Fashioning the Face: Sensorimotor Simulation Contributes to Facial Expression Recognition

Adrienne Wood,¹,* Magdalena Rychlowska,² Sebastian Korb,³ and Paula Niedenthal¹

When we observe a facial expression of emotion, we often mimic it. This automatic mimicry reflects underlying sensorimotor simulation that supports accurate emotion recognition. Why this is so is becoming more obvious: emotions are patterns of expressive, behavioral, physiological, and subjective feeling responses. Activation of one component can therefore automatically activate other components. When people simulate a perceived facial expression, they partially activate the corresponding emotional state in themselves, which provides a basis for inferring the underlying emotion of the expresser. We integrate recent evidence in favor of a role for sensorimotor simulation in emotion recognition. We then connect this account to a domain-general understanding of how sensory information from multiple modalities is integrated to generate perceptual predictions in the brain.

When I wish to find out how wise, or how stupid, or how good, or how wicked is any one, or what are his thoughts at the moment, I fashion the expression of my face, as accurately as possible, in accordance with the expression of his, and then wait to see what thoughts or sentiments arise in my mind or heart, as if to match or correspond with the expression. (Edgar Allan Poe, The Purloined Letter, pp. 215–216 [1]).

The Perceptual Challenge of Recognizing Emotion Expressions

Most people are face perception experts [2]. Faces, especially those expressing emotion, automatically capture our attention [3], and we extract the emotional meaning of those faces in a matter of a few hundred milliseconds [4], even subconsciously [5]. Expressions of intense emotion, such as a wide-eyed expression of fear or a toothy grin, may have evolved to be highly distinguishable signals, easily recognizable even from a distance [6]. However, the majority of expressions we encounter are not prototypical expressions. They are instead fleeting, subtle, and somewhat idiosyncratic facial gestures [7] shaped by learning and culture (Box 1). Different emotions, attitudes, and intentions can be communicated with the slight changes in eyebrow position, head tilt, onset dynamic, or lip press [7]. However, we infer emotions from faces effortlessly; the facial expressions of others are unique and useful sources of information about the social environment.

Part of the ability to extract information from faces can be attributed to visual expertise [8]. Experts in any domain develop heightened perceptual sensitivity to diagnostic features of object categories, improving their ability to detect and discriminate between category instances [9]. Humans are undoubtedly experts in faces, which can be considered an especially relevant class

¹University of Wisconsin – Madison, 1202 West Johnson Street, Madison, WI 53706, USA
²Cardiff University, Tower Building, Park Place, Cardiff CF10 3AT, UK
³International School for Advanced Studies (SISSA), Via Bonomea 265, 34136 Trieste, Italy

*Correspondence: adrienne.wood@wisc.edu (A. Wood).
Box 1. Culture as a Moderator of Emotion Expression and Sensorimotor Simulation

Despite their relative universality, expressions of emotions and associated display rules are subject to considerable cultural variation [113]. Recently, important cultural variation in facial expression has been explained in terms of the diversity of migratory history of a country. In theory, substantial inward migration should create pressures to interact with strangers and clearly communicate one’s intentions. A social environment in which frequently interacting people who lack common emotion language and cultural expectations about which emotion is experienced under a given condition, should, over time, foster a reliance on emotional expression for creating smooth interaction and for building trust.

Consistent with this prediction, two recent studies showed that country-level historical heterogeneity – or the number of source countries that have contributed to the present-day population of a given country over the past 500 years [114] – predicts the endorsement of expressivity norms as well as the production of easily recognizable facial expressions (Figure I). In particular, individuals from historically heterogeneous cultures reported norms that favored the expression of emotion when an emotion was felt, whereas those from historically homogeneous cultures favored the dissimulation of emotional expression [115]. The former also made facial expressions that on average were more easily recognized by individuals from different cultures [116].

The processes recruited by observers to interpret the meaning of emotional expressions may also be determined by the heterogeneity of migratory history of a country. Indeed, the exchange of eye contact, one possible trigger of sensorimotor simulation [36], varies across cultures and is more frequent in some heterogeneous compared to homogeneous cultures [117]. Because social exchange that is reliant on clear emotion communication would presumably also rely more on sensorimotor simulation, a sensible prediction for future research is that historical heterogeneity determines the degree to which people enact sensorimotor simulation of perceived facial expression, particularly when observing individuals with whom one is not well acquainted.

