Context in Emotion Perception

Lisa Feldman Barrett[1,2], Batja Mesquita[3], and Maria Gendron[1]
[1]Department of Psychology, Boston College, [2]Department of Psychiatry and the Martinos Center for Biomedical
Imaging, Massachusetts General Hospital/Harvard Medical School, and [3]Department of Psychology,
University of Leuven, Belgium

Abstract

We review recent work demonstrating consistent context effects during emotion perception. Visual scenes, voices, bodies, other faces, cultural orientation, and even words shape how emotion is perceived in a face, calling into question the still-common assumption that the emotional state of a person is written on and can be read from the face like words on a page. Incorporating context during emotion perception appears to be routine, efficient, and, to some degree, automatic. This evidence challenges the standard view of emotion perception represented in psychology texts, in the cognitive neuroscience literature, and in the popular media and points to a necessary change in the basic paradigm used in the scientific study of emotion perception.

Look at the woman in Figure 1. Is she infuriated? In agony? Actually, this picture shows an ecstatic Serena Williams after she beat her sister, Venus, in the 2008 U.S. Open tennis finals. This picture of Ms. Williams clearly illustrates that context is important for inferring meaning in a face. Yet for much of the last century, the psychology of emotion perception has been guided by the idea that emotions are written on the face as particular arrangements of facial actions and that perceivers can read these actions as easily and effortlessly as they read words on a page. On the basis of experiments in which perceivers across the world have judged apparently context-less faces, many psychologists still operate on the assumption that there are six to ten emotions that are automatically (e.g., Tracy & Robins, 2008) and universally "recognized" in the face (i.e., are perceived as intended by the experimenter; Matsumoto, Keltner, Shiota, O'Sullivan, & Frank, 2008; see also Dolan, 2002). This claim has had pervasive scientific and cultural effects. It sets emotion perception apart from other types of perception in which context effects are well documented. It guides neuroscience studies of emotion perception, as well as investigations of social and emotional deficits in people with mental disorders and neurodegenerative disease. In the United States alone, millions of dollars each year are spent on law enforcement and security training to "read" emotion in the face in the hopes of keeping citizens safe.

In this article, we review recent evidence that automatic and efficient emotion perception is not driven by the structural features of a face alone but also derives from the context in which a face is encoded, even when that context is unintended by the experimenter. It appears that faces carry affective (valence and arousal) information but that the emotional meaning of facial actions is further constructed from the context in which they are embedded. And even affective perceptions are not immune to context.

Three types of context effects are reviewed: (a) *stimulus-based* context, in which a face is physically presented with other sensory input that has informational value; (b) *perceiver-based* context, in which processes within the brain or body of a perceiver can shape emotion perception; and (c) *cultural* contexts that affect either the encoding or the understanding of facial actions.

Stimulus Context in Emotion Perception

In the movie *The Wizard of Oz*, the great and powerful wizard appears to Dorothy and her fellow travelers as a large, disembodied face. But in real life, faces don't appear in isolation. Instead, they appear in a multisensory context that includes a voice, a body, and usually a broader situational arrangement. The emotion seen in the structural configuration of another person's facial actions, even when they are posed in an exaggerated expression (scowling faces for anger, startled faces for fear, pouting faces for sadness, etc.), is influenced by information in the situational context in which a face is embedded.

Corresponding Author:
Lisa Feldman Barrett, Department of Psychology, Boston College, Chestnut
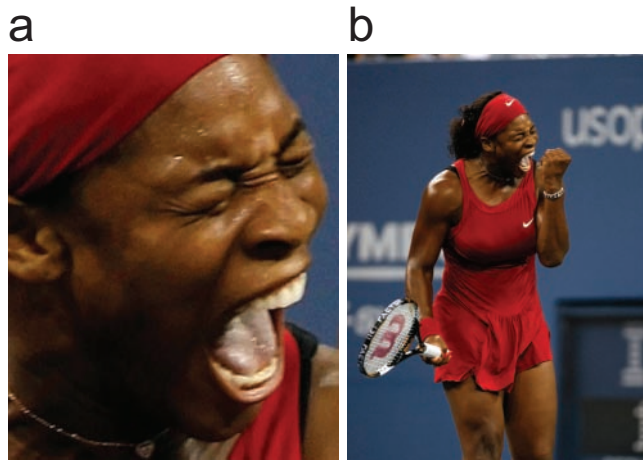Hill MA 02467
E-mail: barretli@bc.edu

