Context Is Routinely Encoded During Emotion Perception

Lisa Feldman Barrett[1,2,3] and Elizabeth A. Kensinger[1,3]

[1]Boston College; [2]Psychiatric Neuroimaging Program, Massachusetts General Hospital, Harvard Medical School; and [3]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School

Abstract

In the present study, we investigated whether context is routinely encoded during emotion perception. For the first time, we show that people remember the context more often when asked to label an emotion in a facial expression than when asked to judge the expression's simple affective significance (which can be done on the basis of the structural features of the face alone). Our findings are consistent with an emerging literature showing that facial muscle actions (i.e., structural features of the face), when viewed in isolation, might be insufficient for perceiving emotion.

Citing Darwin (1872/1965) as inspiration, many scientists believe in the *structural hypothesis* of emotion perception—the idea that certain emotion categories (named by the English words *anger*, *fear*, *sadness*, *disgust*, and so on) are universal biological states that are triggered by dedicated, evolutionarily preserved neural circuits (instincts or affect programs), expressed as clear and unambiguous signals involving configuration of facial muscle activity (*facial expressions*), and recognized by mental machinery that is innately hardwired, reflexive, and universal (e.g., Allport, 1924; McDougall, 1908/1921; Tomkins, 1962, 1963). Several influential models of emotion perception that involve the structural hypothesis now dominate the psychological literature (e.g., Ekman, 1972; Izard, 1971) and are supported by empirical evidence (for a recent review, see Matsumoto, Keltner, Shiota, O'Sullivan, & Frank, 2008). A recent article succinctly summarized the structural hypothesis: "The face, as a transmitter, evolved to send expression signals that have low correlations with one another and . . . the brain, as a decoder, further decorrelates . . . these signals" (Smith, Cottrell, Gosselin, & Schyns, 2005, p. 188).

Even though humans, in the blink of an eye, can easily and effortlessly perceive emotion in other creatures (including each other), there is growing evidence that context acts, often in stealth, to influence emotion perception. Descriptions of the social situation (Carroll & Russell, 1996; Fernandez-Dohls, Carrera, Barchard, & Gacitua, 2008; Trope, 1986), body postures (Aviezer et al., 2008; Meeren, van Heijnsbergen, & de Gelder, 2005), voices (de Gelder, Böcker, Tuomainen, Hensen, &

Vroomen, 1999; de Gelder & Vroomen, 2000), scenes (Righart & de Gelder, 2008), words (Lindquist, Barrett, Bliss-Moreau, & Russell, 2006), and other emotional faces (Masuda et al., 2008; Russell & Fehr, 1987) all influence which emotion is seen in the structural configuration of another person's facial muscles.

Although researchers attempt to remove the influence of context in most experimental studies of emotion perception, one important source of context typically remains: words. A variety of findings support the hypothesis that words provide a top-down constraint in emotion perception, contributing information over and above the structural information provided by a face alone (for a review, see Barrett, Lindquist, & Gendron, 2007; e.g., see Fugate, Gouzoules, & Barrett, in press; Roberson, Damjanovic, & Pilling, 2007). Furthermore, when the influence of words is minimized, both children (Russell & Widen, 2002) and adults (Lindquist et al., 2006) have difficulty with the seemingly trivial task of using structural similarities in facial expressions alone to judge whether or not the expressions match in emotional content (even though the face sets used have statistical regularities built in).

The *conceptual-act model* of emotion (Barrett, 2006, 2009a, 2009b; Barrett et al., 2007) hypothesizes that facial muscle movements alone carry simple affective information

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

(e.g., whether the face should be approached or avoided), and words for emotions increase the accessibility of conceptual knowledge for emotion, which acts as a top-down influence allowing a specific emotional percept to take shape. Within this model, conceptual knowledge is tailored to the specific situation, which leads to the hypothesis that emotion words direct attention to the situation. As a consequence, context is more likely to be encoded (if not consciously recognized) when a person's task is to perceive emotion in the face of another person rather than to judge the face's affective value.

In the present experiment, we tested this hypothesis using a memory paradigm that is sensitive to the way in which processing resources are allocated during encoding. Prior research has shown that context is not readily encoded when people process affectively potent objects (e.g., snakes; Kensinger, Garoff-Eaton, & Schacter, 2007). Yet when it is advantageous for perceivers to attend to the context (e.g., when they must describe the context to the experimenter or remember the context), contexts are better remembered (Kensinger et al., 2007). Perceivers' ability to remember the context can therefore be used as a proxy to understand how resources are devoted toward processing that context. We hypothesized that when asked to perceive emotion (i.e., fear or disgust) in a face, participants would devote more processing resources to encoding and remembering the context than they would when asked to perceive the face's affective value (i.e., whether to approach or avoid it).