![Figure I. Country-Level Historical Heterogeneity and Emotional Expressivity](image)

Figure I. Country-Level Historical Heterogeneity and Emotional Expressivity. (A) A map depicting the number of source countries that contributed to the present-day population of each nation in the past 500 years [114], with more heterogeneous cultures in green. (B) The cross-cultural variance in emotional expressivity explained by different cultural dimensions [115] and references therein. Note that historical heterogeneity, or the number of source countries contributing to a country’s population, explains 18% of the variance (map generated at http://lert.co.nz/map/). Abbreviations: I-C, indicator of collectivism; GLOBE, global leadership and organizational behavior effectiveness survey.

of visual objects; our visual system is honed through experience to be sensitive to subtle differences in relevant facial gestures. In addition to relying on visual expertise, people can infer the emotions felt by another based on context [10]. Emotions tend to occur in specific contexts and in response to particular events. People extract these regularities and develop conceptual emotion knowledge, which may be used to guide emotion recognition (see Glossary) [11].

In addition to visual and contextual routes to emotion recognition, people might also make use of sensorimotor simulation, in which they recreate the motor production of the perceived facial expression in themselves. This subthreshold motor activity in theory triggers partial, often unconscious, activity in other neural systems involved in experiencing the corresponding emotion, and from which the simulator implicitly infers the expresser’s internal state. Recent suggestive evidence indicates further that sensorimotor simulation also feeds back to shape the visual percept itself [12].

In the following we build a case for the sensorimotor simulation model of emotion perception. First, we review findings that people automatically simulate perceived facial expressions, and
that this simulation contributes to their ability to infer emotional states. We combine these observations with evidence that activating one component of an emotion state, such as a facial expression, triggers other emotion components. Drawing on current thinking about how predictions are integrated into both action and perception, we then propose that sensorimotor simulation augments visual perception of facial expressions. In the process of constructing a prediction, inputs to different sensory modalities feed back to influence one another; we can thus think about sensorimotor simulation as a further source of information in constructing a percept. We hope this account of how people recruit their own emotion experience to understand and make predictions about the emotions of others fosters more productive integration of emotion perception research and domain-general work on perception and cognition [13].

Emotions as Patterns of Adaptive Behavioral, Affective, and Physiological Responses

To infer a role for sensorimotor simulation in emotion perception, we need first to define ‘emotion’. Emotions are biologically programmed and functional reactions to relevant stimuli that coordinate brain (e.g., attention, perception, cognition, motivation) and body (e.g., hormones, autonomic nervous system, sensory organs, muscle movements) systems that prepare appropriate behavioral responses [14]. Simply put, if brains evolved to move an organism through space [15], then emotions evolved to organize and direct that movement into an adaptive response [16].

Facial Mimicry as Sensorimotor Simulation

Observers of a facial expression of emotion automatically recreate it (covertly, partially, or completely) [27,28]. We refer to this as ‘sensorimotor simulation’ to highlight the fact that the perceptual process involves activity in somatosensory and motor systems largely overlapping with those that support the production of the same facial expression. Simulation sometimes, but not always (Box 2), results in facial muscle activity in the perceiver, called facial mimicry. Simulation and mimicry can occur even if the expression is irrelevant to the task at hand [29] or is perceived nonconsciously [5,30]. For example, patients with unilateral damage to the occipital visual cortex, who are phenomenally blind in the contralateral visual field, nevertheless show facial mimicry and emotion-congruent changes in physiological arousal when facial expressions are presented to their lesioned hemisphere.

While extant findings point to the existence of imitative capacities early in development [31] and in other primates [32], learning may increase the correspondence between visual input and sensorimotor activity, at least in intentional mimicry [33]. The tendency to simulate or overtly imitate perceived emotion expressions facilitates collective action in a social group and allows an
organism to vicariously learn which stimuli and behaviors are safe, threatening, acceptable, or desirable [34]. Although we argue that sensorimotor simulation is difficult to suppress completely, the probability and intensity of mimicry are sensitive to the perceiver’s current motivation to engage with and understand the expresser [35] (Box 3). In particular, measurable facial mimicry is more likely to occur when the perceiver and expresser achieve eye contact, which signals the relevance of the emotion expression to the perceiver [36]. In the present view, facial mimicry reflects, and sometimes augments, sensorimotor simulation, but without being a necessary step in the emotion recognition process. Change in the intensity of facial mimicry (usually measured with electromyography, EMG) is only a rough indicator of the degree to which people internally simulate perceived expressions – the absence of measurable facial mimicry does not necessarily indicate an absence of sensorimotor simulation (Box 2).