Fig. 1. How is she feeling? Look at the woman depicted in (a). If you had to guess, you might say that she is in pain. But if you look again at (b), when the face is placed in a bodily context, you might now see her as ecstatic.

There is evidence that perceivers' judgments of facial actions are influenced by descriptions of the social situation (e.g., Carroll & Russell, 1996), voices, body postures, and visual scenes (e.g., Aviezer et al., 2008; Righart & de Gelder, 2008; for a review, see de Gelder et al., 2006), and even other faces (e.g., Masuda et al., 2008). For example, scowling faces (posed, exaggerated facial portrayals of anger) are more likely to be perceived as fearful when paired with the description of danger (Carroll & Russell, 1996, Study 1); they are more likely to be perceived as disgusted when paired with a body posture involving a soiled object (Aviezer et al., 2008, Study 1); and they appear determined or puzzled depending on the situation (Carroll & Russell, 1996, Study 2). Situation descriptions exert a strong influence when they are more ambiguous than the exaggerated facial actions being perceived (Carroll & Russell, 1996, Study 3). Context can even influence perceivers' judgments of valence: Faces portraying disgust are perceived as proud when paired with a muscled body whose arms are raised in triumph (Aviezer et al., 2008, Study 2).

These contextual influences are perceived early and automatically and are not disrupted by cognitive load (Righart & de Gelder, 2008). Visual scenes influence the actual structural encoding of posed facial actions, based on electrophysiological recordings of the early N170 event-related-potential component (Righart & de Gelder, 2006). (The N170 occurs 103–200 ms after a face is presented and reflects the structural encoding of faces in visual cortex.) Context also influences how information is sampled from a face, as determined by eye-tracking data (Aviezer et al., 2008, Study 3). When a scowling face (anger) or sneering face (disgust) is placed on a body with fists in the air (anger), perceivers look more to the eye region of a face than to the mouth (as if the face was angry); when these posed faces are placed on a body holding a soiled object (disgust), perceivers look at both the eye and mouth region in equal amounts (as if the face was disgusted). A situation-based context even shapes how a face is processed after the fact, such that briefly presented startled faces (posed

fear) routinely produce increased activation in the amygdala when quickly followed (or masked) by a neutral face but not when masked by visual noise, perhaps because the mask becomes merged with the target face itself (Kim et al., 2010).

Perceiver as Context in Emotion Perception

When seeing emotion in a face, it can feel to perceivers *as if* they are reading a word on a page. And, in fact, they appear to be doing just that—reading words—in many studies of emotion perception. Words constitute a clear example of a perceiver-based context because they provide a top-down constraint in emotion perception, contributing information over and above the affective meaning available in structural information of a face (Barrett, Lindquist, & Gendron, 2007).

As emotion words become more remote from the perception task, people have a more difficult time "recognizing" emotion, even in posed, exaggerated scowls, pouts, sneers, and smiles. When perceivers are asked to spontaneously provide verbal labels for such faces, their accuracy is significantly reduced (57.7%) compared to experiments in which they are asked to match a face to emotion words that are provided by the experimenter (83.4%; Izard, 1971; cited in Russell, 1994). When emotion words are not required in the task and perceivers are simply asked to match the faces on emotional content using the structural similarities in facial actions alone (e.g., two scowling faces would match; a scowling and pouting face would not), accuracy drops to 42%; when emotion words are temporarily made meaningless by a standard laboratory task called semantic satiation (repeating a word over and over again until it sounds like nonsense), accuracy in perceptual-matching task drops even further and this change is not attributable to fatigue (to 36% on average; Lindquist, Barrett, Bliss-Moreau, & Russell, 2006, Study 3). Reducing the accessibility of emotion words experimentally with verbal load (Roberson, Damjanovic, & Pilling, 2007) or naturally via semantic aphasia (difficulty naming words; Roberson, Davidoff, & Braisby, 1999) eliminates the categorical perception of emotion altogether. Categorical perception of emotion can be produced when perceivers learn to pair facial actions with an arbitrary label (Fugate, Gouzoules, & Barrett, 2010, Study 2) in much the same way that even young infants form abstract categories of objects with words.

