Emotion

Emotion Words: Adding Face Value

Jennifer M. B. Fugate, Maria Gendron, Satoshi F. Nakashima, and Lisa Feldman Barrett

Online First Publication, June 12, 2017. http://dx.doi.org/10.1037/emo0000330

Emotion

1528-3542/17/$12.00 http://dx.doi.org/10.1037/emo0000330

Emotion Words: Adding Face Value

Jennifer M. B. Fugate
University of Massachusetts Dartmouth

Maria Gendron
Northeastern University

Satoshi F. Nakashima
Kyoto University

Lisa Feldman Barrett
Northeastern University and Massachusetts General Hospital, Boston, Massachusetts

Despite a growing number of studies suggesting that emotion words affect perceptual judgments of emotional stimuli, little is known about how emotion words affect perceptual memory for emotional faces. In Experiments 1 and 2 we tested how emotion words (compared with control words) affected participants' abilities to select a target emotional face from among distractor faces. Participants were generally more likely to false alarm to distractor emotional faces when primed with an emotion word congruent with the face (compared with a control word). Moreover, participants showed both decreased sensitivity (d') to discriminate between target and distractor faces, as well as altered response biases (c; more likely to answer "yes") when primed with an emotion word (compared with a control word). In Experiment 3 we showed that emotion words had more of an effect on perceptual memory judgments when the structural information in the target face was limited, as well as when participants were only able to categorize the face with a partially congruent emotion word. The overall results are consistent with the idea that emotion words affect the encoding of emotional faces in perceptual memory.

Keywords: emotion perception, emotional faces, emotion words, semantic priming, signal detection theory

Emotion perception happens with ease: humans can effortlessly look at another person's face and see happiness, sadness, anger, or some other emotion. Structural approaches to emotion perception suggest that the information in another person's face informs discrete emotion judgments due to inherent links between facial actions and emotions (e.g., Ekman, 1972, 1992; Ekman & Cordaro, 2011; Izard, 1971, 1994; Matsumoto, Keltner, Shiota, O'Sullivan, & Frank, 2008; Tomkins, 1962, 1963). According to these views, language is either independent to emotion perception (e.g., Brosch, Pourtois, & Sander, 2010; Sauter, LeGuen, & Haun, 2011) or a byproduct of the communication process (i.e., innate categories sediment out into language; Scherer, 2000). As a result, labels applied to faces should be the result of processing of the structural information of the face. An alternative constructionist account suggests that emotion labels are not applied as a result of structural processing, but rather contribute to emotion perception in a predictive manner by bringing online conceptual knowledge associated with an emotion category. Constructionist accounts are similar to structural approaches in that some structural features of facial actions inform perceptions. These approaches are distinct, however, in that perceivers routinely infer emotional states that are informed by other sources of information. Said another way, emotion perception is shaped by the internal context that exists in the mind of a perceiver or additional external context in the environment (Barrett, 2006a, 2006b; 2011, 2012, 2013, 2014).

Jennifer M. B. Fugate, Department of Psychology, University of Massachusetts Dartmouth; Maria Gendron, Department of Psychology, Northeastern University; Satoshi F. Nakashima, Graduate School of Education, Kyoto University; Lisa Feldman Barrett, Department of Psychology, Northeastern University, and Psychiatric Neuroimaging Research Program and Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston, Massachusetts.

Satoshi F. Nakashima is now at Department of Psychology, Hiroshima Shudo University.

Portions of this data were presented in poster format at the meeting for the Society of Personality and Social Psychology, American Psychological Society, and International Society for Research on Emotion, and in oral formats at the meeting for the American Psychological Society and the Society for Experimental Social Psychology.

Experiments were supported by an National Research Service Award (NRSA; F32MH083455) from the National Institute of Mental Health to Jennifer M.B. Fugate and by NRSA (5F32MH105052) to Maria Gendron, as well as a National Institutes of Health Director's Pioneer Award (DP1OD003312) and National Science Foundation Grant BCS 0721260 to Lisa Feldman Barrett. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health.

We thank (in alphabetical order) Eve Barnard, Katie Bessette, Dasha Cherkosav, Colleen Mahrer, and Elisabeth Secor for help with data collection and coding. We also thank Souhee Yang for help with additional data analysis. In addition, we thank Erika Siegel for helpful feedback on an earlier version of this article.

Correspondence concerning this article should be addressed to Jennifer M. B. Fugate, Department of Psychology, University of Massachusetts Dartmouth, Dartmouth, MA 0274. E-mail: jfugate@umassd.edu