**Facial Expressions Modulate Emotion Experience**

How does activity in the somatosensory and motor neural systems of a perceiver contribute to solving the task of inferring an emotion from a perceived facial expression? Empirical evidence confirms that our own facial expressions feed back to modulate our emotional experiences [26]. For instance, covertly inducing participants to smile while they undergo a painful cold-pressor task reduces their physiological arousal and self-reported negative affect [37]. Producing facial expressions can activate other physical components of affect, such as pupil dilation (an indicator of autonomic arousal) [38]. Smiling reduces the rapid neural response to errors (as measured with event-related potentials, ERP) in healthy participants [39] and participants with minor depression [40].

In a complimentary way, inhibiting facial movement can reduce the intensity of emotional experience. For example, paralyzing the corrugator muscle – which produces the furrowed brow of anger and sadness – with injections of botulinum toxin reduces symptoms of depression.
Box 3. Social and Contextual Influences on Facial Mimicry and Sensorimotor Simulation

The automaticity of sensorimotor simulation and its putative role in emotion recognition are sometimes challenged because facial mimicry appears to be sensitive to social context (see [119] for a review). Sensorimotor simulation probably contributes more to emotion processing when the perceived expression is ambiguous, when the context does not clearly predict what the expresser may be feeling, or when the perceiver is motivated to know the emotional state of the expresser. Individual characteristics of the perceiver, such as gender, age, empathic tendency, power status, attachment style, and social anxiety, have been shown to influence facial mimicry (reviewed in [10]). Importantly, the extent to which people enact sensorimotor simulation depends on their relationship with the expresser: in-group members and liked others elicit more facial mimicry than do out-group members and disliked others [119].

Emotion perception relies on limited cognitive resources, such as attention and sensorimotor simulation, and it is therefore unsurprising that people engage in it more with people they want or need to understand. Motivational states may moderate automatic simulation at very early stages in the process [36], and people can further inhibit the process by avoiding eye contact [51]. They may also suppress overt, visible facial mimicry in contexts where mirroring the perceived expression would be inappropriate or incongruent with the emotional state of the perceiver; for instance, people can rapidly react to perceived negative expressions with positive expressions [119]. The flexibility of overt facial mimicry is underscored by evidence that mimicry of anger, but not of happiness or sadness, is reduced in close relationships [120].

Simulation and the Recognition of Facial Expression

We have argued that the sensorimotor processes involved in producing a particular facial expression can trigger other components of the emotion system. This implies that simulating another’s facial expression can fully or partially activate the associated emotion system in the brain of the perceiver. If so, this should be the basis from which accurate facial expression recognition is achieved [46] (the proposed model is presented in Figure 1).

Substantial research evidence suggests that sensorimotor simulation contributes to accurate and efficient recognition of the specific emotion [47,48], valence [49], intensity [50], and intentionality [51,52] conveyed by facial expressions. Targeted disruption of particular facial muscle contractions most affects recognition of expressions that involve those muscles [53,54]. In the cortical blindness study mentioned above [5], the guesses of the patients as to which emotion was present in their “blind” visual field were significantly better than chance, and not significantly worse than their categorization performance for expressions shown to their intact visual field. Because the patients displayed facial mimicry in response to facial expressions they perceived nonconsciously, this study provides suggestive evidence that sensorimotor simulation can aid emotion processing even below awareness.

