The Simulation of Smiles (SIMS) model: Embodied simulation and the meaning of facial expression

Paula M. Niedenthal
Centre National de la Recherche Scientifique (CNRS) and Clermont Université, 63037 Clermont-Ferrand, France
niedenthal@wisc.edu
http://wwwpsy.univ-bpclermont.fr/∼niedenthal/

Martial Mermillod
Centre National de la Recherche Scientifique (CNRS) and Clermont Université, 63037 Clermont-Ferrand, France
martial.mermillod@univ-bpclermont.fr
http://wwwpsy.univ-bpclermont.fr/∼mermillod/

Marcus Maringer
Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands
m.maringer@rug.nl

Ursula Hess
Department of Psychology, Humboldt-Universität Berlin, 12489 Berlin, Germany
Hess.Ursula@psychologie.hu-berlin.de

Abstract: Recent application of theories of embodied or grounded cognition to the recognition and interpretation of facial expression of emotion has led to an explosion of research in psychology and the neurosciences. However, despite the accelerating number of reported findings, it remains unclear how the many component processes of emotion and their neural mechanisms actually support embodied simulation. Equally unclear is what triggers the use of embodied simulation versus perceptual or conceptual strategies in determining meaning. The present article integrates behavioral research from social psychology with recent research in neurosciences in order to provide coherence to the extant and future research on this topic. The roles of several of the brain’s reward systems, and the amygdala, somatosensory cortices, and motor centers are examined. These are then linked to behavioral and brain research on facial mimicry and eye gaze. Articulation of the mediators and moderators of facial mimicry and gaze are particularly useful in guiding interpretation of relevant findings from neurosciences. Finally, a model of the processing of the smile, the most complex of the facial expressions, is presented as a means to illustrate how to advance the application of theories of embodied cognition in the study of facial expression of emotion.

Keywords: amygdala; basal ganglia; Duchenne smiles; eye contact; embodiment; facial expression; mimicry; simulation; somatosensory cortex

A smile is the chosen vehicle for all ambiguities.
— Herman Melville (1852/1996, Pierre, or, The Ambiguities, p. 84)

Smiles can be simple things. Individuals show very high agreement in their assignment of the label “smile” to photographs of facial gestures with certain structural features (Izard 1971). This is true across cultures (Ekman 1994; Haidt & Keltner 1999). Yet, smiles are also complicated things. Although smiles often communicate that the expresser feels “happiness” or “joy” (Frank et al. 1997; Frank & Stennett 2001; Messinger et al. 2001), some smiles signal affiliative intent or a responsiveness to group norms; others express more complex interpersonal or status motivations (Abel 2002; Fogel et al. 2000; LaBarre 1947; Keltner 1995; Tipples et al. 2002).

How do individuals interpret the meaning of a smile? One possibility is that each smile has its own specific facial morphology, which constitutes slightly different visual features. Meanings associated with these configurations could then be learned. Ekman (2001) identified 18 types of smiles and proposed that there might be as many as 50 in all. If visual facial features did all of the work in grounding meaning, we would expect very few errors in interpreting smiles across individuals and cultures. Yet, even though errors in classification of a smile as such are not frequent (Haidt & Keltner 1999), errors of interpretation of specific smile meanings are much more so (Bernstein et al. 2008).
In the present article we argue that observers of smiles sometimes construct an embodied simulation of the nuanced affective state conveyed by the smile that is supported by the brain’s reward, motor, somatosensory, and affective systems. They then use this simulation to represent the smile’s intended meaning. Our approach also outlines the conditions under which simulations, in contrast to other bases of processing, are actually used, and why and when these alternative processes may lead to errors.

To accomplish our goals, we devote the first section of the article to a review of research on the meaning of smiles from the view of the person doing the smiling. Based on existing functional accounts, we characterize smiles as produced by positive emotion (enjoyment smiles), by positive social motives (affiliative smiles), and as a way of communicating and maintaining social status (dominance smiles).

In the second section of the article, we review research on the neural bases of smile processing in the perceiver of the smile. Possible roles of the brain’s reward centers, orbital prefrontal cortex, amygdala, motor regions, and somatosensory cortices are outlined, with accounts of motor processing linked to research on facial mimicry from social psychology and the neurosciences.

In a third section, eye contact is discussed. Our novel proposal is that eye contact automatically triggers an embodied simulation of what a smile means. A large literature in social and developmental psychology supports this claim, and we rely on it to draw conclusions from recent neuroscience findings. Finally, we bring together these summaries to motivate a comprehensive model of smile interpretation for three smile types: enjoyment, affiliative, and dominance smiles.

1. Recognition and access to meaning of expressions

Different processes can support people’s ability to recognize smiles and what they mean (Adolphs 2002; Atkinson 2007; Kirouac & Hess 1999). We begin with visual facial cues and then consider a variety of other cues.

The classification of expressions into basic categories usually relies on a perceptual analysis of the stimuli, sometimes called pattern matching (Buck 1984). Smith et al. (2005) have shown that distinct facial features can be used to classify facial expressions and that these features correspond to a de-correlation model. In their view, configurations of muscles have emerged during phylogenetic development of the human species that maximize differences between the six basic emotional expressions and produce efficient recognition of these expressions. At the neural level, this process seems to be supported by the occipito-temporal cortices (Adolphs 2002).

Whereas the analysis of visual facial features may be sufficient to classify prototypical expressions in simple tasks, this process is unlikely to be sufficient to recognize less prototypic, perhaps more realistic, emotional expressions, or to represent their subtle meanings. In such cases, we propose that perceivers must call on various sources of non-visual information, such as conceptual emotion knowledge about the expresser and the social situation (Kirouac & Hess 1999; Niedenthal 2008). For example, faces provide information about the sex, age, and race of the other person, and individuals possess expectations and stereotypes about how members of these social groups react emotionally (Hess et al. 2005). Such conceptual knowledge about emotion has been shown to exert effects early in the processing of ambiguous facial expressions (e.g., Halberstadt & Niedenthal 2001; Halberstadt et al. 2009; Hess et al. 2009a).

Embodied simulation also supports the recognition and access to meaning of facial expressions (e.g., Atkinson 2007; Decety & Chaminade 2003; 2004; Gallee 2003; 2005; Goldman & Sripada 2005; Keysers & Gazzola 2007; Niedenthal 2007; Niedenthal et al. 2005b; Winkielman et al. 2009). When we use the term “embodied simulation,” we mean that a facial expression has triggered a simulation of a state in the motor, somatosensory, affective, and reward systems that represents the meaning of the expression to the perceiver. In an embodied simulation account, the perception of a facial expression is accompanied by the bodily and neural states associated with the expression and its correspondent emotion. This simulation is then used as the representation of meaning on which an interpretation or judgment is based.

In social psychology, such a view partly involves the marriage of facial feedback theory and affect as information theory. The first holds that facial musculature produces affective feedback that alters subjective state (McIntosh 1996; Zajone et al. 1989). The latter holds that when individuals believe that their affective state was caused by the current object of perception, they use that state to evaluate affective features of the object.

**Paula M. Niedenthal** received her Ph.D. at the University of Michigan and was on the faculty of the departments of Psychology at Johns Hopkins University and Indiana University (USA). She is currently Director of Research in the National Centre for Scientific Research and member of the Laboratory in Social and Cognitive Psychology at Blaise Pascal University in Clermont-Ferrand, France. Her areas of research include emotion–cognition interaction and representational models of emotion. Author of more than 100 articles and chapters, and several books, Dr. Niedenthal is a fellow of the Society for Personality and Social Psychology.

**Martial Mermillod** is Associate Professor of Psychology at the Blaise Pascal University, CNRS (UMR 6024) and Institut Universitaire de France. He has published over 30 peer reviewed articles and book chapters in the domain of emotion and perception, neural computation, and cognitive sciences.

**Ursula Hess** is Professor of Social and Organisational Psychology at the Humboldt-University, Berlin. She has published extensively (over 100 peer reviewed articles and book chapters) in the domain of emotion and perception communication.

**Marcus Maringer** is a lecturer of Psychology at the Hague University for Applied Sciences. He has published several reviewed articles and book chapters in the domain of emotional influences on perception and judgment formation.
contraction, along with orbicularis oculi contraction, is associated with cheeks, a narrowing of the eye opening, and wrinkles (Duchenne 1862). The orbicularis oculi causes a lifting of the eyelids, the orbicularis oculi, and the associated (Duchenne 1862). The Duchenne marker involves the contraction of the muscle around the eye, the orbicularis oculi, and is sometimes indicative of positive emotion (Frank et al. 1993; Soussignan 2002). The perceiver of the “Duchenne” smile also interprets it as expressive of positive emotion (Miles & Johnston 2007) and may respond to it with positive affect (Sarafka & Hietanen 1998). Smiles lacking the marker have been referred to as “false,” “masking,” and “non-Duchenne” smiles (see Figure 1; École Nationale Supérieure des Beaux-Arts 1999). Other research shows that dynamic features of smiles, such as their symmetry, smoothness, duration, and synchrony, may distinguish true and false smiles as well (Cacioppo et al. 1986; Ekman et al. 1980, reviewed in Frank 2002; Hess & Kleck 1990).

Recent research, however, has shown that the utility of the Duchenne versus non-Duchenne distinction is limited (e.g., see Abe et al. 2002, for a review). For instance, some studies have demonstrated that dynamic characteristics of the smile can override the Duchenne marker’s importance in determining judgments of how true (or sincere) the smile is (Hess & Kleck 1990; Krumhuber et al. 2007). Others have shown that non-Duchenne smiles may be associated with self-reported happiness in adults (e.g., Hecht & LaFrance 1998; Hess et al. 1995; Jakobs et al. 1999), and, conversely, that Duchenne smiles may be displayed in situations in which false smiles would be predicted (Bourgeois & Hess 2008). Finally, there is some evidence that the importance of the Duchenne marker varies with culture (Thibaut et al. 2008).

2. What is a smile?

The smile is characterized by the upward turn of the corners of the lips, which is produced by the contraction of the zygomaticus major muscle (Ekman & Friesen 1978). The zygomaticus major, like other muscles involved in the production of facial expression, is activated by the seventh cranial nerve, or the facial nerve (Rinn 1991). The facial nerve can be innervated by one of two motor systems. The subcortical motor system, also known as the extrapyramidal circuit, supports non-voluntary, facial expression. The cortical motor system, also known as the pyramidal circuit, supports learned, voluntarily facial expression, which may vary across cultures and be produced and inhibited intentionally.

Early research on the smile revealed that the frequency, intensity, and duration of the zygomaticus major muscle activity positively predicted self-reported happiness of the smiler (Ekman et al. 1980; Cacioppo et al. 1986). Zygomaticus major contraction, however, is observed not only when positive emotions are experienced, but may also be observed when individuals report feeling negative emotions such as disgust (Ekman et al. 1980), disappointment (Kraut & Johnston 1979), sadness and uncertainty (Klineberg 1940), and general discomfort (see Ekman et al. 1990, for a review).

2.1. True and false smiles

Due to the failure to observe a clear correspondence between activation of the zygomaticus major and positive feelings, several theorists have suggested that the smile should not be treated as a single category of facial expression (e.g., Ekman & Friesen 1982). In the proposed distinctions, “true” or “sincere” smiles were defined as involuntary displays of positive affect, whereas “false” or “insincere” smiles were defined as smiles voluntarily used to communicate that a positive emotion was felt when it was not (and when, in fact, it served to mask negative feelings).

Several morphological and dynamic markers have been proposed to distinguish these two types of smiles. The most frequently cited morphological indicator of a true smile is the Duchenne marker, thanks to Duchenne’s empirical work on smiles and other facial expressions (Duchenne 1862). The Duchenne marker involves the contraction of the muscle around the eye, the orbicularis oculi, pars lateralis. The orbicularis oculi causes a lifting of the cheeks, a narrowing of the eye opening, and wrinkles around the eyes. The combination of zygomaticus major contraction, along with orbicularis oculi contraction, is sometimes indicative of positive emotion (Frank et al. 1993; Soussignan 2002). The perceiver of the “Duchenne” smile also interprets it as expressive of positive emotion (Miles & Johnston 2007) and may respond to it with positive affect (Sarafka & Hietanen 1998). Smiles lacking the marker have been referred to as “false,” “masking,” and “non-Duchenne” smiles (see Figure 1; École Nationale Supérieure des Beaux-Arts 1999). Other research shows that dynamic features of smiles, such as their symmetry, smoothness, duration, and synchrony, may distinguish true and false smiles as well (Cacioppo et al. 1986; Ekman et al. 1980, reviewed in Frank 2002; Hess & Kleck 1990).

Recent research, however, has shown that the utility of the Duchenne versus non-Duchenne distinction is limited (e.g., see Abe et al. 2002, for a review). For instance, some studies have demonstrated that dynamic characteristics of the smile can override the Duchenne marker’s importance in determining judgments of how true (or sincere) the smile is (Hess & Kleck 1990; Krumhuber et al. 2007). Others have shown that non-Duchenne smiles may be associated with self-reported happiness in adults (e.g., Hecht & LaFrance 1998; Hess et al. 1995; Jakobs et al. 1999), and, conversely, that Duchenne smiles may be displayed in situations in which false smiles would be predicted (Bourgeois & Hess 2008). Finally, there is some evidence that the importance of the Duchenne marker varies with culture (Thibaut et al. 2008).

2.2. Functional smiles

The distinction between true and false smiles itself may be largely superseded by another more useful distinction, namely, distinctions based on smile function (e.g., Barrett 2002). We describe three types of smiles that we believe have important and discrete functions, and which may map onto identifiable brain systems that represent different meanings.

As we have already implied, and consistent with experience, many smiles are simply readouts of positive internal states such as happiness (Buck 1984). The “play-face” in primates, such as chimpanzees, held to be a homologue of laughter in humans, corresponds in musculature to the human smile (e.g., Parr & Waller 2006; see our Figure 2, left panel). Humans and some primates smile spontaneously during experiences of pleasure, including visual, auditory, gustatory, and tactile stimulation (Ekman & Friesen 1982). Smiles that are readouts of happy feelings reinforce the behaviors that elicited them in the first place. Thus, the communication of positive emotion through the smile is essential, among other things, for learning in infants, when mothers smile at babies to encourage desired behaviors (Klineberg et al. 1983). We will refer to smiles that express happiness as enjoyment smiles.

Second, smiles can be readouts of positive social intentions that are essential for the creation and maintenance of social bonds, without necessarily being about personal enjoyment (Cashdan 2004; Fridlund 1991; 2002). Such smiles may include the “greeting” smile (Eibl-Eibesfeld 1972), as well as those of appeasement, and perhaps the smile component of embarrassment (Keltner 1995; Hess et al. 2002). In most primates, the silent bare-teeth display (right panel of Figure 2) serves to communicate that the smiler intends no harm and that there is no threat of aggression (van Hooff 1976; Waller & Dunbar
Figure 1. The top two photographs show the Duchenne (left) and non-Duchenne (right) smiles as elicited by Guillaume-Benjamin Duchenne de Boulogne himself, using electrical impulses to manipulate relevant facial muscles. The bottom two photographs show more recent posed versions of the same.

Figure 2. Two chimpanzee facial expressions related to the human smile. The left panel shows a play face believed to be a homologue of laughter and sharing morphological features with the human enjoyment smile (Parr & Waller 2006). The right panel shows a silent bared-teeth display, used in affiliative and appeasement contexts, believed to be homologous with the human affiliative smile and sharing similar musculature (Parr & Waller 2006). Photos courtesy of Dr. Lisa Parr, National Primate Research Center, Emory University, Atlanta, GA. Used with permission.
In primates with relatively egalitarian social systems, including some macaques, mandrills, Gelada baboons, and chimpanzees, the silent bare-teeth display is also seen in grooming, sexual solicitation, and reconciliations (Preuschoft & van Hooff 1997; Waller & Dunbar 2005). Smiles that express positive social motives will be called affiliative smiles. Finally, what we will call dominance smiles are expressions that reflect social status or control, and may include displays that have been called “scheming smiles” (e.g., Öhman et al. 2001; Tipples et al. 2002), “critical smiles,” and perhaps the smile component of the expression of pride (Tracy & Robins 2004; 2008). Darwin also referred to derisive or sardonic smiles in his discussion of sneering and defiance (Darwin 1872, p. 251). Recent analyses of the signals of leadership and dominance in human societies refer to this type of smile. Senior et al. (1999) note that (former Prime Minister of Britain) Tony Blair and (former American President) Bill Clinton have been called “skilled proponents of the dominant smile” (p. 344).

Whether a smile is voluntary on the part of the smiler is probably not a factor that best distinguishes enjoyment, affiliative, and dominance smiles. That is, the involvement of the Duchenne marker will probably not turn out to be diagnostic of one type of smile. Specific instances of the proposed functional types may however be associated with specific postural or other facial features (e.g., the greeting smile contains an “eyebrow flash,” and the embarrassed smile is part of a larger gesture). Thus, it might in principle be possible to construct a visual description of the full bodily and facial characteristics of enjoyment, affiliative, and dominance smiles. But the present issue is: How do perceivers arrive at those meanings? We propose that perceivers use a set of neural and behavioral processes to extract a smile’s meaning that allows them to distinguish between the three functional smile categories in terms of the feelings they generate in the perceiver.

3. What is simulated in the simulation of a smile?

In this section, we explore the neural structures that could play a central role in the embodied simulations that represent different smile meanings. Specifically, we address how the basal ganglia, prefrontal cortex, amygdala, motor system, and somatosensory system all contribute to the experienced meaning of a perceived smile on another person’s face. Later sections will discuss when and how these simulations are triggered, and the implications for emotional experience. Although the focus is on specific neural structures, this does not mean that we endorse a localization approach. It will remain for neuroscientists to fully establish the larger distributed circuits that support the functions we suggest that these structures play.

3.1. Subcortical and cortical affective systems

3.1.1. Basal ganglia. As people perceive smiles, the reward system in their brain may simulate the experience of reward. Research shows that the smile can function as a social reward for both adult and infant humans (Trevarthen 1974; Tronić et al. 1978).

This view is consistent with evolutionary and ecological treatments of the smile (e.g., Buck 1991; Burgoo et al. 1996; Haith 1972; McArthur & Baron 1983). Research on both primates and rats implicates the striatum and the ventral tegmental area (VTA) in reward processing (Kawagoe et al. 1998; Parkinson et al. 2000; Schultz et al. 2000). Further studies indicate that these regions are similarly related to reward processing in humans (Damasio et al. 2000; Davidson & Irwin 1999; Lane et al. 1997; 1999; Mobbs et al. 2003; O’Doherty et al. 2003; Rauch et al. 1999; Redoute et al. 2000). Recent research has also linked the reward areas of the basal ganglia with the perception of smiling faces (Chakrabarti et al. 2006; Lawrence et al. 2004; Lee et al. 2006; Morris et al. 1996; 1998; Okun et al. 2004; Phillips et al. 1998;...
Whalen et al. 1998). When people perceive smiles, they experience them as rewarding.

3.1.2. Prefrontal cortex. Simulations of reward as people perceive smiles may also reflect contributions from prefrontal cortex. The reward circuitry just described in the basal ganglia is associated with what Davidson (1993) has termed pre-goal attainment positive affect. In his view, activation in subcortical reward centers supports an organism’s approach toward an appetitive goal. Davidson contrasts the approach state with post-goal attainment positive affect, which represents a functionally different type of positive emotion, characterized by feelings of affiliation and attachment (biologists refer to the former as incentive value and the latter as reward outcome).

Post-goal attainment positive affect has been linked in a number of studies with activation in the orbitofrontal cortex (OFC) (e.g., Anderson et al. 2003; Elliott et al. 2000; O’Doherty et al. 2001; Rolls 2000). Findings by Davidson and colleagues, for example, demonstrate that smiles displayed by offspring are related to OFC activation in mothers who perceive these smiles (Nitschke et al. 2004). Specifically, the OFC differentiates the sight of one’s own smiling baby from the sight of an unknown smiling baby (who is nevertheless cute and positive in impact). This finding has recently been replicated and extended (e.g., Minagawa-Kawai et al. 2009). The reward outcome system in the brain probably extends to medial orbital and medial prefrontal (dorsal to orbital) cortices, as well as to anterior cingulate cortex (O’Doherty et al. 2003). This system may also be involved in reward learning. In addition, and very importantly, the system is implicated in processing status relations (Zink et al. 2008), and thus may contribute to distinguishing dominant versus submissive smile meanings.

Although the specific terms, pre-goal and post-goal affect, are not necessary for our theorizing, the distinction is relevant for differentiating the perception of enjoyment and affiliative smiles. Because the OFC and contiguous areas process learned emotional responses (Rolls 2004), OFC activation may distinguish the basic rewarding properties of smiles from the reward of experiencing smiles made by people with whom an individual has a significant, previously established emotional relationship (e.g., family members, in-group members). Indeed, other evidence and reasoning suggests that reward systems in prefrontal cortex play a distinct role in responding to affiliative smiles, reflecting their association with attachment information (e.g., Schore 2001).

3.1.3. Amygdala. As people perceive smiles, the amygdala may produce states that further contribute to simulations that underlie how these smiles are interpreted. The amygdala’s involvement in perceiving fear expressions has been considered at length (e.g., Adolphs 2008; Atkinson 2007; Heberlein & Atkinson 2009; Le-Doux 2007). Lesion and neuroimaging studies initially indicated that the amygdala was not only vital for recognizing the emotional expression of fear, but that it further supported the full experience of fear and its behavioral implications (e.g., Calder et al. 2001). These studies did not demonstrate any role for the amygdala in smile processing (e.g., Whalen et al. 1998), and patients with amygdala lesions had not been reported as showing deficits in recognizing happy faces (Adolphs et al. 1994; Hamann et al. 1996; Calder et al. 1996).

Yet, when measuring amygdala activity in human volunteers during rapid visual presentations of fearful, smiling, and neutral faces, Breiter et al. (1996) found that the amygdala also responded preferentially to smiles versus neutral faces (see also a meta-analysis by Fussar-Poli et al. 2009). Other studies supported the conclusion that all emotional expressions can activate the amygdala. For example, a functional magnetic resonance imaging (fMRI) study by Winston et al. (2003) used emotion perception tasks that were either implicit (ratings of maleness) or explicit (deciding whether a face was more or less emotional). Both tasks activated broad cortical and subcortical regions for disgusted, fearful, happy, and sad expressions of either low or high emotion intensity. Most importantly, all expressions activated the amygdala, a finding further supported by Fitzgerald et al. (2006). Additional results demonstrate higher activation of the amygdala for high versus low intensity emotions for explicit and implicit tasks (cf. Surguladze et al. 2003).

Taken together, these results suggest that the amygdala responds to stimuli of motivational significance independently of emotion and processing goals. Indeed, more recent evidence favors an even broader account of amygdala function, namely, that it responds to everything of uncertain meaning to the organism (e.g., Murphy et al. 2003; Sander et al. 2003; Whalen et al. 2001). Applying this hypothesis to the smile, Yang and colleagues proposed that individuals may be uncertain about the meaning of a smile, not merely because of its perceptual features, but also because of its meaning as reflected in the current social or experimental context (Yang et al. 2002). These authors further note that when the meaning of a smile is unclear (such as a smile displayed by an enemy), additional information is needed, similar to our proposal here (see also Lee et al. 2006).

3.2. Facial mimicry in the motor system and shared neural substrates

As people perceive smiles, the motor system may simulate the experience of performing the perceived action, further contributing to how the perceived smile is understood.

3.2.1. Facial mimicry. Facial mimicry is defined as the visible or non-visible use of facial musculature by an observer to match the facial gestures in another person’s facial expression. Perceivers of smiles often automatically mimic these smiles. For instance, electromyographic (EMG) recordings reveal that when individuals view a smile, their zygomaticus major muscle contracts, usually within 500 milliseconds after the onset of the stimulus (Dimberg & Thunberg 1998). Mojzisch et al. (2006) similarly demonstrated that observers automatically mimic smiles expressed by virtual characters in dynamic animations, as did Hess and Bourgeois (in press) in an interactive live setting (see also a review in Hess et al. 1999). Automatically mimicking a smile interferes significantly with simultaneously production of an incongruent facial expression, such as anger (Lee et al. 2007).

3.2.2. Link to corresponding emotions. As anticipated by facial feedback theory, facial mimicry may be accompanied by self-reports of a corresponding emotion, sometimes called emotional contagion (Hatfield et al. 1992; 1993; Laird et al. 1994; Strayer 1993; Wild et al. 2001; Soussignan
2002). For afferent feedback to contribute to an embodied simulation of a perceived smile, however, the perceiver does not necessarily have to experience a conscious change in emotional state – such simulations often appear to have unconscious impact. Findings that mimicry produces emotional effects implicitly are therefore also important.

In an innovative study that assessed the causal relationship between facial mimicry and implicit emotion, Botulinum Toxin (BOTOX) was used to block facial mimicry for expressions of anger (Hennenlotter et al. 2009). Participants were directed to mimic angry and sad facial expressions in still photographs. During the anger (though not sadness) mimicry task, participants whose brows had received BOTOX injections exhibited significantly less activation in the limbic system compared to control individuals who received no injection. The result for anger causally links facial mimicry to emotion, given that disabling the facial musculature decreased emotion activation. When the muscle pattern associated with anger is blocked, part of the embodied meaning associated with anger is lost, such that the emotion is experienced less intensely. Consistent with this experimental finding, and specifically relevant to smiles, is a correlational finding by Lee et al. (2006). In that study, participants mimicked faces expressing smiles, as well as other nonemotional facial movements and expressions. The more participants mimicked the observed smiles, the greater the activations in their striatum and amygdala (see also Schilbach et al. 2006).

A recent paper by Schilbach et al. (2008), who collected fMRI and EMG data simultaneously, is also noteworthy here. Their results showed that spontaneous mimicry occurred during the perception of smiles, which was accompanied by neural activity in the motor system, specifically in the inferior left precentral gyrus. Consistent with previous findings, this study also implicated the medial temporal lobe in the spontaneous mimicry of observed emotional expressions. Finally, the dorsal midbrain was also active, which can be interpreted as signaling arousal induced by direct eye gaze.

In sum, there is evidence that mimicry has a role in causing emotion. Our interest here, however, is not with changes in consciously reported emotional state. Instead, we focus on the fundamental role of mimicry in creating the embodied feeling of a smile, which becomes part of its meaning. In a study motivated by a similar idea, Zajonc et al. (1987) showed that the appearance of spouses who had lived together for at least 25 years and had strong marriages grew more similar in facial appearance over time. Their interpretation was that mimicry supports empathy through afferent feedback, with the incidental effect of producing more similar facial musculature (not through more similar emotion). This preliminary interpretation requires further empirical tests (see Bastiaansen et al. 2009 for further discussion and insight).

3.2.3. Shared neural substrates. Facial mimicry receives considerable attention in the literature on so-called mirror neurons and the notion of emotional resonance. Mirror neurons were first observed in the brains of monkeys in response to limb actions (Gallese et al. 1996; Rizzolatti & Craighero 2004; Rizzolatti et al. 1996). Recordings of mirror neurons in the monkey motor cortex, particularly area F5, associated with the production of hand and mouth movements, were interpreted as constituting a mechanism that could support the implicit comprehension of an action, specifically, the goal of an action (Gallese et al. 1996; Rizzolatti et al. 1996). More recently, a similar mirror system for perceiving and performing action has been described in humans (Fadiga et al. 1995; Gallese et al. 1996; Iacoboni et al. 1999; Rizzolatti et al. 1996). The work on humans points to the premotor and parietal areas as the most likely human homologue of monkey area F5 (e.g., Decety & Grezes 1999).

Inspired by this approach, Carr et al. (2003) found that the passive perception and the intentional mimicry of emotional facial expression activate overlapping sets of brain structures, including the ventral premotor cortex, the insula, and the amygdala (see also Wicker et al. 2003). Mimicry relative to perception, however, was accompanied by greater activation in regions related to emotion processing, in particular, the amygdala and anterior insula, and also the primary motor cortex. Similarly, Hennenlotter et al. (2005) found that voluntary production and perception of smiles activated both the right premotor cortex and the inferior frontal cortex, as well as the right parietal operculum and the left anterior insula. These findings were replicated in a careful study by van der Gaag et al. (2007). Although the reviewed findings are compelling, and do suggest that emotion processing is affected by mimicry, we note that the mirror neuron construct and its roles in human emotion recognition have received considerable criticism (e.g., Turella et al. 2009; Dinstein et al. 2008; Jacob 2008; 2009). A clearer understanding of this construct awaits further research.

The robust activation of premotor areas during the observation of facial expressions (unaccompanied by mimicry) also requires discussion. This finding has been interpreted as demonstrating the presence of an as-if loop, meaning that perception activates the programs for producing facial expressions (e.g., Leslie et al. 2004). The processing of as-if loops has also been called off-line simulation (e.g., Atkinson 2007). One interpretation of these findings is that as-if simulations, when they occur, contribute to smile meaning. In support of this view, research has found that perception-plus-mimicry tends to produce stronger affective responses to smiles than perception alone (e.g., Carr et al. 2003). Motor mimicry seems to play an important role in how smiles and their meanings are simulated.

3.2.4. Roles of mimicry in processing perceived facial expression. If facial mimicry is important in constructing embodied simulations for smiles, then mimicry should affect performance on tasks that measure recognition and access to meaning (Adolphs 2002; 2003; Heberlein & Atkinson 2009; McIntosh 2006). In their review of facial mimicry, however, Hess and colleagues (Hess et al. 1999) did not find evidence that mimicry was causally related to the simple recognition of emotional facial expressions, either directly or as mediated by changes in self-reported emotional state (e.g., Blairy et al. 1999). Hess and Blairy (2001) considered the possibility that failure to support a causal path from facial mimicry to emotion recognition may have resulted from the use of very prototypical facial expressions. When they used naturalistic dynamic stimuli to remedy this problem, they observed that the mirror 

BEHAVIORAL AND BRAIN SCIENCES (2010) 33:6 423
Niedenthal et al.: The Simulation of Smiles (SIMS) model
between the perceived expression and self-reported emotional responses (e.g., happiness when viewing smiles). Nevertheless, they found no evidence of a link from motor mimicry to recognition accuracy, either directly or through changes in emotional responding.