Language is a key part of the internal context which informs emotion perception. In our model of emotion perception, emotion words like "anger," "sadness," and "fear" name folk categories that divide up the continuous and highly variable measurable outcomes (e.g., facial muscle movements, peripheral physiology, behavior; Barrett, 2006a, 2006b; 2011, 2012, 2013, 2014). Although in some instances of an emotion, individuals might produce a canonical "expression" (i.e., widened eyes, mouth agape), what Russell (2003) referred to as "blue-ribbon" instances, there are many more instances in which this structural information is not present, or not sufficient in and of itself. The *language-as-context* hypothesis (Barrett, 2009; Barrett, Lindquist, & Gendron, 2007; Barrett, Mesquita, & Gendron, 2011; for a recent review, see Fugate & Barrett, 2014; Lindquist & Gendron, 2013) suggests that emotion words provide an internal context that helps to constrain the complex, multidimensional flow of information. As a result, bringing an emotion word online might change the nature of the structural information which is encoded from emotional faces. In the current experiments, we build on the *language-as-context* hypothesis to examine *how* emotion words change perceptual memory for emotional faces. Specifically, we investigate how emotion words change participants' likelihood to correctly select a target face from distractor faces. We also investigate how emotion words change perceptual memory biases when the time to encode the face is systematically varied and the emotion words presented are not fully congruent with the face.

Emotion Words and Emotion Judgments

The impact of emotion words on emotion judgments is demonstrated by several lines of research. Making language more accessible impacts emotion judgments. Participants primed with an emotion-related word (e.g., *joyous*; compared with a control word) are quicker to select the correct emotion word (e.g., "happy") to label a smiling face (Carroll & Young, 2005, Experiment 2). Participants primed with a congruent emotion word (compared with a control word) are faster to detect when a facial depiction changes from that emotion category to another emotion category (Fugate, O'Hare, & Emmanuel, in press). Children are also better able to match a facial depiction of emotion to an emotion word (e.g., "angry") than they are at matching two faces posing similar emotions (e.g., two scowling faces; Widen, 2013; Widen & Russell, 2003). Finally, participants have lower sensitivity (d') at identifying whether a second emotional face (of a different identity) belongs to the same emotion category as a target emotional face (face-face trials) compared with when an emotion word belongs to that category (face-word trials; Nook, Lindquist, & Zaki, 2015). Such faciliatory effects of words even impact emotion judgments which do not manipulate language. Simply having an emotion word available in a paradigm forces agreement of emotion judgments compared with when emotion words are not included. For example, when perceivers are asked to match a posed emotional face to a set of emotion words provided by the experimenter, agreement on the meaning is much higher than when perceivers must spontaneously label the faces (cf. Russell, 1994; e.g., Boucher & Carlson, 1980).

Reduced accessibility of emotion word meaning has a similar impact on emotion judgments. When perceivers' accessibility to emotion words is temporarily reduced using a standard laboratory task (via semantic satiation), their ability to judge whether two faces match in emotional content is no better than chance (Lindquist, Barrett, Bliss-Moreau, & Russell, 2006). In addition, patients with semantic dementia (a neurodegenerative disease defined by the loss of abstract words) do not categorize emotional faces into discrete categories. Instead, they sort faces into more basic, affective (positive and negative) piles (Lindquist, Gendron, Barrett, & Dickerson, 2014).

Perceptual memory is also impacted by accessibility of emotion language. Participants remember morphed emotional faces as more "angry" when they are paired with the word "anger" than when paired with no word (Halberstadt, 2003, 2005; Halberstadt & Niedenthal, 2001). Subsequent research suggested that participants do not simply reconstruct the emotional faces at the time of recall, but rather they encode the faces in the presence of words as actually more intense (e.g., *more angry*). To distinguish perceptual encoding effects from postprocessing effects, participants' facial musculature was measured with EMG while encoding the morphed faces with differing category information (e.g., emotion words, personality adjectives, and ideographs from a foreign language). Consistent with the previous demonstrations, participants indicated a more intense exemplar (*more angry*) as the target during recall when encoded with a matching emotion word; moreover, the amount of facial muscle activity in the perceiver's face (assumed to be the product of simulation of the emotion word) during encoding predicted the amount of memory bias. That is, participants who showed more facial activity congruent with the stereotype of the emotion label during encoding indicated targets as containing more of that stereotyped emotion during recall, even when the category information was not paired with the face at recall (Halberstadt, Winkielman, Niedenthal, & Dalle, 2009).

More direct evidence for the *language-as-context* hypothesis comes from experiments which do not rely on explicit judgments of emotion. For example, language effects are present in categorical perception studies, in which perceivers are better able to detect differences among individual emotional faces that are assigned to separate categories than they are at detecting differences of the same magnitude from faces assigned to the same category (for a review, see Fugate, 2013). When perceivers are first placed under verbal load (which makes activating emotion words nearly impossible), categorical perception for emotional faces is eliminated (Roberson & Davidoff, 2000; Roberson, Damjanovic, & Pilling, 2007). In addition, human perceivers show categorical perception for chimpanzee faces (similar to those produced by humans but for which naïve participants do not readily evoke emotion labels) only when first learning the categories of faces with labels (Fugate, Gouzoules, & Barrett, 2010). In that study, participants who learned the same categories (to the same proficiency), but without a label, did not show categorical perception. Thus, having previously learned a label was enough to create the enhanced perceptual ability (Fugate et al., 2010). Perceptual priming of emotional faces is also impacted by the accessibility of emotion language. When the meaning of an emotion word is satiated, participants do not show repetition priming (decreased RTs to recognize a previously seen stimulus) for emotional faces, suggesting that in the absence of emotion words the face is seen as a new stimulus (Gendron, Lindquist, Barsalou, & Barrett, 2012).