Emotion perception also seems impaired in semantic dementia (characterized by loss of word meaning and object knowledge; see Fig. 2). A patient with semantic dementia was asked to free sort a set of 120 posed, stereotyped scowls, pouts, smiles, sneers, startled-looking faces, and neutral faces (20 each), so that all the faces within each pile represented the same emotional feeling. The patient did not produce the five discrete emotion categories (plus neutral) that were portrayed in the set (based on the structural regularities in the faces—all the scowling faces were angry, all the pouting faces were sad, and so on). Instead, the patient sorted on hedonic valence by creating a neutral pile, a positive pile, and a negative pile (the
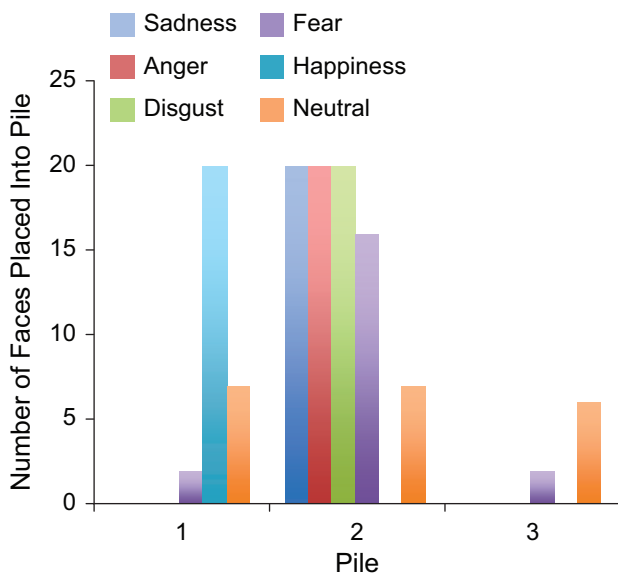
Fig. 2. Emotion perception in a patient with semantic dementia. The x-axis represents the number piles produced by the patient who was asked to look at a series of faces and sort them into piles such that everyone in the same pile felt the same emotion. The y-axis represents the number of faces sorted into each pile. The patient produced three piles: a positive pile (containing mostly smiling faces), a negative pile (containing pouting "sad" faces, scowling "angry" faces, and nose-wrinkled "disgust" faces), and a neutral pile (containing mostly neutral faces). Startled "fear" faces were placed primarily in the "negative" pile, although some were placed in each of the other two piles. Neutral faces appeared in almost equal proportion in all three piles.

last of which contained faces posing anger, sadness, disgust, and fear). Nonetheless, the patient was able to process the structural features in the faces. In a match-to-anchor task, the patient was able to match the scowling tests faces to a scowling anchor, the pouting test faces to a pouting test face, and so on. These findings are preliminary, but they show that the perceptual information from the structural features of the face alone was not sufficient for emotion perception.

Without an emotion word, perceivers notice small changes in facial actions but do not know which ones are psychologically meaningful (i.e., they do not know where to put the category boundary; Fugate et al., 2010, Study 1).

Taken together, these studies indicate that a perceiver's access to conceptual knowledge about emotion, anchored by emotion words, appears to be an essential ingredient in the perception of discrete emotions, even when viewing posed, highly stereotyped faces. Studies that provide participants with words to judge isolated faces might be, however unintentionally, providing a context that shapes the perception process.

