

Two Ways to Facial Expression Recognition? Motor and Visual Information Have Different Effects on Facial Expression Recognition



Stephan de la Rosa¹, Laura Fademrecht¹, Heinrich H. Bülthoff¹,
Martin A. Giese², and Cristóbal Curio^{1,3}

¹Department of Human Perception, Cognition and Action, Max Planck Institute for Biological Cybernetics, Tübingen, Germany; ²Department of Cognitive Neurology, Hertie Institute for Clinical Brain Research, Centre for Integrative Neuroscience, Tübingen, Germany; and ³Department of Computer Science, Reutlingen University

Psychological Science
1–13
© The Author(s) 2018
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797618765477
www.psychologicalscience.org/PS


Abstract

Motor-based theories of facial expression recognition propose that the visual perception of facial expression is aided by sensorimotor processes that are also used for the production of the same expression. Accordingly, sensorimotor and visual processes should provide congruent emotional information about a facial expression. Here, we report evidence that challenges this view. Specifically, the repeated execution of facial expressions has the opposite effect on the recognition of a subsequent facial expression than the repeated viewing of facial expressions. Moreover, the findings of the motor condition, but not of the visual condition, were correlated with a nonsensory condition in which participants imagined an emotional situation. These results can be well accounted for by the idea that facial expression recognition is not always mediated by motor processes but can also be recognized on visual information alone.

Keywords

facial expressions, emotions, visual perception, motor processes, social cognition, open data

Received 12/2/16; Revision accepted 2/2/18

Facial expressions are one of the most important sources of information about the emotional states of other persons, and many people have no problems recognizing facial expressions. Which neural mechanisms allow the visual recognition of facial expressions with seemingly little effort? The recent years have seen a surge of ideas that point to the sensorimotor system as the main contributor to the visual recognition of socially relevant information, such as actions (Friston, 2010; Kilner, Friston, & Frith, 2007; Rizzolatti, Fogassi, & Gallese, 2001; Rizzolatti & Sinigaglia, 2010) and facial expressions. In support of this suggestion, previous research has shown that the execution of facial expressions influences emotion-related judgments of facial expressions (Ipser & Cook, 2016; Maringer, Krumhuber, Fischer, & Niedenthal, 2011; Oberman, Winkielman, & Ramachandran, 2007; Ponari, Conson, D'Amico, Grossi,

& Trojano, 2012; Rychlowska et al., 2014; Wood, Lupyan, Sherrin, & Niedenthal, 2016).

At the core of motor-based recognition accounts is the idea that perceptual processes underlying the visual

Corresponding Authors:

Stephan de la Rosa, Max Planck Institute for Biological Cybernetics, Department of Human Perception, Cognition and Action, Spemannstr. 38, Tübingen, 72076, Germany
E-mail: delarosa@tuebingen.mpg.de

Heinrich H. Bülthoff, Max Planck Institute for Biological Cybernetics, Department of Human Perception, Cognition and Action, Spemannstr. 38, Tübingen, 72076, Germany
E-mail: heinrich.buelthoff@tuebingen.mpg.de

Cristóbal Curio, Max Planck Institute for Biological Cybernetics, Department of Human Perception, Cognition and Action, Spemannstr. 38, Tübingen, 72076, Germany
E-mail: Cristobal.Curio@reutlingen-university.de

recognition of a facial expression activate “somatosensory and motor systems largely overlapping with those that support the production of the same facial expression” (Wood, Rychlowska, Korb, & Niedenthal, 2016, p. 229; for a review, see Goldman & Sripada, 2005). It is believed that these overlapping areas are used for recreating observed facial expressions (Wood, Rychlowska, et al., 2016). The recreation of a facial expression (sensorimotor simulation) is suggested to occur at a subthreshold level, possibly by means of facial mimicry (Korb, Grandjean, & Scherer, 2010; Krumbhuber, Likowski, & Weyers, 2014; Wood, Rychlowska, et al., 2016). The idea is that the sensorimotor simulation of facial expressions activates the associated emotion system in the observer. As a consequence, the observer can experience the internal emotional state of the other person and can use this information to recognize the facial expression (Wood, Rychlowska, et al., 2016).

According to this elegant mechanism, the simulated emotional state is congruent with the visually observed emotional information of the facial expression. Hence, a critical test of the contribution of the sensorimotor system to the visual recognition of facial expressions is to see whether sensorimotor and visual processes provide congruent emotional information about the perceived facial expression.

Relatively little is known about whether sensorimotor and visual processes provide congruent emotional information in facial expression recognition. Several studies have demonstrated that prolonged exposure to a facial expression (adaptation) induces repelling adaptation effects. For example, after participants adapted to a fearful expression, they reported that a face morph between a happy and a fearful expression looked more happy (de la Rosa, Giese, Bülthoff, & Curio, 2013; Fox & Barton, 2007; Luo, Wang, Schyns, Kingdom, & Xu, 2015; Pavlova & Sokolov, 2000; Pell & Richards, 2011; Rutherford, Chattha, & Krysko, 2008; Skinner & Benton, 2010; Webster, Kaping, Mizokami, & Duhamel, 2004; Xu, Dayan, Lipkin, & Qian, 2008). The fact that even low-level visual features, such as a curved line, can induce repelling effects on facial expression recognition (Xu et al., 2008) suggests that visual information might affect facial expression recognition in a bottom-up fashion. Yet the execution of a facial expression seem to have assimilating effects on the visual recognition of facial expressions. Using a different experimental paradigm, Blaes and Wilson (2010) showed that activating smiling muscles led participants to more frequently report a facial expression as happy compared with when participants did not activate these muscles. Such results are often interpreted in terms of the facial-feedback hypothesis. It suggests that the execution of a facial expression induces the executed emotion in the

actor (Darwin, 1872; Laird, 1974; Strack, Martin, & Stepper, 1988). As the overall emotional state of a person affects several different cognitive functions, assimilation effects have been reported for various tasks, including ratings of funniness (Strack et al., 1988) and valence judgments (Hyniewska & Sato, 2015). Yet caution is needed when comparing this literature, as studies reporting repelling effects of visual information on visual expression recognition employed an adaptation paradigm, but studies reporting an assimilating effect on facial expression recognition did not. Hence, the difference between assimilating and repelling effects might be simply due to the use of different experimental paradigms. To overcome this theoretical gap in the current study, we assessed whether sensorimotor and visual processes provide congruent emotional information about a facial expression within the same experimental paradigm.

To ensure that we probed visual recognition underlying facial expression recognition, we used an adaptation paradigm. Adaptation is known for its ability to selectively target perceptual-cognitive processes (Webster, 2015). Adaptation effects are well explained in terms of neural population responses (Giese, 2016; Webster, 2011; Webster & MacLeod, 2011) and agree with physiological measurements (e.g., Barraclough, Keith, Xiao, Oram, & Perrett, 2009). They have therefore been called the psychologist microelectrode (Frisby, 1979). If sensorimotor and visual processes provide congruent emotional information about a facial expression, we would expect that motor and visual adaptation should have similar effects on the subsequent visual recognition of a facial expression.