Together, the results of all these studies suggest that language (specifically emotion words) affect emotion judgments, likely at various levels of processing. The mechanism by which this occurs, however, is not fully understood. One account is that emotion words have their impact via embodiment (Halberstadt, 2003, 2005; Niedenthal, 2007, 2008). According to an embodied theory of emotion, words serve to activate multiple emotional "outputs" (e.g., muscular movement, physiological changes, etc.) emphasized by a particular emotion category. The word leads to simulation (i.e., reenactment) of perceptual, sensory, and behavioral experiences associated with a category-specific emotion, and these simulations guide judgments of category membership. In other variants of embodied emotion, individual instances of affective and contextual information are encoded and bound together by emotion words into meaningful categories. Over time such heterogeneous instances can come to represent new categories (Wilson-Mendenhall, Barrett, Simmons, & Barsalou, 2011; Barrett, Wilson-Mendenhall, & Barsalou, 2014). Such embodied instances can later be simulated to understand one's own or another's emotion state (Barrett et al., 2014).

In other accounts, words serve to sharpen the visual representations of the stimulus itself (see Lupyan, 2012). In this viewpoint, language is an effective means of propagating neural activity because it can activate a distributed representation of related content that can be assigned in multiple categories depending on context and goal-relevancy (Lupyan & Clark, 2015). The prediction that emotion judgments should be more difficult when conceptual information is unavailable to the perceiver than when it is available is consistent with the *language-as-context hypothesis*. (see reviews by Barrett et al., 2007; Fugate & Barrett, 2014; Lindquist & Gendron, 2013; Lindquist, Satpute, & Gendron, 2015). This view is also consistent with those that suggest that words serve to constrain and highlight possible interpretations of a visual stimulus (Bar, 2007; Lupyan, 2012; Lupyan & Clark, 2015; Lupyan & Thompson-Schill, 2012; Lupyan & Ward, 2013; Kveraga, Ghuman, & Bar, 2007).

The present experiments were designed to test how emotion words affect perceptual memory judgments and whether emotion words have a greater effect when the structural information in the face is limited. Although there are a handful of studies that pair emotion words with emotional faces (e.g., Halberstadt & Niedenthal, 2001), none prime the face on a trial-by-trial basis to see how they affect perceptual judgments. Moreover, such studies typically use faces which represent only two emotion categories, which differ in valence (or arousal). We use faces from four emotion categories, each representing a different state of core affect (pleasure-displeasure; high or low arousal; Russell, 1980; Russell & Barrett, 1999) to more thoroughly examine the impact of language across affective space. In addition, the majority of published studies use emotional faces from well-known face sets which are highly caricatured. We use a newer face set created from instructing participants to engage in moving their facial muscles in emotion-defining normative ways, but without explicit reference to an emotion.

Experiments 1 and 2

In Experiments 1 and 2, we presented emotion words and control words (in different trials) prior to an ambiguous target face created from morphing two normative facial depictions of emotion. Participants then saw the target face and two distractor faces. The distractor faces were systematically *more* and *less intense* (with respect to one of the emotions depicted in the target face). For example, participants saw a morphed target face created from *frowning* and *smiling* faces and were asked to select the same face from among a relatively *more frowning* face (*less smiling*) and a relatively *more smiling (less frowning*) face (see Figure 1). If emotion words affect perceptual memory judgments, then when primed with an emotion word (compared with a control word), participants should be more likely to select the distractor face that was *more intense* (with respect to the emotion prime).

In Experiment 2, participants indicated in separate trials whether each distractor face was the target face seen. By separating the actual target face and distractor faces into different trials, we used signal detection theory to explore a potential behavioral mechanism of how emotion words affect perceptual memory judgments. We calculated d' as a standard measure of sensitivity (Macmillan & Creelman, 1991), and c as a measure of bias based on criterion shift that is independent of sensitivity (Macmillan & Creelman, 1990, 1991; Stanislaw & Todorov, 1999). We predicted that participants should have decreased sensitivity (d') to faces when primed with an emotion word compared with a control word. Lower d' would suggest that there is greater perceptual overlap in the categories (target face and distractor faces) when congruent emotion words are evoked (see Stanislaw & Todorov, 1999). Although Nook et al. (2015) found that participants had higher d's when determining whether an emotion word "fit" a target face