Cultural Context in Emotion Perception

The cultural context in which perceivers and targets operate also influences how facial actions are infused with emotional meaning (Elfenbein & Ambady, 2003). Cultures differ in the precise facial actions used to pose discrete emotion categories (called cultural "accents"), and this accounts for a perceiver's in-group advantage when perceiving facial actions as intended

(Elfenbein, Beaupré, Lévesque, & Hess, 2007). While little is known about how culture shapes the processes underlying emotion perception, the degree of psychological similarity between cultures predicts the degree to which their emotion perceptions converge (Elfenbein & Ambady 2003). This finding of cross-cultural variation is not due to linguistic differences per se—perceivers from Quebec (North America) and Gabon (Africa) that share the same language (French) perceive facial actions differently. Neither are cultural differences in emotion perception due to general expressive differences or to some static elements of appearance, in that perceptual differences are not in evidence when perceiving neutral faces; they are specific to the perception of facial actions portraying emotion (Marsh, Elfenbein, & Ambady 2003).

Cultural context also appears to influence how perceivers sample information from a face in a manner that is similar to the influence of situational context. For example, when looking at startled and sneering faces, Western Caucasian perceivers fixate around the eyes, nose, and mouth of a target face, whereas those from an East Asian cultural context fixate primarily on the eyes (Jack, Blais, Scheepers, Schyns, & Caldara, 2009). Because the diagnostic features in posed startles and sneers are centered in the mouth area, the East Asian fixation on the eye region is responsible for their common perception of startled faces as surprise (rather than fear) and sneers as anger (rather than disgust).

Furthermore, Western perceivers (European Americans in the United States in Study 1, Anglo-Saxon visitors to Japan in Study 2) seem to rely primarily on the information in a target face when perceiving emotion, while Japanese make more strategic use of the information in faces surrounding the target (Masuda et al., 2008). These findings are consistent with the fact that individuals from a European American cultural context conceptualize emotions as located within the individual whereas those from a Japanese cultural context conceptualize emotions as reflecting the relationships between people. That being said, even Western perceivers routinely encode the scene context when asked to judge scowling, sneering, or neutral faces in terms of discrete emotions (e.g., "anger" vs. "fear") but not when asked to judge affect (e.g., whether to "approach" vs. "avoid" the face; Barrett & Kensinger, 2010).

Finally, emerging findings in the area of cultural neuroscience point to potentially interesting cultural differences in the brain responses during emotion perception. In Western perceivers, increased amygdala activation is consistently observed at greater-than-chance levels during the perception of posed emotional faces (of mostly Western origin); more recent evidence indicates that perceivers show distinct neural responses to posed emotional faces in ingroup (vs. outgroup) members, with both Japanese and American perceivers showing the most amygdala response to members of their own group (Chiao et al., 2008). Furthermore, a recent study indicated that Japanese perceivers don't show amygdala activity to startled (posed fearful) faces (Moriguchi et al., 2005), calling into question the idea that this expression has preferential, innate signal value.

Conclusions and Future Directions

The findings reviewed in this article clearly indicate that emotion perceptions are routinely influenced by the context in which they occur. Faces appear in a multisensory context that is processed by a brain that is running multiple mechanisms and within a body coursing with hormones and its own sensory signals. Research that is careful to strip away context (in one form or another) shows that facial-muscle actions (i.e., the structural features of a face) might carry affective information, but their discrete emotional meanings are contingent on context. Unbeknown to most perceivers most of the time, context (i.e., the social situation, body postures, voices, scenes, words, and even culture) influences which emotion is seen in the face of another person. The fact that humans easily and effortlessly perceive anger, or sadness, or fear in another person's face is not evidence that facial actions broadcast the internal state of the target person. Simply put, a face does not speak for itself.