Many additional findings further demonstrate that mimicry does not always play a central role in emotion recognition tasks. As noted previously, recognition tasks on prototypic expressions can be accomplished by perceptual analysis alone, without motor mimicry (e.g., Adolphs 2002). Indeed, high-functioning autistic individuals, who do not spontaneously mimic others’ facial expressions (e.g., McIntosh et al. 2006), perform as well as controls when categorizing facial expressions of emotion (e.g., Spezio et al. 2007a). Similarly, individuals with facial paralysis perform quite normally on various recognition tasks (e.g., Calder et al. 2000a; 2000b; Keillor et al. 2002). Finally, actively keeping individuals from mimicking, for example, by asking them to turn the corners of their mouth down while seeing a smile, does not hinder emotion recognition (Blairy et al. 1999).

As all of these findings indicate, facial mimicry is not always required to recognize emotional expressions in simple recognition tasks. In some cases, however, mimicry does facilitate recognition. Niedenthal et al. (2001), for example, observed effects of mimicry when participants had to detect the boundary of facial expression between happiness and sadness. In a more recent study, Stel and van Knippenberg (2008) found that blocking mimicry affected the speed, but not the accuracy, of categorizing facial expressions as positive or negative. Additionally, individuals showing strong automatic mimicry tend to have high levels of empathy (Somby-Borgström 2002; Zajone et al. 1987). These findings point to the possibility that simulation does become important in recognition tasks when they require fine distinctions in smile meaning, such as the processing of different smile types.

**3.2.5. Beyond facial mimicry.** As we noted earlier, postural and other non-facial gestures are also important components of the meaning of a smile. We have discussed facial mimicry so far, but representation of the full meaning of a smile will not be independent of the entire bodily representation of meaning. The sign of appeasement involves a smile, usually a non-Duchenne smile in fact, but also a number of other head and hand gestures (Keltner 1995). Tracy and Robins (2008) have described the expression of pride, which involves a small smile but also a backward tilt of the head, resulting in a lifted chin, as well as a typical posture. We expect the embodiment of the entire expression to be useful in the interpretation of meaning.

**3.3. Simulating embodied experience: Somatosensory cortices**

As people perceive smiles, the somatosensory system may simulate the embodied experience of how the perceived smiles feel, further contributing to representing their meaning. One account of how facial expression could be simulated in an embodied manner is as the output of a simulator (e.g., Barsalou 1999). From this perspective, somatosensory cortices may be involved in simulating the feeling of a perceived smile while processing its meaning (Adolphs 2002; Gallese & Goldman 1999; Keyers et al. 2004).

Right-hemisphere cortices are likely to be involved in simulating emotional expressions (Adolphs et al. 1996; Bowers et al. 1985). For example, patients with lesions in the right somatosensory cortex are poorer at recognizing facial expressions than individuals without such lesions. Specifically, Adolphs et al. (2000) assessed 108 subjects with focal brain lesions and found that the right somatosensory cortex was central for recognizing the facial expressions associated with the six basic emotions. Consequently, these researchers concluded that the right somatosensory cortex generates a representation or an “image” of the felt state, which feeds into the recognition system as a diagnostic cue (e.g., Adolphs 2002; Atkinson 2007).

Following Adolphs et al. (2000), Pourtois and colleagues used transcranial magnetic stimulation (TMS) to selectively interfere with right somatosensory cortex function while participants performed a same/different facial expression-matching task (see Pourtois et al. 2004). This selective interference disrupted task performance. Thus, both lesion and TMS studies implicate somatosensory simulations in the recognition of perceived facial expressions.

Although these previous studies focused on the simple recognition of facial expressions, it is likely that somatosensory simulations also support more subtle interpretation of facial expressions. Not only do somatosensory simulations facilitate recognition, they probably contribute to how perceivers experience the meaning of these facial expressions, specifically, as a felt emotion. An important issue for future research is to assess whether somatosensory simulations indeed play this additional role.

**3.4. What do smiles feel like for the perceiver?**

In the previous sections we suggest that individuals know whether a smile means that the smiler is expressing enjoyment, affiliation, or dominance, because those smiles feel differently in terms of reward, action, and somatosensory experience. Next, we explore how these different experiences help distinguish enjoyment, affiliative, and dominance smiles. From our perspective, we propose that an enjoyment smile involves a basic rewarding feeling of positive affect and that an affiliative smile involves a positive feeling of attachment and intimacy (where the positive feelings of enjoyment vs. attachment/intimacy are distinct). Our analysis of dominance smiles relies on several further assumptions. In hierarchical primate societies such as ours, highly dominant alpha individuals pose a certain threat insofar as they can claim territory or possessions (e.g., food) from lower status group members (Menzel 1973; 1974). Hence, the perceived presence of a dominant other should lead to increased vigilance and preparedness for withdrawal (Coussi-Korbel 1994). Although speculative, we suggest that dominance smiles, like smiles hiding negative intentions or feelings, are associated with the experience of negative rather than positive affect, as indicated by right-lateralized activation (Davidson et al. 1990; Boksem et al. 2009). Thus, the meaning of dominance smiles should not involve the forms of positive emotion associated with the other two smiles.
4. Triggering embodied simulation

For eyes can speak and eyes can understand.
— Chapman

In the previous sections, we saw that embodied simulation is not always implicated in the recognition of emotional facial expressions, indicating that simulation is not always required (for analogous findings and accounts, see Barsalou et al. 2008; Hess & Blairy 2001; Kosslyn 1976; Niedenthal et al. 2009; Solomon & Barsalou 2004; for related situated-simulation assumptions, see Semin & Cacioppo 2008; Smith & Semin 2007; Strack & Deutsch 2004). In this section, we review recent research suggesting that eye contact modulates the presence versus absence of embodied simulation as people perceive smiles.

4.1. Definitions related to eye gaze

Eye gaze is the direction of one’s gaze at another’s eyes, presumably during the search for information useful to attributing the cause of the other’s behavior (von Cranach 1971). Mutual gaze refers to two people gazing at each other’s faces. Eye contact involves two people gazing at each other’s eyes. All such behavior has become of great interest in scientific research lately, usually under the theoretical rubric of “social relevance” (e.g., Adams et al. 2003; Hess et al. 2007; Klucharev & Sanis 2004; Mojzisch et al. 2006; Richeson et al. 2008). Our primary interest here is in eye contact. Specifically, we argue that eye contact counts as a sufficient, but not necessary, trigger of embodied simulation as observers perceive smiles.

4.2. Gaze and simulation of meaning

4.2.1. Eye contact and intimacy. Several sources of evidence suggest that eye contact triggers embodied simulation. In part, the roots of this prediction lie in research on intimacy. Argyle (1972) suggested that increased eye contact during social interaction indicates an increase in intimacy, which is also consistent with Patterson’s (1982; 1983) functional analysis. Additional findings support these proposals: Increased eye contact is associated with increased maternal sensitivity (Lohaus et al. 2001). Individuals make eye contact with people with whom they have close relationships more frequently and for longer durations than with strangers (Russo 1975). Men show more approach behavior after repeated eye contact (Walsh & Hewitt 1985). Dating couples of both sexes tend to look in the eyes of their partners more often than unacquainted couples (Iizuka 1992).

Importantly, the pupils themselves impart no actual descriptive information (beyond dilation and constriction, which inform arousal). Thus, these findings suggest that eye contact triggers something beyond itself that increases intimacy. We suggest that eye contact triggers an embodied simulation of the perceived facial expression and its correspondent feeling for use in interpretation. A classic study by Bavelas et al. (1986) supports this proposal. In their study, a confederate faked experiencing an injury and expressed apparent pain facially. When participants viewed the pained facial expression, they inadvertently mimicked it. The critical manipulation in the experiment was whether the victim of the painful stimulus made eye contact with participants. Analyses revealed that eye contact significantly affected the pattern and timing of participants’ mimicry. Specifically, participants mimicked the perceived expressions of pain most clearly when eye contact was made.

Consistent with this initial finding, Schrammel and colleagues (2009) demonstrated that, when viewing emotional expressions, participants’ zygomaticus major (i.e., smile) activity was higher for happy than for angry or neutral faces, and, most importantly, that this effect was stronger under conditions of eye contact. Furthermore, angry faces elicited more negative affect following eye contact, and happy faces elicited more positive affect, relative to the no-eye contact condition. For both emotions, eye contact modulated the intensity of the experienced emotion. Although both studies just described found that participants in the no-eye contact conditions recognized smiles accurately, the eye contact effects suggest that these participants may not have produced embodied simulations.

It is important to note, however, that drawing a strong conclusion about eye contact’s role as a trigger for the mimicry component of embodied simulation, is difficult. First, as we have seen, some smiles are rewarding stimuli (e.g., the enjoyment smiles used in most research). Therefore, a perceiver of a smile could smile, even without making eye contact with the smiler, simply as a result of the positive emotion associated with the perceived smile (Mojzisch et al. 2006). In this case, the elicited smile would not count as mimicry, but rather as the readout of the perceiver’s positive emotional response to the smile. This interpretation implies that much past research on facial mimicry is not definitive because it does not distinguish mimicking a perceived smile from responding with positive emotion to the sight of a smile. Nevertheless, because a smile can have functionally different meanings, it is an ideal expression for distinguishing mimicry versus emotional responding. Specifically, researchers could determine when perceivers make correct versus incorrect interpretations of different smile types as a function of simply responding with enjoyment to a smile or whether they mimic it. As described next for the SIMS model, we present conditions under which smiles simply produce reward responses versus embodied simulations of the perceived smile and its associated emotion. If the model is correct, then embodied simulation should underlie the correct interpretations of smiles that have specific meanings.

4.2.2. Developmental considerations. The developmental literature offers further support for a relationship between eye contact and embodied simulation of emotional states. In general, eye contact is very important for
infants. For instance, Farroni et al. (2002) demonstrated that three-day-old infants looked longer at faces with direct gaze, as opposed to simultaneously presented faces with averted gaze. Thus, infants prefer faces that establish eye contact. Furthermore, eye contact is associated with stronger neural processing, as demonstrated by analyses of the infant N170 (Farroni et al. 2002; Farroni et al. 2004). Infants also show enhanced neural processing of angry expressions when these expressions are accompanied by direct eye gaze (Striano et al. 2006). Eye contact has thus been called the “main mode of establishing a communicative context between humans” (Farroni et al. 2002, p. 9602).

Infants’ interest in eyes has been interpreted as an interest in the perceptual features of the eyes per se, and as an interest in direct gaze. Because infants benefit when adults understand their internal states and needs, however, and because they have limited means to express those needs, their interest in eye contact could instead reflect an evolutionary-based mechanism for triggering embodied simulation in caretakers. Such an interpretation receives additional support from research showing that infants use behaviors such as smiling, gurgling, and flirting to achieve eye contact with adults who are not looking at them (Blass & Camp 2001). Infants also engage longer and smile more often at individuals who make eye contact (Hains & Muir 1996; Symons et al. 1998).

4.2.3. Role of the amygdala. An increasingly clear connection between eye contact and amygdala activation further supports this account (Dalton et al. 2005; Fourtois & Vuilleumier 2006). Complete amygdala lesions result in a severe reduction in eye contact during actual conversations (Spezio et al. 2007b). Conversely, as neural activity in the amygdala increases, monitoring for emotional gaze events in others increases (Hooker et al. 2003).

A recent study that recorded electroencephalogram (EEG) and EMG while participants viewed smiles and sad expressions showed a similar effect (Achaibou et al. 2008). Peak EEG activation around 100 msec (P1) was associated with greater correspondent EMG activity slightly later, at around 300 msec. Achaibou et al. interpreted the increased P1 amplitude as a signature of the participant’s attentional state, suggesting that deeper visual processing of the facial expression was associated with enhanced mimicry, as indexed by EMG. Although amygdala activation was not measured, all the findings reviewed here on functions of the amygdala suggest that, when the amygdala becomes active, it directs attention to the eyes, and that resulting eye contact elicits greater or more correspondent facial mimicry. The link between amygdala activation and mimicry further follows from involvement of the dorsal midbrain, which has been linked to increases in arousal produced by direct gaze (Donovan & Leavitt 1980). Although we have assumed conservatively here that amygdala activation increases motor mimicry in smiling, it is also possible that it increases embodied simulation as well, thereby producing stronger emotional responses.

4.3. Perceptual information from the eye region versus eye contact

A potential confound exists between looking at the eyes as a way to extract perceptual information and making eye contact as a possible trigger to simulation. Adolphs and colleagues, for example, reported extensive study of patient S.M., who had a bilateral brain lesion that encompassed all nuclei of the amygdala (Adolphs et al. 2005). In initial assessments, S.M. showed a compromised ability to recognize fearful expressions. In subsequent assessments, this failure was attributed to not using information from the eyes. Failure to look at the eye region during smiles may increase compensatory looking at the mouth region. Because S.M. used information from the mouth normally, she was able to correctly distinguish smiles from fear expressions. More recent findings from the same laboratory found that high-functioning autistic children also avoid gazing at the eye region when performing a facial expressions recognition task (Spezio et al. 2007a). Here, too, the autistic group performed as well as the control group in recognition accuracy (based on impoverished information provided in the Bubbles technique), perhaps reflecting the use of other diagnostic perceptual cues, such as the mouth.

The specific ability to distinguish Duchenne from non-Duchenne smiles has been shown to be compromised in autistic adults, who failed to look at the eyes while making judgments (Boraston et al. 2008). More specifically, this result was interpreted as reflecting failure to use information in the perceived expression provided by contraction of the orbicularis oculi muscle in the vicinity of the eyes. Although these findings have typically been interpreted as showing that information from the eye region carries specific perceptual information, they can also be interpreted as demonstrating the importance of the eye region for triggering embodied simulations. As described in the SIMS, we propose that these findings actually demonstrate that eye contact has evolved as a trigger for embodied simulations, which endow smiles with their different functional meanings.

5. Conceptual knowledge about facial expression

We have already mentioned that individuals possess and use stereotyped knowledge and expectations about the meanings of smiles (Halberstadt et al. 2009; Kironac & Hess 1999). A challenge for the present approach is to suggest how this knowledge is represented and also how it relates to embodied simulations of perceived smiles. As will be seen in later sections, conceptual knowledge of smiles plays a central role in the SIMS model of how people interpret their meanings.

The classic view of emotion concepts relies on general models of representation in the cognitive sciences that view concepts as redescriptions of the input from modality-specific systems into an abstract representation language (e.g., Bower 1981; Johnson-Laird & Oatley 1989; Örtony et al. 1987). According to these accounts, higher-order mental content is represented in an abstract, language-like code that has an amodal character (e.g., Fodor 1975). Notably, an amodal representation does not preserve anything analogical about the perceptual experience of the object, event, or state. Instead it is abstracted away from this experience and transduced into some sort of representation language that is abstract and amodal in format. Whether the resulting representation takes the form of something like a logical structure,
a feature list, or vectors on which different values can be positioned, the assumption is that the multi-modal experience of something and its conceptual representation do not take place in the same system (Barsalou et al. 2003; Niedenthal et al. 2005b).

The social psychology and emotion literatures are filled with evidence, now supported by findings from neuroimaging studies, that is consistent with an embodied or grounded account of conceptual processing (Barsalou 1999; 2005; 2008; Gallesse & Lakoff 2005; Niedenthal 2008). From this perspective, the modality-specific states that represent perception, action, and introspection when one interacts with a particular entity or has a particular subjective experience also represent these same stimuli and events when processed in their absence. For example, retrieving the memory of a landscape involves reactivating (simulating) parts of the visual states that were active while perceiving it. From this perspective, then, having a concept is having the ability to reenact experiencing an instance of the respective category.

As can be seen from these two accounts of concepts, the embodied account provides a natural way to link conceptual knowledge about smiles and the related social situation to the actual perception of smiles. Rather than assuming that a smile activates an amodal knowledge structure to represent its meaning, the embodied simulation account proposes instead that a smile triggers a simulation of a smile experience that includes emotion, motor activity, and somatosensory experience (Niedenthal et al. 2005a). Once this simulation becomes active, it provides a conceptual interpretation of the perceived smile, going beyond the information given to place it in a context. Clearly, much about this account remains to be developed, especially the relevant neural systems and their role in social processing.

The grounded cognition position is important in presenting the SIMS model for a specific reason, namely, because an embodied simulation is not always a response to a currently perceived smile. Additionally, an embodied simulation can also be triggered by the activation of related knowledge or expectations (e.g., stereotypes). For instance, when someone activates conceptual knowledge about infants, this could produce an embodied simulation of an infant. Significantly for the SIMS account to follow, these conceptually triggered simulations may not always be relevant or correct in the current situation, when a different embodied simulation triggered by the actual situation and facial expression offer a more correct interpretation. Rather than mimicking and feeling the reward value of an enjoyment smile present in the current situation, for example, an individual who expects to see a dominance smile may experience negative feelings and interpret the smile as expressing superiority, as the result of unjustified conceptual processing (Halberstadt et al. 2009).

6. The simulation of smiles (SIMS) model

We now present a model that seeks to integrate brain and behavior in accounting for the representation of smile meaning. We focus on the possible meanings of a smile as expressing enjoyment, affiliation, or dominance. These judgments can be seen as having correct answers only in so far as smiles with certain identity, postural, morphological, and dynamic markers have been selected or developed as stimuli. In addition, stimulus development would require validation by reference to the intentions and feelings of the person who is smiling, as well as the demonstration of reasonable consensus in interpretation by perceivers. For example, a set of empirically derived enjoyment smiles would be static or dynamic facial displays of people who were experiencing joy, as indicated by self-report, at the time of the smile.

First, we present the core SIMS model, which establishes how grounding smiles differentially in neural systems and behavior causes them to be interpreted differently as meaning enjoyment, affiliation, or dominance (see Figure 4 in sect. 6.1.1). Then, we propose conditions under which perceptual cues, experiential cues, and conceptual knowledge are used to interpret smiles in these different ways (Figures 5–7; Adolphs 2002; Kirouac & Hess 1999). The SIMS model does not attempt to account for the entire neural circuitry that underlies the processing of emotional facial expressions (as discussed, e.g., in Adolphs 2002; 2006; Heberlein & Atkinson 2009), nor does the SIMS model attempt to account for the details of real-time neural time courses. Instead, the focus is on the conditions under which embodied simulation and other cues are used to arrive at judgments of the three smile meanings. Thus, our account aims to represent cognitive-behavioral function, but not the timing of neural activity.

The SIMS model has been largely developed using data collected in Western countries. Nevertheless, it is essential to note that cultural differences may modulate our account. Some clear predictions across cultures can be imagined and have been articulated elsewhere (Niedenthal et al., in preparation).

6.1. Details of the SIMS model

6.1.1. Core SIMS model. We begin by considering how smiles are interpreted in the most ecologically valid situation, that is, in which the smile has an uncertain meaning (in the sense of being unexpected in the context). First, consider the empirically derived enjoyment smile illustrated in the top panel of Figure 4. As the result of eye contact, a reward experience is generated in the basal ganglia, and motor mimicry is generated in the cortical motor regions (in particular, those described by Schilbach et al. 2008). In turn, these two activations produce corresponding bodily experience in the somatosensory cortex. The top panel of Figure 4 represents this overall pattern. Thus, the most straightforward case of processing a smile’s meaning begins with the detection of uncertainty (producing amygdalar activation), which, in turn, directs gaze toward the eyes (eye contact), followed by the generation of reward, motor mimicry, and corresponding somatosensory experience. As this initial example illustrates, an embodied simulation of enjoyment grounds visual perception of the perceived smile, leading to the interpretation that the person is “happy.”

The middle panel of Figure 4 illustrates the same process for the judgment that a smile is an affiliative smile (e.g., that the smile is understood as “friendly” or “conciliatory”). Many of the relevant brain regions are very similar to those just described for the judgment of
Figure 4. The top panel illustrates the case of an enjoyment smile (A) presented such that the meaning is initially uncertain. The perception of the smile is accompanied by activation in the amygdala (B). Research suggests that amygdala activation enhances the probability that eye contact with the smiling person will be made (C). In the SIMS model, eye contact has evolved as a trigger for embodied simulation. Eye contact thus produces increased activation in the reward centers of the basal ganglia (D1) and in motor regions, described by Schilbach et al. (2008) (D2), that support motor mimicry (E). These motor and limbic processes then produce bodily sensations in somatosensory cortex (F). On the basis of these neural activations and behaviors, the smile is judged as indicating that the smiling individual feels happy (G). The middle panel illustrates the process that results in the judgment of a smile as affiliative. The only difference between the content of the two panels (A – G) is the additional OFC (D3) activation, which in theory supports distinctive positive affect related to attachment. The bottom panel shows the processing of a dominance smile (A'). Amygdala activation would again be expected (B') and eye contact would be predicted to occur (C'). Dominance smiles may be associated with a pattern of asymmetrical neural activation related to withdrawal-related negative affect (e.g., Davidson et al. 1990; D1'). Activation in relevant motor regions (D2') would be expected and output resulting in mimicry (E'). Because of the role of prefrontal cortices in processing status, OFC or contiguous regions may also be involved (D3'). Implications of these supported by somatosensory cortices (F1') will ground a judgment of a smile as a smile of superiority of some type (G').
an enjoyment smile. For an affiliative smile, however, processing would also involve OFC activation and perhaps the involvement of closely related prefrontal regions. As discussed previously, these areas may selectively support the distinctive positive feeling of seeing an individual smile with whom one has a close relationship, as in the finding mentioned earlier when OFC differentiates the sight of one’s own smiling baby from the sight of an unknown smiling baby (e.g., Minagawa-Kawai et al. 2009; Nitschke et al. 2004).

The bottom panel of Figure 4 illustrates the processing of a dominance smile. Again, amygdala activation would be expected to signal the expression’s uncertainty and its potential significance. Here, however, negative rather than rewarding affect may be experienced. Although speculative, we have suggested that dominance smiles, like smiles hiding negative intentions or feelings, are associated with the experience of negative rather than positive affect, as indicated by right-lateralized activation (Boksem et al. 2009; Davidson et al. 1990). Because OFC and contiguous regions are central to processing social status (Zink et al. 2008), ventral frontal cortex may also be involved. As before, activation of cortical motor regions and correspondent mimicry also occur and produce corresponding somatosensory experience associated with the feeling of being dominated (e.g., that the smile is experienced as “superior” or “condescending”).

6.1.2. Behavioral constraints in SIMS. Altering the core SIMS model in principled ways facilitates further understanding of smile processing.

6.1.2.1. Experimental inhibition of behavior. In contrast to the three cases in Figure 4, the two cases in Figure 5 assume that embodied simulation is not likely to occur. An enjoyment smile is used as the example stimulus event. As the top panel of Figure 5 illustrates, no eye contact is achieved with the smiler as was the case for Figure 4. This may occur experimentally when the smile has low uncertainty, as in a blocked design when many...
Figure 6. All three panels illustrate situations in which eye contact does not occur, analogous to the three cases in Figure 4 when eye contact does occur. Although the neural activations are very similar, the suppressed social behavior in Figure 6 detaches any emotional processing that occurs from the perceived smile and renders it knowledge-based. Thus, the emotional content is determined by simulation of conceptual knowledge in memory, rather than being driven by the experience of eye contact (Niedenthal 2008). The top panel illustrates the case in which a smile is believed to be a smile of enjoyment (A). Perception of the smile is accompanied by activation in the amygdala (B). We have defined this as a situation in which eye contact is avoided (C). Nevertheless, because the smile is believed for other reasons to be an enjoyment smile, activation of the reward centers of the basal ganglia (D1) occurs, and also the motor brain regions described by Schilbach et al. (2008) (D2). Correspondent smiling (E) occurs, but is determined by simulation of conceptual knowledge and does not count as mimicry (as in Halberstadt et al. 2009). Conceptual implications are represented by somatosensory cortex (F) and confirmation of the smile as an enjoyment smile (G) is made on this basis. Note, of course, that the judgment could be wrong with regard to the smile intention. The middle panel illustrates the same process that results in the judgment that a smile is an affiliative smile (A’–G’), again without eye contact. The only difference with the top panes is the again additional robust OFC activation (D3’), which in theory supports positive affect related to attachment. The bottom panel shows the processing of a smile believed to be a smile of dominance (A’’), where again amygdala activation simply reflects visual cues from the mouth (B’’) and eye contact does not occur (C’’). Withdrawal-related negative affect (e.g., Davidson et al. 1990; D1’’) and activation in relevant motor regions (D2’’) that support smile production is again expected (E’’), as described earlier for Figure 4. Because of the role of prefrontal cortices in processing social status (Zink et al. 2008), OFC or contiguous regions may also be involved (D3’’). The representation of these inputs in somatosensory cortices (F’’) will serve to confirm the interpretation of the smile as an expression of dominance (G’’).
smiles are presented one after the other. As a result, the rich behavior and brain activity associated with eye contact is absent. Instead, recognition of the expression as a smile only reflects processing visual features of the mouth. The matching of this visual input to perceptual representations of stored smiles is indicated in the top panel of Figure 5 by activation in occipito-temporal cortices (Adolphs 2002; Calder & Young 2005). Further interpretation of the smile’s meaning beyond simple recognition could occur if conceptual knowledge produces expectations about smiling and possibly a simulated smile (as described in further detail later). Because eye contact is not available to trigger embodied simulation, some other source of information must be consulted to provide meaning, should additional meaning be relevant.

The bottom panel depicts a second case in which facial mimicry to a smile is inhibited. Inhibition can occur experimentally when spontaneous expression is blocked by mechanical means (e.g., Oberman et al. 2007). In these situations (when uncertainty is again high), perception of the smile will be associated with amygdala activation. Because motor mimicry in inhibited, however, activation in motor systems and emotion systems will be absent or reduced, although premotor regions may be partially active, reflecting the as-if motor loop that simulates motor activity at a relatively abstract level. Because actual motor activity does not occur however, the smile is not felt in somatosensory areas, and recognition occurs on the basis of matching visual input about the perceived face to stored perceptual representations. Again, interpretation of the smile’s specific meaning based on embodied simulation could occur, but would require the use of conceptual knowledge or expectations, given the absence of embodied simulation that results from eye contact.

6.1.2.2. Social inhibition of behavior. The experimental situations represented in Figure 5 also occur in natural social situations, but for far more interesting reasons. Eye contact and mimicry will sometimes be suppressed when a smile is displayed by an individual with whom exchanging eye contact and mimicry is too intimate, risky, or aversive (Lee et al. 2006; Likowská et al. 2008; Mondillon et al. 2007). Unlike in the artificial experimental context, however, it is likely that embodied simulations will become active and contribute to judging a smile’s meaning. Such simulations, however, will again not be responses to the perceived smile, but will result from activating conceptual knowledge.

Emotions caused by conceptual knowledge can indeed affect the interpretation of a perceived facial expression via embodied emotion simulations. Niedenthal et al. (2000), for example, induced states of happiness, sadness, and neutral emotion in participants, triggered by conceptual knowledge. As participants then viewed movies of faces whose expressions changed from happiness or sadness to neutral emotion, their task was to ascertain when the initial emotional expression on the face ended. The induced emotional state of the participants affected their perception of that offset point. In other words, an emotion-congruence effect occurred between a conceptually activated embodied simulation and a perceptually processed face. Participants in a happy state saw smiles persist longer into the ambiguous neutral zone of the expressions, whereas sad participants saw sadness persist somewhat longer. Emotional traits and styles similarly affect performance on this task, further indicating that conceptual knowledge contributes to the interpreted meaning of a perceived facial expression via embodied simulations of emotion (Feldman-Barrett & Niedenthal 2004; Fraley et al. 2006). Again, conceptual knowledge about the smiler (or about the context in which the smile was encountered) triggered embodied emotion simulations that affected judgments of perceived smiles (e.g., Halberstadt et al. 2009).

Figures 6 and 7 illustrate socially inhibited eye contact and mimicry, respectively. For the former, the components are similar to the analogous ones in Figure 4. Here, however, the suppressed social behavior detaches processing from the stimulus and renders it knowledge-based, with embodied simulation and emotional content being provided by conceptual knowledge in memory for
the respective facial expressions (indicated in portion A of each panel; Niedenthal 2008). Facial expression may occur, but is not necessarily related to the perceived smile, therefore not counting as mimicry. Instead, such facial expressions reflect simulation of emotions associated with conceptual knowledge. As a result, the observer’s smile may not constitute mimicry (this may have occurred in Mojszisz et al. 2006), and may obscure the smile’s functional meaning. An interesting question for future research is to examine the neural basis of using one’s own knowledge to generate an embodied emotion simulation versus using facial cues in perceived expressions, such as eye contact (e.g., Decety & Grézes 2006). In Figure 7, facial mimicry is the behavior that is inhibited. As discussed earlier (sect. 3.2.2.), inhibited mimicry reduces emotional processing. Even when a prior expectation for a smile meaning exists (such as the expectation that the person is expressing an enjoyment smile), emotional processes may still be inhibited because the motor system cannot become engaged as usual. Under these conditions, conceptual knowledge about emotion in memory may be used instead to establish the meaning of the smile, rather than embodied simulation triggered by eye gaze. Unlike the process illustrated in Figure 6, however, perceivers here may rely more on “disembodied” linguistically based associations and social concepts. Brain regions involved in the use of semantic associations depend on the task and the way in which the semantic knowledge is called for (e.g., Barsalou et al. 2008).

6.2. Judgments of smiles as true versus false

One particular distinction between types of smiles has received considerable attention: whether a smile is true or false (e.g., Miles & Johnstone 2007). Asking this question already implies a culture-laden value that smiles can be false per se (Fridlund 2002). Some cultures and some individuals seem to mean by false that there is a desire to deceive or manipulate (Frank 2002). This would be akin to saying that a smile put on to hide a negative intent is a false one. Alternatively, individuals sometimes mean that the smile was produced voluntarily and that this act is itself inauthentic or false (Fridlund 2002). This would be akin to saying that any intentional smile is, by definition, false. Although the definitions of true and false smiles will ultimately have to be bootstrapped, at least by culture, we can propose a working definition here: Judging a smile as a true smile is the normative or default judgment, and it means that the smile signifies positive feelings of some kind (e.g., pleasure or joy) or positive intentions. Conversely, judging a smile as false means that the perceiver believes that the smile was motivated by a desire to hide, moderate, or justify something negative (e.g., a lie, a criticism, a feeling of superiority or contempt, a manipulation, an inappropriate affect).