We compared motor and visual facial expression adaptation aftereffects on facial expression recognition. In a control experiment, we further explored whether motor influences are related to emotions, as predicted by the sensorimotor-simulation hypothesis (Wood, Rychlowska, et al., 2016) and the facial-feedback hypothesis. To this end, we used an adaptation paradigm in which participants were exposed to the same emotional face information for a prolonged amount of time. It is typically found that adaptation transiently changes the subsequent percept of an ambiguous stimulus. Here, we were interested in whether the direction of change (i.e., assimilating or repelling effects) in facial expression recognition was the same after participants repeatedly viewed a facial expression (visual adaptation) or repeatedly executed a facial expression (motor adaptation). In the control condition, we compared whether the repeated execution of a facial expression had the same effect on recognition of visual facial expressions as the mere amodal imagination of a situation that elicits the emotion (emotion induction).

Experiment 1

Method

Participants. We recruited participants from the local community in Tübingen, Germany. A total of 12 participants took part in the visual modality-adaptation condition, and another 12 participated in the motor-adaptation condition. The same participants who completed the motor-adaptation condition also participated in the emotion-induction condition. One participant's data in the motor condition were not recorded because of a technical error and were deleted from the analysis. All participants had normal or corrected-to-normal vision. The number of participants was chosen on the basis of previous facial adaptation experiments in our lab that showed that this sample size is sufficient to show reliable effects (de la Rosa et al., 2013). Participants received €8 per hour for their participation in the experiment. All participants gave their written informed consent prior to the experiment. The ethics committee of the Max Planck Society approved this study.

Stimuli. Happy and fearful expressions were recorded from a lay actor using a motion capture system. These facial movements were used to animate a 3-D morphable face model. Because we were interested only in facial movement, the face model did not have teeth or eyes. Animation of the 3-D face model was done by means of scanned 3-D facial action units (AUs) that were chosen to be similar to the Facial Action Coding System (Ekman & Friesen, 1978). The method for the animation of dynamic facial expressions based on the superposition of facial AUs is described elsewhere (de la Rosa et al., 2013). In brief, the expression morph consisted of the calculation of linear interpolations of time-normalized action-unit signals of two facial expressions to generate an ambiguous facial expression. Here, we used a happy and a fearful facial expression. In all of our experimental conditions, we used the following seven morph levels (specifying the amount of happy expression in the morphed facial expression): 0, 0.217, 0.433, 0.650, 0.867, 1.083, 1.30. All stimuli had a duration of 1,042 ms. The morphed facial expressions with a morph weight of 0 and 1.3 were used as adaptors in the visual-adaptation condition. Pilot experiments suggested that a facial expression with a morph-blending weight of 1.3 was more easily identified as happy than a facial expression with a morph-blending weight of only 1. All stimuli served as test stimuli in all adapted-modality conditions (visual, motor, and emotion induction).

Apparatus. All stimuli were presented on a Dell LCD screen with a refresh rate of 60 Hz using a Dell desktop computer. We used a custom-written MATLAB (The

MathWorks, Natick, MA) script that relied on functions provided by the Psychophysics Toolbox 3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997).

Design. The critical manipulation of adapted modality (visual vs. motor adaptation) was tested in a between-subjects manner. The comparison of motor adaptation with adaptation effects induced by asking participants to imagine a situation that elicits the emotion (emotion-induction condition) was tested in a within-subjects manner. Within each adapted modality, we used two facial expressions as adaptors (happy and fearful). For each adaptor, we probed each morph level 15 times (method of constant stimuli). Morph level and adaptor conditions were within-subjects factors. Hence, each adapted-modality condition consisted of 15 (repetitions) \times 7 (morph levels) \times 3 (adaptor levels) = 315 trials. The presentation order of adaptor and morph level was completely random. The testing order of emotion induction and motor modality adaptation was counterbalanced across participants.

Analysis. The data for the analysis can be found at the Open Science Framework (<https://osf.io/rydzg/>). We calculated the proportion of fear responses as a function of morph level for each participant, adaptor, and adapted-modality condition. Using MATLAB's *gfit* function, we then fitted these data with cumulative Weibull functions of the following form:

$$\text{proportion of happy responses} = \gamma + (1 - \gamma - \lambda) \times (1 - \exp(-(x/\alpha)^\beta))$$

Gamma (γ) is the guessing rate, lambda (λ) is the lapse rate, alpha (α) influences the location of the psychometric function along the x -axis, and beta (β) specifies the slope of the psychometric function. We allowed all parameters to vary freely except for γ (Wichmann & Hill, 2001). For each fitted function, we determined the point of subjective equality (i.e., the morph level that corresponded with 50% fearful answers). Data from 1 participant in the motor-adaptation and emotion-induction conditions had to be deleted because of a technical fault. The overall fit was very good (mean explained variance = 0.98, $SD = 0.047$).

Procedure. The following procedure and the choice of experimental parameters were based on a previous study in our lab, which demonstrated that this procedure and parameters induced reliable adaptation effects (de la Rosa et al., 2013). Participants sat in front of the monitor and received the following instructions. At first participants saw a summary of instructions: "Please focus on the nose area. Judge the expression after the beep signal:

Report whether you saw a happy or fearful facial expression. Press any key to start.” The key press initiated the start of an experimental trial, which consisted of a 100-ms black screen. Subsequently, four consecutive presentations of a tone (consisting of a 521-ms 500-Hz tone followed by a 521-ms silence) were presented with an interstimulus interval of 200 ms. These four tones marked the adaptor presentations.

In the visual condition, the tones were accompanied by the presentation of the visual adaptors (happy or fearful expressions). In the motor condition, participants were asked to execute the facial expression whose name was presented on the screen as long as the tone was audible and to return to a neutral facial expression during the silence period of the tone. Hence, participants carried out four facial movements in which they went from a neutral to a smiling facial expression. That is, participants did not hold the facial expression for 4 s. In the emotion-induction condition, we asked participants to vividly imagine a fearful or a happy situation for the entire presentation of the four tones. As a consequence, the number of adaptor repetitions and the adaptor presentation period were the same in these three experimental conditions (motor, visual, and emotion induction). Thereafter followed the presentation of the 100-ms 1000-Hz tone and the test stimulus. Finally, the program presented a 300-ms black screen and the answer screen. The answer screen said, “What did you see? A happy (H) or a fearful (F) expression?” Participants’ task was to report their subjective feeling about the stimulus preceded by the 1000-Hz tone (i.e., test stimulus). Participants pressed the key corresponding to their subjective impression about the test stimulus. The answer screen stayed on until participants had given their response and was then replaced by a black screen. No feedback was given. After participants had given their response, the next trial started when they pressed any key on the answer keyboard. The program offered participants a break after every 45th trial by displaying the following text on the screen: “You may take a break now. Press any key to continue with the experiment.”