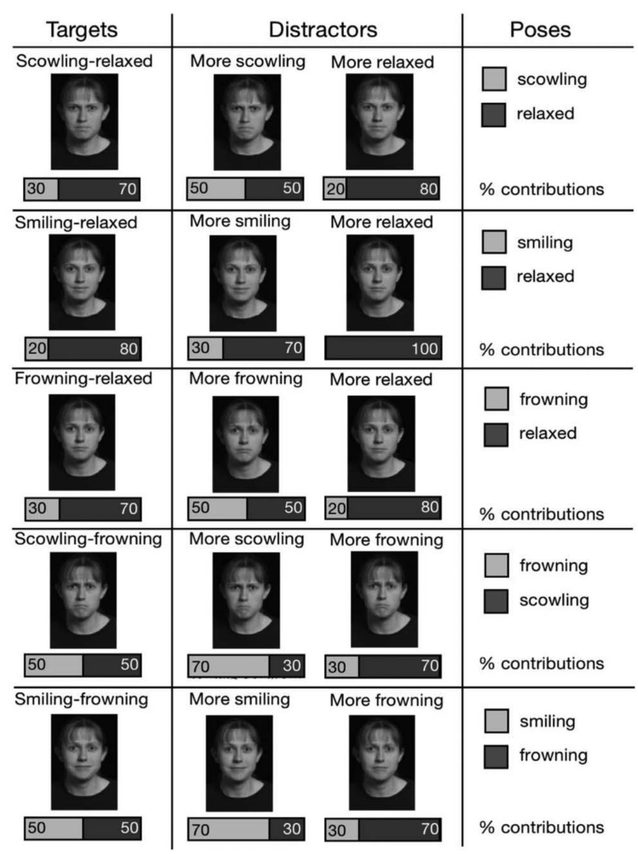

Figure 1. Face set morphs used in Experiments 1 and 2

(face-word trials), this was compared with participants' sensitivity to determine whether a second face was the target (face-face trials). In fact, our current prediction fits nicely with Nook's finding that as the visual similarity of emotional faces increases, participants' sensitivity for judging whether one face matches another faces decreases. We also tentatively predicted that emotion words might change how participants make their decisions. We predicted that participants would have negative response biases (c) to distractor faces when primed with a congruent emotion word compared with a control word. Negative bias represents a greater use of the "yes" key, meaning that participants become more liberal in answering "yes."

Experiment 1

Method

Participants. Thirty-six Boston College undergraduates (8 males, 28 females) between the ages of 18–22 years participated for one departmental research credit or for $10. Four participants identified themselves as "Asian," two as "Black/African American," 23 identified themselves as "White/Caucasian," and five as "more than one race." The remaining two participants did not respond. We did not link demographics to individual participant identifiers, so it was not possible to separate non-White participants' data from those of White participants in this Experiment (but see Experiment 2).

Stimuli. The faces used to create the target morphs were from a standardized set of facial depictions of emotion created and maintained by the Interdisciplinary Affective Science Laboratory.[1] All faces used in these experiments were Caucasian, but both male and female faces were used. This face set has been externally validated and used reliably in several empirical, peer-reviewed articles. We selected the same four faces (*scowling, smiling, relaxed,* and *frowning*) for each of 10 identities (5 male, 5 female). We selected these particular facial depictions because each is associated with a discrete emotion category representing a different combination of core affect (*scowling,* angry: high arousal/negative; *smiling,* happy: high arousal/positive; *relaxing,* calm: low arousal/positive; *frowning,* sad: low arousal/negative). Thus, the selection of these particular faces also allowed us to investigate specific category priming from more generalized affective priming. We morphed two facial depictions from each identity together to create six affective face sets (*scowling–relaxing, scowling–frowning, smiling–relaxing, smiling–frowning, frowning–relaxing, and smiling–scowling*). We created five morphs (80%–20%, 70%–30%, 50%–50%, 30%–70%, and 20%–80%) for each affective face set for each identity using commercially available software (FantaMorph).

Twenty-six pilot participants viewed and labeled each morph (by typing a one-word term) so that we could identify the most ambiguous morph of the five for each identity for each face set (to be used as the target face). Two coders corrected misspellings and typographical errors when the word was reasonably clear and could not be confused for another term. Both coders then recoded synonyms to each of the emotion categories (see Russell, 1994). We defined the most ambiguous morph as the one labeled as both emotions approximately the same number of times (e.g., "angry" and "calm" for the *scowling–relaxing* face set). Thus, we chose the morphed face that was perceptually ambiguous and not necessarily structurally ambiguous (50%–50%). For most identities and affective face sets, the most perceptually ambiguous face was the 30%–70% (or 70%–30%) morph (54% of the time). The two distractor faces in a trial were always the morphs one step greater and one step less than the target face (e.g., target face: 70% *scowling*–30% *relaxing*; distractor face one: 80% *scowling*–20% *relaxing*; distractor face two: 50% *scowling*–50% *relaxing*).[2] The *smiling–scowling* affective face set was not used in the experiment because none of the morphs (for the majority of identities) were identified approximately the same number of times as "happy" and "angry."

Procedure. On a given trial, each target face was primed with two emotion words and a control word (in different trials). The two emotion words were those that corresponded to the normative emotions in the target face (e.g., "anger" and "calm" for the *scowling–relaxing* face set). The control word was an abstract concept word, but not an emotion word (e.g., "belief"). On each trial, a fixation cross appeared centrally on the computer screen for 1,000 ms. Immediately after the cross disappeared, the prime appeared for 500 ms approximately half way down the computer screen. Immediately after the prime offset, the ambiguous target face (3 in. $\times$ 3 in.) appeared in the middle of the computer screen for 1,000 ms. Finally, immediately following the target face offset, three choice faces (the target face and two distractor faces) appeared slightly lower than center on the computer screen (each 3 in. $\times$ 3 in.). Each answer face had a number "1," "2," or "3" underneath the image. The order of the three faces was counterbalanced across trials. The three faces remained on the screen until participants provided a response (see Figure 2). Participants kept the first three fingers of their dominant hand on the numbered keys at all times. We did not direct participants to the primes in any way, except that we asked them to look at the screen at all times. There were a total of 150 experimental trials presented randomly in a single session. We collected accuracy and reaction time (RT) data on each trial.

