether recorded with EMG or machine classifier of videos. This was confirmed in the present report for machine classifier-based analyses. Here, we go beyond these studies in at least two respects. Firstly, we used the machine classifier to assess also the quality of the responses and not just their onset speed. Thus we found that in compatible conditions, smiles are not only faster but also brighter. Second, we investigated the effects of a potential caregiving relationship to
the model presented on the screen. It turned out that the chronometric pattern is very different when frowns are required to infants, both in terms of a strong attenuation of the congruency effect and in the quality of the facial expression. The required frowns showed intrusions that can be interpreted as the result of considering the required expression as unfair. Apart from these concrete findings, the present study may serve as an encouragement that the low-cost recording and machine classification of facial expressions is not only a cheap alternative to EMG recordings but may—in at least some respects—yield even richer information.

Acknowledgements We thank Thomas Pinkpank for technical advice for the video analysis. This study was supported by the Deutscher Akademischer Austauschdienst (DAAD) Grant 57,460,842 to Stephen Katembu, and by the Science Foundation of Ministry of Education of China (Grant No. 18YJC1900027), the Zhejiang Provincial Educational Science Planning Project (Grant No. 2018SCG098), the National Natural Science Foundation of China (Grant No. 31540024) and the K. C. Wong Magna Fund at Ningbo University to Qiang Xu.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data Availability Data that support the findings of this article are available at https://osf.io/7zpbk/files/

Declarations

Conflict of interest The authors declare that there were no financial or other conflicts of interest in conducting this research.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Baltrusaitis, T., Zadeh, A., Lim, Y. C., & Morency, L. P. (2018). Openface 2.0: Facial behavior analysis toolkit. 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), 59–66. https://doi.org/10.1109/FG.2018.00019

Banich, M. T., Mackiewicz, K. L., Depue, B. E., Whitmer, A. J., Miller, G. A., & Heller, W. (2009). Cognitive control mechanisms, emotion and memory: A neural perspective with implications for psychopathology. Neuroscience and Biobehavioral Reviews, 33, 613–630. https://doi.org/10.1016/j.neubiorev.2008.09.010

Beringer, M., Spohn, F., Hildebrandt, A., Wacker, J., & Recio, G. (2019). Reliability and validity of machine vision for the assessment of facial expressions [Article]. Cognitive Systems Research, 56, 119–132. https://doi.org/10.1016/j.cogsys.2019.03.009

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. PLoS ONE, 3(8), e3022. https://doi.org/10.1371/journal.pone.0003022

Bourgeois, P., & Hess, U. (2008). The impact of social context on mimicry. Biological Psychology, 77, 343–352. https://doi.org/10.1016/j.biopsycho.2007.11.008

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. Psychological Bulletin, 115(3), 401–423. https://doi.org/10.1037/0033-2909.115.3.401
Camras, L. A., & Shutter, J. M. (2010). Emotional facial expressions in infancy. *Emotion Review, 2*(2), 120–129. https://doi.org/10.1177/1754073909352529

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology, 76*(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893

Chartrand, T. L., & Lakin, J. L. (2013). The antecedents and consequences of human behavioral mimicry. *Annual Review of Psychology, 64*, 285–308. https://doi.org/10.1146/annurev-psych-113011-143754

Chiew, K. S., & Braver, T. S. (2010). Exploring emotional and cognitive conflict using speeded voluntary facial expressions. *Emotion, 10*(6), 842–854. https://doi.org/10.1037/a0019704

Crivelli, C., & Fridlund, A. J. (2018). Facial displays are tools for social influence. *Trends in Cognitive Sciences, 22*, 388–399. https://doi.org/10.1016/j.tics.2018.02.006

Darwin, C. (1872). *The expression of the emotions in man and animals*. Chicago: University of Chicago Press.

Dimberg, U., & Pettersson, M. (2000). Facial reactions to happy and angry facial expressions: Evidence for right hemisphere dominance. *Psychophysiology, 37*(5), 693–696. https://doi.org/10.1111/1469-8986.3750693

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience, 8*(12), 1784–1790. https://doi.org/10.1038/nn1594

Egner, T., Monti, J. M., Tittschiuh, E. H., Wieneke, C. A., Hirsch, J., & Mesulam, M. M. (2008). Neural integration of top-down spatial and feature-based information in visual search. *The Journal of Neuroscience, 28*(24), 6141–6151. https://doi.org/10.1523/JNEUROSCI.1262-08.2008

Ekman, P. (1985). *Telling lies*. New York: W. W. Norton.

Ekman, P. (1999). Basic emotions. In T. Dalgleish & T. Power (Eds.), *The handbook of cognition and emotion* (pp. 45–60). Chichester: Wiley.

Ekman, P., & Friesen, W. V. (1969a). The repertoire of nonverbal behaviour: Categories, origins, usage, and coding. *Semiotica, 1*, 49–98. https://doi.org/10.1515/semi.1969.1.1.49

Ekman, P., & Friesen, W. V. (1969b). Nonverbal leakage and clues to deception. *Psychiatry, 32*, 88–105. https://doi.org/10.1080/00332747.1969.11023575

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology & Nonverbal Behavior, 1*(1), 56–75. https://doi.org/10.1007/BF01115465

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving emotional conflict: A role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron, 51*, 871–882. https://doi.org/10.1016/j.neuron.2006.07.029

Ferrari, P. F., Gallesse, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience, 17*(8), 1703–1714. https://doi.org/10.1046/j.1460-9568.2003.02601.x

Fölster, M., Hess, U., & Werheik, K. (2014). Facial age affects emotional decoding. *Frontiers in Psychology, 5*, 30. https://doi.org/10.3389/fpsyg.2014.00030

Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. San Diego, CA: Academic Press. ISBN: 0-12-267630-0.