Results

We measured psychometric functions relating the morph level to the recognition performance for both of our experimental manipulations (adapted modality and adapted facial expression). To quantify the differences between the psychometric functions for the different adaptation conditions, we calculated the shift of the point of subjective equality of the psychometric functions between the happy and fearful adaptor conditions. The difference was calculated in such a way that positive differences indicate a repelling effect (e.g.,

adapting to a happy expression made participants perceive the test stimulus as fearful), and negative values suggest assimilation effects (e.g., adapting to a happy expression made participants perceive the test stimulus as happy). Figure 1 shows that adapting the visual modality induces repelling adaptation effects, replicating the well-documented repelling visual-adaptation effects in the literature (Benton, 2009; de la Rosa et al., 2013; Leopold, O’Toole, Vetter, & Blanz, 2001; Luo et al., 2015; Pavlova & Sokolov, 2000; Skinner & Benton, 2010; Xu et al., 2008). In contrast, both the motor modality and nonmodal emotion-induction conditions exhibited assimilation adaptation effects.

To answer the main question of whether motor and visual-adaptation effects induce significant effects in different directions, we compared the overall visual-adaptation effect with the overall motor adaptation effect using a two-sample Welch a priori *t* test. This test showed a large effect, Cohen’s $d = 1.65$, and a significant difference between motor and visual-adaptation effects, $t(10) = 3.81$, $p = .002$. Additional one-sample *t* tests confirmed that the overall adaptation effect within the visual and within the motor adaptation condition was also large and significant—visual-adaptation effect: $t(11) = 2.58$, Cohen’s $d = 0.745$, $p = .026$; motor adaptation effect: $t(10) = 3.14$, Cohen’s $d = 0.947$, $p = .010$. Hence, our results provide strong support for the assumption that prolonged exposure to visual and motor facial expression information induces opposite effects on facial expression recognition.

To examine whether motor-adaptation effects are related to emotion-induction effects, we compared the motor-adaptation condition with the emotion-induction condition. We found no significant differences between the motor and emotion-induction conditions, $t(10) = 1.00$, Cohen’s $d = 0.3$, $p = .342$. More importantly, motor-adaptation effects were significantly correlated with adaptation effects induced by imagination (emotion-induction condition), $r = .723$, $t(10) = 3.14$, $p = .012$ (see also Fig. 2), suggesting a strong relationship between motor and emotion-induction effects. In contrast, adapting the visual modality induced a significantly different effect from nonmodal emotion-induction adaptation, $t(10) = 2.85$, Cohen’s $d = 1.245$, $p = .01$.

Experiment 2 was intended to replicate the findings of Experiment 1 and to exclude an important possible confound. Participants might have executed the facial movement in the emotion-induction condition, which might have led to the high correlation between the results of the motor-modality and the emotion-induction condition. To this end, we replicated Experiment 1 using face tracking to measure the amount of facial movement in all conditions by means of regression (see the Method section). Moreover, in Experiment 2, all conditions were tested in a within-subjects manner,

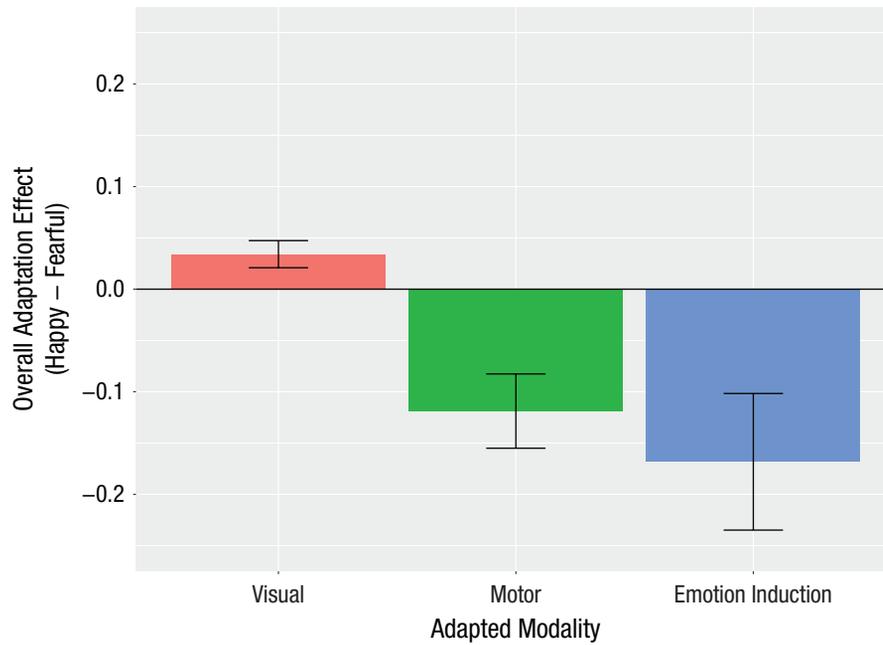


Fig. 1. Overall adaptation effects for each of the three adapted modalities in Experiment 1. Error bars indicate ± 1 SEM. The black horizontal line indicates the absence of an adaptation effect (overall adaptation effect = 0). Participants in the emotion-induction condition imagined a situation triggering the emotion.

and we used another facial expression (disgust instead of fear) to increase the external validity of our results.

Experiment 2

Method

The method was identical to that of Experiment 1 except for the changes noted below. Another 15

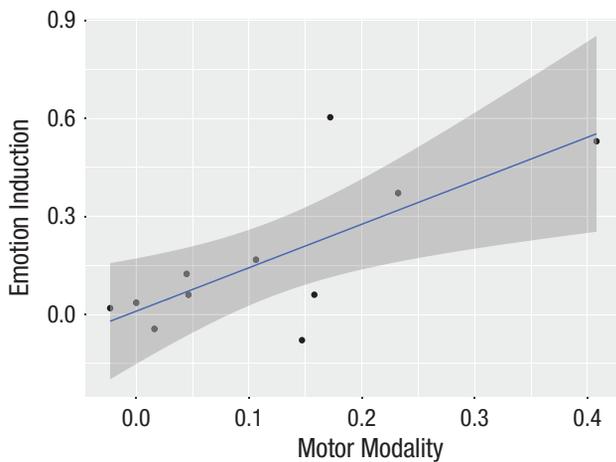


Fig. 2. Overall adaptation effects of the emotion-induction condition as a function of overall adaptation effects of the motor-adaptation condition (Experiment 1). Each dot represents the data from 1 participant. The blue line represents the best-fitting regression, and the gray area is the 95% confidence band.

participants were recruited from the local community of Tübingen.

Apparatus, stimuli, and design. We used a custom-built marker-based facial capture system that has been described in detail elsewhere (Curio, Kleiner, Breidt, & Bühlhoff, 2010; Kleiner, Wallraven, & Bühlhoff, 2004). In brief, this setup consists of five cameras (Basler Vision Technologies, Germany; refresh rate = 60 Hz) pointed at different angles at the participants face. The output of the cameras is synchronized and used to reconstruct the 3-D positions of the markers (Curio et al., 2010).

We used happy and disgusted facial expressions in Experiment 2. All adapted-modality conditions (visual, motor, emotion induction) were tested in a within-subjects fashion. Adapted modality was blocked, and its testing order randomized across participants.