If the coupling between facial expression perception and simulation is consistently increased or inhibited, then individual differences in emotion perception abilities can emerge. For instance, people with mirror-touch synesthesia report experiencing exaggerated sensorimotor simulation and display heightened activity in sensorimotor brain regions when they observe others being touched ([55], see also [56]). These synesthetes perform better than control participants on facial expression recognition tasks, probably because sensorimotor feedback during emotion perception is amplified [57].
On the other hand, a recent study linked deficits in facial mimicry and emotional competence among young and adolescent boys to pacifier use [58]. Specifically, the longer the boys had used a pacifier, the less spontaneous facial mimicry they displayed and the lower their scores on measures of empathy and emotional intelligence years later (Figure 2). No effect of pacifier use on the emotional competence of girls was observed. This difference may be due to greater vulnerability of boys to disruptions of facial expression processing induced by the repetitive inhibition of facial mimicry. By 12 months of age girls on average are superior to boys in their social referencing and use of emotion expressed by caretakers [59,60]. Furthermore, parents talk to girls about emotions more than they talk to boys about this topic [61]. It is thus possible that the dose–response curves are different for boys and girls. Gender and individual differences in emotion socialization notwithstanding, the studies described above suggest that pacifiers can inhibit sensorimotor simulation during an important period in development [62], decoupling perception of facial expressions and automatic mimicry, which has unsurprising consequences for emotion processing (see also [63]).

Other factors that disrupt somatosensory feedback from, or motor output to, the facial muscles are expected to affect emotion recognition. As people age, their sensitivity to bodily feedback declines and may impact on their ability to detect not only their own emotional states [64] but...
perhaps also the expressions of others [65]. As a more extreme example of individual differences in emotion perception as a result of sensorimotor disruption, preliminary evidence suggests that unilateral facial paralysis alters the ability of patients to detect dynamic changes in asymmetric facial expressions [66]. Specifically, facial paralysis had a greater effect on the sensitivity of patients to facial expressions that began on the same anatomical side of the expresser’s face as the paralysis of the patient (see also [45]). At first glance severing or weakening of motor or somatosensory nerves may not be expected to seriously impact on simulation and emotion recognition because we do not consider that overt facial muscle movement is necessary for sensorimotor simulation (Box 2). However, any disruption of the predictive motor-sensory loop could impair sensorimotor simulation. Furthermore, animal models suggest that prolonged facial paralysis can cause the associated motor and somatosensory cortices to shrink or reorganize [67], becoming unavailable for sensorimotor simulation. Whether chronic facial paralysis disrupts emotion recognition or compensatory perceptual strategies can eventually develop is not fully clear (e.g., [68]).
Sensorimotor simulation may be especially useful for emotion recognition when the perceived expression is subtle or ambiguous \[69\], or the judgment to be made is particularly challenging \[46\]. Judging the authenticity of a smile, for instance, is a complex task that relies on detecting subtle differences in the morphological specificities and temporal dynamics of the smile. The intensity of automatic facial mimicry corresponds to the degree of smile muscle contraction in the face of the expresser, suggesting that facial mimicry (and presumably sensorimotor simulation) is sensitive to the particular features of an expression \[52\]. The amount of measurable facial mimicry predicts ratings of the authenticity of a smile \[52\], and disrupting mimicry impairs the ability of observers to distinguish between videos of people producing spontaneous and intentional smiles \[51\]. An interesting and currently unanswered question is whether facial mimicry and simulation, being helpful for emotion recognition, could be trained to improve emotional intelligence. Intentional mimicry may be equally challenging for people as intentional posing of facial expressions, as indicated by the systematic differences in faked and spontaneous expressions (e.g., \[70\]). The level of top-down muscle control involved in deliberate mimicry may disrupt spontaneous sensorimotor influences, and possibly impair or alter emotion recognition, because spontaneous and intentional facial expressions do not recruit the same networks of brain activity \[71\]. To summarize the current evidence, automatic sensorimotor simulation plays a functional role not only in the recognition of actions (\[72\], cf. \[73\]) but also in the processing of emotion expressions.

**Neural Bases of Sensorimotor Simulation**

Sufficient insight into how sensorimotor simulation is implemented in the brain is now available. The neural circuitry involved in perceiving a facial expression is distributed across cortical and subcortical areas, and involves recursive feed-forward and feed-back processes (\[74,75\] for review). In addition to visual cortices, visual recognition of another’s facial expression reliably activates motor and somatosensory cortices, evidencing the link between perception and sensorimotor simulation \[76\]. In particular, somatosensory, motor, and premotor cortices are associated with emotion recognition deficits in research on lesion patients \[77\] and in research using transcranial magnetic stimulation (TMS) \[76,78,79\].