When context is properly acknowledged and assessed, empirical findings run contrary to the notion that facial muscle movements are innate "expressions" of emotion containing all of the information that is necessary and sufficient to recognize emotion in other people. Perceivers routinely make use of whatever context is available when asked to make a specific inference about a target person's emotion from the face. These contextual influences on emotion perception are consistent with evidence showing that context is intrinsically involved in even the most basic aspects of object perception. Context constrains what we expect to see and where we look. It is also used as the basis for predictions about what visual sensations refer to based on past experience and future behavior. Context findings, in and of themselves, don't negate the existence of biologically basic, universal emotions (although other evidence calls this into question). It could be possible that context just helps a perceiver to detect the emotional information encoded in facial muscle actions. Of course, this would require that objective measures of facial muscle activity (independent of a perceiver) verify that specific discrete emotional information is routinely present in a face, and thus far such consistency and specificity has yet to be demonstrated (Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000). For our purposes in this article, however, it is sufficient to say that, to the extent that emotion perception data is used as evidence for the existence of biologically basic emotions, such claims should be reconsidered.

This review also suggests that the standard experimental paradigm for studying emotion perception requires modification. Experiments that are interpreted as providing evidence for universal "emotion recognition" unintentionally structure the context of emotion perception. First, the posed, highly exaggerated face sets that are typically used in studies of emotion perception have tremendous statistical regularities built in that are not present in everyday life. People do not always scowl in anger or pout in sadness. Study participants might form temporary perceptual categories in an experiment by extracting statistical regularities from posed faces. Second, the emotion words routinely present in experiments stealthily constrain the meaning of ambiguous facial muscle configurations, leading to the mistaken conclusion that emotion perception is automatic and completely bottom-up (i.e., driven by the structural configuration of facial muscles alone). Third, the presentation of isolated faces highlights the dominant Western notion of emotions that is not necessarily representative of other cultures. Taken together, these observations suggest that Western scientists have created an artifact: In our efforts to study emotion perception in an elemental way, we have created an experimental method that has strong but hidden context effects. Having study participants judge disembodied faces by choosing from a small selection of emotion words produces a psychological model of emotion perception that is intellectually pleasing, and true in context, but it has little to do with reality, because the context does not actually happen in real life. It will not be possible to understand the brain basis of emotion perception in any ecologically valid way by presenting posed, stereotyped faces in isolation. Future research should move beyond the study of one-shot, posed faces to study emotion perception in real-time, contextualized facial acts. Context in all of its various forms should also be explicitly modeled.

Finally, findings showing the importance of context in emotion perception are important for national security and safety. When training lie-detection experts, perhaps law enforcement and safety officials should consider how perceivers can better use context to help read emotions in a face. Training programs should take advantage of the human brain's remarkable ability to integrate multiple forms of context when inferring the mental states of others from a face.

Recommended Reading

Aviezer, H., Hassin, R.R., Ryan, J., Grady, C., Susskind, J., Anderson, A., et al. (2008). (See References). A set of experiments demonstrating that body postures serve as a powerful context influencing how information is sampled from faces posing facial actions, as well as which emotions are perceived in those actions.

Barrett, L.F., & Kensinger, E.A. (2010). (See References). The first experiment showing that context is routinely encoded during emotion (vs. affective) perception in Western perceivers.

Barrett, L.F., Lindquist, K., & Gendron, M. (2007). (See References). A highly accessible overview of the Language as Context Hypothesis.

Carroll, J.M., & Russell, J.A. (1996). (See References). A representative set of experiments that demonstrate how social situations serve as a context for emotion perception in the face.

Masuda, T., Ellsworth, P., Mesquita, B., Leu, J., Tanida, S., Van de Veerdonk, E. (2008). (See References). A representative set of experiments demonstrating cultural variation in the use of context during emotion perception.

Declaration of Conflicting Interests

References

Aviezer, H., Hassin, R.R., Ryan, J., Grady, C., Susskind, J., Anderson, A., et al. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science, 19*, 724–732.

Barrett, L.F., & Kensinger, E.A. (2010). Context is routinely encoded during emotion perception. *Psychological Science, 21*, 595–599.

Barrett, L.F., Lindquist, K., & Gendron, M. (2007). Language as a context for emotion perception. *Trends in Cognitive Sciences, 11*, 327–332.