Procedure. A total of 17 markers were attached to participants' shaved skin using Mastix glue for theater production. Markers were manually associated with AUs of a 3-D morphable face model. Activation for each AU was calibrated within the tracking software by having participants perform 17 facial movements. In addition, participants were also told to look neutral. Participants then made several facial expressions to train a classification algorithm (for details, see the Analysis section). Specifically, participants made the two expressions (happy and disgusted) repeatedly in randomized order and in order

of increasing intensity, namely none (neutral; 0% of the facial expression), weak (33% of the facial expression), medium (66% of the facial expression), and strong (100% of the facial expression). Participants were informed about the percentage numbers and instructed to produce an expression with the intensity that corresponded best to the percentage. After this training phase, the actual experiment started, as described in Experiment 1.

Analysis.

Facial expression analysis. The input for the classification of the facial expression was activation in AU units. Specifically, for each trial, we first searched for the frame associated with the overall highest activation. Classification of facial expressions was done on these peak frames. Classification was based on a feature vector assembled via the current distribution over AU weights at these peak frames. We used the library for support vector machines (LIBSVM) toolbox (Chang & Lin, 2011) in MATLAB for the classification of facial expressions. The classifiers for the classification of three expression classes (neutral, happy, and disgusted) were trained using three separate support vector machines. The training was done on the training data for each participant separately. Tenfold cross-validation (done for each participant separately) showed

a very good classification performance of the trained classifier with an average cross-validation classification accuracy of 84.22% ($SE = 1.91$). These individually obtained classifiers were then used to classify the facial expression on each experimental trial during the actual experiment.

Facial intensity analysis. We conducted a principal component analysis (PCA) on the AU activations in the training phase. We then determined the linear first-order link function between the PCA scores and the instructed intensity using the *fittype* function of MATLAB. This link function described the relationship between AU activations and facial expression intensities in the training phase using a first-order polynomial. The link function was then used to predict the intensity of the facial expression from the AU activations during the experiment. Because of technical problems, we monitored the expressions of only 10 participants.

Results

Facial expression classification. The recognized facial expressions as identified by the facial expression classifier are shown for each experimental condition in Figure 3. Participants by and large stuck to the instructions: On more

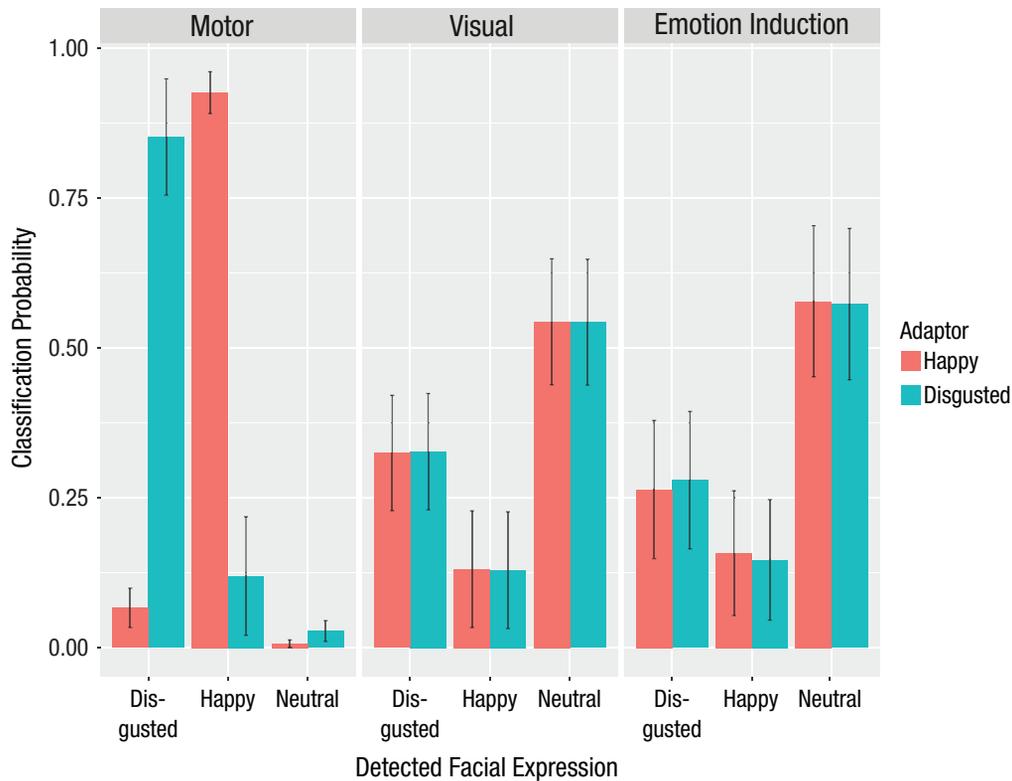


Fig. 3. Classification probability for each detected facial expression and adaptor, separately for each of the three conditions in Experiment 2. Error bars indicate $\pm 1 SEM$. The classification probability indicated how often the classifier detected the expression in the corresponding experimental condition.

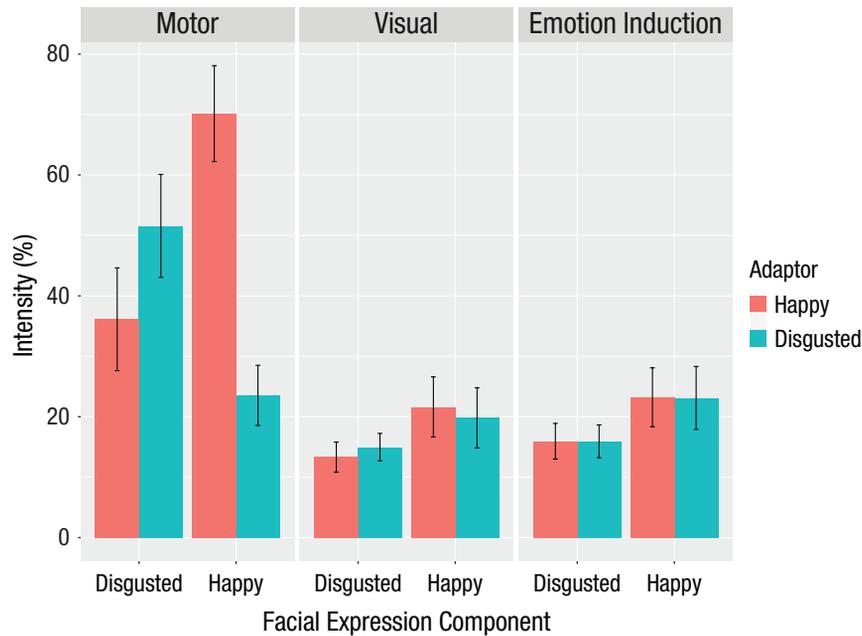


Fig. 4. Intensity of happy and disgusted facial expressions for each adaptor, separately for each of the three conditions in Experiment 2. Error bars indicate ± 1 SEM. On the y-axis, 0% refers to the neutral face and 100% to the most intense expression that the participant was able to carry out during the preexperiment training phase.

than 85% of the trials, participants executed the facial expression in the motor-adaptation condition, and in over 50% of the trials, participants did not make any facial expression in the visual-modality and emotion-induction condition. In the latter two conditions, the classifier was still able to detect a disgusted facial expression on 30% of the trials of each condition. Importantly, the proportion of detected facial expressions between the visual and emotion-induction conditions was not significantly different, as indicated by a three-factor within-subjects analysis of variance (ANOVA). For this ANOVA, we used square-root-transformed classification probability as a dependent variable (to meet the normality assumption) and adapted modality, adapted facial expression, and detected expression as factors. None of the factors or interactions were significant ($p > .05$).