Rather than a clear dichotomy, the true-false distinction is a continuum that is most often represented in the English language by a value on a scale of “genuineness,” “authenticity,” or “sincerity” (Hess & Kleck 1990). Different measurable features of smiles have been claimed to be associated with these judgments. Indeed, Duchenne smiles are often judged as more genuine than non-Duchenne smiles. Individuals also judge smiles with a slow onset, a slow offset, and a shorter apex as more genuine than smiles that have a sudden onset, a sudden offset, and a long apex (Krumhuber & Kappas 2005). The former smiles are further associated with more positive personality judgments (e.g., Krumhuber et al. 2007). Notably, however, these features do not in any way map isomorphically to the true/false judgment. For instance, as already mentioned, the importance of Duchenne markers for judging smiles in still photos varies with culture (Thibault et al., submitted). In addition, the Duchenne marker is far less important in dynamic as compared to still smile stimuli (e.g., Krumhuber et al. 2007).

The SIMS model predicts when a smile will be judged as more or less genuine and on what basis. For example, enjoyment and affiliative smiles with initially uncertain meaning should always be judged as true smiles. This is because the default judgment of a smile being genuine is based on the presence of positive affect and motor mimicry in an embodied simulation associated with it. In contrast, dominance smiles should be more likely to be judged as false smiles, due to the greater presence of negative affect and withdrawal in their associated simulations (e.g., Davidson et al. 1990).

In contrast, when eye contact is not made, or when mimicry is inhibited by social factors, true/false judgments should be guided by cultural beliefs and stereotypes stored in conceptual knowledge. As a result, true/false judgments may be much less systematic than when eye contact and mimicry both support the embodied simulation.

A recent study by Maringer et al. (in press) supports this proposal. Participants were given empirically derived “true” dynamic smiles or “false” dynamic smiles to see and asked to rate them on scales of genuineness (Krumhuber et al. 2007). Orthogonally crossed with the true-false manipulation, half of the participants were able to freely mimic the smiles, and the remaining half held a pencil in their mouths so as to block facial mimicry. The results were clear: In the mimicry condition, participants who saw true smiles rated them as more genuine than participants who saw false smiles. In the mimicry-blocked condition, participants did not distinguish between the two types of smiles. Thus, the ability to mimic facial expressions was essential for distinguishing true from false smiles, implicating embodied simulation in performing these judgments accurately.

In a second study, participants viewed true smiles and were told either that these smiles occurred in a context in which a true smile would be expected by cultural stereotypes or that these smiles occurred in a context in which a false smile would be expected. Mimicry was blocked in half of the participants, and all participants rated the smiles on genuineness. Results showed that when mimicry could occur, cultural beliefs and stereotypes did not affect the perceived genuineness of the smiles, but when mimicry was blocked, smiles that occurred in contexts associated with true smiles were judged as more genuine than those that occurred in contexts in which false smiles could also occur. Again, embodied simulation appears to be a critical cue for establishing that a smile is genuine.

7. Beyond smiles

The present article and model focus on the smile because it is probably the most complex of all emotional
expressions. Not only are smiles highly diverse in the conditions under which they occur, they are also highly diverse in their possible meanings. Significantly, however, similar embodied processes should support interpreting the expression of other basic emotions (e.g., anger, sadness), as well as the expression of secondary emotions (e.g., shame, embarrassment). In general, the perception of any facial expression should trigger affective, motor, and somatosensory experience that gives the expression meaning. Thus, the structure, but not the content, of the SIMS model can be used analogically to investigate the interpretation of facial expressions besides the smile.

As an illustrative example, consider anger. Expressions of anger are used to establish and maintain social hierarchies, among other social functions (Keltner & Haidt 1999). Anger expressions are slightly different than those of contempt (Ekman & Heider 1988), but may occur in similar situations. Indeed a family of anger-related emotions exists, ranging from frustration to outrage. Given their different social functions, it would be essential for socially competent individuals to correctly experience their subtle meanings in the respective situations (e.g., Fischer & Roseman 2007).

A functional analysis of anger expressions is likely to suggest different social determinants of eye contact and mimicry than those for smiling. For instance, there is evidence that anger expressions are mimicked less to the degree that a situation is social, unless anger is expressed with averted gaze indicating that the source of the anger is not the perceiver (Bourgeois & Hess 2008; Hess et al. 2009a). Furthermore, because anger is statistically less frequent in most environments than the smile, and has different implications for the observer, some neural circuits, such as those involving the amygdala may be activated more often and be less sensitive to environmental determinants of uncertainty. Perceiving and mimicking an anger expression may also have different somatosensory experiences than perceiving and mimicking related emotions, such as frustration, contempt, and outrage. Finally, anger has been linked to neural structures that do not appear relevant for interpreting smiles, including the ventral striatum (Calder et al. 2004) and dopamine receptors (Lawrence et al. 2002).

This example is not intended to be comprehensive. Rather, it simply suggests that the present model can be extended to other emotions, but cannot be extended effectively without due consideration of the specific circumstances, together with the accompanying behavioral and neural responses. We believe, however, that the basic structure of the model, with its emphasis on integrating specific social behaviors with related forms of embodied simulation, is viable in modeling the meaning of all emotional facial expressions.

8. Conclusions

The impact of Darwin’s (1872) writing on facial expression and its centrality in more recent theories of emotion (e.g., Ekman 1989) has pushed the study of emotion expression toward a dictionary of facial muscles and their combinations as signs to internal states. The notion that a physical description has access to disembodied amodal knowledge, which can, in turn, be used to interpret the perceptual input, has been a natural way to think about the processing of emotional facial expressions in the context of the Cognitive Revolution. Given more recent developments in grounded cognition and neuroscience, however, it seems increasingly likely that a construct along the lines of embodied simulation is necessary for a full account of how people establish the meaning of facial expressions. Increasingly, new empirical techniques provide additional ways of measuring embodied simulation and establishing causal evidence for its roles in social processing (e.g., the use of transcranial magnetic stimulation [TMS] in Pitcher et al. 2008). The SIMS model of the smile is an attempt to show how behavioral and neuroimaging findings can be integrated to generate new and productive process models of facial expression more generally. Solving the riddle of the smile will, we believe, provide important groundwork for understanding the full array of emotional facial expressions.

ACKNOWLEDGMENTS

The authors thank Ralph Adolphs, Anthony Atkinson, Markus Brauer, Julie Grézes, and Piotr Winkielman for comments on an earlier draft of the paper. We extend a particularly grateful thanks to Lawrence Barsalou for providing very specific and detailed feedback on multiple drafts. The work of Paula M. Niedenthal and Martial Mermillod was supported by a grant (FaceExpress – Blanc CSD9 2006) from the Agence National de Recherche (ANR), France; the work of Ursula Hess by a grant from the National Science Foundation (NSF) and the Social Sciences and Humanities Research Council of Canada.

Open Peer Commentary

Mimicry and simulation in gesture comprehension

doi:10.1017/S0140525X10001445

Martha W. Alibali and Autumn B. Hostetter*

*Department of Psychology, University of Wisconsin—Madison, Madison, WI 53706; ‡Department of Psychology, Kalamazoo College, Kalamazoo, MI 49006.

mwalibali@wisc.edu Autumn.Hostetter@kzoo.edu

http://psych.wisc.edu/alibali/home.html

Abstract: According to the SIMS model, mimicry and simulation contribute to perceivers’ understanding of smiles. We argue that similar mechanisms are involved in comprehending the hand gestures that people produce when speaking. Viewing gestures may elicit overt mimicry, or may evoke corresponding simulations in the minds of addressees. These real or simulated actions contribute to addressees’ comprehension of speakers’ gestures.

In their target article, Niedenthal et al. present the SIMS model, which provides a theoretical account of how embodied simulation is involved in interpreting smiles. We suggest that mechanisms like those proposed in the SIMS model are also involved in interpreting the spontaneous hand gestures that people produce when speaking. In past work, we have argued that gesture production is based on simulated actions and perceptions (Hostetter & Alibali 2008). In this commentary, we make the
case that embodied simulation is also involved in gesture comprehension.

The Gesture as Simulated Action (GSA) framework (Hostetter & Alibali 2008; 2010) explains how representational gestures might arise from an embodied cognitive system. Aligned with theories of embodied cognition (e.g., Barsalou 1999; Glenberg 1997), the GSA framework holds that language processing activates perceptual and motor states. Speakers form mental images of information they are expressing; these mental images activate the same perceptual and motor systems that are involved in perceiving and interacting with physical objects in the world. This activation of the motor system is sometimes overtly expressed as representational gestures.

Here, we consider whether simulated actions could also be involved in gesture comprehension. Niedenthal et al. argue that when people perceive smiles, they engage their motor systems, either to mimic the perceived smiles or to simulate the experience of perceiving the perceived smiles. This mimicry or simulation contributes to perceivers' understanding of the perceived smiles.

We suggest that the same mechanisms—mimicry and simulation—might also be involved in comprehending representational gestures. Imagine a speaker describing how he makes pizza, who uses gestures to depict how he uses his fingertips and his kneckles to flatten the dough. Imagine further that the speaker's addressee has never made pizza crust before. In comprehending the speaker’s description and gleaning information from his gestures, the addressee might overtly mimic the speaker’s actions, or she might mentally simulate such actions.

Do speakers overtly mimic other speakers’ gestures? A growing body of evidence indicates that they do (e.g., Kimbara 2006; 2008; McNeill, in press). Mimicry does not arise as a simple coincidence of people using similar gestures when they talk about similar things. Instead, similarity of gesture form is more likely when interlocutors can see one another than when they cannot. Suggesting that mimicry is purposeful (Kimbara, 2008). McNeill (in press) observes that researchers who code gestures often use mimicry as part of their effort to make sense of a speaker’s gestures. He argues that addressees use mimicry to understand gestures that may not be readily interpretable otherwise.

Although overt mimicry of gestures does occur, it is not as widespread as might be expected if it were the primary mechanism of gesture comprehension. Recall that Niedenthal et al. emphasize that smile perceivers sometimes simulate the experience of the perceived smile, without overtly mimicking it. Along similar lines, it seems likely that gesture comprehenders sometimes simulate the gestures they perceive, without overtly producing them. These embodied simulations of gestural actions may contribute to comprehenders’ interpretation of the perceived gestures. Support for this mechanism comes from studies that have shown activation of premotor areas when people observe actions (for reviews, see Jeannerod 2001; Rizzolatti et al. 2001). Premotor cortex is also activated when people observe gestures (e.g., Montgomery et al. 2007), and this activation is modulated by semantic information from the accompanying speech (Willems et al. 2007). It has been argued that, when addressees perceive gestures, they interpret them, at least in part, by activating the cortical networks involved in producing gestures, via an “observation-execution matching system” (e.g., Holle et al. 2008; see Dick et al. 2009, for discussion). This idea is similar to the one we advance here.

The SIMS model holds that, in the case of smiles, eye contact triggers embodied simulation. It is possible that eye contact may also initiate embodied simulation of the co-occurring gestures; however, no studies to date have specifically examined the relation between eye contact and gesture comprehension. Instead, the little research that has explored the role of eye gaze in gesture comprehension suggests that addressees are most likely to attend to speakers’ gestures when speakers direct their gaze to the gestures (e.g., Gullberg & Holmqvist 1999). Further, addressees are more likely to incorporate information from the gesture into their comprehension when speakers have fixated on the gestures (Gullberg & Kita 2009).

One explanation for these findings is that speakers’ gaze to their own gestures signals that the gestures are important; addressees then directly attend to those gestures and run “as-if” simulations of the gestures that facilitate comprehension. Speakers’ shift in eye gaze may be particularly important for initiating addressees’ simulations of speech-accompanying gestures, because addressees typically focus on speakers’ faces, making it difficult to attend to detailed features of gestures produced in peripheral space.

Could mimicry and simulation account for individual and situational differences in addressees’ reliance on gesture for comprehension? Research has shown that gestures make a greater contribution to comprehension when the material to be comprehended is complex, ambiguous, or challenging relative to the addressee’s skills (Graham & Heywood 1976; McNeill et al. 2000). It also seems likely that individuals differ in their reliance on gesture for comprehension, although there is little data directly addressing this point. We suggest that addressees vary in how likely they are to engage in overt mimicry and/or simulation. Across contexts, individuals may be more likely to engage in mimicry and/or simulation in situations where language comprehension is more difficult, as suggested by McNeill (in press).

In sum, building on our past work about simulated action as an explanation for why speakers produce gestures, we suggest that similar mechanisms may be involved in comprehension of gestures. Viewing a speaker’s overt simulation of an action or event may elicit overt mimicry, or it may evoke a corresponding simulation in the addressee’s mind. In the SIMS model, mimicry and simulation contribute to perceivers’ understanding of smiles. We suggest that mimicry and simulation also contribute to addressees’ comprehension of speakers’ gestures.

Emotion simulation and expression understanding: A case for time

doi:10.1017/S0140525X10001457

Marian Stewart Bartlett
Institute for Neural Computation, University of California, San Diego, 92093-0440.
marni@salk.edu http://mplab.ucsd.edu/~marni

Abstract: Niedenthal et al. present a model for embodied emotion simulation and expression understanding that spans multiple brain systems. This commentary addresses the potential role of time in this model, and its implications for understanding social dysfunction.

Niedenthal et al. diverge from an increasing trend for localizationism in cognitive neuroscience, and present a model of emotional expression and understanding that spans the somatosensory system, motor system, reward system, amygdala, basal ganglia, and prefrontal cortex, as well as systems involved in eye gaze and body posture. This model enfoldes ideas relating to mirror neurons and the link between perception and production, but as a subset of a larger system for understanding emotions in others.

Not only is this model of emotion unlocalized in space, it is also not localized in time. The systems involved must communicate and pass activation through time in order to fully engage this emotion system, akin to a closed-loop dynamic process. It may be crucial that the timing of the interactions of these
“Smile down the phone”: Extending the effects of smiles to vocal social interactions

doi:10.1017/S0140525X10001469

Frédéric Bassoa and Olivier Oullierb

a University of Rennes 1, CREM CNRS UMR 6211, IGR-IAE de Rennes 1, 35000 Rennes, France; bLaboratoire de Psychologie Cognitive (UMR 6146), Université de Provence & CNRS, Aix-Marseille Université, 13331 Marseille Cedex3, France.

frederic.basso@univ-rennes1.fr olivier@oullier.fr
http://www.oullier.fr

Abstract: The SIMS model offers an embodied perspective to cognition and behaviour that can be applied to organizational studies. This model enriches behavioural and brain research conducted by social scientists on emotional work (also known as emotional labour) by including the key role played by body-related aspects in interpersonal exchanges. Nevertheless, one could also study a more vocal aspect to smiling as illustrated by the development of “smile down the phone” strategies in organizations (e.g., call centres). Because of the absence of face-to-face interactions during phone conversations, those “vocal smiles,” including the tone of voice, constitute one of the keys to understanding emotional work in call centres (Belt et al. 2002; Taylor & Bain 1999). These centres are the illustration of the marketing logic known as Customer Relationship Management, aiming at developing long-term relationships between companies and their customers (Gans et al. 2003).

This managerial framework clearly constitutes an expression of emotional work centered around smiles down the phone to create empathy between the client and the sales representative (Richardson & Howcroft 2006). It appears that this vocal aspect to smiling is not considered in Niedenthal et al.’s SIMS. For instance, they focus on face-to-face interactions and do not take into account, so far, how efficient smiles can be in voice-to-voice ones.

Previous studies on emotional work complement the perspective offered by SIMS, especially concerning affiliative smiles (see sects. 2.2 and 6.1.1 of the target article). Hence, we suggest gathering face-to-face and voice-to-voice interactions in an embodied perspective – that is, still in opposition to an “amodal” view of knowledge (Lakoff & Johnson 1999). In our view, Barsalou’s (1999) simulator would be enhanced by Lakoff and Johnson’s theory of conceptual metaphors for it would ground SIMS in natural language – although this theory is not limited to the study of words.

Emotional work sheds new light on the metaphors used by actors in organizations. According to Froggatt (1998, p.332), the metaphorical language employed by nurses (e.g., draining and burden) reflects the emotional aspect of their work. Along the same line, Rees et al. (2007) argue for the existence of metaphors in describing the relationship between patients and physicians. Some therefore consider the emotional work of the physician towards the patient as a “metaphoric framework of clinical empathy” (Larson & Yao 2005, p. 1104). Here too, the challenge of creating empathy with others can be satisfied by using metaphors given that empathic processes at stake are at the core of the emotional work of the physician (Larson & Yao 2005).

Hence, metaphors are a way to access the emotional work of the actors in that the expression of emotions is metaphorical by essence. In line with Lakoff and Johnson (1980), Hochschild (2005, p. 344) stresses how metaphors “guide feeling, and, of course, feelings also guide metaphors.” Similarly, when addressing emotional concepts, Lakoff (1987, p. 377) points that: “When we act on our emotions, we act not only on the basis of feeling but also on the basis of that understanding. Emotional concepts are thus very clear examples of concepts that are abstract and yet have an obvious basis in bodily experience.”

Overall, it is noteworthy in this context that the three key elements used in SIMS to decrypt what lies beneath smiling face-to-face (perceptual cues, experiential cues, and conceptual knowledge; see sect. 6) can be found in voice-to-voice contexts.
In our view, perceptual cues can be embedded in tone of voice (see, e.g., Morris & Keltner [2000] about aggressive behaviours in negotiation) and experiential ones in the words employed (e.g., Rafaeli & Sutton 1990). Despite the absence of eye contact, other components of emotional labour (greeting, thanking, and smiling; Sutton & Rafaeli 1985) are clearly observable in voice-to-voice interactions.

Conceptual metaphors being essential to express emotions (Kövecses 1990; 2008; Lakoff 1987; 1993), a conceptual knowledge of emotions is metaphorical. Of course, a conceptual metaphor such as FACE STANDS FOR FEELINGS (Yu 2008, p. 251; see also, Lakoff 1993, pp. 34–35) illustrates the importance of face-to-face interactions in emotions. But smiles participate to the bodily experience of emotional concepts (e.g., smile and happiness, Lakoff & Johnson 1980, p. 18; see also, Peña Cervel 2001, p. 258). To a larger extent, *The body is a container for the emotions* (Lakoff 1987, p. 383).

Conceptually, the link between the metaphorical approach to emotions and emotional work becomes tighter when one considers the role of relationship metaphors in voice-to-voice interactions. Although one has to be careful not to conclude that the emotional and the human relationship domains can be confounded, human relations are conceptually close to that of emotions (Kövecses 2008, p. 387).

Hence, regardless the modality (visual or vocal), the perception and the understanding of a smile remain embodied. Moreover, the emotional work perspective provides another way to identify “whether a smile is true or false” (sect. 6.2).

For if the issue of emotional work is to make the sales representative’s smile sincere to satisfy the consumer (Grandey et al. 2005), this smile is not always perceived as such. The employee is not necessarily experiencing the emotion conveyed by the expressed smile (Pugh 2001) as it sometimes seems insincere (and therefore false; sect. 2.1) in the customer’s eyes. Over the phone, an insincere tone of voice can even lead to a “phone rage” of the customers, i.e., “people losing their temper on the telephone” (see Deery & Kinnie 2002, p. 8, reporting a study conducted in 1997 by Reed Employment Services on more than 500 organizations in the United Kingdom).

In summary, emotional work permits one to introduce concepts like *sentimental work* (how people use their own emotions to influence someone else’s emotions; Zapf 2002) and *emotional dissonance* (how people experience the gap between emotions that are felt and those which are displayed in a given situation; Hochschild 1983, p. 90). Together with SIMS, emotional work therefore leads to a more complete approach to smiling that encapsulates the vocal modality.

**Expressive smiles or leucosignals?**

doi:10.1017/S0140525X10001627

**Paul Bouissac**

French Linguistics, University of Toronto, Toronto, Ontario M5S 1K7, Canada.

[Email address: paul.bouissac@gmail.com](mailto:paul.bouissac@gmail.com)

**Abstract:** The assumption that a complex and fuzzy notion like *smile* can be the basis of a scientific, rather than semantic, inquiry can only lead to confused and inconclusive results. It would be more productive to start with the well-defined and measurable patterns of the clearly visible contrasts that are produced on the human face by various muscular contractions around the white patches formed by the sclera and the teeth. These features are universal, whereas a common word, in whatever language, is necessarily ambiguous, culture-dependent, and historically rather than biologically determined.

To investigate facial signaling, the best strategy is to focus on the visual features that are obviously adaptive in a species in which territoriality and ranking are prominent. What is adaptive in facial signaling is the correct reading of conspecifics’ intentions at a safe distance. The signals must be robust if they are to be adaptive for both the emitter and the receiver. One cannot afford to hesitate whether an approaching outsider or a group member is friend or foe. But jumping too fast to the defensive in the presence of ambiguity can be equally ill-adaptive in species depending on cooperation which must have evolved unambiguous peace-making signals. Natural selection has favored signals based on the chronic opposition: white versus any color of the spectrum. White ensures optimal reflectance even in reduced luminosity; the color upon which white patterns are produced by muscular contractions can vary indefinitely under other evolutionary forces, from camouflage to ostentatious handicap. Chemical and acoustic signaling is efficient in some environments, but, for open space species, chromatic signaling is the medium of choice, as it covers at the speed of light the distance required for fight or flight decisions. Tigers sport tufts of white hairs on the back of their ears, which they twitch to produce white flashes toward approaching conspecifics to warn them of their alertness to the situation. Similar contrasts are exploited in the baring of teeth in canines and in the permanent status display of the tusks in elephants. The rich muscular system of the face in the primate makes it possible to vary the chromatic patterns, thus generating a range of *leucosignals* (from Greek *leukos* = brilliant white).

The human face makes constant use of the contrasts between the white patches of the sclera and the teeth, and the color of the facial skin (Bouissac 2001; 2005). These signals are well-defined, measurable, and subject to experimental manipulations. Muscular contractions create typical configurations of highly visible white patterns, which combine information pertaining to the gaze with information on the status of the jaws (relaxed vs. tense or clenched). The foldings of the skin can be relevant indicators, but can be only visible at a much closer range than the salient leucosignals and their well-defined morphologies. The range of data processed by Niedenthal et al. encompasses close range perception, while an inquiry into smiles as leucosignals is congruent with interaction at greater distances. The primate brain’s face detectors are fast processors that must lead to instant decisions based on unambiguous signals (Tsao & Livingstone 2008).

It can be predicted, as Niedenthal et al. show, that leucosignals involve the amygdala, since previous evidence indicates that its activation correlates with processing social signals such as facial fear-generating information, including threats linked to social ranking (Buchanan et al. 2009, pp. 289–318). Neurons in the amygdala are indeed responsive to faces, particularly in a hierarchical group (Leonard et al. 1985; Whalen et al. 2009), in which the white of the eyes and the teeth form distinct visual parameters (compare the ratio holding between sclera, iris, and eyelids in the fear face with the same ratio in the angry face). In all social species, the most dangerous enemies are conspecifics. Other predators of humans can be more easily manipulated or controlled, because their semiotic mismatch offers more windows of opportunity once it is deciphered. But in the human primates, peace-making is equally vital. It could thus also be predicted that the same neuronal systems would instantly switch from high alert to pleasure arousal or, at least, be processed by contiguous neuronal systems. Indeed, nothing is more threatening for a human than a new face on which intentions cannot be read. The production of friendly leucosignals may suddenly defuse the tense state of fight readiness and trigger the dopamine flow that stimulates the reward centers (Schultz 2007). All this is consistent with the data that are marshalled in Niedenthal et al.’s article without having to hypothetically construct the cognitive and emotional embodiment detailed by Niedenthal (2007). This is not to deny that such embodiment may occur as a secondary effect with some adaptive consequences, but simply to point out that it is too costly in time and energy to be the primary source of vital decisions such as to
trust, fight, or flee. Leucosignals can provide the fast and frugal information that triggers adaptive behavior without having to be emotionally processed.

Building from the basis of leucosignals, a scientific inquiry of the many facial configurations which are covered in English by the term smile would avoid getting conceptually entangled in the quagmire of semantic and literary interpretations, whether embodied or not. Niedenthal et al. rightly point out that Ekman (2001) suggests that there might be as many as fifty kinds of smiles. The enigmatic or ironic smile (e.g., Mona Lisa with leucosignals reduced to the very minimum), the so-called Duchesne smile, the seductive smile (e.g., Barack Obama showing only his upper teeth), and the provocative triumphal smile in the face of adversity (e.g., Tony Blair showing both upper and lower teeth), to name only a few examples from contemporary icons, are most likely governed by neuro-behavioral systems that have evolved separately under a variety of selective constraints, and have in common only the fact that they end up on the display board of the face in the form of leucosignals whose meanings depend on the raw social context in which they occur. Whether they should all be labelled with the same term is highly questionable. Some are obviously gestures aimed at influencing conspecifics; others are probably mere leakages of a dopamine flooded inner state; and still others are likely the results of various interferences of any of these.

However, dealing with smiles from a purely biological evolutionary point of view definitely would miss the gene–culture coevolution, which cannot have failed to impact this most important tool of sociality. This latter dimension deserves more scrutiny than the lip service paid by Niedenthal et al. to the cross-cultural investigation of smiles.

The role of embodied change in perceiving and processing facial expressions of others

doi:10.1017/S0140525X10001639

Pablo Brin˜ol,a Kenneth G. DeMarree,b and K. Rachelle Smithb

aDepartment of Psychology, Universidad Autónoma de Madrid, 28049 Madrid, Spain; bDepartment of Psychology, Texas Tech University, Lubbock, TX 79409.

Abstract: The embodied simulation of smiles involves motor activity that often changes the perceivers’ own emotional experience (e.g., smiling can make us feel happy). Although Niedenthal et al. mention this possibility, the psychological processes by which embodiment changes emotions and their consequences for processing other emotions are not discussed in the target article’s review. We argue that understanding the processes initiated by embodiment is important for a complete understanding of the effects of embodiment on emotion perception. For example, smiling often makes us feel happier and can increase our liking of everything around us. In accord with multi-process theories of judgment, such as the elaboration likelihood model (Petty & Cacioppo 1986), smiling and other positive behaviors, such as head nodding, can influence judgments and emotions by serving as a simple positive cue, or by affecting either the amount (less thought) or direction (biased positive thoughts) of thinking (for a review, see Brin˜ol & Petty 2008). If smiling reduces thinking, then embodied mimicry could influence subsequent judgments by reducing the discrimination between true and false smiles or between different types of smiles. Mimicking a smile might also facilitate the processing of other smiles when it makes us feel happier by biasing our thoughts or serving as a positive cue, whereas the very same action might increase the difficulty in processing other, subsequent smiling faces when it operates by reducing our amount of thinking.

As these examples illustrate, understanding the processes underlying embodied perception and change are essential to fully understand whether, when, and how individuals process emotional information. Recently, we have proposed that embodiment can not only influence what people think, but can also impact what people think about their own thoughts (i.e., a meta-cognitive process called self-validation; Petty et al. 2002). Meta-cognition refers to thoughts about thoughts (see Brin˜ol & DeMarree, in press, for a review). The main idea behind the concept of embodied validation is that people’s own behaviors can impact their judgments by affecting thought confidence. The confidence that emerges from behavior can magnify the effects of any available mental contents (Brin˜ol & Petty 2003). In a recent study of embodied validation, Brin˜ol et al. (2009) asked participants to think about and write down their best or worse qualities while they were sitting with their back erect and while pushing their chest out (confident posture) or slouched forward with their back curved (doubt posture). Then participants reported their self-esteem. In line with the self-validation hypothesis, thoughts generated about the self only affected self-attitudes in the relatively high confidence posture. Conceptually similar, DeMarree et al. (2010) conducted research in which participants were presented with different facial expressions of emotion on a computer screen after generating positive or negative thoughts. Consistent with the notion that merely perceiving facial expressions of emotion can influence thought-confidence, people relied on their thoughts more when exposed to facial expressions depicting emotions associated with confidence (e.g., happiness, anger) than when exposed to facial expressions depicting emotions associated with doubt (e.g., sadness, surprise). Importantly, this research suggests that smiling can lead to negative feelings when it increases the confidence in previously generated negative thoughts. If smiling validates negative thoughts making people feel bad, it might produce a number of different consequences for processing subsequent smiles (e.g., introducing a negative bias, increasing attention and therefore discrimination between different types of smiles).

Furthermore, this research suggests that the confidence constraint associated with smiles can be important, because happy and angry emotions produce similar outcomes by validating thoughts. Importantly, although the types of smiles discussed by Niedenthal and colleagues all have in common the presence of a positive valence, they might differ in their associated level of confidence, and these variations in confidence can also moderate some of the results described in the review. For example, smiles associated with more confidence (e.g., dominance smiles) might be perceived more easily than less confident...
Beyond smiles: The impact of culture and race in embodying and decoding facial expressions

doi:10.1017/S0140525X10001470

Roberto Caldara
Centre for Cognitive Neuroimaging (CCNi), Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, G12 8QF, United Kingdom; and Département de Psychologie, Université de Fribourg, 1700 Fribourg, Switzerland.

r.caldara@psy.gla.ac.uk

http://www.psy.gla.ac.uk/~roberto/

Abstract: Understanding the very nature of the smile with an integrative approach and a novel model is a fertile ground for a new theoretical mechanism and insights. However, from this perspective, I challenge the authors to integrate culture and race in their model, because both factors would impact upon the embodying and decoding of facial expressions.

The central idea of this stimulating target article is that embodied simulation represents a critical feature to effectively decode the nature of a smile. I endorse the argument and also agree with Niedenthal et al. that the integrative approach they propose is timely and necessary to understand the decoding of smiles (and facial expressions). The model posited here by the authors summarizes their view, offering a novel vision and prompting many questions for future research. Niedenthal et al. state that:

'The SIMS model has been largely developed using data collected in Western countries. Nevertheless, it is essential to note that cultural differences may modulate our account. Some clear predictions across cultures can be imagined and have been articulated elsewhere.' (sect. 6, para. 13)

The authors cite their ongoing work in this regard (Niedenthal et al., in preparation). However, I think it is necessary that they clarify their view regarding the impact of culture and race from their theoretical position in the target article.