For example, inhibiting the right primary motor (M1) and somatosensory (S1) cortices of female participants with TMS reduced spontaneous facial mimicry and delayed the perception of changes in facial expressions \[79\]. This finding suggests that disrupting the ability of observers to recruit sensorimotor networks during emotion perception reduces their sensitivity to facial expression dynamics. The application of TMS to the same locations had no effect on mimicry and perception of facial expressions by male participants, representing an interesting gender difference that deserves further investigation \[80\]. Other studies implicate the premotor cortex \[81\] and the pre-supplementary area \[82\] in emotion detection and recognition.

Many simulationist accounts (e.g., \[46\]) postulate a linear sequence of events in which coarse visual processing is followed by spontaneous facial motor output, which then generates somatosensory processing of facial feedback (see Box 2 for a discussion of whether facial muscle contractions are necessary for sensorimotor simulation). It is, however, well established that the motor and somatosensory systems are interconnected, and that without the somatosensory system motor output is ‘blind’ and highly ineffective \[83\].

Current models of motor functioning reserve a crucial role for sensorimotor integration in the planning and executing of motor acts \[84,85\]. Instead of solely relying on the time-consuming readout of sensory feedback, the brain uses efference copies of the outgoing motor commands. These corollary discharges of the motor output constitute a form of internal monitoring of the body’s state, and actively suppress expected sensory input. Predicted and obtained somatosensory input are continuously compared and mutually subtracted at multiple levels: from the cortex, over subcortical structures, down to the spinal cord. This supports fast and
adaptive regulation of movement and, for example, the ability to distinguish sensory stimulation caused by self-generated movements (e.g., when hitting somebody) from that occurring due to external stimuli (e.g., when being hit).

One of the appeals of simulationist accounts is that they do not require additional systems to the one underlying an organism’s own actions. In fact, shared representations (and overlapping neural activation patterns) appear to exist for the production, observation, and imagination of action, as well as for the experience, observation, and imagination of touch [15,86,87]. Simulation of observed action and touch therefore competes with the person’s own motor and somatosensory states to influence motor control, physiological responses, and conscious representations [86]. Of relevance for emotion perception, the brain needs to distinguish between simulation of the facial actions of an expresser and its own actual motor output. In the healthy brain, simulation triggered from observation, leading to vicarious activation of motor and somatosensory systems, may be partially inhibited and/or properly attributed to other individuals, through frontal and parietal cortical areas [88]. It has also been suggested that simulation is inhibited by the body’s own sensorimotor feedback, based on the observation that limb loss can magnify vicarious motor and somatosensory activation [86]. Inhibitory mechanisms such as these allow a perceiver to recruit preexisting sensorimotor processes for the purpose of understanding the facial expressions of others without leading to widespread and maladaptive mirror-touch synesthesia [55].

Future research will need to isolate the cortical and subcortical motor areas involved in the production of motor signals leading to facial mimicry, and the somatosensory areas involved in processing and comparing motor efference copies and incoming signals from facial feedback. In part, these questions can be tackled by inhibiting specific cortical areas using methods of TMS (e.g., [79]). In the meantime, we can already argue for crosstalk in neural systems that may explain how sensorimotor activity influences visual perception of facial expressions and actions more generally [89].

**Emotion Perception and Cross-Modal Sensory Integration**

Perceptual inputs in any given modality (e.g., sight, sound, proprioception) are often degraded, unclear, or incomplete, and the brain therefore generates predictions to complete the initial input, sometimes by integrating information from other sensory modalities [90]. Indeed, sensory inputs from one modality (e.g., hearing a scream) feed back to influence perception in another modality (detecting fear in a face), compensating for perceptual ambiguities [91–94]. This is because particular sounds, smells, sights, and tactile cues reliably co-occur in the environment of an individual [95]. A realistically synchronized input to one sensory modality can affect a percept in another modality [92]. For instance, experimental participants are faster to consciously perceive moving visual objects when the direction of the stimuli’s movement is congruent with the participants’ vestibular input (which informs them about their position in space and co-occurs with changes in their visual field) [91].

Cross-modal perceptual integration is often observable in laboratory experiments involving ambiguous perceptual tasks [94]. Facial expressions are often fleeting and difficult to discriminate; this, coupled with the fact that in many interactions prolonged looking at the face of a social partner is aversive [96], can make the visual signal challenging to process. Emotion perception should therefore benefit from cross-modal integration which allows the formation of a more robust and unambiguous percept [97]. Consistent with this prediction, perceptual adaptation to angry facial expressions transfers to the perception of angry vocalizations, providing suggestive evidence for cross-modal integration of emotion percepts [98].