Results

Calculation of bias scores. Overall, participants selected the correct target face 47.5% of the time. For each participant, we coded each key press into a "bias score" for each target face. If participants selected the *more intense* distractor face with respect to the emotion word, we coded the response as +1 (e.g., the participant was primed with "anger" and selected the *more scowling* distractor face). If participants selected the *less intense* distractor face with respect to the emotion word prime "anger," we coded the response as −1 (e.g., the participant selected the *less scowling*

[1] Development of the Interdisciplinary Affective Science Laboratory (IASLab) Face Set was supported by the National Institutes of Health Director's Pioneer Award (DP1OD003312) to Lisa Feldman Barrett. Gendron, M., Lindquist, K. A., & Barrett, L. F. (unpublished data). Ratings of IASLab facial expression stimuli available through http://www.affective-science.org/ Participants gave explicit consent to be photographed and their likness reproduced.

[2] For trials in which the target face was the 30%–70% and 20%–80% morph, the physical distance between the correct and incorrect faces was unequal (always 10% and 20%, respectively). For trials in which the target face was the 50%–50% morph, however, the physical distance between all three faces was equal.

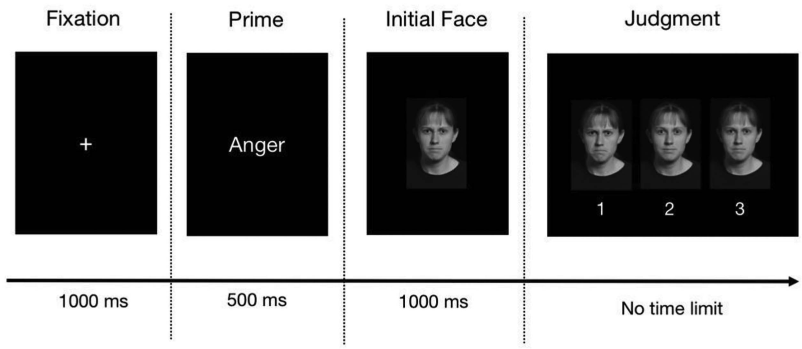

Figure 2. Procedure for Experiment 1

face). If participants selected the correct target face on a trial, we coded the key press as 0 (no bias). In the trials primed with a control word, we coded the selection of a distractor face for each of the two emotions to which the trial served as a control. For example, when a *smiling–scowling* face was primed with a control word, a selection of a *more smiling* face was coded as +1 to compare with the answer selected when this trial was primed with the emotion word "happy," and as a −1 to compare with the answer selected when this trial was primed with the emotion word "anger."

Analyses. We report the analyses for bias score and reaction time separately.

Bias score. We first checked for significant main effects of target face identity and individual face set. There were no main effects for target face identity or face set, so we aggregated across both prior to analysis. Although not necessarily anticipated, we found a main effect of emotion category (*angry, sad, happy,* and *calm*). Therefore, to test our hypothesis about the effect of words, we also retained emotion category as a factor in our analysis. We used a two factor repeated measures analysis of variance (rmANOVA). Factor 1 was the word type (emotion vs. control word). Factor 2 was emotion category. Overall, we found a main effect of word type: $F(1, 35) = 3.520$, $p < .05$; Power = 0.742; $\eta_p^2 = 0.091$; a main effect of individual emotion category: $F(3, 105) = 29.209$, $p < .001$; Power = 1.000; $\eta_p^2 = 0.455$. Moreover, there was a significant interaction: $F(3, 105) = 2.913$, $p < .05$; Power = 0.963; $\eta_p^2 = 0.077$. Because there was no limit to respond to trials, we checked whether the pattern of data changed when participants who responded more than three SDs outside the average range of RTs were removed from analyses, as is standard for similar types of studies (e.g., Adams & Kleck, 2003). No differences were found in the data when these participants ($n = 3$) were removed. Therefore, we kept all 36 participants in for the rest of the analyses.

To follow up the interaction, we performed separate paired-sample *t* tests between emotion and control words for each emotion category. For all four emotion categories, participants had larger, positive biases when primed with an emotion word. For the majority of emotion categories, this positive bias was caused by participants selecting the *more intense* face when primed with the congruent emotion word, but also selecting the *less intense* face when primed with the control word. The biases produced between emotion and control words were not significant for all emotion categories, however (see Table 1 for means and standard deviations). Because we were interested in the overall effect of emotion words (compared with control words) we optimized our paradigm for this analysis. We include Table 1 simply for those readers who are interested in the effects by emotion category, although we caution against drawing conclusions at the level of an individual emotion category.