Human populations are marked by a variety of practices, beliefs, norms, and forms of organization. The term "culture" is typically used to describe the particular behaviors and beliefs that characterize a social or ethnic group, often located in a particular geographical location. Thus, by definition, culture represents a powerful deterministic and invisible force, which is responsible for shaping the way people think and behave. Importantly, culture also shapes the way people express their emotion. Westerners live in individualistic societies and tend to express their emotions explicitly. By contrast, Easterners live in collectivist societies and tend to suppress their emotions to maintain harmony amongst the group (Markus & Kitayama 1991). More importantly, it has recently been shown that culture impacts not only upon visual perception, but also, critically, upon the extraction of information from faces.

My colleagues and I (Blais et al. 2008) have recently shown that culture shapes how people deploy eye movements to sample information from the facial input space. Western Caucasian adults employ a well-established triangular fixation pattern during face learning, recognition, and categorization by race (Blais et al. 2008). Contrary to intuition, East Asians direct fixations to the central area of the face, around the nose, for the very same visual categorization tasks, while reaching a comparable behavioral performance. The eye movement strategy deployed by the East Asian observers might not straightforwardly relate to gaze avoidance, since this cultural contrast in eye movement generalizes to other visually homogenous categories (Kelly et al. 2010) and is abolished in constrained viewing conditions (Caldara et al. 2010). However, these observations do not rule out the possibility that the eye movement strategy used by Easterners to process visually homogenous objects, might arise from facing a continuous pressure from the Eastern cultural norm promoting gaze avoidance during human face-to-face interactions (Knapp & Hall 2005).

In stark contrast with those previous findings in face recognition, East Asian observers, compared to Western Caucasian observers, oversample information from the eye region during the categorization of facial expressions (Jack et al. 2009), ignoring information from the mouth region, even for the "happy" expression. Besides showing that the eye movement strategies deployed to decode facial expressions are culturally specific, our data also suggest that the transmission of facial expressions is not universal. Easterners focus uniquely on the eye region to decode signals, as they might expect expressive diagnostic signals coming from this facial region. Indirect evidence for this position is provided by the emoticons. While Westerners use a change in the mouth to convey a change from "happy" to "sad" with emoticons (i.e., :-) vs. :-( ), Easterners emphasize information from the eye region (i.e., ^_^ vs. ;_; ), with a neutral mouth. Critically, this perceptual bias towards the eyes for the Easterners generates confusions in facial expression decoding, particularly for the expressions of surprise, fear, anger, and disgust. Therefore, our eye movement data on face recognition showing the lack of direct gaze contact in Easterners, coupled with those on expressions showing that face expressive signals are culturally tuned, point to culturally specific decoding errors and perhaps the necessity of a culturally specific SIMS model. The confusion shown by East Asian observers in decoding particular facial expressions might arise from a lack of embodiment of Westerner expression signals, offering a novel interpretation to the cross-cultural impairment in facial expression decoding.

An interesting side note on this point: East Asian observers should potentially be better than Westerners in categorizing fake and genuine smiles, as they concentrate (and perhaps develop a particular expertise for) their fixations on the eyes only.

Human populations are also marked by physiognomic variations. Race is a universal, socially constructed concept used to rapidly categorize humans originating from different geographical locations by some physiognomic variations (i.e., skin tone, eye shape, etc.). Humans are markedly better at recognizing same- compared to other-race faces. Our studies advocate the existence of finely tuned mechanisms to process same-race (SR) faces, probably developed as a by-product of visual experience (e.g., Caldara & Abdi 2006; Michel et al. 2006a; 2006b), which drive the rapid categorization of other-race faces (e.g.,
Honest smiles as a costly signal in social exchange

doi:10.1017/S0140525X10010287

Samuele Centorrino,a Elodie Djemai,a Astrid Hopfensitz,a Manfred Milinski,a and Paul Seabrightb

aToulouse School of Economics, 31000 Toulouse, France; bMax Planck Institute for Evolutionary Biology, 24306 Plien, Germany

Abstract: Smiling can be interpreted as a costly signal of future benefits from cooperation between the individual smiling and the individual to whom the smile is directed. The target article by Niedenthal et al. gives little attention to the possible mechanisms by which smiling may have evolved. In our view, there are strong reasons to think that smiling has the key characteristics of a costly signal.

When considering the role of smiles in social interactions, it is important to consider situations involving incomplete information about our partner in the interaction. In many social interactions uncertainty typically exists about the extent of cooperation to be expected from the interaction, and there may also be uncertainty about the stakes of the interaction and the possible choices available to the partner. A self-interested individual will never trust her partner if she believes that the partner will act selfishly in a situation that calls for cooperation. Meanwhile, mutual trust among a group of individuals might be socially beneficial, a situation typical for public goods. Given that trust and cooperation are widely observed in humans and other animals, we have to ask which signals might be used to identify trustworthy partners. Smiling may be just such a signal.

The evolution of such signals is not easy to explain. If trustworthy individuals have observable characteristics that distinguish them from non-trustworthy individuals, these signals must have an inherent cost that make them impossible or difficult to fake. Because a non-trustworthy individual would profit from interacting with a trustworthy individual, he has a large incentive to fake the signal. As has been known since the work of Zahavi (1975) in biology and Spence (1973) in economics, a signal can remain reliable in the face of incentives for mimicry only if its inherent cost is larger than this potential gain to the mimic – but it must still be a cost worth bearing for a genuine bearer of the signaled trait. Since trustworthiness might depend on the situation at hand, it seems unlikely that a fixed physiological characteristic would be an appropriate signal, but it is likely that behaviors revealing internal emotional and psychological states could be effective. Smiles are likely to be a signal identifying possibly valuable interaction partners via their correlation with psychological states that reliably indicate trustworthiness.

The information content of the signal might concern either personal characteristics of the partner (the degree of altruism or strong reciprocity) or unobservable characteristics of the situation (possible gains to be made or the choice options available to the partner). There is a literature on smiling as a signaling device that has traditionally focused mainly on the former, but both might be equally important and influential.

The idea that emotions can serve as commitment and signaling devices has been developed in economics since Frank (1988) and Hirshleifer (1987). In particular, negative emotions can trigger costly punishment behavior of non-cooperative actions (Fehr & Gächter 2000). Smiling has been identified as a signal interpreted as representing cooperative tendencies (Krumhuber et al. 2007). When participants in an economic trust game are presented with pictures showing their partner either smiling or with a neutral face, they are more likely to trust the smiling partner. Whether this is an adaptive strategy is a separate question. In particular, the reliability of the picture as a guide to the partner’s behavior might depend on when the picture was taken. Photos taken independently from the decision situation seem to contain little information, while photos taken when the partner was in the process of the relevant decision seem to contain nonverbal cues that allow subjects to deduce the other players’ strategy. Indeed, it is not smiles as such, but rather smiles that are perceived as “honest” that are conducive to trust. “Genuine” or Duchenne smiles are reportedly difficult to fake and most easy to produce when the person is in a positive and sharing mood (Mehu et al. 2007).

Indeed, lying is known to induce a significant level of cognitive load on the liar, which makes it difficult to react spontaneously, and which reduces subtle movement (Vrij et al. 2006). Honest smiling, which involves a contraction of the orbicularis oculi, further has the obvious cost of reducing the smiler’s visual field. Doing so might have fitness costs for the smiler, and thus should be done only sparingly, which can explain why smiling convincingly is cognitively costly. Genuine smiles can therefore be taken as an honest signal of the partner being in a positive mood, one that makes deception of the interaction partner less likely. In a situation concerning trust, this might mean that the trustee has already made a trustworthy decision, or that he is aware that he has something interesting to offer to the partner.

We can relate this idea to the smile types identified by Niedenthal et al. “Enjoyment smiles” are smiles elicited by a general positive atmosphere, and which are seen as honest signals that the partner feels comfortable in the current situation. “Affiliative smiles” are smiles that are used as a signal to the partner. Niedenthal et al. claim that affiliative smiles will be initially considered as “true smiles” (section 6.2) and as “false” when: “the perceiver believes that the smile was motivated by a desire to hide, moderate, or justify something negative” (section 6.2, para. 1, emphasis ours). We believe that it will depend partly on the specific situation of the smile whether the smile is perceived as false or not. Whether an affiliative smile has characteristics that lead it to be judged true or false is therefore influenced not only by the situational knowledge of the observer, but also by unobservable characteristics of the smiler. Therefore, the smile can be seen as a valuable signal of this unobservable information.

Indeed, we have observed in a recent experimental study (Centorrino et al. 2010) that honest smiles lead to higher levels of trust, and that a smile is rated as more convincing when the trustee has larger stakes at his disposition. Therefore, honest smiles can be seen as a costly signal that can only be produced when in a situation potentially beneficial to the observer of the smile, and can thus serve as a signal valuable to the observer and to the smiler alike.
Eyes, amygdala, and other models of face processing: Questions for the SIMS model

Bhismadev Chakrabarti
School of Psychology and Clinical Language Sciences, University of Reading, Whiteknights, Reading RG6 6AL, United Kingdom.
b.chakrabarti@reading.ac.uk http://sites.google.com/site/bhismachakra

Abstract: This commentary raises general questions about the parsimony and generalizability of the SIMS model, before interrogating the specific roles that the amygdala and eye contact play in it. Additionally, this situates the SIMS model alongside another model of facial expression processing, with a view to incorporating individual differences in emotion perception.

The target article proposes a new model for processing smiles, as a specific case within the broader domain of facial expressions of emotion. Niedenthal et al. suggest that three types of smiles – dominance, affiliative, and enjoyment – are distinguished from one another using an embodied simulation (ES) based framework. ES, according to the authors, can be driven both by eye contact as well as by activation of conceptual stereotypes or memory. This work extends current simulation-theory based models of facial expression perception (Goldman & Sripada 2005), by suggesting a key role for eye contact in triggering ES. This is a novel suggestion, and one that will be examined later in this commentary.

In order to understand the SIMS model, it is important to first evaluate the exact nature of the differences between the three kinds of smiles. The authors point out that Duchenne markers are not sufficient to discriminate between these three smiles, and it is unclear whether there are physical differences in facial musculature among these expressions (e.g., as coded for by facial action units; Ekman & Friesen 1978). It is also not clear if this has been tested empirically, using a standard emotion recognition paradigm, where facial expressions are presented with a view to incorporating individual differences in emotion perception.

The three types of smiles that the authors suggest can be regarded as examples of smiles in three among many possible contexts. Indeed, Bharata’s treatise on the dramatic arts (Nātyaśāstra, c. 200 A.D.) discusses at least six different kinds of smiles and more than eight different elicitors for them (Vatsyayan 1996). More recent work suggests at least 30 different emotion states to be linked with smiles, at varying levels of social complexity (Golan & Baron-Cohen 2006). In view of this, one could think of emotion recognition as a two-component process, consisting of (a) perception of the physical properties of the face, and (b) processing of the contextual information (including conceptual stereotypes), which together lead to recognizing an emotion expression. These components might be processed in parallel. The question for the proposed model in its current form, therefore, is: Does it offer the most parsimonious explanation for this two-component process? Or would this model require additional “modules” in order to explain processing of expressions other than smiles, or smiles of a type that has not been explicitly discussed in the model (e.g., a mischievous or a derisive smile)?

Keeping questions about the parsimony of the model to one side, the current form of the model raises two questions about the proposed role of eye contact. The first is anatomical: Is there any support for the proposition that the amygdala promotes eye contact? While it is known that eye contact often results in amygdala activation in the observer, the evidence in support of the suggested causal link in the opposite direction appears limited. Hooker et al. (2003) reported higher amygdala activation for trials with no direct gaze compared to those with a 40 percent probability of a direct gaze. It is not clear if this reflected a generalised increase in arousal for monitoring any salient event, or whether it was specific to monitoring eye contact. Amygdala response in people with Autism Spectrum Conditions (ASC) was found to correlate positively with fixation on the eyes region of the stimuli, but this is possibly due to social anxiety in people with ASC, experienced particularly while looking at someone’s eyes. Moreover, this relation between gaze fixation time and amygdala activation was not found for the controls in the same study (Dalton et al. 2005). In summary, it is not clear whether the proposed model predicts higher or lower amygdala activation in response to eye contact, and how that relates to the extent of ES observed.

The second question asks: How crucial is eye contact for ES to occur, and is its role emotion-specific? ES has been shown to occur while viewing body parts only (e.g., studies on perception of pain [Singer et al. 2004] and touch [Keysers et al. 2004]) in the absence of any eye contact. This raises the question of whether the role of eye contact in the proposed model is anything more than a device to capture spatial attention. Indeed, it has been shown that attention can influence ES during action perception (Bach et al. 2007; Chong et al. 2008). It is reasonable to hypothesise that the same mechanisms for ES operate during

Figure 1 (Chakrabarti). A generalized model of facial expression (Chakrabarti et al. 2006). Left of the dotted line indicates systems involved in perception of all emotion expressions. Right of the dotted line indicates structures that may selectively respond to specific emotions. IOG = Inferior Occipital Gyrus, FG = Fusiform Gyrus, STS = Superior Temporal Sulcus, IFG = Inferior Frontal Gyrus, PM = Premotor Cortex, IPL = Inferior Parietal Lobule.

doi:10.1017/S0140525X10001482
Because smile types can be identified based on either physical or contextual cues, this suggests that any information provided by simulation simply serves to enhance the identification of the smile’s meaning (e.g., by speeding up the recognition process), rather than being critical in recognizing the smile.

One study which suggests that simulation may be critical in recognizing smiles was conducted by Maringer et al. (in press; cited in the target article). This study showed that mimicry can facilitate the discrimination of true from false dynamic smiles. However, this advantage may be due to the fact that mimicry provides an additional representation of a smile stimulus, which allows participants who mimic to receive more input about a smile than participants who do not mimic. Future experiments can control for the number of stimulus representations between the mimicry and no-mimicry conditions by adopting a yoked design. This would involve videotaping the faces of participants in the mimicry condition when they view a sequence of face stimuli, and playing these recordings to yoked participants in the no-mimicry condition while they view the same stimulus sequence. In this way, any differences between the two conditions can be attributed to qualitative differences in stimulus processing, rather than to differences in the number of stimulus representations received.

According to the SIMS model, smiles elicited when viewing positive (i.e., enjoyment) smiles indicate mimicry. In this case, it is the *current* expression and feeling of the *expresser* that is simulated. Niedenthal et al. argue that, by contrast, people are less likely to mimic negative (e.g., dominant) smiles, but instead, they simulate the feeling of being dominated. In this case, it is the *previous* experience of the *perceiver* that is being simulated. It is not clear why the different smiles should invoke simulations with reference to different people- and time-perspectives. Instead, it would be more parsimonious to claim that the same mechanism (i.e., mimicry) occurs for both types of smiles, and in general, when viewing pictures of angry faces (e.g., Cannon et al. 2009; Dimberg et al. 2000). The question of whether mimicry occurs similarly for both positive and negative smiles may be answered by analyzing the results of Maringer et al. (in press) in more detail. If Niedenthal et al. consider that people should inhibit mimicry for false smiles as they would for dominant smiles (because, to extend the authors’ reasoning, it is not socially appropriate to mimic false smiles), then people should not mimic false smiles in the no-mimicry condition of Maringer et al.’s study, but they should nevertheless be able to simulate reactions that they have previously felt in response to false smiles. As such, participants in the mimicry and no-mimicry conditions should recognize false smiles with equal accuracy. However, if participants spontaneously mimic false smiles, then false smiles should be recognized more accurately in the mimicry than in the no-mimicry condition.

Niedenthal et al. claim that eye contact increases emotional intensity in response to angry and happy faces (Schrammel et al. 2009) because it triggers the simulation of expressions and their associated emotions. However, the relation between eye contact and emotional intensity may be explained by other mechanisms. Eye contact may enhance the emotional experience of the perceiver because it implies that the perceiver is the target of the expression, which in turn increases the relevance of the expressed emotion for the perceiver (e.g., Cacioppo 1978). Alternatively, emotional intensity may increase under eye contact because angry and happy faces both signal an approach motivation, which is congruent with the approach motivation signaled by a direct gaze, but incongruent with the avoidance motivation signaled by an averted gaze (see Argyle & Cook 1976; Davidson & Hugdahl 1995; Harmon-Jones & Seligman 2001). Thus, happy
faces with an averted gaze may be considered to be more ambiguous than happy faces with a direct gaze, causing them to elicit less emotion, or ambivalent emotion. In support of this, studies show that when the approach/avoidance motivation of the gaze direction matches the approach/avoidance motivation of the emotion expression (e.g., when happy and angry faces are presented with a direct gaze, and sad and fearful faces are presented with an averted gaze), the emotion is identified more quickly (Adams & Kleck 2003; 2005), and the perceived emotional intensity of the expression also increases (Adams & Kleck 2005). These results with happy and angry faces are consistent with the SIMS model. However, the model assumes that similar processes should underlie the simulation of all types of expressions, and thus it cannot account for why sad and fearful faces should be identified more quickly, and considered to be more emotionally intense, when presented with averted rather than with direct gazes.

The SIMS model represents an ambitious integration of social, cognitive, and neuro-psychology. However, one aspect that needs to be clarified is whether simulation operates similarly for different types of smiles, such as true versus false smiles, and smiles associated with physical markers versus more ambiguous smiles. It also remains to be demonstrated whether simulation is indeed critical for the interpretation of smiles, or whether physical and contextual cues play a greater role in this process. Nevertheless, the model raises many testable hypotheses, which provide a fruitful starting point for exploring the contribution of simulation to the processing of facial expressions.

ACKNOWLEDGMENT
The writing of this commentary was facilitated by a grant from the “Action de Recherche Concertée”: No. ARC06/11-337.

Eye gaze and conscious processing in severely brain-injured patients

doi:10.1017/S0140525X10001494

Camille Chatelle, a Steven Laureys, a Steven Majerus, a and Caroline Schnakers a

a Coma Science Group, Cyclotron Research Centre, University of Liège, 4000 Liège, Belgium; b Department of Cognitive Psychopathology, University of Liège, 4000 Liège, Belgium.

camille.chatelle@ulg.ac.be steven.laureys@ulg.ac.be
smajerus@ulg.ac.be C.Schnakers@ulg.ac.be
http://www.ulg.ac.be
www.coma.ulg.ac.be

Abstract: Niedenthal et al. discuss the importance of eye gaze in embodied simulation and, more globally, in the processing of emotional visual stimulation (such as facial expression). In this commentary, we illustrate the relationship between oriented eye movements, consciousness, and emotion by using the case of severely brain-injured patients recovering from coma (i.e., vegetative and minimally conscious patients).

The vegetative state (VS) is characterized by a preserved arousal level but also by the absence of any sign of consciousness of the environment and of oneself (e.g., absence of oriented responses to environmental stimulation) (The Multi-Society Task Force on PVS 1994). The patient in a minimally conscious state (MCS), on the contrary, demonstrates inconsistent, basic but reproducible signs of consciousness of the environment and of oneself (e.g., oriented responses to environmental stimulation suggesting an interaction between the patient and his/her surroundings) (Giacino et al. 2002). Both populations may be particularly useful for studying the concept of consciousness via the progressive recovery of the most primary aspects of consciousness, and the behavioral and cerebral correlates that accompany them.

Niedenthal et al. discuss the importance of eye gaze in embodied simulation and, more globally, in the processing of emotional visual stimulation (such as facial expression). Eye gaze involves interaction between the subject and his/her surroundings. At a neurobehavioral level, oriented eye movements are usually associated with conscious processing (Giacino et al. 2002). MCS patients are able to track a person moving in front of them, whereas VS patients are not. Only MCS patients can therefore detect emotional visual stimuli present in their surroundings. According to a recent study, visual pursuit is more frequently detected in response to emotional stimulation (e.g., the appearance of one’s face in a mirror) than to non-emotional stimulation (e.g., a person moving) (Vanhaudenhyse et al. 2008).

MCS patients also show more complex emotional responses than VS patients. For instance, VS patients may smile or grimace, shed tears, or make grunting or groaning sounds, but for no discernible reason (The Multi-Society Task Force on PVS 1994; Working Party of the Royal College of Physicians 2003), similarly to what can be observed in anencephalic infants lacking a functional neocortex (Massimelli 2007). MCS patients can manifest appropriate smiling or crying in response to specific linguistic or visual content of emotional stimuli, suggesting an appropriate interaction between the patient and his/her environment (Giacino et al. 2002). Here, consciousness is reflected by the presence of basic but contingent social and emotional behaviors, those requiring a complex cortical activity. Using auditory (Laureys et al. 2000) and noxious stimuli (Boyl et al. 2008; Laureys et al. 2002), previous studies in functional neuroimaging have shown that, while brain activation isolated to primary cortices is observed in VS patients, an activation in associative areas as well as preserved functional connectivity between cortico-cortical areas (i.e., between primary and associative cortices) is observed in MCS patients (see our Fig. 1). MCS patients seem, therefore, to present a more complex social and emotional pattern at a behavioral level and a more integrative brain processing than VS patients.

Even if oriented eye movements, such as visual pursuit, are usually considered a sign of consciousness, controversies remain concerning visual fixation. Visual fixation includes eye gaze and can be described as the eyes changing from their initial fixation point and refixating on the new target location (Giacino et al. 2004). As eye gaze involves interaction between the subject and his/her surroundings, visual fixation would have to be considered as a sign of consciousness. However, whereas the Aspen Workgroup considers this behavior as a sign of consciousness (Giacino et al. 2002), the College of Physicians considers its presence as reflecting a reflexive activity (Working Party of the Royal College of Physicians 2003). Using active paradigms (i.e., the patients were asked to mentally perform a task), voluntary brain activity has been obtained in patients showing mainly low-level behavioral responses such as visual fixation, suggesting it may be a sign of consciousness (Owen et al. 2006; Schnakers et al. 2008). However, we have recently shown no differences in brain activity and connectivity in VS patients presenting or not presenting visual fixation, suggesting that it does not necessarily reflect consciousness and higher-order cortical functioning (Bruno et al. 2010). According to these results, it is not clear whether visual fixation reflects conscious interaction between the patient and his/her surroundings.

In conclusion, consciousness level is associated with social and emotional behaviors of different complexity. Conscious patients (i.e., MCS) show high-level behavioral responses and brain activation as compared to non-conscious patients (i.e., VS). Nevertheless, the particular case of visual fixation would have to be further investigated in order to better define its relationship to consciousness.
How does perceiving eye direction modulate emotion recognition?

doi:10.1017/S0140525X10001305

Laurence Conty,a Julie Grèzes,a and David Sanderb,c

aCognitive Neuroscience Laboratory, INSERM U960, Ecole Normale Supérieure, 75005 Paris, France; bLaboratory for the Study of Emotion Elicitation and Expression (E3 Lab), Department of Psychology, Swiss Center for Affective Sciences, University of Geneva, CH-1205 Geneva, Switzerland; cCentre Interfacultaire en Sciences Affectives (CISA) – University of Geneva, CH-1205 Geneva, Switzerland.

ACKNOWLEDGMENTS

Our research is funded by the Belgian Fund for Scientific Research (FNRS), European Commission, James McDonnell Foundation, Mind Science Foundation, French Speaking Community Concerted Research Action (ARC-06/11-340), Fondation Médecine Reine Elisabeth, University of Liège, and the CNRS/FNRS-CGRI collaboration funds.

Abstract: Niedenthal et al. postulate that eye contact with the expresser of an emotion automatically initiates embodied simulation. Our commentary explores the generality of such an eye contact effect for emotions other than happiness. Based on the appraisal theory of emotion, we propose that embodied simulation may be reinforced by mutual or averted gaze as a function of emotional context.

We congratulate Niedenthal et al. for their outstanding contribution to both empirical and conceptual approaches in psychology and neuroscience. As reported by the authors, smiling faces are judged happier when gazing at the participant as compared to gazing sideways (e.g., Adams & Kleck 2005). The authors advance that such an increase in positive intensity could reflect a particularly accurate embodied simulation during smile perception. Here, we explore whether such a model could account for recognition of all emotions. Alternatively, the triggering function of eye contact may be specific to a subset of emotions that includes happiness.

Although facial expressions are efficient emotional signals, eye gaze direction is important for signaling the referent of an expression and, therefore, the attended object of the elicited emotion (George & Conty 2008). Converging evidence suggests that these two signals are integrated during the perception of facial emotion. Further, behavioral studies have reported that the perception of facial expressions can be modulated by eye-gaze direction. However, the effect of gaze on emotion recognition depends on the type of expression. For example, it has been shown that angry faces are perceived to express more anger with direct than averted gaze, whereas fearful faces are perceived to express more fear with averted than direct gaze (Adams & Kleck 2003; Sander et al. 2007). These results can be explained within a self-relevance framework (Sander et al. 2003). Indeed, it has been proposed that angry expressions with averted gaze or fearful expressions with direct gaze have less relevance for oneself than these expressions with a direct or averted gaze, respectively. The rationale is that the aversive dimension of anger is higher when one is the target of it (and hence gazed at); similarly, the threatening content of fearful expressions is higher when the face is looking away from the observer, as the object of fear for another agent can also reflect a potential menace for the observer. So far, these results suggest that emotion perception is influenced by the self-relevance of expression based on gaze direction, and this process appears to be emotion-dependent.

Niedenthal et al. propose that the amygdala may produce states that further contribute to simulations that underlie how smiles are interpreted. A recent brain imaging experiment is particularly relevant to this view and the prediction for the role of gaze contact in the processing of happiness: Sato et al. (2010) showed that dynamic happy faces elicit more amygdala activation when gazing at – as opposed to away from – the observer. Yet, the amygdala does not always show greater response to direct as compared to averted gaze for all emotions, but rather supports its involvement in the appraisal of self-relevance (N’Diaye et al. 2009). Brain-imaging and patient studies suggest that the amygdala deals with the integration of emotional expression and gaze direction (Conty et al. 2010; Cristinizio et al. 2010), with typical differential amygdala activity for anger versus fear processing as a function of direct versus averted gaze (Adams et al. 2010; Hadjikhani et al. 2008; Sato et al. 2010). Therefore, amygdala responses to a

Figure 1 (Chatelle et al.). Cerebral activation to noxious stimulation. In the darker brain regions (which appear in red in the online version, available at www.journals.cambridge.org/bbs) that activated more during noxious stimulation in healthy controls (top), in vegetative state (bottom, left side) and in minimally conscious state (bottom, right side) as compared to rest (adapted from Boly et al. 2008 and Laureys et al. 2002).
Cultural variations on the SIMS model

doi:10.1017/S0140525X10001317

Christine M. Covas-Smith, Justin Fine, Arthur M. Glenberg, Eric Keylor, Yexin Jessica Li, Elizabeth Marsh, Elizabeth A. Osborne, Tamer Soliman, and Claire Yee

Department of Psychology, Arizona State University, Tempe, AZ 85282.

ccovas01@gmail.com jjfine@asu.edu
Glenberg@asu.edu ekeylor@asu.edu
yjli1@asu.edu Elizabeth.R.Marsh@asu.edu
Elizabeth.A.Osborne@asu.edu tamer.soliman@asu.edu
ciyee@asu.edu

http://psychology.clas.asu.edu/glengberg

Abstract: Niedenthal et al. recognize that cultural differences are important when interpreting facial expressions. Nonetheless, many of their core observations derive more from individualistic cultures than from collectivist cultures. We discuss two examples from the latter: (1) lower rates of mutual eye contact, and (2) the ubiquity of specific “functional smiles.” These examples suggest constraints on the assumptions and applicability of the SIMS model.

Niedenthal et al. posit an embodied model for perceiving the meaning of smiles through implicit simulation. As noted in section 6, the model is grounded on insights taken from predominately individualistic cultures. This commentary focuses on two culturally limited aspects of the SIMS model. First, the model emphasizes that eye contact serves as a triggering mechanism for the activation of embodied simulations. Second, the model explicates three principal functions that smiles and the simulation of those smiles serve in interpersonal settings. There are compelling reasons to believe that differences in social and environmental context will modulate these particular components of the model.

Eye contact may not provide the primary route toward successful interpersonal simulation in all cultures. (Indeed, as noted in section 4.1, eye contact is presumed to be a sufficient but not a necessary trigger.) Some collectivist cultures, like East Asian cultures, tend to avoid eye contact when processing facial expressions. Children in these cultures are taught that direct eye contact is a sign of disrespect (Argyle & Cook 1976). Such cross-cultural differences suggest that eye contact is a socially established cue that varies with particular cultural norms. How might such a socially refined behavior affect the SIMS model? Assuming that eye contact is the most frequent triggering mechanism for interpersonal simulation, East Asians would be predicted to often miss this trigger. The result would be little direct simulation or mediated interpretation of the facial expressions of others. Instead, East Asians would need to rely on conceptual or contextual knowledge when reading facial expressions. Although possible, this conclusion is suspicious. Why should it be the case that any culture would develop in such a way as to negate the occurrence of direct affective resonance? This prediction seems especially peculiar given that “collectivist” cultures are defined by their strong tendency toward shared social intentions and responsibilities.

A reasonable alternative is that the triggering mechanism of affect simulation is not solely based on the activation of a focal facial component, but rather on a gestalt of facial muscular configuration encompassing the supra- and infra-orbital, paranasal, and peribuccal regions. The targeted relational contour could be detected by a fixation that lands centrally at (or slightly above) the tip of the nose, hence making it possible to peripherally process the relevant broader area at once. Evidence for such an alternative relational account is suggested by the findings of a cross-cultural face-processing study (Blais et al. 2008). East Asian participants did not exhibit the fixation triangle of eyes and nose adopted by their American counterparts, but they preferred instead to fixate a narrow nasal locus.

The relational alternative could extend beyond a biological (facial or otherwise) narrow scope. There is evidence that, in making decisions, people in collectivist cultures forge a broad relational model that gives primacy to social and physical cues assimilated from the environment. Kitayama et al. (2003) demonstrated that American participants were accurate when attempting to replicate the absolute length of a line drawn in a square frame, whereas East Asians were more accurate when replicating relative line length, taking into account both the length of the line and its relation to the size of the frame. Similarly, people from individualistic cultures are more inclined to make person-centered attributions of an event’s cause, whereas people from collectivist cultures are more inclined to identify a situation-centered cause (Morris & Peng 1994). This idea is consistent with Niedenthal et al.’s suggestion in section 6.1.1 that context, such as the strength of a personal relationship, is likely to play a role in simulation.