Cacioppo, J.T., Berntson, G.G., Larsen, J.T., Poehlmann, K.M., & Ito, T.A. (2000). The psychophysiology of emotion. In R. Lewis & J.M. Haviland-Jones (Eds.), *The handbook of emotion,* (2nd ed., pp. 173–191). New York: Guilford.

Carroll, J.M., & Russell, J.A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70*, 205–218.

Chiao, J.Y., Iidaka, T., Gordon, H.L., Nogawa, J., Bar, M., Aminoff, E., et al. (2008). Cultural specificity in amygdala response to fear faces. *Journal of Cognitive Neuroscience, 20*, 2167–2174.

de Gelder, B., Meeren, H.K.M., Righart, R., Van den Stock, J., van de Riet, W.A.C., & Tamietto, M. (2006). Beyond the face: Exploring rapid influences of context on face processing. *Progress in Brain Research, 155*, 37–48.

Dolan, R.J. (2002). Emotion, cognition, and behavior. *Science, 298*, 1191–1194.

Elfenbein, H.A., & Ambady, N. (2003). When familiarity breeds accuracy: Cultural exposure and facial emotion recognition. *Journal of Personality and Social Psychology, 85*, 276–290.

Elfenbein, H.A., Beaupré, M.G., Lévesque, M., & Hess, U. (2007). Toward a dialect theory: Cultural differences in the expression and recognition of posed facial expressions. *Emotion, 7*, 131–146.

Fugate, J.M.B., Gouzoules, H., & Barrett, L.F. (2010). Reading Chimpanzee faces: A test of the structural and conceptual hypotheses. *Emotion, 10*, 544–554.

Jack, R.E., Blais, C., Scheepers, C., Schyns, P.G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology, 19*, 1–6.

Kim, M.J., Loucks, R.A., Neta, M., Davis, F.C., Oler, J.A., Mazzulla, E.C., et al. (2010). Behind the mask: The influence of mask-type on amygdala responses to fearful faces. *Social, Cognitive and Affective Neuroscience, 5,* 363–368.

Lindquist, K., Barrett, L.F., Bliss-Moreau, E., & Russell, J.A. (2006). Language and the perception of emotion. *Emotion, 6*, 125–138.

Marsh, A.A., Elfenbein, H.A., & Ambady, N. (2003). Nonverbal "accents": Cultural differences in facial expressions of emotion. *Psychological Science, 14*, 373–376.

Masuda, T., Ellsworth, P., Mesquita, B., Leu, J., Tanida, S., & Van de Veerdonk, E. (2008). Placing the face in context: Cultural differences in the perception of facial emotion. *Journal of Personality and Social Psychology, 94*, 365–381.

Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M., & Frank, M. (2008). Facial expressions of emotion. In M. Lewis, J.M. Haviland-Jones, & L.F. Barrett (Eds.), *The Handbook of Emotion* (3rd ed., pp. 211–234). New York: Guilford.

Moriguchi, Y., Ohnishi, T., Kawachi, T., Mori, T., Hirakata, M., Yamada, M., et al. (2005). Specific brain activation in Japanese and Caucasian people to fearful faces. *Neuroreport, 16*, 133–136.

Righart, R., & de Gelder, B. (2006). Context influences early perceptual analysis of faces. An electrophysiological study. *Cerebral Cortex, 16*, 1249–1257.

Righart, R., & de Gelder, B. (2008). Rapid influence of emotional scenes on encoding of facial expressions. An ERP study. *Social Cognitive and Affective Neuroscience, 3*, 270–278.

Roberson, D., Damjanovic, L., & Pilling, M. (2007). Categorical perception of facial expressions: Evidence for a "category adjustment" model. *Memory & Cognition, 35*, 1814–1829.

Roberson, D., Davidoff, J., & Braisby, N. (1999). Similarity and categorisation: Neuropsychological evidence for a dissociation in explicit categorisation tasks. *Cognition, 71*, 1–42.

Russell, J.A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115*, 102–141.

Tracy, J.L., & Robins, R.W. (2008). The automaticity of emotion recognition. *Emotion, 8*, 81–95.