Another study used a constrictive gel facemask, which provides the wearer with distorted sensory feedback from the face, to alter the typical somatosensory-motor feedback loop of the
participants [12]. While wearing the facemask, participants completed a facial expression discrimination task that did not rely on their conceptual emotion knowledge (as do tasks that have been used in related research) but only on their ability to detect small differences between facial expressions. It was found that facial expression discrimination accuracy was worse for participants wearing the facemask, compared to control participants. This difference did not extend to accuracy on a discrimination task that used non-face stimuli (Figure 3). Speculatively, participants may build a representation of the target facial expression that is partially grounded in sensorimotor activity, and which extends the otherwise limited working memory capacity of the visual system. They can then refer to this representation during the discrimination stage of the task. The representation generated by the individuals wearing the gel facemask may however be distorted and ultimately reduce accuracy. Irrespective of the exact mechanism, the findings suggest that, in addition to informing the judgments of the perceivers about a facial expression (Figure 1C), sensorimotor simulation also contributes directly to visual perceptual processes (Figure 1F).

Ultimately, emotion recognition is a game of prediction because we cannot directly access another’s experience. In addition to facilitating accurate emotion perception, simulation may also generate predictions about the future actions and emotions of others [99–102]. Prior expectations about the emotional state of an expresser [103,104], as well as the ongoing affect of the perceiver [105], may also contribute to this predictive process, coloring the visual percept via sensorimotor activation. In this way, simulation becomes another contributor to the multi-modal
percept of the expressed affective state, constructed by predictively combining sensory information with prior experience [106,107].

**Concluding Remarks**

We have surveyed the substantial evidence that (i) activation of one component of an emotion state (such as a facial expression) often triggers other emotion components, and (ii) perceiving emotion expressions relies in part on automatic simulation of emotion expression production. Connecting these two lines of evidence, we argued that the often-observed sensorimotor simulation process allows a perceiver to access stored knowledge, grounded in the distributed emotion system, about the emotional and motivational states associated with the facial expression. During emotion perception, sensorimotor simulation may partially or fully reactivate related concepts, affective states, and autonomic and behavioral changes, which recursively shape and are shaped by visual information. This form of multi-modal integration can influence the visual percept [12] and generate predictions, a necessary function of all cognitive processes [106].

When considered in the context of current understanding by psychologists of how conceptual knowledge is distributed in the perceptual networks of the brain [108], and how this knowledge is accessed to generate predictions and cross-modal percepts [90], simulationist accounts of emotion recognition do not require any special neural ‘hardware’. Visual emotion perception is augmented, as is performance in other perceptual tasks, by recruiting input from other modalities and recreating emotion states [109]. The process is highly flexible but automatic [35].

Deeper understanding of the complex interplay of vision processing, sensorimotor simulation, learning, and evolved emotion states will inform interventions for people suffering from central [110] or peripheral [45] motor diseases, as well as for disorders associated with mimicry and/or emotion recognition impairments [111]. While many questions remain unanswered (see Outstanding Questions), research continues to generate a progressively more precise and elaborate understanding of how humans can infer complex mental states from even the slightest facial expression.

**Acknowledgments**

We thank two anonymous reviewers and Lawrence Barsalou for their helpful feedback. A.W. was supported by National Science Foundation (NSF) Graduate Research Fellowship 2012148522, M.R. by European Science Research Council (ESRC)-ORA Grant 505465, and P.N. by NSF BCS-1251101.

**References**

Trends in Cognitive Sciences


47. Künstele, J. et al. (2014) Facial EMG responses to emotional expressions are related to emotion perception ability. PLoS ONE 9, e84053


51. Rychlikova, M. et al. (2014) Blocking mimicry makes true and false smiles look the same. PLoS ONE 9, e89076


67. Urgesi, D. et al. (2015) Layer 5 pyramidal neurons’ dendritic remodeling and increased microglial density in primary motor...
68. Bate, S. et al. (2013) First report of generalized face processing difficulties in Möbius sequence. PLoS ONE 8, e62656