The functional roles of smiling might also be culture-specific. Niedenthal et al. make the cogent suggestion that smiles are distinguished based on the part they play in facilitating interpersonal actions and reactions. Therefore, changes in social structure should coincide with changes in the function of smiles that operate within that structure. Accordingly, some functional
Abstract: We outline three possible shortcomings of the SIMS model and specify these by applying the model to autism. First, the SIMS model assigns a causal role to brain processes, thereby excluding individual and situational factors. Second, there is no room for subjective and high-level conceptual processes in the model. Third, disentangling the different stages in the model is very difficult.

The Simulation of Smiles (SIMS) model holds that observing a smile leads to the extraction of meaning via different behavioral and neural processes on in the perceiver’s part. In what they call their core model, Niedenthal et al. propose that the same smile is interpreted in three different ways, according to its functional meaning as enjoyment, affiliation, or dominance, by engaging different brain areas (see their Figure 4). By doing so, the SIMS model subscribes to some kind of mind-brain identity theory, establishing a one-to-one mapping of mental states to brain processes. Although such a philosophy of mind is implicit in most of modern cognitive neuroscience, SIMS is very explicit in assigning a unique, causal role to the brain processes and leaving no room for individual and situational factors which may also affect how a particular smile is interpreted. Moreover, the SIMS model appears to ignore the role of subjective and high-level conceptual processes which are crucial in emotion processing (e.g., using one’s knowledge about the person who is smiling or the context in which the smile takes place). Finally, disentangling the different stages in the model is difficult, which complicates the process of deriving specific hypotheses from it and testing them empirically. Niedenthal et al. provide a fascinating synthesis of the huge literature on embodied simulation and emotion, both in psychology and the neurosciences, but distinguishing between perceptual processes, motor mimicry, and empathizing in particular cases of emotion understanding turns out to be quite a challenge. In the remainder of this commentary, we will make these arguments more specific by trying to apply the SIMS model to autism.

Autism is an early-onset neurodevelopmental disorder characterized by the co-occurrence of impairments in social reciprocity and communication, accompanied by stereotyped and rigid patterns of interests and activities. Early behavioral signals of autism, such as atypical eye contact, absence of social smiles, and absence of gaze following, are closely related to the processing of facial expressions (including smiles). At the cognitive level, children and adults with autism often have difficulties in ascribing mental states to others, referred to as mentalizing or Theory of Mind problems (e.g., Tager-Flusberg 2007). This means that they experience problems in placing themselves in the mental worlds of others and in empathizing with others. Therefore, autism provides an ideal test case to evaluate the SIMS model. Applying SIMS to autism (SIMS-A) should yield additional insight or at least testable hypotheses, if the model pretends to be more than hocus-pocus with component processes and neural circuits (cf. magic words “Sim Sala Bim”).

First, linking psychological processes to neural mechanisms is definitely a way to make progress, also in research aimed at understanding the causal pathways behind a complex disorder like autism. There is a significant genetic component to autism, and some of the genes involved are clearly affecting the developing brain (e.g., Bourgeron 2009). There is a huge body of literature aimed at discovering structural and functional differences between brains of people with and without autism (e.g., Amaral et al. 2008). However, our understanding of the vagaries and complexities of information and emotion processing in autism cannot be reduced to looking only inside the brain. If people with autism have difficulties in picking up subtle cues from faces, differentiating relevant from irrelevant stimulus changes in faces and facial expressions, integrating local details with more global context information and bottom-up perceptual information with top-down knowledge about the situation, and especially in doing all of this automatically and without painstaking effort, it is clear that an exclusive focus on neural circuits alone will not tell the whole story. At least not at the present stage of our scientific endeavor, when linking functional explanations at the psychological level to possible neural correlates is still a tricky business as far as complexity, multi-factorial phenomena are concerned.

Second, autism researchers often take individual and situational factors into account when trying to explain the inconsistencies between studies finding anomalies in emotion processing in children and adults with autism and studies that do not. Clinical samples differ in factors such as age, general language skills, general intelligence, subset and severity of symptoms, etc. Moreover,
Researchers have used a huge variety of stimuli (e.g., faces in isolation vs. embedded in complex scenes and events, static vs. dynamic facial expressions, uni- vs. multi-modal, brief vs. unlimited presentations), tasks (e.g., explicit vs. implicit, focused vs. distributed attention, sorting, matching, labeling), and measurements (e.g., performance levels, response times, eye movements). All of these factors have shown to influence the outcome of studies, but none of them is currently incorporated into the SIMS model.

Third, the SIMS model claims to be able to distinguish the role of processes like establishing eye contact, facial mimicry, simulating embodied experience, using conceptual knowledge, social inhibition, and so forth; but we are not convinced that it succeeds in doing so. For instance, can SIMS distinguish, functionally or neurally, cases of embodied emotion simulation triggered by bottom-up facial cues or by top-down knowledge? How do we differentiate the SIMS model from the many different models making use of components from the so-called mirror neuron system (e.g., Rizzolatti & Fabbri-Destro 2008)? If all of these crucial yet relatively basic distinctions appear to be difficult, what about all the subtleties which appear to be involved in the literature on autism? Will it be possible to specify the SIMS model sufficiently so that testable hypotheses can be derived from it? We hope so, but there is clearly still a very long way to go.

Addressing these three general issues regarding the SIMS model and their specific elaboration regarding autism, will clarify what kind of model SIMS is, and what it can and cannot offer.

Le bon dieu est dans le detail: Is smiling the recognition of happiness?

doi:10.1017/S0140525X10001512

José-Miguel Fernández-Dols and Pilar Carrera
Facultad de Psicología, Universidad Autónoma de Madrid, 28049 Madrid, Spain.
jl.dols@uam.es pilar.carrera@uam.es

Abstract: We question two conceptual assumptions made by Niedenthal et al.: the dichotomy between true and false smiles and the close tie between recognition and experience of emotion. An excessive dependence on everyday language suggests overly parsimonious accounts of a complex set of relations between smile, experience, and context.

Recognition is a fundamental topic in the study of emotion, central to the claim of universality (Darwin 1872/1965), and is the subject of great controversy. For some theorists (e.g., Russell 1994), recognition is thought of as the attribution of a semantic label to another’s state, whereas for basic emotion theorists (e.g., Ekman 1994), recognition is thought of as part of an emotion signaling system grounded in a specific brain region.

Niedenthal et al. advance the study of how emotion is recognized in others. First, they offer new data on the brain regions involved in recognition of emotion from facial expression. Second, they go beyond the isolated, still, photographed face devoid of context to the facial expression as dynamic muscular movements embedded in a context of interaction between the sender and the perceiver (see Fernández-Dols & Carroll 1997; Fernández-Dols et al. 2008). Third, they propose that recognition is a form of emotion knowledge based on embodied simulation of the sender’s emotional state.

Nevertheless, we see two main problems: (1) We do not agree with Niedenthal et al.’s distinction between true and false smiles; and (2) we question the assumption, common to basic emotion theory and Niedenthal et al.’s model, that recognition of emotion is closely tied to the experience of emotion.

True and false smiles. Niedenthal et al.’s model is focused on the detection of true smiles of enjoyment, dominance, or affiliation. We are glad to see that Niedenthal et al. have acknowledged that the Duchenne smile (Ekman et al. 1990) is not the marker of a true smile. Duchenne smiles can be effortlessly posed, do not predict senders’ affective state (Krumhuber & Manstead 2009), and can be caused by negative experiences (e.g., losing a game; Schneider & Josephs 1991; pain: Kunz et al. 2009). Nevertheless, Niedenthal et al. maintain the original theoretical assumption behind the Duchenne smile, namely, the true versus false dichotomy. According to Niedenthal et al., true smiles are honest displays that reveal positive feelings, whereas false smiles are manipulative displays that hide negative feelings.

In our view, any model aimed at detecting the “truth” or “honesty” of an expression is, ultimately, scientifically unsound. Smiles are not essentially true or false. Indeed, we doubt that any dichotomous classification of smiles will survive scientific scrutiny for long. Smiles are too morphologically diverse for a dichotomy: they can be located on a continuum based on their tempo, intensity, asymmetry, and concomitant facial movements (Krumhuber & Kappas 2005; Krumhuber & Manstead 2009). This dynamic diversity has consequences for the receiver’s judgments (e.g., perception of trustworthiness; Krumhuber et al. 2007) and is connected to the sender’s relational strategies (e.g., Vigil 2009). Thus, each particular facial expression is too complex in its meaning to be either true or false. A consciously exaggerated “false” smile is “true” – and extremely functional – in the context of baby talk (Chong et al. 2003). The equally functional spontaneous “true” smile of a two-month-old baby is clearly “false” in that it lacks any of the “honest” emotional or cognitive correlates of a “true” adult smile.

Simulation as emotional experience. Niedenthal et al.’s simulation-as-emotional-experience theory sees emotions as “natural kinds” in which experience and expression belong to a tight cluster of properties caused by the same neural structures (Barrett 2006; Fernández-Dols & Russell 2003). On this view, emotion knowledge is expression (mimicry), experience, and recognition of basic emotion. There are two apparent sources of evidence in support of this view: experimental and clinical.

We question the experimental support (e.g., Halberstadt et al. 2009; Niedenthal et al. 2009). Electromyography shows a weak or even contradictory relation between emotion knowledge and facial mimicry (e.g., activity of the levator labii when processing joy or nonspecific activity of the corrugator when processing disgust; Niedenthal et al. 2009). The claim that such mimicry causes the experience of the corresponding basic emotion is not scientifically justified on the basis of current evidence. We see no reason to assume that the link between emotion knowledge and facial mimicry is different from the observed link between nodding your head and emotion knowledge (Niedenthal 2007) – that is, culture-specific and learned. No evidence forces us to the additional assumption of an innate connection between mimicry, experience, and recognition of emotion. Mimicry could be the outcome of several parallel processes (Lishner et al. 2008; Yabar et al. 2006).

We also question the clinical support for the recognition-as-experience view. Case studies show paired deficits in both experience and recognition of emotions (e.g., fear or disgust) apparently caused by damage in one specific brain region (e.g., the amygdala or the insula and basal ganglia; for a review, see Goldman 2006). Nevertheless, further research on the shared neural substrate of such paired deficits has raised doubts as to whether these regions are organized for experiencing and recognizing basic emotions. The functions of these regions are broader (e.g., Heberlein et al. 2008), more diverse (e.g., Heberlein & Altmann’s distinction between true and false smiles), and more complex than originally thought (e.g., Chen et al. 2009; Straube et al. 2010). Better experimental procedures and measures challenge the original interpretation of the paired deficits (Milders et al. 2003; Rapcsak et al. 2000). Understanding recognition requires a more conceptually sound analysis of the dynamics and functions of specific brain regions (see Kagan 2007).
**Conclusion.** The study of the recognition of emotion must move beyond everyday language. The assumption that words such as “enjoyment” or “smile” carve nature at its joints prompts us to unhelpful dichotomies, such as “false versus true” smile or “basic versus non-basic” emotion. Moving beyond overly parsimonious assumptions would open the way to more complex but realistic models in which the perception of an expression is a consequence of not only affect or emotion, but also of diverse variables, such as, for example, stimulus acuity (e.g., Bohn et al. 2010), sender’s gender (e.g., Vazire et al. 2009), facial structure (e.g., Zebrowitz et al. 2010), and culture (e.g., Ellenhbein & Ambady 2002).

ACKNOWLEDGMENT
The authors are supported by grant PSI2008-0849 from the Spanish Government.

**Does motor mimicry contribute to emotion recognition?**

doi:10.1017/S0140525X10001524

Cindy Hamon-Hill and John Barresi
Department of Psychology, Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada.
cindy.hamon-hill@dal.ca jbarresi@dal.ca
http://jbarresi.psychology.dal.ca/

Abstract: We focus on the role that motor mimicry plays in the SIMS model when interpreting whether a facial emotional expression is appropriate to an eliciting context. Based on our research, we find general support for the SIMS model in these situations, but with some qualifications on how disruption of motor mimicry as a process relates to speed and accuracy in judgments.

We applaud the efforts of Niedenthal et al. to integrate multiple processes underlying recognition of smiles in facial expression into a model that can generalize to the recognition of other emotions. However, we detect a need for greater clarity regarding the role of motor mimicry in a complex judgment task. In our paradigm for recognizing emotion in a naturalistic facial expression, a static contextual image eliciting one of eight possible emotions is presented simultaneously with a five-second facial expression that is either static or dynamic. Participants judge when the facial expression is the genuine response to the paired contextual image (“match”) and when it is not (“mismatch”) (Hamon-Hill & Barresi 2008a, 2008b; 2009; Ouellette et al. 2010). To test for reliance on motor mimicry during this task, participants hold a pen horizontally in their mouth throughout alternating blocks of trials. In agreement with the SIMS model, we recognize embodied simulation as a key component in the underlying processes activated during this task (Barresi & Moore 1996; 2008). We propose the observers, while perceiving the paired stimuli, generate a first-person embodied response to the contextual image, which includes autonomic, somatic, and semantic aspects. The observers also generate a third-person representation of the target facial expression and search for a match between their own embodied response to the contextual image and their understanding of the target facial expression. For instance, observation of an elderly man holding the hand of his dying wife may elicit a response of “feeling sad,” prompting the observers to search for signs of sadness in the target facial expression.

Earlier simulation models for emotion recognition propose facial mimicry contributes to recognition of facial expressions (Goldman & Sripada 2005); however, there is limited support for this claim (Blairy et al. 1999; Hess & Blairy 2001), and imaging studies measuring motor activity seldom test for accuracy in judgments (Carr et al. 2003). Using our paradigm, we tested for an effect of disruption to motor mimicry on judgments about a dynamic facial expression as genuine to a given context. We tested the effect of two experimental manipulations, one to inhibit mimicry (Niedenthal et al. 2001) and the other to interfere with mimicry (Oberman et al. 2007). These two manipulations, which require holding a pen in one’s mouth, were cited indiscriminately in the target article, suggesting both inhibit motor mimicry. However, Oberman et al. (2007) established a difference between the two: Whereas holding a pen between the lips and teeth (Niedenthal et al. 2001) holds the muscles in a neutral state, thereby “inhibiting” a motor response, “biting” down on a pen held horizontally between one’s teeth, without the lips touching the pen, “interferes” with spontaneous motor mimicry due to “noise” generated by consistent activation of multiple facial muscles (Oberman et al. 2007). During this manipulation, activation in the motor systems is neither absent nor reduced as suggested in the target article.

Like Oberman et al. (2007), we saw no effect of disruption to mimicry on accuracy resulting from the inhibition manipulation. However, observers in the inhibition condition and the no-pen control condition both showed a significant response bias to judge the target face as a match to the paired context image. Surprisingly, the response bias was absent in the interference condition, where participants correctly discriminated when faces were not a genuine response to the paired contextual image. When we repeated the experiment applying only the interference manipulation to test for a differential effect on processing dynamic or static facial expressions, once again accuracy was not altered, but processing was affected. When mimicry was disrupted, observers were slower to process static, positive faces, yet were faster to process dynamic, positive faces.

The SIMS model offers an explanation for a response bias to facial expressions based on conceptual knowledge. According to the model, an embodied response to a context stimulus should prime the observer’s judgment about the emotion perceived in a target facial expression. Additionally, disruption to motor mimicry should increase the priming effect due to greater reliance on this embodiment of conceptual knowledge. Our paradigm emphasizes the incorporation of embodied responses to context into judgments made about facial expression. However, the response bias was strongest when participants were free to mimic and while mimicry was “inhibited”, yet it was significantly diminished during “interference” with mimicry. Niedenthal et al.’s account of processing facial expression when motor mimicry is disrupted does not adequately explain what we have observed. We suggest the interference generated by the bite manipulation causes a mismatch state in the self with respect to the embodied response to the context image, and, consequently, the observer searches for a “mismatch” in the other. This leads to a reduction in the bias to see a match and increases the likelihood of improved performance on mismatch trials.

Another limitation to the model is the lack of adequate discussion regarding the quality of stimulus, in particular the difference due to static and dynamic emotional expressions. Neuroimaging studies indicate greater motor activity during observation of static compared to dynamic faces (Enticott et al. 2008; Kilts et al. 2003). Moreover, we would expect less reliance on embodied simulation involving motor mimicry during perception of dynamic faces because of the rich quality of information available in a dynamic facial expression (Ambadar et al. 2005). Dynamic facial displays can reveal blended emotions that change over time, shifting in degrees of expressivity. Perception of static faces, which typically present the apex of the expression, may require the observer to enrich the information by simulating the full expression. One consequence of this difference between static and dynamic expressions is that we would expect greater interference with processing static faces when motor mimicry is disrupted. That is indeed what we observed. Disruption to mimicry resulted in slower processing of static but not dynamic faces.
In sum, motor mimicry, while a contributor to a fully embodied form of emotion recognition, may not be an essential component when interpreting rich facial expressions in naturalistic contexts, but rather may primarily serve social functions such as indicating sympathy and understanding.

No mirrors for the powerful: Why dominant smiles are not processed using embodied simulation

doi:10.1017/S0140525X10001536

Li Huang and Adam D. Galinsky
Department of Management and Organizations, Northwestern University, Evanston, IL 60208.

Abstract: A complete model of smile interpretation needs to incorporate its social context. We argue that embodied simulation is an unlikely route for understanding dominance smiles, which typically occur in the context of power. We support this argument by discussing the lack of eye contact with dominant faces and the facial and postural complementarity, rather than mimicry, that pervades hierarchical relationships.

Not all smiles are equal. As indicated by Niedenthal et al., some smiles express joy and affiliation, whereas others express dominance. A complete model of smile interpretation needs to examine the different aspects of social relationships that each type of smile facilitates. Armed with this understanding, one can determine how likely it is that embodied simulation would occur in any given context. This commentary proposes that dominance smiles are categorically different from enjoyment and affiliative smiles in terms of their social contexts and implications and are therefore less likely to be processed through embodied simulation.

Social relations vary in both the “horizontal” and “vertical” dimensions (Fiske 1992). The “horizontal” dimension describes the affective or socioemotional dimension, encompassing the valence of feelings and emotional closeness of interpersonal relations (Berger 1994; Osgood et al. 1957). Enjoyment and affiliative smiles, which communicate positive emotions and cooperative social motives, tend to occur when social relations are characterized horizontally.

The “vertical” dimension, in contrast, describes the fact that interpersonal relations can be organized along a low-to-high continuum of dominance, power, status, or hierarchy (Hall et al. 2005). Dominance smiles, which express status and control, facilitate social relations that take place in the vertical dimension of sociality. In this dimension, individuals seek to identify and/or establish their rank in a hierarchy rather than to build social bonds. As a result, eye contact tends to be suppressed, and complementarity of facial and postural expressions between dominant and non-dominant individuals is prevalent. These processes should render embodied simulation a less frequently used method for understanding a dominance smile.

With regard to eye contact, Niedenthal et al. argue that it is a key trigger for embodied simulation. Eye contact is a nonverbal behavior made possible only when two people are gaz ing into each other’s eyes. However, there is significant difference in the frequency of eye contact within vertical and horizontal relationships. Horizontal relationships between mutually dependent individuals facilitate eye contact, which helps communicate positive emotions, seek reconciliation, solicit positive attitudes, and establish intimacy (Edinger & Patterson 1983; Mehraban 1969). On the other hand, because dominant individuals are less dependent, they tend not to reciprocate the gaze, making it less possible for eye contact to occur.

Niedenthal et al. also suggest that risk or aversiveness is one of the reasons individuals may suppress eye contact. Numerous findings suggest that exchanging eye contact with individuals who display dominance smiles may be deemed too risky and that subordinate individuals may avert their gaze to prevent inappropriately challenging their dominant counterparts. Dominance faces are more similar to threatening or angry faces than they are to happy or sad faces (Lipp et al. 2009; Öhman et al. 2001). Because individuals avoid eye contact with those who display aggression (Ellsworth & Carlsmith 1973), they would presumably avoid eye contact with those who display dominance as well. Similarly, a variety of primate species respond to threat from dominant animals by averting eye contact and even the face all together (Bertrand 1969; de Waal 1989; Van Hooff 1967). Avoiding eye contact removes the threatening stimulation, creating a “cut-off” to reduce tension in the recipient of the threat while sending out a submissive signal (Chance 1962; Redican 1975). Thus, both the individuals producing dominance smiles and the perceivers of those smiles tend to avoid eye contact.

With regard to mimicry and complementarity, we propose that when individuals do smile in response to a dominance smile, they are likely to produce a knowledge-based “complimentary or submissive” smile, instead of an “exact replica” of the dominance smile. Niedenthal et al. state that embodied simulation may still occur when eye contact is suppressed. They argue that individuals can use conceptual knowledge to create facial mimicry, which then activates an off-line simulation, or an as-if loop.

However, there is ample reason to believe that facial mimicry does not typically occur in the presence of a dominance smile. Our argument is bolstered by the frequently observed complementarity, as opposed to mirroring, of bodily expressions in vertical relationships between humans or primates. Humans use smiles not only to express enjoyment, affiliation, or dominance, but also to produce submissiveness (Whalen & Kleck 2008). Submissive smiles appease and show deference. Darwin misinterpreted this expression as a true smile, but it is arguably the “descendant” of fear grimace seen in monkeys – subordinate monkeys offer grimace toward dominant monkeys as a signal of submission. Similar findings are found in humans with regard to posture – dominance postures tend to induce submissive postures in others, a process called dominance complementarity (Tiedens & Fragale 2003). This research suggests that, when individuals smile in response to a dominance smile, they will produce a submissive smile. To the extent that the entire embodied simulation process relies on facial mimicry, a lack of facial mimicry suggests the absence or partial absence of embodied simulation. The absence of facial mimicry in vertical relationships implies that dominance smiles are less likely to be processed through embodied simulation.

We do, however, acknowledge that in certain contexts, presumably what Niedenthal et al. have called high-uncertainty situations, dominance smiles may be processed through embodied simulation. When the perceiver is competing for a higher rank, instead of adhering to an established social hierarchy, eye contact and facial mimicry may be observed, and embodied simulation of dominance smiles may occur.

In conclusion, embodied simulation is less likely to be used to process dominance smiles because they typically occur in a social context where eye contact and motor mimicry are minimal. Future research should explore the precise processes used to understand dominance smiles. More broadly, future research on the SIMS model should investigate the role of social context in the use of embodied simulation. It should also discuss whether social contexts affect the types of smiles perceivers produce. Given that hierarchy is the dominant form of social organization (Leavitt 2005; Magee & Galinsky 2008), considering this context will not only help expand an understanding of when embodied simulation is used to understand facial expressions, but will also produce a more complete taxonomy of smiles.
Could embodied simulation be a by-product of emotion perception?

doi:10.1017/S0140525X10001548

Julian Kiverstein\textsuperscript{a} and Edvaro Zamuner\textsuperscript{b}

\textsuperscript{a}School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, EH8 7PU, Scotland, United Kingdom; \textsuperscript{b}Department of Philosophy, LaTrobe University, Victoria 3086, Australia.

\textsuperscript{a}j.kiverstein@ed.ac.uk \textsuperscript{b}e.zamuner@latrobe.edu.au

http://www.philosophy.ed.ac.uk/people/research-staff/

julian-kiverstein.html


Abstract: The SIMS model claims that it is by means of an embodied simulation that we determine the meaning of an observed smile. This suggests that crucial interpretative work is done in the mapping that takes us from a perceived smile to the activation of one’s own facial musculature. How is this mapping achieved? Might it depend upon a prior interpretation arrived at on the basis of perceptual and contextual information?

Smiles can be used to communicate a range of different psychological phenomena, including positive emotions, complex social intentions, and even a person’s social status. Given this diversity of possible meanings, how do we succeed on a given occasion in working out what a particular smile means? To do so, we must single out which of these many nuanced psychological phenomena the smile expresses. Niedenthal et al. say we solve this problem in part by “simulating” the nuanced states that we observe in others. We agree that embodied simulation may make an important contribution to the type of understanding we have of a smile. However, the exact nature of this contribution seems to us to remain an open question. In the spirit of friendly critics, it is this question we will take up in our commentary.

The SIMS model claims that we work out what a smile means in three interrelated stages. Given the work that motor mimicry and its effects is being asked to do in the determination of a smile’s meaning, something important is clearly happening in the transition from perception at Stage 1 to motor mimicry or action at Stage 2. There are a number of possible affective states that the smile you are producing might be expressing: but when I copy your facial expression, this results in me expressing the very same affective state. Thus, the process that allows me to map the perception of your behaviour onto the activation of my own facial musculature must somehow be singing out the meaning of the smile I am seeing. We don’t wish to deny that this might be possible, but the SIMS model, so far as we can tell, doesn’t tell us how this mapping is supposed to be effected.

The authors appeal to eye contact to explain how an embodied simulation gets triggered. Hence, they clearly think that at least sometimes the embodied simulation is sufficient for us to arrive at an interpretation of a smile. Supposing this is so, this makes it all the more urgent to know how the problem of determining what a smile means is solved by producing an embodied simulation. In order for me to mimic a smile that is affective, mustn’t I have already somehow worked out that the smile is affective? If so, how?

Consider, now, the studies in which facial mimicry is blocked or socially inhibited (see sects. 6.1.2.1. and 6.1.2.2. of the target article). The absence of motor mimicry in these studies has the consequence that “activation of motor systems and emotion systems will be absent” (sect. 6.1.2.1. para. 2). Hence, if recognition is achieved, it must be by some means other than embodied simulation. Niedenthal et al. suggest this could be achieved by matching visual input to a stored perceptual representation. If we sometimes have recourse to this strategy, why don’t we always use this strategy? Niedenthal et al. go on to allow that embodied simulation could still occur in this scenario, but it would have to be triggered by the use of conceptual knowledge since it does not arise from eye contact. However, if an interpretation of a smile has already somehow been achieved by matching visual input to a perceptual representation, what work is left for the embodied simulation to do?

Furthermore, how is the perceptual representation selected that is used to give meaning to the visual input? Niedenthal et al. have endorsed an embodied or grounded account of perceptual processing. Therefore, when they talk about conceptual knowledge triggering an embodied simulation, they must mean that some reactivated multi-modal representation is what triggers an embodied simulation. However, they don’t explain how visual input leads to the reactivation of the specific multi-modal representations that provide us with the interpretation of a smile. Once again, an appeal is made to a mapping from visual input but this time to a multi-modal representation, and it is by means of this mapping that we come to assign a meaning to a smile. However, there is no account given of the mechanisms that might bring about such a mapping.

Could it be that the problem of interpreting the smile is already settled at Stage 1 in perception, and this perceptual interpretation is what subsequently causes motor mimicry and its associated effects? Consider a parallel problem of determining the goal of an instrumental action. Csibra (2007) has argued that prior to an embodied simulation of an instrumental action is a stage of processing in which a visual analysis is formed of an action and the context in which the action is taking place. He hypothesises that it is on the basis of this visual analysis that the goal of the action is determined. Perhaps a comparable visual analysis takes place in the case of smiles in which contextual information is combined with information gathered from complex visual cues to arrive at an interpretation of a smile’s meaning.

This is not to say that embodied simulation makes no contribution to our understanding of expressive behaviour. It might make it possible for us to respond to a smile warmly and share in the emotion the smile expresses. In the absence of an embodied simulation, our response to and understanding of an emotion is by comparison “pale, colourless and destitute” to paraphrase William James. The Hennenlotter et al. (2005) study (cited by Niedenthal et al. in sect. 3.2.2) would seem to provide some support for this suggestion. Subjects prevented from mimicking expressions of sadness and anger by means of a BOTOX injection exhibited less limbic system activation than controls. Thus, mimicry certainly plays a causal role in generating emotion; and that may, in turn, affect the character of a person’s affective understanding. What remains unclear, however, is whether we understand smiles by sharing another’s feelings, or whether we can share in another’s feeling only because we have already understood the meaning behind their smile.

The dynamic interaction of conceptual and embodied knowledge

doi:10.1017/S0140525X10001329

Daniël Lakens\textsuperscript{a} and Kirsten I. Ruys\textsuperscript{b}

\textsuperscript{a}Faculty of Industrial Engineering & Innovation Sciences, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands; \textsuperscript{b}Psychology Department, Utrecht University, 3584 CT Utrecht, The Netherlands.

D.Lakens@tue.nl \textsuperscript{b}K.I.Ruys@uu.nl

web.tue.nl/staff/nl/d.lakens http://www.goallab.nl

Abstract: We propose the SIMS model can be strengthened by detailing the dynamic interaction between sensorimotor activation and contextual conceptual information. Rapidly activated evaluations and contextual knowledge can guide and constrain embodied simulations. In addition, we stress the potential importance of extending the SIMS model to dynamic social interactions that go beyond the passive observer.
In describing the processes that contribute to the meaning of a perceived smile, Niedenthal et al. astutely capture the complex nature of emotion processing. We applaud the approach taken by the authors in the current article: Embodied perspectives on cognition are in need of exactly such a detailed investigation of the interaction between conceptual and embodied knowledge. The question is not whether, but how: sensorimotor processes and cognition interact (e.g., Fodor 1985). The SIMS model illustrates how embodied responses to facial emotional expressions facilitate recognition of these expressions. We believe, however, that a SIMS model that explicitly addresses the dynamic interaction between conceptual and embodied knowledge would provide an even more useful framework to develop and test specific predictions.

Given that an embodied response to smiles (e.g., motor activation in the facial muscles) emerges around 500–1000 msec after observation of a smile (Dimberg & Pettersson 2000; Dimberg et al. 2000), it seems important to integrate more rapid cognitive processes in the SIMS model. First, before a smile is even observed, contextual information might play an essential role in determining which brain areas are activated when eye contact is established (e.g., OFC activation when talking to a loved one), and therefore, the type of smile that is expected. Second, the evaluative connotation of a smile and more specific conceptual information are activated rapidly at the moment that a smile is observed (within 120 msec; Ruys & Stapel 2008b); see also Palermo & Rhodes 2007.

Considering that information processing occurs as efficiently as possible, it seems unlikely for embodied simulations to emerge for the purpose of recognition when associative knowledge structures rapidly enable the recognition of functional smiles. Although Niedenthal et al. acknowledge that contextually driven, rapid processing occurs, they also argue that in many circumstances embodied simulations allow observers to infer the correct meaning of ambiguous smiles, when conceptual or stereotypical knowledge would lead to errors of interpretation.