Participants viewed objects or facial expressions (fearful, disgusting, or neutral) in a neutral context and judged either their willingness to approach or avoid the objects or faces (an affective categorization) or whether the faces were fearful or disgusting, using words presented to them during the task (an emotion categorization). We predicted that when participants were asked to label the facial expression with an emotion word, they would show better memory for the context in which the face was presented (even though the context itself was neutral) than they would when asked to make an affective judgment of the face. We reasoned that this would be true because the structural features of the face alone, even in a caricatured face (such as those typically used in studies of emotion recognition), are not typically sufficient to allow an emotion perception. As a consequence, perceivers attempt to use whatever context information is available, no matter how impoverished. In contrast, when perceivers are asked to make an affective judgment of a face, the information contained in the structural aspects of the face is more likely to be sufficient. Furthermore, we expected that labeling the emotion elicited by an object would not alter the memory for the context; there would be less disambiguation required with an object than with a face, and thus little need to devote processing resources to the context.

Method

Participants

Participants were 36 students at Boston College (23 males and 13 females). All were native English speakers with normal or corrected-to-normal vision. No participant reported a history of neuropsychological or psychiatric disorder, nor did any participant report taking medication that would affect the central nervous system.

Materials

Study materials were images of 42 neutral contexts, such as a lake or a patio; images of 42 individuals, each of whom portrayed fear, disgust, and neutral facial expressions (from Tottenham et al., 2009); and images of 126 objects (42 fearful, such as a snake; 42 disgusting, such as vomit; and 42 neutral, such as a canoe; from Kensinger et al., 2007). Six versions of 42 scenes were created by placing one instance of each kind of facial expression or one instance of each kind of object against each scene background (see Fig. 1). Foils for the recognition test were images of an additional 18 neutral contexts; 18 individuals with fear, disgust, and neutral expressions (6 instances of each); and 18 fearful, disgusting, or neutral objects (6 of each type). Normative ratings of valence and arousal provided assurance that the neutral contexts were perceived as neutral, and that the objects were perceived either as negative and arousing or as neutral (Kensinger et al., 2007; Waring & Kensinger, in press). Ratings for the facial expressions had been gathered by the investigators who created the face database (Lyons, Akamatsu, Kamachi, & Gyoba, 1998; Tottenham et al., 2009).

Procedure

Participants were randomly assigned to one of four experimental conditions: label-object, label-face, approach-object, or approach-face. In each condition, participants were presented with 42 scenes, each presented for 2 s. In the object conditions, each scene incorporated an object and a background (14 scenes included an object that elicited fear, 14 included an object that elicited disgust, and 14 included a neutral object). In the face conditions, participants viewed scenes that incorporated a face and a background (14 scenes included a fearful expression, 14 included a disgusted expression, and 14 included a neutral expression). In the approach conditions (approach-object or approach-face), participants were instructed to indicate whether they would want to approach, stay at the same distance from, or back away from the scene if they were to encounter it in their everyday life. In the emotion-label conditions (label-object or label-face), participants were asked to label the facial expression of the person or to indicate the emotion evoked by the object in the scene, selecting from the following options: "disgust," "fear," or "neutral."

After this study phase, participants performed a short distractor task (completion of Sudoku puzzles for 3 min) intended to eliminate recency effects from memory. Participants then performed a recognition memory test. They were shown a series of 60 neutral contexts and either 60 faces or 60 objects (depending on the study condition); contexts were presented separately from the objects or faces, and all images were

Fig. 1. Examples of the experimental stimuli. In each scene, the image of a fearful, neutral, or disgusting object or (from top to bottom) a fearful, neutral, or disgusted facial expression was superimposed on a neutral context, such as a supermarket, a desert, or insects on a log in a lake.

shown one at a time. Of the 60 contexts, 42 had been studied, and 18 were novel foils. Similarly, of the 60 faces or objects, 42 had been studied, and 18 were foils. For each context, face, or object, participants indicated whether they (a) had not seen the image earlier, (b) were unsure but thought they had seen it earlier, (c) were sure they had seen it earlier, or (d) were very sure they had seen it earlier. In all reported analyses, we considered responses of "sure" or "very sure" to indicate the endorsement of an item as studied. False alarm rates for the

unstudied items were low (under 10%) and did not differ as a function of emotion category, task instruction (label vs. approach), or task type (face vs. object). Therefore, we discuss only responses to the studied images.