Facial expression intensity. The facial expression intensities are shown separately for each facial expression component (happy and disgusted) and condition in Figure 4 (note that there was no neutral component, as neutral was defined as a facial expression without movement). Clearly, only in the motor-adaptation condition, the happy and disgusted facial expression components were largest in the happy and disgusted adaptation conditions, respectively. The intensity values indicated that participants made at least one expression in these conditions that corresponded to about 50% of

the most intense facial expressions that participants made in the training phase prior to the main experiment. In the emotion-induction and visual-adaptation condition, both facial expression components were activated very little and approximately to the same amount. We analyzed this pattern statistically in a repeated measures ANOVA with intensity as a dependent variable and component (happy vs. disgust), modality (motor, visual, and emotion induction), and adaptor (happy vs. disgust) as within-subjects factors. We found a significant three-way interaction, $F(2, 18) = 24.32$, $\eta_p^2 = .73$, $p < .001$, suggesting that intensity of happy components was specific to the particular combination of adaptor and modality.

To see whether this three-way interaction resulted from intensity patterns in the motor condition, we conducted the same ANOVA with only the motor condition data removed. In this case, none of the effects (main effects and interactions) were significant (all $ps > .23$). In contrast, a two-factor within-subjects ANOVA with component and adaptor as factors and intensity as a dependent variable, which was conducted on the motor-adaptation condition only, revealed a significant two-way interaction, $F(1, 9) = 27.83$, $\eta_p^2 = .76$, $p < .001$. These results demonstrate that participants carried out statistically distinguishable facial expressions only in the motor condition but not in the visual-adaptation and emotion-induction condition.

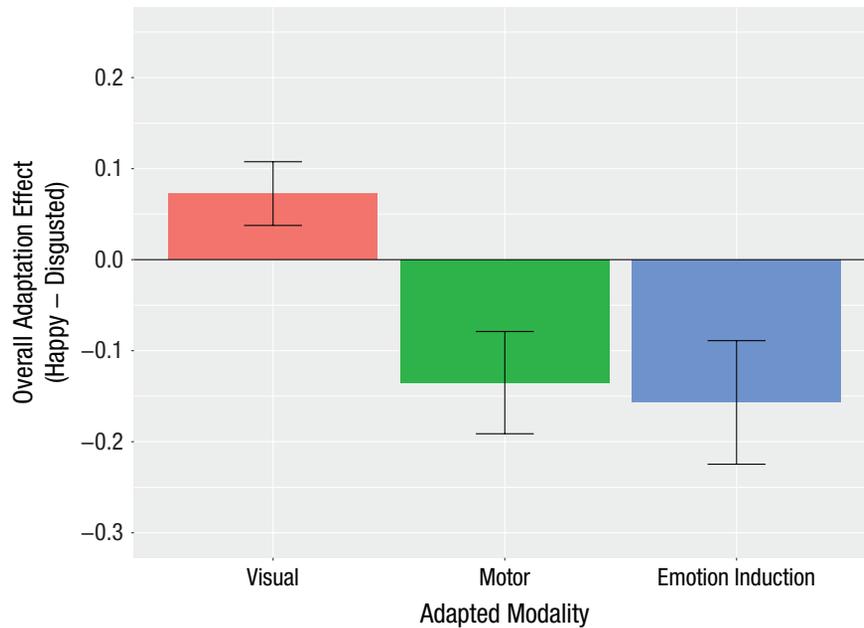


Fig. 5. Overall adaptation effects for each of the three adapted modalities in Experiment 2. Error bars indicate ± 1 SEM. The black horizontal line indicates the absence of an adaptation effect (overall adaptation effect = 0). Participants in the emotion-induction condition imagined a situation triggering the emotion.

Adaptation effects. The adaptation effects are shown in Figure 5 and seem to replicate those of Experiment 1 (see Fig. 1). Again, repelling perceptual effects are indicated by values larger than zero, and assimilation perceptual effects are indicated by values smaller than zero. Prolonged visual exposure to a facial expression caused a repelling effect, whereas prolonged execution of a facial expression or imagining emotions (emotion-induction condition) caused an assimilation effect. Again, we found a significant difference between the visual- and motor-adaptation conditions, $t(14) = 3.36$, Cohen's $d = 0.946$, $p = .002$, and between the visual and emotion-induction condition, $t(14) = 3.38$, Cohen's $d = 0.871$, $p = .005$. The motor-adaptation and the emotion-induction conditions were nonsignificantly different, $t(14) = 0.39$, Cohen's $d = 0.1$, $p = .699$. As in Experiment 1, only adaptation effects between the motor and emotion-induction condition were significantly related, $t(13) = 2.85$, $p = .013$, $r = .61$ (for all others, $p > .05$). Hence, both Experiment 1 and Experiment 2 demonstrate that adapting the visual modality is different from adapting the motor modality or amodal adaptation to the emotion (emotion-induction condition).

Can facial movement in the emotion-induction condition explain the similar effects between the motor and emotion-induction conditions? We do not think so because the facial movement patterns of the classification data between these two experimental conditions were significantly different. Specifically, a three-way

ANOVA with adapted expression (happy vs. disgusted), adapted modality (motor vs. emotion induction), and detected expression (happy vs. disgusted) revealed a significant three-way interaction, $F(1, 9) = 68.16$, $\eta_p^2 = .883$, $p < .001$. This three-way interaction indicated that happy and disgusted adaptors induced different facial expressions in the motor-adaptation condition and emotion-induction condition. Despite the differences in facial movement in these two conditions, the perceptual effects are strikingly similar and significantly related. Likewise, we found statistically significant behavioral differences between the visual and the emotion-induction conditions. Yet the amount and type of facial movement was not statistically significantly different between these conditions (see the ANOVA results in the Facial Expression Classification section). Hence, the behavioral differences between the visual-adaptation and emotion-induction conditions cannot be explained by differences in facial movement.

General Discussion

Theories of motor-based facial expression recognition suggest that facial recognition is primarily based on sensorimotor simulation mechanisms. Accordingly, sensorimotor and visual processes should provide congruent emotional information about a facial expression. Here, we set out to test this prediction and found results inconsistent with this hypothesis. Specifically, we

showed in two experiments that prolonged exposure (adaptation) to motor and visual facial expression information alters the perception of a facial expression in very different ways. Participants who visually adapted to one facial expression more frequently saw the non-adapted facial expression in the subsequently presented ambiguous test stimulus. In contrast, participants who motorically adapted to a facial expression through repeated execution of a facial expression were more likely to see the adapted facial expression in the subsequent ambiguous test stimulus. Hence, visual and motor processes provide different emotional information about a facial expression.