The second possibility is that rapidly activated contextual and embodied knowledge would provide an even more useful framework to develop and test specific predictions. When we see our own children smile different functional smiles. Although this is an empirical question, the current literature suggests the OFC is activated when seeing our offspring display any emotional reaction, such as smiling, but also crying (Noritsui et al. 2008). We believe the SIMS model needs to incorporate the interaction between contextual conceptual information and embodied simulation to account for how people differentiate between the different meanings of functional smiles.

What is the function of the embodied simulation of an affiliative smile, when contextual conceptual information and rapid evaluative processing can often determine the meaning of such a smile much more efficiently? We believe that a SIMS model that goes beyond the passive observer can provide an answer by examining the role embodied simulations play in the social dynamics in which smiling behavior typically occurs. The simulation of smiles is not only a motor process that contributes to understanding, but also a process that serves a communicative function (Ruys & Stapel 2008a) involving social cognitive mechanisms (Schilbach et al. 2008).

As described by Niedenthal et al., social contextual variables can sometimes inhibit the simulation of an observed smile. Indeed, the automatic imitation or representation of the actions of others has been shown to be moderated by social information (e.g., Longo & Bertenthal 2009; Ruys & Aarts, in press; van Baaren et al. 2003). Although speculative, OFC activation and rapid processing of the evaluative connotation of a smile may together determine whether social inhibition of the motor simulation occurs or not. By determining whether corresponding somatosensory experiences are produced on a dyadic level, embodied simulations can contribute to the meaningfulness of social interactions (e.g., Parkinson 1996).

Understanding the meaning of smiles can be further determined by a range of bodily and cognitive factors, such as the observer’s somatosensory experience at the time a smile is observed, or the observer’s expectations and goals. We believe that understanding in this emergent process, which is guided and constrained by the dynamic interaction between embodied and conceptual information (see also, Mesquita et al. 2010). Specifying this interaction should be the focus of future theoretical and empirical work.

**Show your teeth or not: The role of the mouth and eyes in smiles and its cross-cultural variations**

doi:10.1017/S0140525X10001263

Chao Liu,a,b Yue Ge,a Wen-Bo Luo,b and Yue-Jia Luoa

aState Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, 100875, China. bDepartment of Psychology, University of Michigan, Ann Arbor, MI 48109-1109.

Liuchao@umich.edu

guye818@163.com wenbo9390@sina.com Luoyj@bnu.edu.cn

http://www-personal.umich.edu/~liuchao/index.html

http://psychnbrain.bnu.edu.cn/teachcms/luoyuejia.htm

Abstract: Previous studies with Westerners have found that both the mouth and eyes are crucial in identifying and interpreting smiles. We proposed that Easterners (e.g., Chinese and Japanese) evaluate the role of the mouth and eyes in smiles differently from Westerners. Individuals in collectivist Eastern society heavily rely on information from the eyes to identify and interpret the meaning of smiles.

A major potential limitation of the SIMS model is, as authors Niedenthal et al. have explicitly stated, “The SIMS model has been largely developed using data collected in Western countries. Nevertheless, it is essential to note that cultural differences may modulate our account” (sect. 6, para. 3, emphasis ours). Here we echo this proposal and discuss how people in Eastern cultures
differently evaluate the role of the mouth and eyes in smiles, and how such variations might modulate the SIMS model.

In the target article, the authors concentrate on the eyes and eye contact; however, in Western cultures, the mouth is as crucial as the eyes in identifying and interpreting facial expressions, especially happiness. The most unequivocal evidence comes from patient S.M., who has severely impaired detection of anger and fear after losing the bilateral amygdala, but still maintains an entirely normal identification of happiness (Adolphs & Tranel 2000; Adolphs et al. 1994). This dissociation is due to his inability to use information from the eyes but an intact ability to use information from the mouth (Adolphs et al. 2005). It therefore provides direct evidence that the mouth alone could supply sufficient information for the recognition of happiness and smiles, which has also been confirmed by the computational analysis of FACS-coded faces (Smith et al. 2005). The second vein of evidence derives from facial mimicry and the mirror neuron system. Facial mimicry is critical in constructing embodied simulations for smiles, and thus plays an important role in the SIMS model, yet the majority of studies on facial mimicry manipulate only the mouth and related zygomatic muscles (Blairy et al. 1999; Maringer et al., in press). In addition, mimicry and embodied simulation have been closely linked to the mirror neuron systems (Keysers & Gazzola 2007), whereas studies of the mirror neuron system are essentially focused on actions from the mouth and hands, such as drinking, eating, and grabbing (Iacoboni et al. 1999; 2005). Thus, it is not surprising that in the SIMS model, the mouth is the apt alternative of the eyes for observers in the social inhibition condition (see Figures 3 and 6 of the target article).

However, so far most data have been collected from Western cultures. There is evidence that Easterners evaluate the role of the mouth and eyes differently from Westerners. For example, “do NOT show ones’ teeth when smiling” was a strict rule of discipline for ancient Chinese women, who even used adornments (e.g., fake dimples) to make up for the scarcity of emotional information conveyed by the mouth during their closed-mouth smiles (see our Fig.1, Top). Another example is the smile emoticons used by Easterners and Westerners. While Westerner use emoticons exaggerating the mouth with a crimped line and simplifying the eyes as two dots, e.g., :-) or :), Japanese use emoticons with simplified mouth but crimped eyes, e.g., (^_^) or (^_^) (Jack et al. 2009; Pollack 1996; Yuki et al. 2007). Chinese, especially females, go even further by evoking the ancient tradition of attaching fake dimples, for example, (^^) (Marshall 2003).

Figure 1 (Liu et al.). **Top**: Ancient Chinese women used fake dimples to embellish their closed-mouth smiles (Xinhuanet 2008). **Bottom**: Chinese participants who focused on the eyes in a real versus fake smiles identification task showed a negative correlation between accuracy and individualism scores, but a positive correlation for collectivism scores.
Motivational aspects of recognizing a smile

doi:10.1017/S0140525X10001330

Janek S. Lobmaiera and Martin H. Fischerb

aDepartment of Psychology, University of Bern, CH-3012 Bern, Switzerland;
bSchool of Psychology, University of Dundee, Dundee, DD1 4HN, Scotland, United Kingdom.

janek.lobmaier@psy.unibe.ch m.h.fischer@dundee.ac.uk

http://www.kog.psy.unibe.ch/content/ueber_uns/lobmaier/index_ger.html
http://www.dundee.ac.uk/psychology/people/academics/mhfischer/

Abstract: What are the underlying processes that enable human beings to recognize a happy face? Clearly, featural and configural cues will help to identify the distinctive smile. In addition, the motivational state of the observer will influence the interpretation of emotional expressions. Therefore, a model accounting for emotion recognition is only complete if bottom-up and top-down aspects are integrated.

Most of the research on facial expression has focused on negative emotions, such as anger, fear, or disgust (e.g., Adams et al. 2003; Mathews et al. 2003; Putman et al. 2006), while positive emotions have received relatively less attention. The attempt to examine the nature of smiles by Niedenthal et al. is therefore highly welcome. To assist with further development of SIMS, we highlight its omission of observer bias and the importance of measuring the time course of motor resonance.

Niedenthal et al.’s proposal as to how a smile may be processed by an observer is rather one-sided. Importantly, they do not take into account the observer’s motives and suppositions. As the authors point out, analysis of the facial features may not always be sufficient to classify realistic emotional expressions, and then the perceiver must call on alternative sources of information, such as the social situation in which the expression occurs. More than that, the perceiver will evaluate and interpret the expression on the basis of his or her emotional state and social setting. Specifically, if an observer is in a happy mood, then an emotion expressed by another person will be interpreted differently compared to when the observer is sad or angry.

According to the affect-as-information theory (e.g., Schwarz 1990), mood is used to make social judgments. When a person is in a certain mood, he or she may inadvertently attribute that mood state to other people, which in turn will influence the interpretation of an observed facial expression. In fact, Niedenthal herself (Niedenthal et al. 2000) found an effect of emotional state in a task where participants had to detect at which frame of a movie a facial expression (e.g., happiness) became neutral. When the participant’s mood was congruent with the initially displayed emotion, this expression was perceived to persist longer than when participant mood and initial stimulus expression were incongruent. Hence, a happy person will more readily interpret emotion is not too extreme) a given emotional display as happy (at least as long as the display was available). In the 1950’s psychologists argued that what we perceive is influenced by our motivations or “set” (the so-called “New Look”). For example, the size of more valuable coins was overestimated, and hungry participants were quicker than satiated participants to identify food-related words (for critical discussion, see Erdelyi 1974). Applying this insight to our present task, the motivational tendency, such as what we wish to see, might also influence the recognition of smiles. Supporting this claim, we recently found a self-referential positivity bias when judging another person’s happiness. In the context of the SIMS model, expressions may be interpreted as more positive, just because this is what we wish the expression to be: not only will we perceive a smile as directed to us, we might also interpret a facial expression to be more positive than it really is.

The SIMS model explains recognition of smiles as a mainly stimulus-driven process. This approach is in contrast with various current models of emotion recognition. For example, the appraisal model of emotion (e.g., Scherer 2001) postulates that stimuli undergo evaluation checks on the basis of various dimensions: novelty (is a stimulus novel, predictable?), intrinsic pleasantness (positive or negative evaluation), goal significance (is it relevant to pursue my goal?), coping potential (can I deal with it?) in combination with social and personal standards (is it acceptable for me and others?). In the context of happy emotions, this means that (unconscious) cognitive evaluation of

ACKNOWLEDGMENTS

This work was supported by the NSF China (30930031); the Ministry of Education, China (PCSIRT, IRT0710); National Key Technologies R&D Program (2009BAI77B01); and by the Global Research Initiative Program, NIH, USA (1R01TW007897) to Yue-Jia Luo.
Grounding the meaning of non-prototypical smiles on motor behavior

doi:10.1017/S0140525X1000155X

Timothy A. Mann and Yoonsuck Choe

Department of Computer Science and Engineering, Texas A&M University, College Station, TX 77843-3112.
mann23@tamu.edu choe@tamu.edu
http://faculty.cse.tamu.edu/choe/

Abstract: We address how the motor system can contribute to the "meaning" component of smile perception. A smile perceiver can ground the meaning of non-prototypical smiles by interacting with the presenter to maintain the presenter's type of smile. In this case, the meaning of that smile is congruent with the motor behavior that elicits that smile (such as a funny gesture).

Niedenthal et al. define embodied simulation as "a simulation of a state in the motor, somatosensory, affective, and reward systems that represents the meaning of the expression to the perceiver" (sect. 1, para. 4, italics in the original). In this sense, the state of the sensory-motor system classifies the smile ("represents the meaning"), but the meaning of the classification to the perceiver must also be grounded. Although prototypical facial expressions may be grounded in intrinsic reward systems, the subtleties of non-prototypical smiles likely must be learned and, therefore, may be grounded in the socio-behavioral context surrounding the smile: Smiles are not presented or perceived in isolation; they are surrounded by rich socio-behavioral context. For example, it is difficult to consider the meaning of dominance smiles without reference to more than facial expressions.

Like other forms of communication, smiles can have profound indirect consequences. For example, Sharlemann et al. (2001) investigated how smiles can be used to negotiate cooperation between humans. Smiles (and, more generally, facial expressions) provide clues about the values and objectives of other agents. Because smiles have only negligible direct physical impact, the meaning of a smile cannot be fully realized by considering only facial sensory-motor systems. Instead, we propose that an important component of the meaning of a smile must be grounded in high-level behaviors. But this raises a question: How can the meanings of smiles be grounded in high-level behaviors?

Choe and Smith (2006) present a framework that explains the process of learning the meanings of internal sensory states by grounding them in terms of specific patterns of motor action, actions that maintain invariance in the internal sensory state. Our Figure 1 shows the framework, and Figure 2 shows how invariance in internal state can link up with motor-based meaning. The action that maintains invariance in the agent's internal sensory state (e.g., 45° movement in Figure 2) exactly corresponds to the meaning conveyed by the encoded internal state.

In Choe and Smith (2006), a visuomotor learning system was presented with the problem of grounding a set of visual states on its motor system, when the meaning of the visual states and motor primitives were initially unknown. A useful criterion in maintaining sensory invariance while acting. The sensory invariance criterion defines a probabilistic mapping from sensory states to motor primitives. In this experiment, some motor primitives maintained sensory invariance (e.g., by moving the agent parallel with the current sensory pattern), while other motor primitives violated sensory invariance (e.g., by moving perpendicular to the current sensory pattern). The agent learned to maintain sensory invariance while viewing a scene (see Fig. 2), which resulted in actions that moved the agent's visual field parallel to the current visual stimulus. The learned meaning of each sensory state was the set of motor primitives that resulted in maintaining that sensory state. Learning was accomplished with a standard reinforcement learning algorithm, and produced a mapping from visual states to congruent motor primitives (see sect. 3). This pairing between sensory and motor primitives grounds sensory representations using action. The meaning of a sensory state is given by the actions that maintain it, and this meaning can be learned through interactions with the environment.

This approach can be extended to external sensory stimuli, such as the smiling face of a conspecific. When viewing a non-prototypical smile, a perceiver (visuomotor learner) can interact with the smile presenter, with the objective of maintaining the type of smile on the smile presenter's face. For example, when a perceiver sees an enjoyment smile, the objective is to maintain the enjoyment smile (i.e., sensory invariance). By interacting with the presenter (e.g., telling a joke, duplicating the perceived smile, or frowning), the perceiver can learn a mapping from the particular smile and his/her own motor behavior that are likely to maintain the smile type. Like the mapping shown in Figure 2, the meaning of a non-prototypical smile can be grounded by the actions (of the perceiver) that maintain it (in the presenter).

Unlike the computational experiment from Choe and Smith (2006), where the initial mapping is randomly generated, some meaningful structure for smile interpretation may already be innate or learned from demonstrations of other people interacting in response to smiles. For example, prototypical smiles may be innately grounded, increasing likelihood of certain interactions when those smiles are observed. This prior knowledge about the range of interactions with a smile presenter could help reduce the search space needed to learn the meaning of non-prototypical smiles.

While the target article explains that the meanings of smiles are grounded in embodied cognition, the authors are not specific about which sensory-motor systems ground the meanings of smiles. We propose that, although some meaning can be captured by the facial sensory-motor system, important aspects of smile meaning cannot be captured without grounding smiles with respect to high-level actions making up the behavioral context. We have also described how a simple criterion (i.e.,
maintaining sensory invariance) can be used to ground the meanings of smiles on high-level actions.

The proximate mechanisms and ultimate functions of smiles

doi:10.1017/S0140525X10001561

Marc Mehu\textsuperscript{a} and Karim N'Diaye\textsuperscript{b}

\textsuperscript{a}Swiss Center for Affective Sciences, University of Geneva, CH-1205 Geneva, Switzerland; \textsuperscript{b}Research Center of the Brain and Spine Institute, Université Pierre et Marie Curie (UPMC), CNRS UMR 7225, INSERM UMR S 975, 75006 Paris, France.

marc.mehu@unige.ch http://www.affective-sciences.org
karim.ndiaye@upmc.fr

Abstract: Niedenthal et al.’s classification of smiles erroneously conflates psychological mechanisms and adaptive functions. This confusion weakens the rationale behind the types of smiles they chose to individuate, and it obfuscates the distinction between the communicative versus denotative nature of smiles and the role of perceived-gaze direction in emotion recognition.

Although we commend Niedenthal et al. for their effort to bring together the ethological functions of smiling and the neurocognitive mechanisms involved in its perception, we remain sceptical about their success, for the model they propose rests on a classification of smiles that is inconsistent with primate ethology and evolutionary theory.

Our reading of the ethological literature leads us to question the three types of smiles that Niedenthal et al. individuate (i.e., enjoyment, affiliative, and dominance smiles). In particular, the case made for the dominance smile category is based on anecdotal evidence and post hoc rationalization of disparate research.
A better coverage of the abundant literature on the relationship between smiling and social status would have made a stronger argument in favor of, or against, a “dominance” smile category. Although dominant individuals have also been observed smiling during interactions with subordinates (LaFrance & Hecht 1999), smiling has often been associated with submission (Hall 2002; Schmid-Mast & Hall 2004). It is surprising to see in this target article a category of dominance smiles while the psychological and ethological literature could equally well support arguments in favor of a submission smile category (e.g., Mehu & Dunbar 2008), which would also be distinct from the affiliative smile that Niedenthal et al. define. Finally, we doubt that smiling functions to establish dominance. The ethological literature rather suggests that smiling mitigates the deleterious effect dominance may have on the maintenance of egalitarian relationships (Flack & de Waal 2007; Preuschoft & van Hooff 1997).

A major asset of the target article is to propose a classification of smiles based on their function, rather than one based on their morphological features (sect. 2.2). However, from an evolutionary viewpoint, this functional typology is problematic because it mixes proximate and ultimate explanations of behavior (Tinbergen 1963). For example, the enjoyment smile is a category based on one of the proximate factors involved in smile production (i.e., positive emotion), whereas the “affiliative” and “dominance” smile categories reflect the ultimate function of smiles. As these criteria pertain to two different levels of explanation, they are not exclusive of each other. Indeed, research has shown that Duchenne smiles (also called enjoyment smiles) could be reliable signals of cooperative (and therefore affiliative) intentions; the honesty of the signal would be ensured by the emotional component that is difficult to strategically manipulate (Brown et al. 2003, Mehu et al. 2007). The alternative view that enjoyment smiles are merely cues, that is, symptoms of positive emotional experience without communicative function (Ekman 1997), could not be invoked to justify the distinction between enjoyment smiles and affiliative smiles for the reason that cues are not functionally related to a receiver’s response (Hasson 1994) and are, therefore, difficult to integrate within the sender-receiver process that Niedenthal et al. describe.

The confusion of proximal and ultimate factors leads to misunderstanding the crucial issue about the nature of smiles as communicative (smiles as social signals) versus denotative (smiles as symptoms of emotions) facial expressions. This is also illustrated in the case of dominance smiles in which Niedenthal et al. conflate the bodily expression of a positive feeling of superiority and the communicative intentions directed at social challengers. This conflation bears direct consequences on the embodied simulation processes that Niedenthal et al. claim to be central for interpreting the meaning of smiles. For example, their model could with difficulty explain how the perceiver of a dominance smile experiences negative feelings that have no obvious correspondence with the feelings actually experienced by the smiler, whereas alternative accounts regarding the functional role of such smiles (e.g., defusing social conflict) escape this contradiction.

We believe that an appropriate classification of smiles should be based on EITHER proximate causes (smiles with different motivational basis) OR ultimate causes (smiles with different adaptive functions). In the latter case, the classification is more laborious because the function of smiling is poorly documented in the literature (a situation that may explain Niedenthal et al.’s difficulty in achieving a satisfactory functional classification of smiles). Alternatively, a classification based on morphological and dynamic aspects of smiles may be useful (e.g., Duchenne vs. non-Duchenne). Such a categorization would be more relevant to the SIMS model, since the processing of a smile’s physical features is central to embodied simulation (see sect. 3.2 of the target article).

Because gaze represents an important vector of communicative intentions, the overlook of communicative/denotative distinction also weakens the theory regarding the role of eye contact. According to Niedenthal et al., eye contact would be a sufficient, but not necessary, trigger of the embodied-simulation process: When eye contact occurs, it triggers simulation of the current expression leading to embodied simulation and a better understanding of smiles’ meaning. On the other hand, when eye contact is obstructed, the simulation is based on conceptual knowledge and is more prone to error (sect. 5, para. 5 and sect. 6.2, para. 4). This implies that a simulation triggered by eye contact is more likely to result in accurate interpretations of smiles, making eye contact slightly more critical (if not necessary) than presented by the authors.

We are therefore surprised that Niedenthal et al. make little reference to the rich literature on eye contact and gaze direction on emotion recognition. Various studies on the effect of gaze direction on emotion recognition have shown that eye contact may indeed increase perceived emotion (relatively to averted gaze), but it may also decrease it, depending on the nature of the emotional expression (see Bindemann et al. [2008] for recent research and references therein). These data hardly fit with a general enhancing role of eye contact in emotion perception.

We propose that Niedenthal et al. consider an alternative account based on the functional role of eye contact according to which the effect of eye direction on emotion perception reflects the congruency of the sender’s gaze direction with the relevance of the emotional signal for the receiver: Angry (and happy) faces are more relevant to the receiver when she is gazed at (i.e., direct gaze), while fearful faces with averted gaze may signal a relevant threat for the receiver (Sander et al. 2007). This view implies that the role of eye contact occurs at a later stage of emotional expression recognition than that posited by Niedenthal et al. Instead of having a ubiquitous enhancing effect, eye contact would thus increase or decrease the perceived intensity of the emotional display, depending on its relevance for the observer. Therefore, the involvement of brain structures such as the amygdala need not reflect the uncertainty associated with smiles (as Niedenthal et al. presuppose; see sect. 6.1.1, para. 1) but, instead, the self-relevance of emotional expressions in their context of occurrence (N’Diaye 2009).

Conceptual knowledge: Grounded in sensorimotor states, or a disembodied deus ex machina?

doi:10.1017/S0140525X10001573

Ezequiel MorSELLA, a,b Carlos Montemayor, b Jason Hubbard, a and Pareezad Zaroliaa

a Department of Psychology; San Francisco State University, San Francisco, CA 94132-4168; b Department of Philosophy, San Francisco State University, San Francisco, CA 94132-4168; c Department of Neurology, University of California, San Francisco, CA 94143.

morsella@sfsu.edu cmontema@sfsu.edu jashubbard@gmail.com pzarolia@gmail.com

http://bbs.sfsu.edu/emorsella/

Abstract: If embodied models no longer address the symbol grounding problem and a “disembodied” conceptual system can step in and resolve categorizations when embodied simulations fail, then perhaps the next step in theory-building is to isolate the unique contributions of embodied simulation. What is a disembodied conceptual system incapable of doing with respect to semantic processing or the categorization of smiles?

One challenge encountered when reverse engineering the brain is that the same function (e.g., categorizing an enjoyment smile) can be carried out by vastly different mechanisms (e.g., visual
feature analysis, conceptual processing, or embodied simulation). This gives the brain tremendous flexibility and the ability to compensate for failed mechanisms. To the detriment of the scientist, it also makes it difficult to experimentally isolate the mechanism(s) underlying a given function (e.g., the categorization of smiles) and identify the conditions under which a given function will be carried out by one mechanism versus another. With this in mind, Niedenthal et al. should be commended for their trailblazing contribution that specifies both the conditions under which embodied simulation is likely to occur (e.g., following eye contact) and the neural-physiological components of this intricate and counterintuitive mechanism.

When appreciating such an advancement, it is worth keeping in mind that the primary explanatory value of embodied approaches to knowledge representation (e.g., Barsalou 1999; Glenberg 1997) rested in their ability to address the symbol grounding problem (Harnad 1990) faced by "propositional" (aka "amodal") accounts of mental representation (e.g., Landauer & Dumais 1997), in which meaning is represented by arbitrary symbols (e.g., binary digits) that do not retain any of the properties of the sensorimotor states that gave rise to them (see review in Markman & Dietrich 2000). Similar to "analogical" and "modal" accounts, embodied models propose that meaning is constituted by (or "grounded in") sensorimotor information, such as that furnished by facial expressions. Since their advent, embodied models have been challenged on empirical grounds. Regarding emotion, for example, an extensive study revealed that people with bilateral facial paralysis stemming from Moebius syndrome can categorize the facial expressions of others without difficulty (Bogart & Matsumoto 2010). Moreover, the ability of subjects to express facial expressions does not influence accuracy. Similarly, in semantics, it remains unclear whether lesions of action-related regions of the brain lead to impairments regarding the semantic-conceptual knowledge of action; Patients suffering from deficits due to brain lesions can show impairments in the way they use certain objects but, revealing the abilities of a "embodied" conceptual system, can nonetheless name those objects, or even recognize the pantomimes associated with them (Mahon & Caramazza 2008).

In response to such challenges, embodied approaches now propose that sensorimotor information plays a more humble role: In the target article, Niedenthal et al. portray sensorimotor information not as that which constitutes meaning, but merely as auxiliary information that is "referred to" when semantic processing is challenged, as when categorizations are difficult to make. In these accounts, sensorimotor information is demoted to a kind of supplementary information that, just like any kind of information, can at times facilitate or interfere with semantic processing. When the strategy of embodied simulation fails, it is the conceptual system that arrives at the scene to solve the categorization, much like a deus ex machina.

The question now is whether the conceptual system is always necessary for categorization. From one standpoint, the answer is yes: Until a token facial expression is tipped as a member of a given category (an inherently conceptual process), the organism cannot know (consciously or unconsciously) the meaning of the observed expression. If the conceptual system must always be at the scene, then what is its nature and how does it solve the symbol grounding problem? Embodied approaches were expected to answer this by demonstrating, for example, that disrupting the sensorimotor states in which a conceptual process is grounded will not just modulate the process, but also eradicate it. It seems that, apart from modulations that could be construed as priming or its opposite (i.e., interference), there is no evidence of conceptual processing requiring sensorimotor states. Apart from the aforementioned neurological evidence, this is evident with respect to abstract concepts (e.g., that of if and justice) and everyday phenomena such as imageless thought (e.g., a state of uncertainty or readiness; cf. Hilgard 1987). Regarding imageless thought, Woodworth (1915) concluded, "I know of no reason in neurology or psychology for supposing that the elements of conscious content are contributed solely by the sensory receiving centers" (p. 137).

A way forward: Identifying the limitations of "deembodied" processing. Perhaps a critical next step in theory-building for embodied approaches is to identify the unique contribution(s) of embodied simulation. What is it that a disembodied conceptual system is incapable of doing with respect to semantic processing or the categorization of smiles? Emotion research has broached this topic by revealing some limits of language, a form of propositional processing. Evidence suggests that linguistic processing does not lead to the same kinds of emotional effects and forms of affective learning as modal stimuli (LeDoux 1996). Olsson and Phelps (2004) demonstrated that, through vicarious classical conditioning, subjects can acquire a learned fear response toward a subliminal stimulus by perceiving someone else receive a shock after being presented with that stimulus. However, simply telling subjects about the contingency between stimuli did not lead to this form of vicarious conditioning to subliminal stimuli. This suggests that the phylogenetically old systems mediating fear conditioning may not understand language, though they can process the meaning of basic perceptual events: "classical conditioning and observational learning . . . might be supported by an evolutionarily old system that predates the emergence of language" (Olsson & Phelps 2007, p. 1099).

That language fails to yield the vicarious conditioning afforded by observational learning may be because language comprehension yields "faint" activations of analogical, sensorimotor-like symbols (Boroditsky & Ramscar 2002; Zwaan 2008) or, at odds with embodied accounts, because linguistic-semantic knowledge is simply distinct from sensorimotor knowledge. Regarding the pioneering investigation by Niedenthal et al., the next advance may be to isolate that which only embodied simulations can achieve. It may be that aspects of social mimicry and empathy require embodied simulation, or that the three smile-categories may operate differently with respect to subliminal or supraliminal stimuli, with the conceptual deus ex machina being capable of arriving at the scene only when the theater lights of consciousness are on.

Abstract: Many species’ non-threat facial expression involves an open mouth and retracted lip corners—the smile. This served to make an accompanying vocalization sound like it originated from a smaller signaler undermines the notion that the perception of the smile employs embodied simulation of the smiler’s state.

The nature and function of the smile—indeed, of any facial expression—is of interest not only in psychology and the neurosciences, but also in ethology, which is the study of behavior, especially that exhibited in similar form cross-species and which can be explained by Darwinian natural selection (e.g., Andrew 1963; Bolwig 1964; Eibl-Eibesfeldt 1971).

For example, Morton (1977) showed that avian and mammalian threat and non-threat (approasement, submission) vocalizations use fundamental frequency (F0) in remarkably similar ways: threat vocalizations have low F0, whereas non-threat vocalizations have high F0. A dog’s low-pitched threatening growl and its high-pitched submissive yelp are familiar examples. Morton

What’s behind the smile?

doi:10.1017/S0140525X10001585

John J. Ohala
Department of Linguistics, University of California, Berkeley, CA 94720-2650. ohala@berkeley.edu
http://linguistics.berkeley.edu/phnlab/users/ohala/index3.html

Abstract: Many species’ non-threat facial expression involves an open mouth and retracted lip corners—the smile. This served to make an accompanying vocalization sound like it originated from a smaller signaler. That such signals are deceptive and benefit primarily the signaler undermines the notion that the perception of the smile employs embodied simulation of the smiler’s state.

The nature and function of the smile—indeed, of any facial expression—is of interest not only in psychology and the neurosciences, but also in ethology, which is the study of behavior, especially that exhibited in similar form cross-species and which can be explained by Darwinian natural selection (e.g., Andrew 1963; Bolwig 1964; Eibl-Eibesfeldt 1971).

For example, Morton (1977) showed that avian and mammalian threat and non-threat (approasement, submission) vocalizations use fundamental frequency (F0) in remarkably similar ways: threat vocalizations have low F0, whereas non-threat vocalizations have high F0. A dog’s low-pitched threatening growl and its high-pitched submissive yelp are familiar examples. Morton
argued that the modification of F0 conveys to the perceiver an estimate of the body size or maturity of the vocalizer and thus the degree of threat because, other things being equal, body size and maturity are correlated with the mass of the vibrating membrane (vocal cords or syringal membranes). These deceptions benefit both parties by averting injury (or worse) resulting from actual combat.

I demonstrated that this “frequency code” applies to human speech as well: the degree of perceived dominance is inversely correlated with voice F0 (Ohala 1984). Furthermore, this accounts for well-documented common cross-linguistic pattern (some would say “universal”) that questions, uncertainty, and pleading are marked by a high F0 somewhere in the utterance (when these intentions are not indicated by lexical, morphological, or syntactic means). In addition, I hypothesized that the frequency code helps to resolve a long-standing puzzle about the mouth shapes accompanying threat and non-threat displays in species having plastic (i.e., movable) facial masks: canids and primates, including humans. An open-mouth and retracted lip corners is found in non-threat displays, as opposed to more protruded, rounded lip shape in threat displays. Why in non-threat should more teeth – potential weapons – be displayed in contrast to fewer teeth showing in a threat display (Izard 1971)? The frequency code explains these dichotomous facial expressions, too: These different lip configurations will modify the resulting resonances of an accompanying vocalization in a way to convey to the perceiver an impression of a vocalizer with shorter or longer vocal tract This, I argue, is the origin of the smile. Via the well-known process of ritualization, the smile can be given without an accompanying vocalization or even with a closed mouth. This analysis meets the condition of coherence and generality in ethological explanations advocated by Darwin (1872, p.15): “whether the same principle by which one expression can, as it appears, be explained, is applicable in other allied cases, and especially, whether the same general principles can be applied with satisfactory results, both to man and the lower animals.” Thus, consider a case most of us are familiar with: a submissive dog. The vocalization has a high F0, high resonances (due to mouth corner retraction), ears flattened, no bristling of the hair on the back, tail between the legs, crouching posture – the latter elements all designed to make the animal seem as small as possible. A dog’s threat display is just the opposite.