Results

Memory for context

We examined the recognition of the neutral contexts on trials with faces by conducting an analysis of variance (ANOVA) with judgment type (emotion vs. affective categorization; i.e., "label" vs. "approach" decision) as a between-subjects factor and item type (affective vs. neutral faces) as a within-subjects factor. This analysis revealed no main effects of judgment type or item type, but did reveal an interaction between the two factors, $F(1, 16) = 6.94$, $p < .05$, $\eta_p^2 = .30$. Consistent with our prediction, post hoc t tests confirmed that when the context was presented with an affectively potent facial expression, the context was remembered better in the emotion-categorization condition than in the affective-categorization condition, $t(17) = 2.67$, $p < .05$. In contrast, the neutral context was remembered equivalently regardless of the judgment task when paired with a neutral face ($p > .25$; see Table 1). The divergent effects of emotion versus affective categorization on memory for context were not related to differences in the time that it took participants to make their decision about each face. Also as predicted, enhanced memory for the context during emotion categorization was specific to those scenes that included faces. The results of an ANOVA conducted on the recognition of the neutral contexts paired with objects revealed only a main effect of item type, $F(1, 16) = 12.1$, $p < .01$, $\eta_p^2 = .43$; participants had worse memory for contexts paired with an affective object than for contexts presented with a neutral object (see Table 1).

Memory for faces and objects

We also examined recognition scores for the faces themselves, conducting an ANOVA with judgment type (emotion vs. affective categorization) as a between-subjects factor and item type (affective vs. neutral) as a within-subjects factor. This analysis revealed no main effects or interactions (all Fs < 1, ps > .4), indicating that recognition of the faces was not affected by

Table 1. Mean Recognition Memory for Contexts as a Function of Judgment Type and Item Type

Judgment type	Contexts with a face		Contexts with an object	
	Affective	Neutral	Affective	Neutral
Emotion	.67 (.07)	.54 (.07)	.40 (.04)	.52 (.05)
Affect	.44 (.05)	.52 (.06)	.38 (.01)	.52 (.05)

Note: Standard errors are given in parentheses.

Table 2. Mean Recognition Memory for Faces and Objects as a Function of Judgment Type and Item Type

Judgment type	Faces		Objects	
	Affective	Neutral	Affective	Neutral
Emotion	.39 (.04)	.40 (.04)	.67 (.04)	.48 (.05)
Affect	.41 (.05)	.44 (.05)	.67 (.04)	.48 (.04)

Note: Standard errors are given in parentheses.

either of these variables (see Table 2). The results of an ANOVA conducted on the recognition scores for the objects themselves revealed only a main effect of item type, $F(1, 16) = 14.6$, $p < .01$, $\eta_p^2 = .48$; participants had better recognition for affective objects than for nonaffective ones (see Table 2).

Discussion

The present study is the first to clearly show that when a person looks at a human face with the goal of perceiving emotion, the perceiver encodes the face in context. As predicted, participants remembered contextual information better when they were required to perceive emotion in a face (whether it was fearful or disgusted) than when they were asked to make an affective judgment about the face (either to approach or to avoid it). The goal to categorize the face as fearful or disgusted required that perceivers use all information available to them—both the information contained within the structural configuration of facial muscles themselves and the information contained in the broader context. The structural features of the face were not irrelevant to the task, however, because merely having a goal to categorize emotion was not, in and of itself, sufficient for encouraging the encoding of context. Perceivers did not show enhanced memory for contexts paired with structurally neutral faces during the emotion-categorization task.

These findings are consistent with the conceptual-act model of emotion, in which emotion perceptions are situated conceptualizations of affective information that is available in the sensory world (Barrett, 2006, 2009a, 2009b; Barrett et al., 2007). They are also consistent with an emerging scientific literature suggesting that the structural configuration of a face might be sufficient for perceptions of affect, but insufficient for perceptions of emotion. When understood in the larger literature on emotion perception, our findings suggest that facial muscle movements might sometimes provide a beacon to indicate the affective significance of target person's mental state, but perceivers routinely encode the context when asked to make the more specific inference about a target person's emotion. Future research should determine whether emotion words routinely lead people to sample the context to disambiguate the meaning of all facial actions, or of only those facial actions that are perceptually similar (such as those of disgust and fear).

Our findings highlight the limitations in the methods used in many studies of emotion perception. Studies routinely use targets that pose an exaggerated configuration of muscle movements for each emotion category and present faces in isolation, on the assumption that all of the relevant information about a person's internal state is being carried in the face. Experiments presenting faces with no situational context might be omitting an important factor that normally influences the emotion perception process, however. Furthermore, by ignoring the fact that a perceiver's goals are manipulated by the judgment task at hand (i.e., by whether or not the perceiver is asked to infer an emotion), researchers might mistakenly assume that emotion categorization is a process that is automatically triggered in a bottom-up fashion on the basis of the information available in the structural configuration of facial muscles alone, although more recent evidence suggests that this is not the case (e.g., Herba et al., 2007). Studies that provide participants with words to judge isolated faces may, however unintentionally, be providing a context that influences the perception process.