How can the results be understood within sensorimotor-based theories of facial expression recognition? In facial expression recognition, the goal of the observer is to determine which emotional state of a person is associated with a facial expression. The observer faces the problem that the other person's emotional state is internal and therefore hidden from the observer. Hence it cannot directly be observed. Recently, Wood, Rychlowska, and colleagues (2016) suggested that the observer solves this problem by reconstructing the facial movement using the sensorimotor and motor system that is shared between the processes involved in the observation and the execution of a facial expression (for a review of similar simulation accounts, see Goldman & Sripada, 2005). Specifically, the observer uses these shared neural mechanisms to covertly or overtly simulate the observed expression. By means of this simulation and with the help of the emotion system, the observer can experience the emotional effect of the observed facial expression and thereby infer the internal emotional state of the other person. Note that according to this suggestion, visual recognition of a facial recognition is supported by the emotional system and not performed by the sensorimotor and motor systems alone. Because this explanation presupposes a coupling of motor and emotional system in facial expression recognition, making a happy facial expression should induce a happy feeling in the observer. Wood, Rychlowska, and colleagues' (2016) suggestion can therefore easily explain why conducting a facial movement induces the executed emotion and hence why motor adaptation leads to assimilation effects. The coupling of motor and emotion system in Wood and colleagues' explanation would also predict a correlation between the motor and emotion-induction effects in the current experiments. However, a challenge for this and other simulation accounts is the finding that visual adaptation produces clearly different effects from motor adaptation.

We suggest that the clearly different visual-adaptation effects indicate another vision-based route

to facial expression that does not necessarily rely to the same extent on sensorimotor simulation. Note that this suggestion does not call into question the theoretical proposals of previous simulation accounts but rather extends them. The idea of vision-based facial expression recognition is long standing and has been explained by several influential models (Bruce & Young, 1986; Haxby, 2001; Haxby, Hoffman, & Gobbini, 2000). The different visual and motor adaptation effects can be well understood if visual recognition of facial expression can be achieved in two ways: by visual information or by sensorimotor simulation.

What would be the advantage of having two ways of recognizing facial expressions? We suggest that vision-based recognition of facial expressions could be useful in situations in which the observer is interested in determining the emotional expression without re-experiencing the associated emotion. An example situation is where one sits in a street restaurant and observes other people walking by. Here, simulating the multitude of facial expressions of the passersby would induce many emotions in the observer, interfering with the possible intention to relax and merely observe other people. In contrast, simulating facial expressions can be useful in situations when one would like to relate to an emotional state of another person, for example, when showing empathy in a social interaction. Hence, humans might employ different strategies of facial expression recognition depending on the emotional involvement of the observer in a situation. Yet future research is needed to scrutinize this suggestion, for example, by determining whether humans use discrete or continuous transitions between vision-based and simulation strategies for facial expression recognition.

In our study, we chose not to physically constrain facial movement in the visual-adaptation and emotion-induction condition to keep the experimental conditions as similar and therefore as comparable as possible. A physical constraint of facial movements, such as by taping the face, in the two nonmotor conditions would have had the advantage of minimizing facial expression artifacts. On the other hand, it would have introduced one more factor by which the nonmotor- and motor-adaptation conditions differed in addition to the main manipulation of modality. We therefore would have introduced a possible confound that could explain differences between the motor and nonmotor conditions. For this reason, we refrained from constraining facial movement in the nonmotor-adaptation conditions.

Nevertheless, it is possible that these subtle facial expressions in the nonmotor conditions introduced a possible confound. Specifically, it is possible that participants executed more subtle facial movements in the emotion-induction than in the visual-adaptation

condition. As a result, participants would have been able to experience the associated emotion to a larger degree in the emotion-induction condition compared with the visual condition. This could explain the different effects between the emotion-induction and visual-adaptation condition. Yet our analysis of the intensity and type of facial movement shows that both the magnitude and the type of facial movement was very similar in both nonmotor conditions. In other words, because subtle facial movements were very similar across both nonmotor conditions, they are unlikely to explain the differences between these two conditions. Hence, we do not think that subtle facial expressions in the motor condition are at the core of our main finding.

Can our results be explained by the idea that intentional and automatic facial expression have a different neural basis? Specifically, participants made intentional expression in the motor-adaptation condition and automatic facial expressions in the nonmotor-adaptation conditions. If the origin of the behavioral differences in the three conditions were a different neural bases for intentional and automatic facial expressions, one would expect that motor-adaptation and emotion-induction conditions should be associated with different results. This is so because in the motor-adaptation condition, participants executed intentional facial expression, and in the emotion conditions, participants executed automatic facial expressions. Yet contrary to this expectation, we found results in the motor-adaptation and emotion-induction conditions to be quite similar in our two experiments. Moreover, one would expect automatic facial expressions to occur in the visual and the emotion condition alike. Therefore, performance in these conditions should be very similar. Again, our results do not support this hypothesis. Visual and emotion adaptation elicit opposite effects. Taken together, because intentional and automatic expression patterns do not correspond with the occurrence of behavioral effects, we think that it is difficult to explain our main results by the fact that intentional and automatic facial expressions tap into different neural substrates.

Both motor as well as visual adaptation were undoubtedly able to influence facial expression recognition in the current experiment. So far, motoric influences on visual facial expression recognition have been interpreted in favor of a motor-based account of facial expression recognition (Blaesi & Wilson, 2010; Ipser & Cook, 2016). Our research points to the necessity to consider not only the presence of a motoric effect but also its direction for making inferences about the understanding of the underlying psychological processes.

An important future question is to investigate to what degree the priming effects in the emotion-induction and motor-adaptation condition are the result of a change in perception or response bias. One way to dissociate these two options is by means of signal detection theory, which allows the calculation of discrimination performance and response bias from target and nontarget trials. As we did not have nontarget trials in the current experiment, we were not able to calculate these measures in the current data. However, a slight change of participants' subjective report task to an absolute identification task would allow the calculation of both d' and response bias (for an example of such a task, see de la Rosa, Gordon, & Schneider, 2009). In any case, the observation that visual adaptation had a different effect on participants' behavior compared with motor adaptation or emotion induction is important for the current study.

We would like to point out that despite the relatively small sample size in the two experiments, the effect sizes are large (the smallest effect size was Cohen's $d = 0.745$). This can be partly explained by the high reliability of visual-adaptation effects, which reduces the noise level in the data and therefore leads to larger power. The high reliability of the adaptation effect on a participant level can be seen in Figure 6. More than 80% of the participants showed the effect in the same direction in all of the experimental conditions.