Our data show that emotion words affected perceptual memory judgments by creating biases toward the *more intense* distractor face (with respect to the emotion prime). We view these findings as a necessary first step to understanding the effect of emotion words on emotion judgments.

Reaction times. We also performed a similar set of analyses using participants' RTs. We did not find any significant main effects or a significant interaction. Thus, the speed with which participants made perceptual memory judgments was the same regardless of the type of word prime and emotion category. We interpret this to mean that participants' biases were not the result of a speed–accuracy trade-off.

Experiment 2

The purpose of Experiment 2 was to better understand the nature of the biases that emotion words created in Experiment 1. We wanted to see whether the biases produced in Experiment 1 could be attributed to participants' decreased sensitivity to discriminate between target and distractor, and/or changing the criteria to respond. In this experiment, we presented the target face and the two distractor faces in separate trials rather than as an array. This

Table 1
Mean Differences on Bias Scores, Confidence Intervals, Effect Sizes, and Statistical Power (Achieved Post Hoc) Between Emotion Words and Control Words in Experiment 1

Category	Difference, *M* (*SD*)	Confidence intervals	Effect size *dz*	Power
Angry	.167 (.343)*	[.051, .283]	.486	.889
Happy	.043 (.263)	[−.046, .132]	.164	.247
Sad	.069 (.334)	[−.044, .181]	.207	.334
Calm	.065 (.340)	[−.050, .180]	.191	.302

* significantly different, $p < .05$, based on dependent *t* tests.

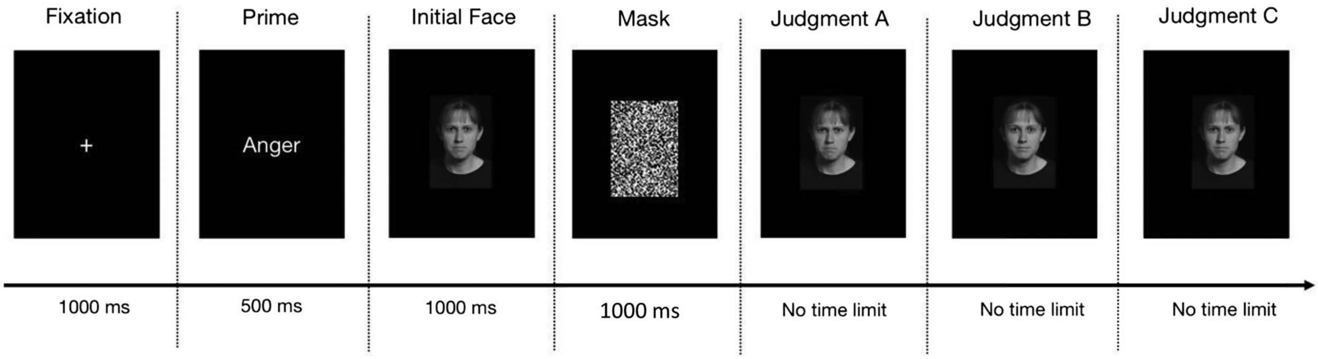

Figure 3. Procedure for Experiment 2

change allowed us to look at participants' likelihood to accept or reject both the *more* and *less intense* distractor faces (rather than just which of the two was selected). It also allowed us to use signal detection theory to determine whether emotion words changed participants' sensitivity and bias, thereby informing us of a potential mechanism for emotion words' effect on perceptual memory judgments of emotion.

Method

Participants. Thirty-six Boston College undergraduates (12 males, 24 females) between the ages of 18–22 years participated for one departmental research credit or for $10. Three participants identified themselves as "Asian," five as "Black/African American," 25 as "White/Caucasian," and three as "more than one race." In this study, we kept race at an individual identifier so it was possible to separate "White" from "non-White" responders in the analysis (see below).

Stimuli. Half of the stimuli used in Experiments 1 and 2 were used as target faces in this experiment. Instead of 10 identities, we used two male and two female faces to limit the number of trials (see procedure).

Procedure. We followed the same procedure as in Experiment 1, with the exception that the each of the three answer faces (the target face and the two distractor faces) was shown in separate trials. All other aspects of the study remained the same to those in Experiment 1, with the additional exception that we added a static mask for 1,000 ms to disrupt the processing of the initial image, and both the initial and answer faces were presented larger (approximately 4 in × 4 in) and in the middle of the screen (see Figure 3). There were a total of 180 experimental trials presented randomly in a single session. Participants were instructed to press the "1" key if the answer face was the same as the target face and the "2" key if it was a different face. There was no time limit to respond.

Results

Coding of false alarms and missed detections. Overall, participants answered correctly 33.3% of the time, thus ensuring that the task was difficult enough to prevent ceiling effects and provide somewhat nonoverlapping distributions of target and distractor faces. We first coded all correct answers with a 0 and all incorrect answers with a 1. Participants were incorrect if they answered "no" to the answer face when it was the target face (e.g., missed detection) or answered "yes" to the answer face when it was a distractor face (e.g., a false alarm). Participants were correct if they answered "yes" when the answer face was the target (e.g., correct detection) or answered "no" to the answer face when it was a distractor face (e.g., correct rejection). Second, we used a full signal detection theory (SDT) analysis to calculate sensitivity (d') and bias (c) to assess how emotion words affected the judgments.