That modification of the vocal anatomy is important in conveying an impression of size is reinforced by the fact that, besides the plastic modification of the vocal structures, there are implastic modifications evident in the sexual dimorphism in the adult vocal tract. The mature human male has a longer vocal tract than the female (by some 15 to 17%) – this is accomplished by the familiar descent of the larynx starting at puberty – and the male’s vocal cords become fully 50% longer – but probably 100% more massive than the female’s. The first modification accounts for the quantitatively different resonances of the “same” vowels as spoken by adult males and females. The latter modification accounts for the typical one octave difference in males’ and females’ average speaking F0. The longer vocal cords, by the way, are bought at the cost of enlargement of the laryngeal cartilages in the male, resulting in a prominent “Adam’s apple.”

Numerous other species – from whooping cranes, geese, and the howler monkey – show sexually dimorphic aspects of the vocal anatomy which testify to the importance of vocalization in agonistic encounters, in sexual selection, and so on. Thus, enhancement of various aspects of adult males, human and nonhuman, gives rise to such sexually dimorphic aspects as the human male’s beard, the male lion’s mane, and the size dimorphism between male and female elephant seals, and the like. Size does matter.

What does all this imply for the target article by Niedenthal et al.? Just this: If, as I have argued, a smile is induced by the smiler to induce a reaction or response in the perceiver that is favorable to the smiler, then the authors’ contention that “observers of smiles sometimes construct an embodied simulation of the nuanced affective state conveyed by the smile” (introduction, para. 3) is undermined. My view is similar to that espoused by the “functionalist” approach to the analysis of the outwardly visible states – call them “emotions” if one will (Campos et al. 1994). One must differentiate between (a) inner neurophysiologic states – which, to further the interests of the one experiencing these states, often would be best kept secret – although in some cases uncontrollable cues to such states may inadvertently “leak out,” for example, sweating, tremor, blushing, sexual arousal; and (b) signals such as smiles, frowns, threatening looks, and vocalizations, which are made in order to induce a targeted reaction/response in the perceiver – that is, a response favorable to the one emitting the signal. Ethology teaches us: stereotyped behavior – and this includes signaling – is generally done to benefit the individual or group exhibiting it. Such signals are not, strictly speaking, cooperative.

What’s embodied in a smile?

doi:10.1017/S0140525X10001597

Disa A. Sauter and Stephen C. Levinson
Max Planck Institute for Psycholinguistics, 6500AH Nijmegen, The Netherlands.
disa.sauter@mpi.nl stephen.levinson@mpi.nl
http://www.mpi.nl/people/sauter-disa
http://www.mpi.nl/people/levinson-stephen

Abstract: Differentiation of the forms and functions of different smiles is needed, but they should be based on the empirical data on distinctions that senders and receivers make and on the physical cues that are employed. Such data would allow for a test of whether smiles can be differentiated using perceptual cues alone or whether mimicry or simulation are necessary.

The target article proposes an ambitious model of the perception of facial signals, specifically the interpretation of smiling expressions of amusement, dominance, and appeasement. Niedenthal et al. rightly draw attention to the need for differentiation in the forms and functions of different types of smiles, but the distinctions the authors propose lack support from empirical findings, as they themselves concede. This limitation undermines the proposed model. We propose that firm foundations for a model of the perception of emotional signals should be based on the distinctions that senders and receivers make, the cues that are actually employed, and the correspondence between physical cues and subjective experience.

One study has specifically investigated the perception of the three types of smiles that Niedenthal et al. suggest (Hess et al. 2002). When asked to pair facial expressions with vignettes describing dominance, appeasement, and amusement scenarios, Canadians of European extraction associated the amused vignettes with strong Duchenne smiles, but no difference was found between the smiles that were selected for the dominance and appeasement scenarios. For both dominance and appeasement vignettes, responses were distributed across a range of weak- and medium-intensity smiles. These findings demonstrate that amusement is associated with a smile expression, a finding that has recently been replicated cross-culturally and extended to the auditory domain (Sauter, in press; Sauter et al. 2010). However, Hess et al.’s (2002) results suggest that viewers do not reliably differentiate between dominance and appeasement smiles, or, alternatively, that the facial expressions associated with these states are not well established. Niedenthal et al. base their model on the three functionally defined smile
categories – amused, dominant, and appeasement smiles – but as we have illustrated, it is not established that viewers can differentiate between these three expressions.

There seem, however, to be other distinctions in positive emotion expressions for which there is better empirical grounding. Recently, a growing number of studies have started to distinguish between signals of a range of positive emotions (see Sauter 2010), providing evidence for smile categories other than those suggested by Niedenthal et al. For example, in a production study of posed positive emotion displays, Shiota et al. (2003) found that amusement was expressed via open-mouthed smiles, whereas pride was associated with smiles with compressed lips. This finding suggests that displays of some positive affective states are signalled by physically distinct smile configurations, although it did not establish whether viewers are sensitive to these cues. A recent study investigated both the production and perception of spontaneous amused, embarrassed, nervous, and polite smiles, by investigating physical cues and human judgments (Ambadar et al. 2009). Ambadar et al. showed that viewers use variation in morphological and dynamic characteristics of different kinds of smiles. For example, in comparison with smiles perceived to signal politeness, smiles that were perceived by viewers as amused more often included open mouth, larger smile amplitude, larger maximum onset and offset velocity, and longer duration. This demonstrated that viewers’ judgments were directly related to the physical cues that differentiated between these expressions. Together, these studies suggest that distinct physical cues are associated with smiles signalling different positive emotions, and that viewers are sensitive to these cues, but the categories are not coincident with those employed in the target article.

Regardless of the specific smile types employed, participants can typically use perceptual, conceptual, and embodiment processes in making these judgments. Additional empirical evidence is necessary for assessing whether an embodiment model actually fits the data. Specifically, Niedenthal et al.’s SIMS model proposes that smiles activate neural regions that cause motor mimicry and somatosensory experience, which form the basis of the viewers’ interpretation of the smile they see. Some of the data cited by Niedenthal et al. in support of their model are suggestive of emotional mimicry. However, these data are also compatible with the possibility that motor cortex activation in emotion perception tasks reflects downstream associations, rather than playing a primary perceptual role. The proposed model does little to rule out the possibility that viewers may differentiate between different smiles by using perceptual cues alone (perhaps together with conceptual information), before (or in the absence of) any mimicry or simulation.

The authors cite evidence for a causal role of somatosensory cortex in emotion perception from a study showing that transcranial magnetic stimulation (TMS) of this area selectively interferes with performance in a facial expression matching task (Pourtois et al. 2004). However, the manipulation selectively affected the reaction time of judgments of fearful expressions, but did not affect the judgments of happy expressions. Furthermore, no interaction was found with the gaze direction of the faces, and no effect was found on participants’ accuracy. The results of the study therefore do not seem to demonstrate a causal role for somatosensory cortex in emotion perception.

In sum, we argue that differentiations of smiles should be based on findings establishing which distinctions senders and receivers make, and what physical cues are actually used by viewers in their judgments. Furthermore, convincing demonstrations that motor cortex activation plays a causal role in emotion perception are lacking, and the possibility remains that viewers can differentiate between smiling expressions from perceptual cues alone, without recourse to simulation.

**Is eye contact the key to the social brain?**

doi:10.1017/S0140525X10001275

Atsushi Senju and Mark H. Johnson
Centre for Brain and Cognitive Development, Birkbeck College, University of London, London WC1E 7HX, United Kingdom.
a.senju@bbk.ac.uk
mark.johnson@bbk.ac.uk

**Abstract:** Eye contact plays a critical role in many aspects of face processing, including the processing of smiles. We propose that this is achieved by a subcortical route, which is activated by eye contact and modulates the cortical areas involved in social cognition, including the processing of facial expression. This mechanism could be impaired in individuals with autism spectrum disorders.

---

![Figure 1](image.png)

**Figure 1.** An illustration of the fast-track modulator model (FTM). Perceived eye contact (upper left) is initially detected by a subcortical route that projects to various regions of the social brain network (thick black lines). This signal from the subcortical route then interacts with contextual modulation based on the task demands as well as the social context (thick grey lines) to modulate the response of these regions to the following input from a cortical route (thin black lines). These pathways are based on previous analyses of cortical and subcortical face processing, as well as top-down voluntary attention. OFC = orbitofrontal cortex, PFC = prefrontal cortex, STS = superior temporal sulcus. Reproduced with permission from Senju and Johnson (2009b).
The Simulation of Smiles (SIMS) model proposed by Niedenthal et al. emphasizes the core role of eye contact, which is hypothesized to trigger embodied simulation of the perceived smile. The authors also speculated that the same mechanism may also mediate the processing of other facial expressions. However, eye contact is known to modulate a far wider range of cognitive processes, such as the encoding of gender, identity, and gaze (Senju & Johnson 2009b). We recently reviewed this phenomenon, which we termed the “eye contact effect,” and proposed the fast-track modulator (FTM) model to explain its neural and developmental basis (Senju & Johnson 2009b). In this commentary, we present a brief overview of the FTM model and discuss several areas in which the FTM model complements the SIMS model, and thus would facilitate further exploration of the neural, cognitive, and developmental mechanisms underlying the effect of eye contact on face processing.

The FTM model proposes that the eye contact effect is mediated by a subcortical face detection pathway hypothesized to include the superior colliculus, pulvinar, and amygdala. This route is fast, operates on low spatial frequency visual information, and modulates cortical face processing (see our Figure 1). Evidence that the fast route is fast comes from event-related potential and magnetoencephalographic studies showing that components associated with this pathway can occur at shorter latencies than those usually associated with the cortical processing of faces (Bailey et al. 2005). In addition, evidence that the subcortical route modulates cortical processing comes from several functional imaging studies indicating that the degree of activation of structures in the subcortical route (amygdala, superior colliculus, and pulvinar) predicts or correlates with the activation of cortical face processing areas (George et al. 2001, Kleinhans et al. 2008). It has also been proposed that the subcortical route is also responsible for face preference in newborn infants (Johnson 2005) and even in adults (Tomalski et al. 2009). We hypothesize that the combination of this subcortical pathway and contextual modulation given by the task demands and social context directly or indirectly modulates key structures involved in the cortical social brain network.

The FTM model shares several key features with the SIMS model. However, there are several differences between these two models, wherein the FTM model expands and broadens the SIMS model. First, the FTM model proposes the neural mechanism linking eye contact and facilitation of cortical face processing, including the embodied simulation. The FTM model proposes that perceived eye contact directly activates a subcortical route, which then modulates the cortical areas involved in different aspects of social cognitive processing. Thus, it is possible to incorporate the SIMS model by arguing that the subcortical route also modulates the motor cortex, which controls mimicry. The FTM model also provides new predictions about the effect of eye contact on the processing of smiles: It should be fast and operate on low spatial frequency visual information.

Second, the FTM model can also provide alternative hypotheses about the mechanism by which eye contact facilitates the processing of smiles. The FTM model hypothesizes that the subcortical route receiving input from perceived eye contact directly modulates the cortical face processing areas. This contrasts with the SIMS model which assumes that eye contact must elicit the embodied simulation first in order to facilitate the processing of smiles. As we discussed above, the FTM does not rule out the possibility that the subcortical route mediates the embodied simulation in response to eye contact. However, the FTM also leads us to propose a more parsimonious hypothesis: that the subcortical route directly modulates visual cortical areas, which then facilitates the processing of facial expression, including smiles. For example, the FTM model predicts that eye contact modulates the processing of smiles even when the activation of the motor cortex is experimentally suppressed. By contrast, the SIMS model would not predict that eye contact facilitates the processing of smiles under this condition, because embodied simulation is suppressed.

Third, the FTM model presents a unique perspective on the development of the eye contact effect. In the target article, Niedenthal et al. have suggested an interesting hypothesis that the preference for eye contact in infants reflects an evolutionary-based mechanism for triggering embodied simulation, even though they do not discuss how such a mechanism develops. By contrast, the FTM model assumes that infants are born with widespread connections between the subcortical route and cortical structures. As a consequence, input from perceived eye contact initially activates widespread cortical structures, which combines with architectural bias to form specific connections between the subcortical “eye contact detector” and relevant cortical and subcortical structures during the course of development. Interestingly, recent studies on the early development of autism spectrum disorder (ASD), with manifest atypical patterns of eye contact behaviour, are consistent with the predictions based on the FTM model. Even though infants and young children with autism show apparently typical eye contact behaviour (Chawarska & Shic 2009), neuroimaging studies demonstrate more widespread and non-specific cortical activation in response to eye contact (Elsabbagh et al. 2009), and behavioural studies demonstrated that eye contact does not facilitate cognitive processing in children with ASD (Senju et al. 2003). These studies suggest that infants and young children with ASD are sensitive to eye contact, but that it fails to modulate cortical face processing in the same specialized way as typically developing children (Senju & Johnson 2009a). Future studies will need to test whether eye contact elicits facial mimicry and affects the processing of smiles in individuals with ASD, especially since current evidence is inconsistent as to whether individuals with ASD show spontaneous facial mimicry (Magnée et al. 2007) or not (McIntosh et al. 2006, Oberman et al. 2009).

Eye contact plays a critical role in face-to-face communication, and we propose it is the key to adaptively modulate many aspects of social cognition, including the processing of smiles. We hope the areas of overlap and contrast between the SIMS and the FTM models will generate empirical studies, and help further understand the neural, cognitive, and developmental mechanisms underlying human social behaviour.

**ACKNOWLEDGMENTS**

Atsushi Senju was supported by an ESRC Research Fellowship (RES-063–27–0207), and Mark H. Johnson was supported by the UK Medical Research Council (G0701484).

**Can we really leave gender out of it? Individual differences and the Simulation of Smiles model**

doi:10.1017/S0140525X10001342

Elizabeth Simpson and Dorothy Fragaszy

Department of Psychology, University of Georgia, Athens, GA 30602-3013.
simpson@uga.edu doree@uga.edu
http://simpson.myweb.uga.edu
http://psychology.uga.edu/people/bios/faculty/Dorothy_M_Fragaszy.php

**Abstract:** Gender differences in face-based emotion recognition, notably differential use of mimicry, may compromise the extent to which the Simulation of Smiles (SIMS) model can be generalized to populations besides the adult females on which it has been tested. Much work indicates sex differences in face-based emotion recognition, including smile recognition.
One problem for the Simulation of Smiles Model (SIMS) proposed by Niendenthal et al. is that there are individual differences in face-based emotion recognition (hereafter emotion recognition), mimicry, and embodied cognition. Given that these are important features of the model, the model's predictive power may vary as a function of these features. Sex is one important individual difference that needs to be considered (Cahill 2006; Vigil 2009).

Though facial mimicry appears to improve emotion recognition (Niendenthal et al. 2001; Oberman et al. 2007; Stel & van Knippenberg 2008) – consistent with the SIMS model – the link between mimicry and emotion recognition has only been reported in females. This could be because women, compared to men, rely more on facial feedback for emotion recognition, at least for some expressions. That is, when facial mimicry is inhibited, women's emotion recognition suffers, while men's does not (Stel & van Knippenberg 2008). Whereas women appear to rely more on this fast and efficient "hot" (embodied) simulatory route for emotion recognition, men may rely more on one or more "cold" (rule-based, non-embodied) strategies. This may account for the finding that women are better than men at distinguishing a number of facial expressions, including smiles (Krumhuber et al. 2007).

Behavioral and neuroimaging evidence is congruent with the notion of sex differences in emotion recognition strategies. For example, men and women have different face viewing patterns during emotion recognition (Hall et al. 2010). While women look more at the eyes, compared to men, men have longer first fixations to the eyes and mouth, perhaps reflecting men's slower processing speed for information contained in those areas (Hall et al. 2010). Additionally, there are anatomical differences between men and women in brain regions containing mirror neurons, which underlie simulation during emotion recognition (Goldman & Sripada 2005). During empathetic face-to-face interactions, women, compared to men, experience more activation in brain regions containing mirror neurons (Schulte-Rüther et al. 2008). Moreover, brain regions containing mirror neurons differ in men and women, both structurally and functionally (Cheng et al. 2009).

Researchers examining sex differences in emotion recognition should use caution, for the complexity of measuring emotion recognition may not be apparent. For example, stimulus materials should be selected carefully because women are more sensitive to dynamic, compared to static expressions, while men are not (Biele & Grabowska 2006). The gender of face stimuli also influences the perceived emotion: both men and women are better at detecting negative expressions on male faces and positive expressions on female faces (Becker et al. 2007). Also, males and females are not mimicked to the same degree: females are mimicked more often than males (e.g., Lakin et al. 2008).

To complicate things further, the likelihood of mimicking depends on top-down influences, including one's conscious or unconscious goals, such as to recognize a face, to attract a potential mate, to understand a spoken language, or to affiliate (e.g., after social rejection). These different motivations alter our viewing patterns of faces and the feature to which we attend (e.g., Saether et al. 2009; Schyns et al. 2002). If we are attending to a non-emotional aspect of a face – such as skin color – we mimic less (Cannon et al. 2009). Men may attend more to non-emotional aspects of faces – such as attractiveness – while women may attend more to emotional aspects. Men's empathy also appears to be mediated by social-cognitive factors that do not influence women's empathy. For example, similar brain regions are active when experiencing pain and when viewing pain in another individual, for both sexes (Singer 2006); however, men no longer display the same neural pattern when the other person is perceived as behaving unfairly (Singer 2006). This suggests that men, compared to women, may more readily inhibit embodied emotion recognition strategies in such circumstances.

One encounters difficulties when trying to interpret the literature on sex differences in emotion recognition. Studies of mimicry often: (a) use only female participants, (b) use samples that are too small to allow comparisons of sex differences, (c) fail to test whether there are gender differences, or (d) fail to report the proportion of males and females in the sample. If there are only female participants, it remains untested whether effects would likewise occur in males. For example, females' embodiment during facial expression viewing has not been tested in male participants (Efron et al. 2006). Although others have assumed no gender differences, the body of work reported here suggests this assumption may be incorrect.

It would be fruitful to know whether encouraging men to mimic – rather than relying on their natural tendency to do so – improves their emotion recognition in either speed or accuracy. Interestingly, administering testosterone reduces facial mimicry in females (Hermans et al. 2006); however, we know of no work which has examined whether increased testosterone impairs emotion recognition in either sex. This also raises the question of whether men's emotion recognition would be improved if they were instructed to attend to the eyes – something that women do naturally – since the frequencies of looks to the eyes is positively correlated with emotion recognition accuracy (Hall et al. 2010). Developmental work is crucial for identifying the emergence of sex differences in face viewing patterns, mimicry, and emotion recognition, to determine the extent to which such behaviors and skills are continuous or changing throughout life.

We agree with Niendenthal et al. that embodied simulation is an important strategy for emotion recognition. We also concur that it is wise to examine each facial expression of emotion separately when modeling the processes underlying emotion recognition. In addition, we believe it is of utmost importance to consider individual differences – and especially sex differences – in the use of simulatory processes for each type of expression.

---

**Abstract:**

The parent-infant dyad, characterized by contingent social interactions that develop over the first three months postpartum, may depend heavily on parental brain responses to the infant, including the capacity to smile. A range of brain regions may subserve this social key function in parents and contribute to similar capacities in normal infants, capacities that may go awry in circumstances of reduced care.

Perhaps preceding smiling is crying, a signal believed to notify parents to attend to safety, acquire additional resources, and honestly indicate need and vigor (Soltis 2004). Parents will then interact with their infants in cry-care loops (Swain et al. 2004) in order to negotiate their return to contentment. Thus, crying and smiling may be considered as unconditioned responses in habit, to which various forms of conditioning are introduced by the environment – the parents’ behaviors in particular – in a gradual, increasingly contingent manner.

The parental capacity to adapt to the postpartum may depend on highly conserved hypothalamic-midbrain-limbic-paralimbic-cortical circuits acting in concert to support parent-response to...
Commentary/Niedenthal et al.: The Simulation of Smiles (SIMS) model

Embodied simulation and the search for meaning are not necessary for facial expression processing

doi:10.1017/S0140525X10001603

Jacob M. Vigil and Patrick Coulombe
Department of Psychology, University of New Mexico, Albuquerque, NM 87131-1161.

Niedenthal et al. emphasize a functional perspective on facial expression processing of variant forms of smiling behaviors. We remain skeptical, however, of their central thesis, that simulative sensitivities module the autonoetic representation or epistemic “meaning” of other people’s facial expressions. The authors provide no direct support for how simulation or mimicry of musculature movement patterns is necessary for effectively responding to, and hence processing of, human facial expressions. Rather, facial expression processing can be achieved through lower-cognitive, heuristical perceptual processing and expression of prototypical morphological musculature movement patterns that communicate discrete trustworthiness and capacity cues to conspecifics.
judgments. Moreover, we feel that the authors overemphasize the role of higher-order cognitive abilities such as the need to infer meaning of self- and other-generated expressive behaviors. Certainly, all socially-embedded species as well as human infants are able to mimic and effectively respond to (e.g., with avert gaze) their social surroundings via heuristical processes that do not require epistemic-supporting mental faculties. More parsimonious conceptualizations of facial expression processes suggest cognitive representations of meaning play only a supplementary role compared to perceptual and behavioral heuristics in producing and responding to discrete facial musculature movement patterns such as smiles.

One set of models that may alternatively explain how observers process different types of smiles is the Socio-Relational Framework of Expressive Behaviors (SRFB; Vigil 2009). According to this paradigm, both the appraisal and the expression of nonverbal behavior can be understood in terms of two primary dimensions of affective responses. The first dimension is the heuristical motivation to either attract or dissuade interactions and relational bonding with others, depending on whether other people signal the potential to be a reciprocating social partner or whether they signal interpersonal threat. The second dimension is the appraisal (in the perceiver) and the phenotypic display (in the expresser) of what has been referred to as capacity and trustworthiness cues. Capacity cues signal one’s aptitude or ability to inflict beneficial or detrimental consequences on others and are usually associated with the impression of dominance. Trustworthiness cues, in contrast, signal one’s interpersonal intentions and motivations and are usually manifested as appeasement or vulnerability displays that are associated with impressions of submissiveness. Several scientists have insisted that the perceptual processing of capacity and trustworthiness cues in others is both necessary and sufficient for predicting interpersonal (e.g., affiliative or avoidant) dispositions and discrete emotive responses toward others (e.g., Fiske et al. 2006; Rosenberg et al. 1968; Todorov 2008; Vigil 2009; Wojciszke 2005). As shown in Figure 1, trustworthiness perceptions are parsimonious predictors of the basic desire to either affiliate with, or to avoid, social objects (see also Wojciszke 2005). As hypothesized by the SRFB (Vigil 2009, p. 379), simultaneous perceptions of high- versus low-capacity levels can be used to predict specific and discrete affiliative (admiration vs. sympathy) and avoidant (disgust vs. fear) reactions toward social objects (Fig. 1).

According to the SRFB, because disgust and admiration project the impression of dominance, while fear and sympathy project the impression of submissiveness (see Leary 1957; Marsh et al. 2005; Montepare & Dobish 2003), these emotions reflect the proclivities to advertise the capacity and trustworthiness components of one’s own reciprocity potential to others (Vigil 2009). Thus, it is possible to interchange socially relevant information through heuristical perceptual and behavioral algorithms, independent of executive mental faculties. Rather, simulative and representational abilities probably evolved to anticipate the outcomes of social situations that can result in
variable consequences such as discussion of third-party affiliates or monitoring future patterns of others’ behaviors (Geary 2005). Specific and functional responding to, and hence processing of, other people’s facial expressions does not require the monitoring of phrasing consequences, but rather instant and uniform (approach/avoidant) responses to the immediate behavioral states of others, irrespective of representational meaning.

In this sense, discrete facial expressions, including smiling, should correspond to features of the mouth and eye regions that signal capacity and trustworthiness cues to others (Vigil 2009). For example, people may signal capacity cues through the baring of teeth (e.g., Duchenne smiles that demonstrate bilateral symmetry, healthiness, and weaponry) and through direct threat stares, which may operate to conceal eye sclera (e.g., demonstrating the capacities to concentrate on the environment, to endure the risk of inciting interpersonal evaluation and potential conflict, and the motivation to protect the eye). Likewise, trustworthiness cues can be demonstrated through the concealment of teeth (e.g., closed-mouth appeasement smiles) and exaggeration of displayed sclera, which may instead operate to explicate one’s vulnerability (e.g., from dust particles and sun glare) and allow observers to easily monitor the signaler’s attention, thereby reducing one’s threat impression to others (fears are also produced from wide-eyed positioning). Humans display the teeth and conceal the sclera in correspondence with dominant emotions (e.g., joy and anger), and they conceal the teeth and expose the sclera in correspondence with submissive emotions (e.g., sadness, fear, surprise, and sympathy). Thus, heuristical presentation and concealment of teeth and sclera across discrete emotive gestures may have evolved in part to signal capacity and trustworthiness cues to others, and ultimately promoting interpersonal affiliation and avoidance. These possibilities support an alternative model to Niedenthal et al.’s proposal, by showing that the cognizant awareness of and the epistemic search for the meaning of emotional sensations may not be necessary for effective responding to different types of smiles and nonverbal expressive behaviors more generally.

Embodied and disembodied processing of emotional expressions: Insights from autism spectrum disorders

doi:10.1017/S0140525X10001640

Piotr Winkielman
Department of Psychology, University of California, San Diego, La Jolla, CA 92039-0199, and Department of Psychology, Warsaw School of Social Sciences and Humanities, 03-815 Warszawa, Poland.
pwinkiel@ucsd.edu
http://psy.ucsd.edu/~pwinkiel

Abstract: Processing of facial expressions goes beyond simple pattern recognition. To elucidate this problem, Niedenthal et al. offer a model that identifies multiple embodied and disembodied routes for expression processing, and spell out conditions triggering use of different routes. I elaborate on this model by discussing recent research on emotional recognition in individuals with autism, who can use multiple routes of emotion processing, and consequently can show atypical and typical patterns of embodied simulation and mimicry.

Niedenthal et al. tackle the classic problem of how the mind processes the ubiquitous and important social stimuli – facial expressions. The problem is tricky because most expressions carry multiple social and emotional meanings. As the authors say, visually recognizing that an expression is a smile (as opposed to a frown) can be relatively easy. But deciphering its social relevance is often difficult. After all, to distinguish among an affiliative, enjoyment, and dominance smile, for example, one needs not only perform a finer visual discrimination, but also integrate one’s own emotional response with the information about the intent, social status, and veracity of the expressor. To advance research on this tricky problem, Niedenthal et al. present an integrative model, which, in a welcome opposition to the extant approaches, does not attempt to brush those complexities under the carpet. The authors’ model of expressions processing identifies various embodied and disembodied ways of accomplishing the task. The ways range from simple perceptual processes, such as pattern-matching, to complex conceptual processes, such as mentalizing. The most generative part of the model spells out some conditions that promote the use of embodied simulation, which relies on somatosensory and motor processes in the brain and the body. The authors point out the role of perceptual factors (e.g., difficult tasks benefit from motor simulation) but also social ones (e.g., eye gaze).

The goal of my commentary is to extend Niedenthal et al.’s proposals by discussing the issue of recognition and mimicry of facial expressions in individuals with autistic spectrum disorders (ASD). Although the target article touches on this issue (sect. 3.2.4), the treatment is necessarily brief and does not capture the full complexity of the empirical data. This is unfortunate because the ASD literature offers good support for the authors’ thesis that the engagement of embodied simulation processes is conditional, and is triggered by a mix of perceptual, conceptual, and social factors.

When do individuals with ASD mimic? Passive observation versus task engagement. Several studies examined facial mimicry in high-functioning individuals with autism spectrum disorder (ASD). Some found deficits in spontaneous mimicry. That is, in an unprompted setting, ASD individuals do not smile to a smile or frown to a frown. In one experiment, participants’ task was to “just watch” large pictures of happy and angry expressions presented for several seconds (McIntosh et al. 2006). Facial electromyography revealed congruent facial reactions only in control, but not ASD, participants. Importantly, ASD participants reacted typically when mimicry was explicitly encouraged, thus controlling for nonspecific motor or perceptual deficits. In another experiment, participants’ task was to “just watch” a video in which a student described his adventures in an amusement park, displaying happy expressions (Stel et al. 2008). Analysis of experimenter-coded facial expressions revealed less spontaneous mimicry in ASD than in control participants. This effect was observed despite no differences in the amount of time spent looking at the screen and no differences in voluntary mimicry. In yet another study, the task was a passive observation of 2-second presentations of several expressions (Dapretto et al. 2006). IMRI analysis showed that ASD individuals had less activity in several areas, including the prefrontal cortex (inferior frontal gyrus), amygdala, insula, and the striatum. In short, “just watching faces” does not trigger embodied simulation in ASD individuals. Interestingly, similar ASD impairments were found in studies that explored spontaneous mimicry of non-emotional actions (e.g., hand movements) with EEG measures of motor cortex activation, such as mu-wave suppression. Thus, ASD participants show a typical level of mu-wave suppression when performing an action, but less mu-wave suppression (i.e., less motor cortex activation) to observation of an action, as compared to typical participants (Oberman et al. 2005).

But the story is more complex. Studies of non-emotional mimicry suggest that ASD participants occasionally show “spontaneous mirroring.” For example, reduction in action imitation in ASD is found only for actions displayed from an egocentric perspective (away from the observer), but not from an allocentric perspective (toward the observer) (Theoret et al. 2005). In another example, ASD individuals show little EEG mu-wave suppression when watching an action performed by a stranger, but normal mu-wave suppression when watching an action by a family member (Oberman et al. 2008). In the domain of emotional mimicry, one study found no impairments of spontaneous mimicry, measured with EMG, when ASD individuals were given the task of integrating visual and auditory cues to
emotion (Magnee et al. 2007). A recent study showed that giving ASD participants a task of recognizing expressions (rather than just "watching them") led to a comparable level of mimicry as the typical group, even though ASD mimicry was slower (Oberman et al. 2009). In short, it appears that "task engagement" is necessary to trigger mimicry in ASD individuals.