Our results are broadly congruent with the existing evidence for the influence of the motor system on the visual recognition of actions using adaptation studies. Researchers in several functional MRI (fMRI) studies examined cross-talk between motor-visual areas in a methodological rigorous way using fMRI adaptation paradigms with inconclusive results. While some scholars using this paradigm were able to identify cortical areas sensitive to both motor and visual input (Chong, Cunnington, Williams, & Kanwisher, 2008; de la Rosa, Schillinger, Bülthoff, Schultz, & Uludag, 2016; Kilner, Neal, Weiskopf, Friston, & Frith, 2009), others were not (Dinstein, Hasson, Rubin, & Heeger, 2007; Lingnau, Gesierich, & Caramazza, 2009). A recent physiological study with macaque monkeys pinned the fMRI adaptation effect to the adaptation of local field potentials rather than a neuronal firing rate change (Caggiano et al., 2013). In humans, physiological evidence exists for neurons that are sensitive to both motor and visual stimulation (Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010). Although these fMRI and physiological results speak in favor of the existence of motor-visual linkages in humans, they do not directly demonstrate the influence of the motor system during visual recognition. Here, behavioral adaptation complement these results by demonstrating that motor adaptation is able to

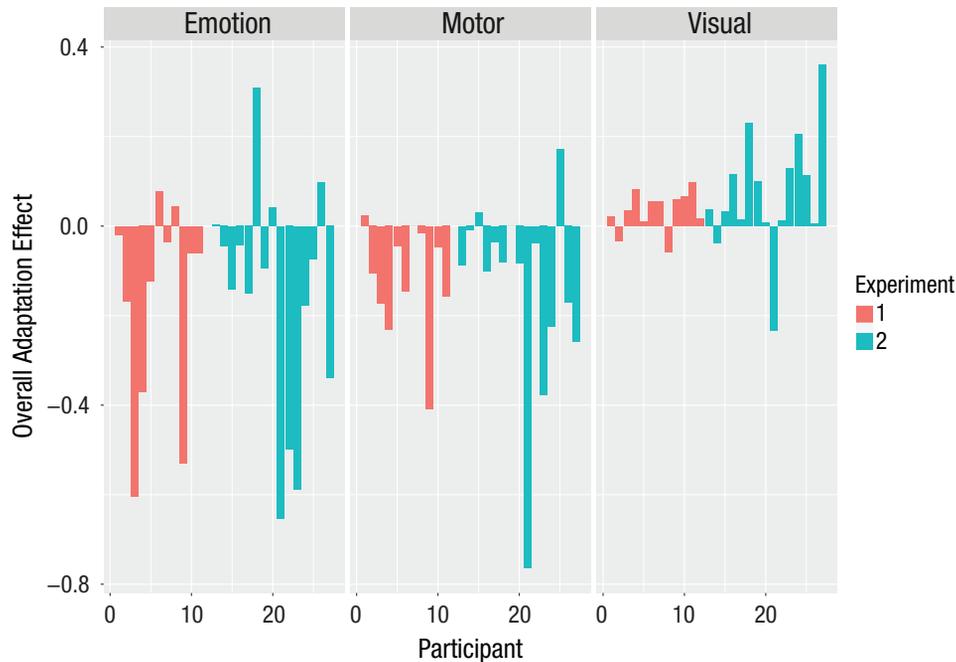


Fig. 6. Individual overall adaptation effects shown for each participant and adaptation type in each experiment. Values smaller than zero indicate assimilation effects, and values greater than zero indicate repelling effects.

influence visual recognition in direction discrimination (Barchiesi, Wache, & Cattaneo, 2012) and categorization tasks (de la Rosa, Ferstl, & Bülthoff, 2016). Yet in the action-recognition domain, this influence seems to be restricted to action observation and does not extend to social interactions in which action are simultaneously observed and executed (de la Rosa, Ferstl, & Bülthoff, 2016).

In conclusion, we demonstrated that visual and motor adaptation have different effects on facial recognition. These results challenge the view that facial expression recognition relies solely on motor-based processes.

Action Editor

Alice O'Toole served as action editor for this article.

Author Contributions

S. de la Rosa, M. A. Giese, and C. Curio developed the study idea, S. de la Rosa and C. Curio designed and set up the study, S. de la Rosa conducted Experiment 1 and analyzed the data, and all authors wrote the manuscript. All the authors approved the final version of the manuscript for submission.

Acknowledgments

We thank Alexander Bauer for collecting the data for Experiment 2.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

C. Curio was supported by Perceptual Graphics Project PAK38 CU 149/1-1/2 funded by the Deutsche Forschungsgemeinschaft (DFG) and by Bundesministerium für Bildung und Forschung (BMBF) project KollRo 4.0 (13 FH 049 PX 5). C. Curio and M. Giese were also supported by European Union Future and Emerging Technologies–Open Project TANGO. M. Giese was supported by Human Frontier Science Program RGP0036/2016, DFG GZ: KA 1258/15-1, European Commission H2020 project CogIMon H2020 ICT-644727, and BMBF FKZ 01GQ1704.

Open Practices



All data have been made publicly available via the Open Science Framework and can be accessed at <https://osf.io/rydzg/>. Materials have not been made publicly available, and the design and analysis plans for the experiments were not preregistered. The complete Open Practices Disclosure for this article can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797618765477>. This article has received the badge for Open Data. More information about the Open Practices badges can be found at <http://www.psychologicalscience.org/publications/badges>.

References

- Barchiesi, G., Wache, S., & Cattaneo, L. (2012). The frames of reference of the motor-visual aftereffect. *PLOS ONE*, *7*(7), Article e40892. doi:10.1371/journal.pone.0040892