Finally, the results of the present study suggest numerous avenues for future research focused on understanding how and why a perceiver's goals influence the way in which the context is sampled during emotion perception. For example, do the goals of the perceiver influence when the face and its context are processed in a configural versus a holistic fashion? When does incorporating context help or hurt emotion perception? These are not questions that traditionally have been asked in the scientific study of emotion perception, in part because it has been assumed that facial recognition relies on a set of processes that are not influenced by these extrinsic factors. Yet the present results suggest that these are exactly the types of questions that must be answered in order to understand facial recognition in a more ecologically sensitive fashion.

References

Allport, F. (1924). *Social psychology*. New York: Houghton Mifflin.

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. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, *10*, 20–46.

Barrett, L.F. (2009a). The future of psychology: Connecting mind to brain. *Perspectives in Psychological Science*, *4*, 326–339.

Barrett, L.F. (2009b). Variety is the spice of life: A Psychologist Constructionist approach to understanding variability in emotion. *Cognition & Emotion*, *23*, 1284–1306.

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

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.

Darwin, C. (1965). *The expression of emotions in man and animals*. Chicago: University of Chicago Press. (Original work published 1872)

de Gelder, B., Böcker, K.B., Tuomainen, J., Hensen, M., & Vroomen, J. (1999). The combined perception of emotion from voice and face: Early interaction revealed by human electric brain responses. *Neuroscience Letters*, *260*, 133–136.

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition & Emotion*, *14*, 289–311.

Ekman, P. (1972). Universal and cultural differences in facial expressions of emotions. In J.K. Cole (Ed.), *Nebraska Symposium on Motivation*, 1971 (pp. 207–283). Lincoln: University of Nebraska Press.

Fernandez-Dohls, J.M., Carrera, P., Barchard, K.A., & Gacitua, M. (2008). False recognition of facial expressions of emotion: Causes and implications. *Emotion*, *8*, 530–539.

Fugate, J.M.B., Gouzoules, H., & Barrett, L.F. (in press). Reading chimpanzee faces: A test of the structural and conceptual hypotheses. *Emotion*.

Herba, C.M., Heining, M., Young, A.W., Browning, M., Benson, P.J., Phillips, M.L., & Gray, J.A. (2007). Conscious and nonconscious discrimination of facial expressions. *Visual Cognition*, *15*, 36–37.

Izard, C.E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.

Kensinger, E.A., Garoff-Eaton, R.J., & Schacter, D.L. (2007). Effects of emotion on memory specificity: Memory trade-offs elicited by negative visually arousing stimuli. *Journal of Memory and Language*, *56*, 575–591.

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

Lyons, M., Akamatsu, S., Kamachi, M., & Gyoba, J. (1998). Coding facial expressions with Gabor wavelets. In *Proceedings of the 3rd International Conference on Face & Gesture Recognition* (pp. 200–205). Washington, DC: IEEE Computer Society.

Masuda, T., Ellsworth, P.C., 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.G. (2008). Facial expression of emotion. In M. Lewis, J.M. Haviland-Jones, & L.F. Barrett (Eds.), *The handbook of emotion* (pp. 211–234). New York: Guilford.

McDougall, W. (1921). *An introduction to social psychology*. Boston: John W. Luce. (Original work published 1908)

Meeren, H.K.M., van Heijnsbergen, C.C.R.J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences, USA*, *102*, 16518–16523.

Righart, R., & de Gelder, B. (2008). Recognition of facial expressions is influenced by emotion scene gist. *Cognitive, Affective, & Behavioral Neuroscience*, *8*, 264–278.

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

Russell, J.A., & Fehr, B. (1987). Relativity in the perception of emotion in facial expressions. *Journal of Experimental Psychology*, *116*, 223–237.

Russell, J.A., & Widen, S.C. (2002). A label superiority effect in children's categorization of facial expressions. *Social Development*, *11*, 30–52.

Smith, M.L., Cottrell, G.W., Gosselin, F., & Schyns, P.G. (2005). Transmitting and decoding facial expressions. *Psychological Science*, *16*, 184–189.

Tomkins, S.S. (1962). *Affect, imagery, and consciousness: Vol. 1. The positive affects*. New York: Springer.

Tomkins, S.S. (1963). *Affect, imagery, and consciousness: Vol. 2. The negative affects*. New York: Springer.

Tottenham, N., Tanaka, J., Leon, A.C., McCarry, T., Nurse, M., Hare, T.A., et al. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*, 242–249.

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, *93*, 239–257.

Waring, J.D., & Kensinger, E.A. (in press). Effects of emotional valence and arousal upon memory trade-offs with aging. *Psychology and Aging*.