



## UvA-DARE (Digital Academic Repository)

### Comment: What's embodied in a smile?

Sauter, D.A.; Levinson, S.C.

**DOI**

[10.1017/S0140525X10001597](https://doi.org/10.1017/S0140525X10001597)

**Publication date**

2010

**Published in**

Behavioral and Brain Sciences

[Link to publication](#)

**Citation for published version (APA):**

Sauter, D. A., & Levinson, S. C. (2010). Comment: What's embodied in a smile? *Behavioral and Brain Sciences*, 33(6), 457-458. <https://doi.org/10.1017/S0140525X10001597>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

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 syringeal 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.18): “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 non-human, 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 intended 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.

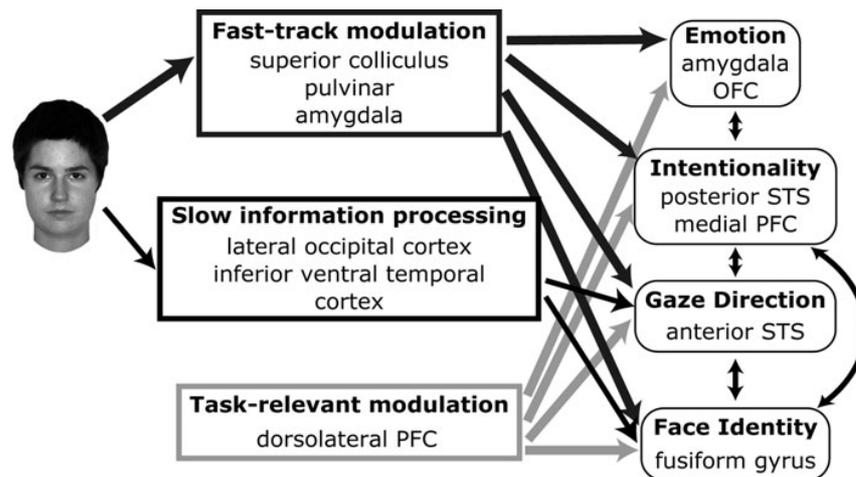


Figure 1. (Senju and Johnson). An illustration of the fast-track modulator model (FTM). Perceived eye contact (upper left) is initially detected by a subcortical route that projects to various regions of the social brain network (thick black lines). This signal from the subcortical route then interacts 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).