Glocker, M. L., Langleben, D. D., Ruparel, K., Loughead, J. W., Gur, R. C., & Sachser, N. (2009). Baby schema in infant faces induces cuteness perception and motivation for caretaking in adults. *Ethology, 115*(3), 257–263. https://doi.org/10.1111/j.1439-0310.2008.01603.x

Hess, U., & Fischer, A. (2014). Emotional mimicry: Why and when we mimic emotions. *Social & Personality Psychology Compass, 8*, 45–57. https://doi.org/10.1111/spc3.12083

Jens, S. A. (2017). Can you see it? Facial expression leakage in response to emotional intensity. *Undergraduate Honors Theses*. Paper 1124. https://scholars.wm.edu/honorstheses/1124

Korb, S., Malsert, J., Strathearn, L., Vuilleumier, P., & Niedenthal, P. (2016). Sniff and mimic - Intra-nasal oxytocin increases facial mimicry in a sample of men. *Hormones and Behaviour, 84*, 64–74. https://doi.org/10.1016/j.yhbeh.2016.06.003

Künecke, J., Wilhelm, O., & Sommer, W. (2017). Emotion recognition in nonverbal face-to-face communication. *Journal of Nonverbal Behavior, 41*, 221–238. https://doi.org/10.1007/s10919-017-0255-2

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion, 24*(4), 1377–1388. https://doi.org/10.1080/02699930903485076

Lee, V., & Wagner, H. (2002). The effect of social presence on the facial and verbal expression of emotion and the interrelations among emotion components. *Journal of Nonverbal Behavior, 26*(1), 3–25. https://doi.org/10.1023/A:1014479919684
Lorenz, K. (1943). Die angeborenen Formen möglicher Erfahrung. [The innate conditions of the possibility of experience.]. *Zeitschrift Für Tierpsychologie, 5*, 235–409. https://doi.org/10.1111/j.1439-0310.1943.tb00655.x

Otte, E., Habel, U., Schulte-Rüther, M., Konrad, K., & Koch, I. (2011a). Interference in simultaneously perceiving and producing facial expressions - evidence from electromyography. *Neuropsychologia, 49*(1), 124–130. https://doi.org/10.1016/j.neuropsychologia.2010.11.005

Otte, E., Jost, K., Habel, U., & Koch, I. (2011b). Exploring cross-task compatibility in perceiving and producing facial expressions using electromyography. *Acta Psychologica, 138*(1), 187–192.

Preston, S. D., & de Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*(1), 1–71. https://doi.org/10.1017/s0140525x02000018

Quaglia, J. T., Zeidan, F., Grossenbacher, P. G., Freeman, S. P., Braun, S. E., Martelli, A., Goodman, R. J., & Brown, K. W. (2019). Brief mindfulness training enhances cognitive control in socioemotional contexts: Behavioral and neural evidence. *PLoS ONE, 14*(7), e0219862. https://doi.org/10.1371/journal.pone.0219862

Recio, G., & Sommer, W. (2018). Copycat of dynamic facial expressions: Superior volitional motor control for expressions of disgust. *Neuropsychologia, 119*, 512–523. https://doi.org/10.1016/j.neuropsychologia.2018.08.027

Rosenberg, E. L. (2005). The study of spontaneous facial expressions in psychology. In P. Ekman & E. L. Rosenberg (Eds.), *What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (FACS)* (2nd ed., pp. 3–18). New York: Oxford University Press.

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition & Emotion, 23*, 1307–1351. https://doi.org/10.1080/02699930902928969

Scherer, K. R., & Ellgring, H. (2007). Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? *Emotion, 7*(1), 113–130. https://doi.org/10.1037/1528-3542.7.1.113

Scherer, K. R., Mortillaro, M., & Mehu, M. (2013). Understanding the mechanisms underlying the production of facial expression of emotion: A componential perspective. *Emotion Review, 5*(1), 47–53. https://doi.org/10.1177/1754073912451504

Song, S., Zilverstand, A., Song, H., d’Oleire Uquillas, F., Wang, Y., Xie, C., Cheng, L., & Zou, Z. (2017). The influence of emotional interference on cognitive control: A meta-analysis of neuroimaging studies using the emotional Stroop task. *Scientific Reports, 7*(1), 2088. https://doi.org/10.1038/s41598-017-02266-2

Tian, Y. L., Kanade, T., & Cohn, J. F. (2005). Facial expression analysis. In S. Z. Li & A. K. Jain (Eds.), *Handbook of face recognition* (pp. 247–276). New York: Springer.

Tottenham, N., Hare, T. A., & Casey, B. J. (2011). Behavioral assessment of emotion discrimination, emotion regulation, and cognitive control in childhood, adolescence, and adulthood. *Frontiers in Psychology, 2*, 39. https://doi.org/10.3389/fpsyg.2011.00039

van Baaren, R., Janssen, L., Chartrand, T. L., & Dijksterhuis, A. (2009). Where is the love? The social aspects of mimicry. *Philosophical Transactions of the Royal Society of London Series b, Biological Sciences, 364*(1528), 2381–2389. https://doi.org/10.1098/rstb.2009.0057

Williams, P., & Aaker, J. L. (2002). Can mixed emotions peacefully coexist? *Journal of Consumer Research, 28*(4), 636–649. https://doi.org/10.1086/338206

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.