But what is "task engagement"? What factors turn on and off embodied simulation? This is not simply the question of attention and general motivation. After all, studies established that individuals stay "on task." One idea is that factors like allocentric versus egocentric perspective, "family versus stranger," and "emotion identification versus observation" work similarly to "eye gaze," discussed in the target article. Specifically, all these factors switch the processing strategy from the disembodied, "pattern-matching" strategy, to one that presumably requires access to the underlying states of the individuals modeling the expressions or actions.

Of course, this idea needs to be tested directly. Here is one possibility. It is known that in perceiving expressions, ASD individuals rely more heavily on a "cold" rule-based strategy (Rutherford & McIntosh 2007). As a result, they are more likely than controls to accept as realistic the exaggerated images of expressions (smiles that are stronger than real smiles), presumably because those expressions represent "best fits" to the rule. However, this "rule-based" processing should diminish when they are dealing with faces that smile and look "at them" and are produced by family members, rather than total strangers. Importantly, the task needs to be realistically challenging – it cannot be easily solvable by a simple pattern-matching strategy (e.g., recognizing a distinct expression containing pure happiness, or a simple emoticon, like the Walmart's smiley). Under such conditions, the embodiment (mimicry) should contribute to emotion recognition. This is critical, because as Niedenthal et al. note, mimicry contributes to performance of typical participants only when the perceptual task is hard, the participant cares, and the subjective emotional response is a useful signal for the recognition task (Oberman et al. 2007).

I hope that this exchange inspires fruitful research that will benefit not only theoretical understanding of emotion processing, but also better research on atypical social functioning (see Winkielman et al. 2009 for further discussion).

Authors’ Response

The future of SIMS: Who embodies which smile and when?

doi:10.1017/S0140525X10002748

Paula M. Niedenthal,a Martial Mermillod,a Marcus Maringer,b and Ursula Hessc

aCentre National de la Recherche Scientifique (CNRS) and Clermont Université, 63037 Clermont-Ferrand, France; bDepartment of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands; cDepartment of Psychology, Humboldt-Universität Berlin, 12489 Berlin, Germany.

niedenthal@wisc.edu
http://wwwpsy.univ-bpclermont.fr/~niedenthal/
martial.mermillod@univ-bpclermont.fr
http://wwwpsy.univ-bpclermont.fr/~mermillod/
m.maringer@rug.nl
Ursula.hess@psychologie.hu-berlin.de
http://www.psychology.hu-berlin.de/staff/1683737

Abstract: The set of 30 stimulating commentaries on our target article helps to define the areas of our initial position that should be reiterated or else made clearer and, more importantly, the ways in which moderators of and extensions to the SIMS can be imagined. In our response, we divide the areas of discussion into (1) a clarification of our meaning of "functional," (2) a consideration of our proposed categories of smiles, (3) a reminder about the role of top-down processes in the interpretation of smile meaning in SIMS, (4) an evaluation of the role of eye contact in the interpretation of facial expression of emotion, and (5) an assessment of the possible moderators of the core SIMS model. We end with an appreciation of the proposed extensions to the model, and note that the future of research on the problem of the smile appears to us to be assured.

R1. Introduction

One of the problems with researching smiles in humans is that humans use language that communicates their complex causal attributions for smile behavior, and thus for its meaning. This fact, we believe, has strongly influenced the history of the scientific study of the smile in social psychology. This fact has also led us astray in some ways. The prime example of a complex attributional judgment is that a smile is "false" (versus "true"). The spontaneous use of these labels by observers has strongly influenced the course of social psychological research on the smile. Yet, even though researchers can create stimuli in the laboratory that have the properties of smiles that are labeled, perhaps even quite universally in laboratory experiments, as "true" versus "false," (or more or less genuine) it is probable that these are not the expressions occurring in daily life that are judged "true" and "false." Hence, we agree strongly with Fernández-Dols & Carrera that the "true" versus "false" distinction is not the best way to carve up the smile landscape (see also Hess & Kleck 1990). Bouissac clearly agrees with this point.

Indeed, when Ekman and Friesen first studied the importance of the Duchenne marker (e.g., Ekman & Friesen 1982), they defined as "false" smiles the expressions made by individuals who were effortfully trying to deceive. Specifically, the individuals whose smiles were ultimately the subject of study were trying to cover up distress in response to an anxiety-producing film or situation, with the goal of convincing perceivers that they were actually experiencing positive affect. When individuals were engaged in this deception, there was little evidence of the Duchenne marker. Ekman et al. (1990) further showed that EEG recordings of individuals who expressed non-Duchenne smiles revealed evidence of negative affect or of conflict (Wacker et al. 2010). In any event, the conclusion was that smiles called "false" were sometimes associated with the subjective experience of negative affect in the smiler.

But there would probably be no agreement about the attributed causes of "true" and "false" universally. In recently collected cross-cultural data, one of our North American female respondents reported that even if she is having a bad day when at her job, she wants people to feel good, and so she smiles. Is this a false smile or a true smile? It is false in so far as it does not reflect how she "feels in general today," and so perhaps some individuals or cultures would call her smile "false." But it is true in that she has a positive social motivation. She thinks those smiles are "true."

Lurking under the "true" and "false" distinction are therefore three plausible categories of attribution upon
which the judgment relies: the feeling state of the smiler (positive, neutral, negative), the intended outcome of the smile interaction for the smiler (positive, neutral, negative) and the intended outcome of the smile interaction for the perceiver of the smile (positive, neutral, negative). The relative weight of each of these dimensions in determining the judgment that a smile is “false” versus “true” probably varies across culture. Our instinct, based on responses to some of our open-ended questions, is that in North American culture the attribution that the smiler intends harm to the perceiver of the smile is the dominant information that drives the “false” judgment. However, in other cultures, perhaps in Europe, attributions based on other dimensions are also likely to be equal grounds for the judgment of “false.”

So, all we really know about the terms “true” and “false” as regards smiles, is that they are used. We also know that we can construct smiles that look more or less true and false, and receive those of meaning judgments judgments. But we do not know more than that, and perhaps that is not the best way to go. Indeed, the SIMS was not developed to distinguish between the type of person or culture that would judge a smile to be “true” or “false.” That is what our own cross-cultural data collection is about. The SIMS model is an attempt to define the behavior and brain processes that support certain general classes about smiles. As regards the judgment that a smile is “false” in the core model, we argued that dominance smiles are not likely smiles of positive affect; and so, depending upon the culture, the simulation of a dominance smile may usually support a judgment of “false” (i.e., in particular, when the judgment scale that is used is a scale of “authenticity” or “genuineness”). But any attention we paid to the “true” and “false” distinction was meant to lend coherence to the literature. As stated, we believe that a different approach to carving up the smiles space is more useful, and that is why we promote a functional account of smiles. A parallel semantic analysis would yield the most convincing knowledge base, of course (as suggested by Bouissac).

R2. Functional arguments

Our ethologist commentators such as Mehu & N’Diaye, Ohala, as well as Sauter & Levinson and Centorrino, Djemai, Hopfensitz, Milinski, & Seabright (Centorrino et al.), contest in different ways our use of “functional” language in defining the three major smile types of interest to the SIMS. They also bemoan our lack of data (so do we, but we are working on it). It is therefore our burden, first, to define “functional” in our account. Further discussion of our typology follows in section R3.

For ethologists, functional is synonymous with “adaptive.” But it is also the case that the discussion of the relationship between function and form, among ethologists and evolutionary theorists, is not settled or consensual. The debate is, in fact, a quite lively one (e.g., Allen & Bekoff 1995; Ovren & Rendall 2001). We are aware that it is reckless to make functional (i.e., adaptive) arguments about any species that is living in a habitat in which it did not evolve, such as human beings living in suburban environments (Moffatt & Nelson, 1992). Hence, we have been most influenced by emotion researchers who use the language of functionalism in a way that is relevant to humans and was initially motivated by responses to the idea that emotions disrupt reason and are generally harmful psychological events (Niedenthal et al. 2006). These researchers include Frijda, Scherer, Oatley, Fischer, Haidt, Tracy, and Keltner, among many other modern emotions researchers.

These researchers’ definition of function with regard to emotion and emotional expression is consistent with Keltner and Gross (1999) or Scherer (1987), who hold that emotions are solutions to problems and opportunities related to physical and social survival. Those authors further state “Functional ascriptions, therefore, refer to the history of a behavior, trait, or system, as well as its regular consequences that benefit the organism, or more specifically, the system in which the trait, behavior, or system is contained” (Keltner & Gross 1999, p. 469).

Social functionalist accounts propose specifically that two very important challenges for human survival include (1) attachment to caretakers and potential mates, and (2) integration into groups. Another important challenge is the establishment and maintenance of social hierarchies. The assumptions are that human survival depends upon group membership and that long-term group functioning requires effective leadership (Keltner & Haidt 1999). The importance of the smile as a foundational social glue used in the solution of these problems is made even more precisely by some of our own commentators. For example, as Swain & Ho note, the baby’s smile is a highly physiologically rewarding stimulus to a parent. It helps the parent hang in there. Additional research, reported by Swain and colleagues (e.g., Lenzi et al. 2009; Strathearn et al. 2009; Swain 2008), provides even more information about the circuits by which the brain supports responding to attachment objects in particular.

A social functional account does not shun comparative psychology or ethology, however, and we rely on it for our account of the landscape of the human smile. Indeed, as de Waal (2003) has suggested, Darwin was on to something when he focused on facial expression of emotion as a good candidate test of his evolutionary theory. Along with de Waal, we consider the primate and comparative psychology literatures to be useful for motivating our distinction among smiles as conforming to enjoyment smiles, affiliative smiles, and dominance smiles (noting that these are names for expressions that yield classes of human attributional judgments that can be summarized as reflecting enjoyment, affiliation, and dominance).

R3. Our smile typology

But there are a number of conceptual problems to be resolved, especially about the proposed category of dominance smiles. Some of those issues require collaboration with ethologists. One question has to do with the evolution of displays of teeth in primates and the role of teeth display in human smiles. There has been confusion in this regard because of the initial reliance on the study of rhesus macaques, which have a strong linear, hierarchical societal structure. Referring to Darwin, de Waal (2003) writes:

For example, he noted that the bared-teeth expression ... by a black Sulawesi macaque, occurs when the animal is pleased to
be caressed. Retraction of the lips to expose both rows of teeth is indeed a relaxed, friendly expression in this species as opposed to the same expression in most macaques, in which it signifies submission. How do we know this? Quantitative analysis of natural social interaction sequences among Sulawesi macaques demonstrates that the bared-teeth display predicts the onset of affiliative contact between sender and addressee, hence that it likely is associated with a positive social attitude. In these macaques, teeth-baring often occurs mutually between individuals. In the better known rhesus macaque, in contrast, teeth-baring is given exclusively by subordinate to dominant individuals – hence never mutually – and is a common response to threats and intimations ...
The colloquial term “fear grin” for all teeth-baring expressions derives from the familiarity of researchers with the rhesus macaque – the most common laboratory primate in the West – rather than from a comprehensive look at the primate order, in which this expression has a variety of meanings. (de Waal 2003, p. 10)

One of the important points and reasons for quoting this passage at length is to note that the use of teeth in a gesture such as the smile, and therefore the meaning(s) of the smile, varies even within closely related primates and depends upon the type of social structure in which they live. We can be motivated by this comparison, noting that the showing of teeth in the gesture called the smile is very complex. This variation exists even across human cultures. For instance, some Asian cultures teach the covering of the teeth during smiling and laughing, whereas other cultures do not.

The possibility that the same expression, such as the smile or the laugh, can have different meanings across and within cultures is further expressed by this quote:

The laughing expression of apes is clearly homologous with that of our own species; the laugh derives from a widespread mammalian play expression. As we have seen, however, homology does not necessarily imply that the expression functions in the same way in all hominoids (i.e., humans and apes). In bonobos and chimpanzees laughing is closely tied to play, whereas in our own species it occurs under a much wider range of circumstances. Playful interaction is obviously included and can be considered the original laughing context, but we use the same expression also in bonding (i.e., “laughing with”) and, sometimes, as a hostile signal (i.e., “laughing at”) (de Waal 2003, p. 17).

These three categories of laughter conform to our own typology for the smile. Furthermore, we suggest that the relationship between laughing and smiling in humans may be a closer one than it is for other primates. Indeed, there may be a relationship between the human smile and the show of teeth as in threat, as suggested by Ohala; however, it seems likely, as we think may be implied by the previous quote, that the different human smiles have their basis in laughter (which also shows teeth). This does not deny the possibility, discussed by Liu, Ge, Luo, & Luo (Liu et al.), that some cultures hide the teeth possibly to avoid miscommunication of threat or else a violation of an undesired association to animals.

Among the many ways of validating this smile typology, we have recently begun a reanalysis of data collected for other theoretical purposes (Deborah Prentice, personal communication). The experiments in which the data were collected involved the presentation of norm-conforming and norm non-conforming oral statements by individuals who were ostensibly peers (students at the same university) of the participants. As the participants listened to the norm conforming and non-conforming statements, their faces were recorded by video, and four facial muscles were recorded with electromyography techniques. In addition to measures of the face, measures of the participants’ moods and their reactions to the peers whose comments they were listening to were also taken. One of the most striking things is that when listening to both conforming and non-conforming statements, participants smiled. The difference was that their perceptions of the non-conforming peers were negative. Their smiles were smiles of derision, not affiliation. We can test this impression by relating the use of specific muscles to the feelings and the judgments of the participants to ask whether individuals smile both out of affiliation and superiority and whether those smiles have different properties. Maybe different muscles are employed, and maybe the body is used differently (which we can code from the videos).

Evaluating the characteristics of smiles associated with judgments of enjoyment, affiliation, and dominance, will force us to push the SIMS to be more specific in its statements of the brain systems that ground smile meaning of course. In their commentary, Mann & Choe argue that, although SIMS holds that the meaning of smiles can be grounded in embodied simulation, it is not specific enough about which sensory-motor systems ground the meanings of smiles (a related argument is presented by Chang & Vermeulen). Mann & Choe propose that, although some meaning can be captured by the facial sensory-motor system, important aspects of smile meaning cannot be captured without grounding smiles with respect to high-level actions making up the behavioral context.

Mann & Choe indeed suggest an interesting way to ground the meaning of smiles in higher-level actions of the perceiver of the smile. They propose that when a perceiver sees an enjoyment smile, the objective is to maintain the enjoyment smile (maintain sensory invariance), for instance, by telling a joke. The meaning of a smile can then be grounded in the (higher-level) actions of the perceiver, which maintain the smile in the sender. In our target article, we focused mainly on the affective feedback of the facial sensory-motor system as a consequence of facial simulation. In that case, the meaning of the smile is grounded in the affective output of the simulation. In addition, we argue that the meaning of a smile can be grounded in the social and behavioral context in which a smile occurs. For example, a smile shown by a salesclerk who tries to sell a pair of shoes is seen as less genuine compared to a smile that is shown by a person who just sold a pair of shoes (Maringer et al., in press). Hence, similar to what Mann & Choe propose, we do argue that the meaning of smiles can be grounded in high-level actions of the sender (rather than the perceiver). Sensory invariance in this case might relate to the different contexts and behaviors that are likely to be accompanied by similar facial expressions. In this way, the commentary by Mann & Choe has extended our link between meaning of smiles and behavioral context of the sender to the link between meaning of smiles and the behavioral context of the perceiver. We believe that both behaviors (perceiver’s and sender’s behaviors) can be used to ground meaning of facial expressions.

Along with refining and validating the smile typology, it will be pertinent to specify the nature and role of
facial mimicry in classification, recognition, and interpretation of these facial expressions. Although we do believe that the existing data suggest a role for mimicry in these processes, along with Hamon-Hill & Barresi we feel that the role is far from clear. In addition to the problems raised by their commentary, even the question of the potential roles of motor efference versus reafference is still an open one.

**R4. Top-down processes in smile interpretation**

Commentators Evers, Noens, Steyaert, & Wagemans (Evers et al.); Lakens & Ruys; Lobmaier & Fischer; and Morsella, Montemayor, Hubbard, & Zarolia (Morsella et al.) worry in different ways about integrating top-down processes and individual differences in SIMS. Fortunately, the former is an important element of the model, and we have already advanced some ideas about the latter, which we elaborate upon in the section 4.

Top-down processes include the application of beliefs, stereotypes, expectations, and motivated biases. We have noted that such processes can provide important input into the interpretation of smile meaning. For example, Maringer et al. (in press) have shown that when people believe smiles are not genuine read-outs of positive feeling, these beliefs influence judgments of how genuine a given smile is (in a particular culture) when mimicry is inhibited. In fact, Hess et al. (1998) found that just evoking the notion that expressions may not be genuine, by asking participants to rate their genuineness, eliminated mimicry to the same expressions that were mimicked when participants rated the emotion expressed. Further, Halberstadt et al. (2009) have shown that when individuals believe that an ambiguous facial expression is actually communicative of a specific emotion, they later mimic the ambiguous expression in terms of that specific emotion.

Granted that top-down processes affect the lower-level perceptual and motor processes involved in encoding facial expression and deducing meaning, in our view the problem is how top-down and bottom-up processes communicate. Our solution was to suggest that both concepts and emotions can be grounded in states of sensory-motor systems and that interactions between these procedures arise naturally from this account. As we noted in section 5, paragraph 4,

the embodied account provides a natural way to link conceptual knowledge about smiles and the related social situation to the actual perception of smiles. Rather than assuming that a smile activates an amodal knowledge structure to represent its meaning, the embodied simulation account proposes instead that a smile triggers a simulation of a smile experience that includes emotion, motor activity, and somatosensory experience [...]. Once this simulation becomes active, it provides a conceptual interpretation of the perceived smile, going beyond the information given to place it in a context.

Of course, when an individual is not behaviorally interacting with a smile (eye contact and mimicry are inhibited), it is certainly possible that conceptual knowledge represented in language does all of the work (Boroditsky & Prinz 2008). For example, if one believes that Tony Blair’s smile is a smile of dominance, then without paying any attention to the smile whatsoever, one is quite capable of applying that concept as expressed in language.

In fact, as we note in discussing the SIMS model, there are a number of different ways that can be used to decode the meaning of expressions, and situational knowledge, or the knowledge about the emotionality of members of a group (see Hess & Kirouac 2000; Hess et al. 2009b), can replace simulation in situations where gaze contact is inhibited. Yet, even in such cases, simulation may still play a role. For example, Houde et al. (2009) have demonstrated that providing observers with verbal labels describing the emotional state of a person who shows a neutral expression, elicits emotion congruent facial expressions – in essence, mimicry – in the observer.

This issue is also addressed in research on disorders such as autism and of culture, where it is pertinent to establish which of the critical behaviors – eye contact and facial mimicry – is present or absent for those populations, and what beliefs are held, ready to be applied instead of or together with an embodied simulation. Note also that eye contact was proposed as a sufficient and developmentally primary trigger to embodied simulation of the smile. Research on individual differences will also need to determine other sufficient triggers of embodiment (such as motivation; see Mondillon et al. 2007). This is particularly true if it turns out to be the case that perceivers of dominance smiles tend not to make eye contact (even if ever so fleetingly) with those who display such expressions, as is asserted by Huang & Galinsky.

Winkielman, who focuses specifically on research on autism, rightfully points out that we are a long way from understanding the conditions that tend to favor an embodied simulation strategy for processing facial expression of emotion. We agree that the importance of successful communication, as sometimes determined by the significance of the relationship between the expresser and the perceiver, is a factor that determines the type of processing that grounds judgments of smile meaning. Individuals may devote more resources to understanding those with whom they have an intimate or interdependent relationship (e.g., Zajonc et al. 1988). But a rigorous and theoretically concise account of why that is so will be required. One avenue is to integrate some of the insights from the commentary by Vigil & Coulombe, particularly as regards the role of attachment in these processes. We do note that according to Vigil & Coulombe our central claim is that “simulation sensitivities modulate the autonoetic representation or epistemic ‘meaning’ of other people’s facial expressions”; however, we would not articulate the central claim of our target article in that way as the statement only addresses a small part of the model. Furthermore, like Winkielman, we do believe that the role of mimicry in interpreting the meaning of facial expression has received some impressive supportive evidence (e.g., from Maringer et al., in press; discussed earlier and in the target article).

**R5. What is it about eye contact?**

Aside from our proposed functional categorization of smiles and its links to social-attributional judgments, the idea that eye contact has developed as a developmentally
basic and subsequently sufficient trigger of embodied simulation gave rise to the most energetic discussion. There is much to work on here.

Senju & Johnson, for instance, have suggested that the SIMS model be integrated with their fast-track modulator (FTM) model. The FTM model holds that eye contact is mediated by a subcortical face detection pathway hypothesized to involve the superior colliculus, pulvinar, and amygdala (Senju & Johnson 2009). The root of this subcortical pathway is assumed to be constituted of alpha ganglion cells, at the origin of magnocellular layers and LSF subsequent cortical pathways, which are very fast but provide only low spatial frequency (LSF) visual information. This subcortical pathway has been confirmed for auditory information (Campeau & Davis 1995; Doron & Ledoux 1999) and visual information (Doron & Ledoux 1999; Linke et al. 1999; Shi & Davis 2001) in rats.

Concerning humans, different studies suggest the existence of a preferential link between LSF information and the emotional system, particularly threat detection. This plausible preferential link was obtained on the basis of neuroimaging (Morris et al. 1999; Pourtois et al. 2005; Vuilleumier et al. 2003), neural-network modeling (Mermillod, in press a; Mermillod et al. 2009), and behavioral experiments (Bocanegra & Zelzenberg, 2009; Holmes et al. 2005; Mermillod et al., in press b). However, whereas these studies hint at a preferential link between LSF visual information and emotional processes, possibly occurring at the level of the amygdala, they do not constitute formal evidence for the subcortical pathway assumed in Ledoux’s (1996) model. Therefore, it is speculative to incorporate this pathway in the SIMS model at this time.

However, this is obviously a very interesting way to develop the SIMS model and, more specifically, the complementarities between fast and automatic processes of relevant emotional stimuli in the environment, potentially operating at the level of subcortical structures and more complex associative processes occurring at the level of cortical areas. The interesting contribution of the FTM model is that it specifies the relationship between early visual processes and subsequent emotional processes related to eye contact. This link between perception and emotions should therefore be investigated.

Similarly, the question of automatic and unconscious response to eye gaze was raised by Chatelle, Laureys, Majerus, & Schankers (Chantelle et al.) with regard to severely brain-injured patients. These commentators propose that minimally conscious state (MCS) patients could be differentiated from vegetative state (VS) patients based on eye gaze. They propose that intentionality and different level of consciousness could be assessed through eye gaze. For our model, this raises the question of the consciousness of different level of processing. In relationship with the previous comment proposed by Senju & Johnson, we suggest the possibility of a direct, automatic, but also unconscious subcortical pathway for eye gaze orientation. This hypothesis can be supported by research on emotional blindsight (de Gelder et al. 1999; Pegna et al. 2004), showing that hemianopic patients might be able to detect emotions presented in their blind visual field above chance (de Gelder et al. 1999). An important point is that these patients are not conscious of this blind perception and feel that they respond at chance level. In other words, this subcortical pathway (if it exists) seems to operate beyond the scope of consciousness. Combining this finding with the FTM model (Senju & Johnson 2009), we can assume that automatic but uncontrolled eye gaze direction could be directed by subcortical structures like the superior colliculus, the pulvinar, and possibly the amygdala for automatic responses to environmental stimuli in VS patients, whereas MCS patients might be able to use more neural resources involving cortical and other subcortical structures. As suggested by Chatelle et al., we believe that investigation of the neural underpinnings of directed versus automatic eye gaze in VS versus MCS patients could be a very interesting way to investigate the neural basis of consciousness.

Finally, as highlighted by Chakrabarti’s comments, a model of smile perception involving the question of eye contact necessarily raises the question of parsimony. More than 30 emotional states can be associated with smiles (Golan & Baron-Cohen 2006), and it would be rather implausible to associate one neural module per type of emotional feeling. The goal of the SIMS model, however, was to propose general principles, associated with different neural pathways, to account for the emotional processing of smiles. We hope that this parsimonious approach, based on scientific evidence in support of our new theoretical framework, will be able to account for the larger set of empirical data relating to smiles. Among them, and as proposed by Chakrabarti, we will have to determine whether not only does eye contact act as a trigger for subsequent amygdala activity, but, conversely, amygdala activity can boost the search for relevant social cues and, therefore, eye contact. Chakrabarti also raises the question of determining whether embodied simulation processes are emotion specific. The SIMS model focuses on smiles, but as noted in the target article, we assume that similar processes are likely to occur in the recognition of other emotional expressions.

R6. Moderators and individual differences

The SIMS model was developed with the goal of defining general processes by which individuals attribute meaning to a smile. In particular, the model highlights the role of eye contact, mimicry, and the induction of specific brain states. However, along with our commentators, we encourage discussion of the potential impact of individual differences. Foremost amongst these are the differences of gender and culture, as emphasized by several commentators (Bouissac; Caldara; Grézes, & Sander (Conty et al.); Liu et al.; and Simpson & Fragaży).

The current development of the SIMS model does not discuss those influences in detail, even though we specifically note that “it is essential to note that cultural differences may modulate our account.” Clearly, we agree with the notion that individual differences have a role to play and can be expected to moderate some of the processes we discuss. As Covas-Smith, Fine, Glenberg, Keylor, Li, Marsh, Osborne, Soliman, & Yee (Covas-Smith et al.) note, cultural differences in eye contact preferences should entrain differential predictions based on the SIMS model. The question then arises – do we know enough about these differences to accurately predict their effect? And here the issue becomes rather
more complex. Caldara provides evidence that members of collectivist cultures fixate faces around the nose area—
that is, do not actually seek full eye contact. By contrast, Yuki et al. (2007) note that Japanese participants weigh
eye information more heavily than mouth information, and similar data are reported by Liu et al. as being
under review. Just this short summary of data presented in the commentaries shows that at this point we may not
be able to predict the impact of culture on eye gaze and, consequently, on its role for smile interpretation. What
is needed are intercultural studies focusing not only on emotion decoding accuracy but also more precisely on
the process of interpreting facial expressions. The SIMS model provides a framework for such research.

For example, Hess and Kirouac (2000) have proposed that when individuals do not know each other, they tend
to resort to stereotype knowledge about social group members when decoding facial expressions. We have
already considered whether this is a general rule or whether this fact indeed varies across culture; and how
this fits with the SIMS. One of our major cultural hypotheses relates to the processing of in-group versus out-group
facial expressions (e.g., Niedenthal & Maringer 2009). The foundational history of some societies—sometimes called
“settler societies” where the focus is on the fact that a land occupied by an indigenous peoples was taken over and
settled by people from other cultures and nationalities—may strongly influence the key behaviors of SIMS.

The other individual differences highlighted by commentators is that of gender differences (Simpson & Fra-
gaszy). Should we expect embodiment to play a different role for men and women? Simpson & Fragaszy point out
that there are differences between men and women in emotion recognition accuracy in some studies and these
are positively correlated with looks to the eye (Hall et al. 2010). They also note that women who have been given
testosterone show reduced mimicry (Hermans et al. 2006). In fact, a number of top-down processes, such as
motivation and the observers’ own beliefs and values, are likely to modulate behaviors linked to mimicry and conta-
gon (Hess & Fischer, under review). In this context it is also interesting to note that men and women differ in
their capacity for interoception (e.g., Harver et al. 1993); and it has been suggested that women are more likely to
base at least the perception of their own emotions on social context cues than men are (Pennebaker & Roberts
1992). This opens the door to the possibility that women may also use social context cues and emotion knowledge
to a larger extent when considering the meaning of smiles. These issues are certainly important and should
be investigated in the framework of the SIMS model. Yet, the factors discussed above only affect quantitative
aspects of the process and do not require the addition of new processes into the SIMS model.

For instance, Bartlett argues for the need to consider time. By “time,” she means that part of the development
of a complex model that integrates behaviors and multiple brain circuits involves a specification of the timing of these
interactions and the neural activations. Silvan S. Tomkins was an early advocate of the notion that emotions are
defined in part in terms of the timing of their onset and their duration (Tomkins 1962; 1963). These temporal
components constitute the very form of the emotion. We agree that timing principles will be an important part of
the model and that timing gone awry will be at the basis of dysfunction in emotional information processing.

Another extension is proposed by Alibali & Hostetter, and also by Basso & Oullier, who point to the fact that
emotions are also expressed in gesture and in the voice. Embodied simulation of these parts of emotional experi-
ence will also play a role in emotional information processing per se and will interact with the simulation, and
therefore, the understanding of facial expression. Different weights and roles of the many cues to emotion can
in part be determined by examining the cases of mismatch between one (facial expression) and the other (emotional
gesture). Another important question will be why emotion is communicated often in more than one expressive
channel. Why do we have emotional gesture and prosody if we already have the exquisitely manced face? We
feel that the SIMS model presents a valuable framework for the study of these and other pertinent processes.

NOTE

1. An intended negative outcome for the smiler would be unusual and probably related to a non-typical state, as one
smiles before committing an act that one knows will be punished or will lead to self-harm.

References


BEHAVIORAL AND BRAIN SCIENCES (2010) 33:6 469
References/Niedenthal et al.: The Simulation of Smiles (SIMS) model


Ekman, P. (2001) Telling lies: Clues to deceit in the marketplace, politics, and marriage. WW. Norton. [PBri, aPMN]


Emotion 16:193–201. [aPMN]

Emotion 16:193–201. [aPMN]

Emotion 16:193–201. [aPMN]

Emotion 16:193–201. [aPMN]

Emotion 16:193–201. [aPMN]
References/Niedenthal et al.: The Simulation of Smiles (SIMES) model


Thibault, P., Levesque, M., Gosselin, P. & Hess, U. (submitted) The Duchenne marker is NOT a universal signal of smile authenticity: But it can be learned! [APMN]


BEHAVIORAL AND BRAIN SCIENCES (2010) 33:6

479
References/Niedenthal et al.: The Simulation of Smiles (SIMS) model

amygdala activity without explicit knowledge. *Journal of Neuroscience* 18:411–18. [aPMN]