- Barraclough, N. E., Keith, R. H., Xiao, D., Oram, M. W., & Perrett, D. I. (2009). Visual adaptation to goal-directed hand actions. *Journal of Cognitive Neuroscience*, *21*, 1805–1819.
- Benton, C. P. (2009). Effect of photographic negation on face expression aftereffects. *Perception*, *38*, 1267–1274. doi:10.1068/p6468
- Blaesi, S., & Wilson, M. (2010). The mirror reflects both ways: Action influences perception of others. *Brain and Cognition*, *72*, 306–309. doi:10.1016/j.bandc.2009.10.001
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327. doi:10.1111/j.2044-8295.1986.tb02199.x
- Caggiano, V., Pomper, J. K., Fleischer, F., Fogassi, L., Giese, M. A., & Thier, P. (2013). Mirror neurons in monkey area F5 do not adapt to the observation of repeated actions. *Nature Communications*, *4*, Article 1433. doi:10.1038/ncomms2419
- Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3), Article 27. doi:10.1145/1961189.1961199
- Chong, T. T., Cunnington, R., Williams, M. A., & Kanwisher, N. (2008). fMRI adaptation reveals mirror neurons in human inferior parietal cortex. *Current Biology*, *18*, 1576–1580. doi:10.1016/j.cub.2008.08.068
- Curio, C., Kleiner, M., Breidt, M., & Bühlhoff, H. H. (2010, January). *The virtual face mirror project: Revealing dynamic self-perception in humans*. Poster presented at the 4th International Conference on Cognitive Systems, Zürich, Switzerland. Retrieved from http://cogsys2010.ethz.ch/doc/cogsys2010_proceedings/cogsys2010_0137.pdf
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London, England: J. Murray.
- de la Rosa, S., Ferstl, Y., & Bühlhoff, H. H. (2016). Visual adaptation dominates bimodal visual-motor action adaptation. *Scientific Reports*, *6*, Article 23829. doi:10.1038/srep23829
- de la Rosa, S., Giese, M. A., Bühlhoff, H. H., & Curio, C. (2013). The contribution of different cues of facial movement to the emotional facial expression adaptation aftereffect. *Journal of Vision*, *13*(1), Article 23. doi:10.1167/13.1.23
- de la Rosa, S., Gordon, M., & Schneider, B. A. (2009). Knowledge alters visual contrast sensitivity. *Attention, Perception, & Psychophysics*, *71*, 451–462. doi:10.3758/APP.71.3.451
- de la Rosa, S., Schillinger, F. L., Bühlhoff, H. H., Schultz, J., & Uludag, K. (2016). fMRI adaptation between action observation and action execution reveals cortical areas with mirror neuron properties in human BA 44/45. *Frontiers in Human Neuroscience*, *10*, Article 78. doi:10.3389/fnhum.2016.00078
- Dinstein, I., Hasson, U., Rubin, N., & Heeger, D. J. (2007). Brain areas selective for both observed and executed movements. *Journal of Neurophysiology*, *98*, 1415–1427. doi:10.1152/jn.00238.2007
- Ekman, P., & Friesen, W. (1978). *Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.
- Fox, C. J., & Barton, J. J. S. (2007). What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Research*, *1127*, 80–89. doi:10.1016/j.brainres.2006.09.104
- Frisby, J. P. (1979). *Seeing: Illusion, brain, and mind*. Oxford, England: Oxford University Press.
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*, 127–138. doi:10.1038/nrn2787
- Giese, M. A. (2016). Face recognition: Canonical mechanisms at multiple timescales. *Current Biology*, *26*, R534–R537. doi:10.1016/j.cub.2016.05.045
- Goldman, A. I., & Sripada, C. S. (2005). Simulationist models of face-based emotion recognition. *Cognition*, *94*, 193–213. doi:10.1016/j.cognition.2004.01.005
- Haxby, J. V. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*, 2425–2430. doi:10.1126/science.1063736
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233. doi:10.1016/S1364-6613(00)01482-0
- Hyniewska, S., & Sato, W. (2015). Facial feedback affects valence judgments of dynamic and static emotional expressions. *Frontiers in Psychology*, *6*, Article 291. doi:10.3389/fpsyg.2015.00291
- Ipsier, A., & Cook, R. (2016). Inducing a concurrent motor load reduces categorization precision for facial expressions. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 706–718. doi:10.1037/xhp0000177
- Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processing*, *8*, 159–166. doi:10.1007/s10339-007-0170-2
- Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *The Journal of Neuroscience*, *29*, 10153–10159. doi:10.1523/JNEUROSCI.2668-09.2009
- Kleiner, M., Brainard, D., & Pelli, D. G. (2007). What's new in Psychtoolbox-3? [Abstract]. *Perception*, *36*(Suppl.), 14.
- Kleiner, M., Wallraven, C., & Bühlhoff, H. H. (2004). *The MPI VideoLab - A system for high quality synchronous recording of video and audio from multiple viewpoints*. Retrieved from http://www.kyb.tue.mpg.de/fileadmin/user_upload/files/publications/pdfs/pdf2774.pdf
- Korb, S., Grandjean, D., & Scherer, K. R. (2010). Timing and voluntary suppression of facial mimicry to smiling faces in a go/nogo task—an EMG study. *Biological Psychology*, *85*, 347–349. doi:10.1016/j.biopsycho.2010.07.012
- Krumhuber, E. G., Likowski, K. U., & Weyers, P. (2014). Facial mimicry of spontaneous and deliberate Duchenne and non-Duchenne smiles. *Journal of Nonverbal Behavior*, *38*, 1–11. doi:10.1007/s10919-013-0167-8
- Laird, J. D. (1974). Self-attribution of emotion: The effects of expressive behavior on the quality of emotional experience. *Journal of Personality and Social Psychology*, *29*, 475–486.

- Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience*, *4*, 89–94. doi:10.1038/82947
- Lingnau, A., Gesierich, B., & Caramazza, A. (2009). Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. *Proceedings of the National Academy of Sciences, USA*, *106*, 9925–9930. doi:10.1073/pnas.0902262106
- Luo, C., Wang, Q., Schyns, P. G., Kingdom, F. A. A., & Xu, H. (2015). Facial expression aftereffect revealed by adaption to emotion-invisible dynamic bubbled faces. *PLOS ONE*, *10*(12), Article e0145877. doi:10.1371/journal.pone.0145877
- Maringer, M., Krumhuber, E. G., Fischer, A. H., & Niedenthal, P. M. (2011). Beyond smile dynamics: Mimicry and beliefs in judgments of smiles. *Emotion*, *11*, 181–187. doi:10.1037/a0022596
- Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*, *20*, 750–756. doi:10.1016/j.cub.2010.02.045
- Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2007). Face to face: Blocking facial mimicry can selectively impair recognition of emotional expressions. *Social Neuroscience*, *2*, 167–178. doi:10.1080/17470910701391943
- Pavlova, M., & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics*, *62*, 889–899. doi:10.3758/BF03212075
- Pell, P. J., & Richards, A. (2011). Cross-emotion facial expression aftereffects. *Vision Research*, *51*, 1889–1896. doi:10.1016/j.visres.2011.06.017
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Ponari, M., Conson, M., D'Amico, N. P., Grossi, D., & Trojano, L. (2012). Mapping correspondence between facial mimicry and emotion recognition in healthy subjects. *Emotion*, *12*, 1398–1403. doi:10.1037/a0028588
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, *2*, 661–670.
- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*, 264–274. doi:10.1038/nrn2805
- Rutherford, M. D., Chattha, H. M., & Krysko, K. M. (2008). The use of aftereffects in the study of relationships among emotion categories. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 27–40. doi:10.1037/0096-1523.34.1.27
- Rychlowska, M., Cañadas, E., Wood, A., Krumhuber, E. G., Fischer, A., & Niedenthal, P. M. (2014). Blocking mimicry makes true and false smiles look the same. *PLOS ONE*, *9*(3), Article e90876. doi:10.1371/journal.pone.0090876
- Skinner, A. L., & Benton, C. P. (2010). Anti-expression after-effects reveal prototype-referenced coding of facial expressions. *Psychological Science*, *21*, 1248–1253. doi:10.1177/0956797610380702
- Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *54*, 768–777.
- Webster, M. A. (2011). Adaptation and visual coding. *Journal of Vision*, *11*(5), Article 3. doi:10.1167/11.5.3
- Webster, M. A. (2015). Visual adaptation. *Annual Review of Vision Science*, *1*, 547–569. doi:10.1007/bf00535812
- Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, *428*, 557–560. doi:10.1038/nature02361.1
- Webster, M. A., & MacLeod, D. I. A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *366*, 1702–1725. doi:10.1098/rstb.2010.0360
- Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*, 1293–1313.
- Wood, A., Lupyan, G., Sherrin, S., & Niedenthal, P. (2016). Altering sensorimotor feedback disrupts visual discrimination of facial expressions. *Psychonomic Bulletin & Review*, *23*, 1150–1156. doi:10.3758/s13423-015-0974-5
- Wood, A., Rychlowska, M., Korb, S., & Niedenthal, P. (2016). Fashioning the face: Sensorimotor simulation contributes to facial expression recognition. *Trends in Cognitive Sciences*, *20*, 227–240. doi:10.1016/j.tics.2015.12.010
- Xu, H., Dayan, P., Lipkin, R. M., & Qian, N. (2008). Adaptation across the cortical hierarchy: Low-level curve adaptation affects high-level facial-expression judgments. *The Journal of Neuroscience*, *28*, 3374–3383. doi:10.1523/JNEUROSCI.0182-08.2008