Analysis 1—Incorrect responses. The majority of incorrect responses in our paradigm were false alarms (M = 67.2%; SE = 2.4%), with relatively few missed detections. Therefore, we analyzed only false alarms for which we had sufficient data. To do so, we performed a three-factor rmANOVA, with word type (emotion vs. control words), emotion category, and face strength (*more intense* distractor and *less intense* distractor) as factors. We found a main effect for word type: $F(1, 35) = 8.374$, $p < .001$; Power = 0.981; $\eta_p^2 = 0.193$, and emotion category: $F(3, 105) = 11.781$, $p < .001$; Power = 0.999; $\eta_p^2 = 0.252$. Moreover, there was one significant two-way interaction between emotion category and face strength: $F(3, 105) = 38.761$, $p < .001$; Power = 1.00, $\eta_p^2 = 0.525$, and one marginally significant two-way interaction between word type and face strength: $F(1, 35) = 3.337$, $p < .10$; Power = 0.982; $\eta_p^2 = 0.087$. When we removed the 10 "non-White" participants' data and reran the analysis, all effects were the same. Therefore, we did not treat "non-White" and "White" participants' data differently in the rest of the analyses reported. We also checked whether any participant responded outside the normal range of RTs, as we did in Experiment 1. Only one participant's overall RT qualified them as an outlier (>3 *SD*s). When we removed this participant to make sure his or her data did not affect the overall pattern of results, no differences in the data were found. Therefore, we kept this participant's data in for the rest of the analyses.

To follow up the interaction, while preserving word type (central to our hypothesis), we analyzed each emotion category separately. For all emotion categories, participants made more false alarms when primed with an emotion word compared with a control word. They also made more false alarms to the distractor face that was *more intense* (with respect to the emotion prime; see

Table 2
Mean Differences on Bias Scores, Confidence Intervals, Effect Sizes, and Statistical Power (Achieved Post Hoc) Between Emotion Words and Control Words to More Intense and Less Intense Distractor Faces in Experiment 2

Category	Difference, *M* (*SD*)	Confidence intervals	Effect size *dz*	Power
More intense				
Angry	.094 (.228)*	[.017, .171]	.412	.782
Happy	.097 (.248)*	[.013, .181]	.391	.744
Sad	.026 (.171)	[−.325, .083]	.152	.227
Calm	.072 (.158)*	[.018, .125]	.455	.850
Less intense				
Angry	.024 (.223)	[−.051, .100]	.108	.156
Happy	.007 (.194)	[−.059, .072]	.036	.076
Sad	.039 (.160)	[−.015, .093]	.244	.417
Calm	.069 (.177)*	[.010, .129]	.390	.742

* significantly different, $p < .05$, based on dependent *t* tests.

Table 2 for *Ms* and *SDs*).[3] Not all interactions between face strength and word type were significant when separated by emotion category, however. We include Table 2 for those interested in the breakdown of results by emotion category for the *more* and *less intense* faces.

Analysis 2—Signal detection theory (SDT). For our SDT analyses, we calculated both d' (as a measure of sensitivity) and c (as a measure of bias, independent of sensitivity; Macmillan & Creelman, 1990, 1991; Stanislaw & Todorov, 1999). d' measures the distance between the mean of the signal distributions and the mean of the noise distribution in standard deviation units. A value of 0 indicates an inability to distinguish signal from noise; the larger the value the greater the ability to distinguish between signal and noise (see Stanislaw & Todorov, 1999). Although β has historically been the most common measure of response bias, most investigators now prefer to use c which assumes that participants respond "yes" when the decision variable exceeds the criterion and "no" when it does not. Therefore, responses are based directly on the decision variable rather than a likelihood ratio (Richardson, 1994; cf. Stanislaw & Todorov, 1999, p. 140). c is also not affected by changes in d', as β typically is affected. c is defined as the difference between the criterion and the *neutral point* where neither response is favored. Negative values of c signify a bias toward responding "yes" (the criterion lies to the left of the neutral point), whereas positive values signify a bias toward the "no" response (the criterion lies to the right of the neutral point; cf. Stanislaw & Todorov, 1999, p. 140). We used the following formulas for calculating d' and c (Macmillan & Creelman, 1990, 1991; Stanislaw & Todorov, 1999):

$$d' = \text{norminv (hit)} - \text{norminv (false alarm)}$$

$$c = -[\text{norminv (hit)} + \text{Norminv (false alarm)}]/2$$

***d′* analyses.** We now performed a three-factor rmANOVA on d'. We found a main effect of word type: $F(1, 35) = 30.208$, $p < .001$; Power = 1.00; $\eta_p^2 = 0.463$, as well as main effects for emotion category and face strength: $F(3, 105) = 13.319$, $p < .001$; Power = 0.999; $\eta_p^2 = 0.276$ and $F(1, 35) = 8.284$, $p < .01$; Power = 0.980; $\eta_p^2 = 0.191$, respectively. Moreover, we found significant effects for all two-way interactions,[4] and a significant three-way interaction: $F(3, 105) = 2.976$, $p < .05$; Power = 0.999; $\eta_p^2 = .078$.

Despite the three-way interaction, we found that the effects of word type were similar for all four emotion categories. Therefore we aggregated across emotion category to simplify the interaction above. We found that both word type and face strength produced significant main effects: $F(1, 35) = 30.208$, $p < .001$, Power = 1.000; $\eta_p^2 = 0.463$, and $F(1, 35) = 8.284$, $p < .01$, Power = 0.980; $\eta_p^2 = 0.191$, respectively. Moreover, the interaction between word type and face strength was significant: $F(1, 35) = 11.002$, $p < .01$; Power = .897; $\eta_p^2 = 0.239$. Despite the interaction, participants had less sensitivity (lower d') on both types of distractor faces when primed with an emotion word compared with a control word, $t(36) = -6.428$, $p < .001$ (*more intense* with respect to the emotion prime) and $t(36) = -2.675$, $p = .01$ (*less intense* with respect to the emotion prime; see Table 3 for *Ms* and *SDs*). We include Table 3 for those interested in the breakdown of results by emotion category for the *more* and *less intense* faces.

Overall, emotion words (compared with control words) lowered sensitivity between a *more intense* distractor and a target, as well as between a *less intense* distractor and a target. We interpret this to mean that emotion words decreased the distance between the peaks of the target and distractor category distributions (i.e., there is more overlap of the distributions in the presence of emotion words). Therefore, participants' biases in Experiment 1 were not due to the fact that they were simply matching the most intense face to the emotion word (i.e., a response bias). If this had been the case, we would have expected participants to show lowered d' between the target and *more intense* face only. We maintain, however, that ordering the faces and selecting the face that was most like the emotion word would have been difficult to do in this experiment because the faces were shown separately. Experiment 3 addresses this issue ever further to show that the effects are more than response biases.

c analyses. Although sensitivity to both types of distractor faces was decreased in the presence of emotion words (compared with control words), we next wanted to assess whether there was also a shift in response bias, independent of sensitivity. To this end, we performed the same three-factor rmANOVA as above, but now on c. We found main effects of word type and emotion category: $F(1, 35) = 15.254$, $p < .001$; Power = 1.00; $\eta_p^2 = 0.304$, and $F(3, 105) = 4.747$, $p < .01$; Power = 0.961; $\eta_p^2 = 0.119$, respectively. Moreover, we found one significant two-way inter-

[3] Main effects for word type for each emotion category "angry": $F(1, 35) = 4.074$, $p = .05$; Power = 0.501; $\eta_p^2 = 0.104$; "happy": $F(1, 35) = 3.514$, $p < .05$; Power = .446, $\eta_p^2 = 0.091$; "calm": $F(1, 35) = 10.220$, $p < .01$; Power = 0.875; $\eta_p^2 = 0.226$; and "sad": $F(1, 35) = 2.556$, $p < .10$; Power = 0.343; $\eta_p^2 = 0.068$. Main effects for face strength for each emotion category "angry": $F(1, 35) = 34.737$, $p < .001$; Power = 1.0; $\eta_p^2 = 0.498$); "happy": $F(1, 35) = 35.309$, $p < .001$; Power = 1.0; $\eta_p^2 = 0.502$; "calm": $F(1, 35) = 65.500$, $p < .001$; Power = 1.0; $\eta_p^2 = 0.652$; "sad": $F(1, 35) = 14.646$, $p < .001$; Power = 0.961; $\eta_p^2 = 0.296$.

[4] Significant two-way interactions are as follows: *wordtype * face strength:* $F(1, 35) = 11.002$, $p < .01$, Power = 0.897; $\eta_p^2 = .239$; *word type * emotion category:* $F(3, 105) = 22.154$, $p < .001$, Power = 1.0, $\eta_p^2 = .388$; *emotion category * face strength:* $F(3, 105) = 15.965$, $p < .001$, Power = 1.0, $\eta_p^2 = .313$.

action[5] and a significant three-way interaction: $F(3, 105) = 4.851$, $p < .01$; Power = 1.00; $\eta_p^2 = .122$.

To keep the follow-up analyses consistent with those for d', we next aggregated over emotion category to simplify the interaction. We found only a significant main effect for word prime: $F(1, 35) = 15.254$, $p < .001$; Power = 1.00; $\eta_p^2 = 0.304$. The effect of face strength and the interaction were not significant. We found that the effects of word type were similar for all four categories, although only three reached statistical significance (with the fourth being marginally significant). Participants had more "yes" responses when primed with an emotion compared with a control word (see Table 4 for *Ms* and *SDs*). We include Table 4 for those who are interested in seeing the results by emotion category for the *more* and *less intense* faces.

Analysis summary. Overall, the results show that emotion words (compared with control words) caused participants to make more false alarms, especially to distractor faces which were *more intense* (with respect to the emotion prime). Using signal detection theory, however, we find that participants' sensitivity was lowered for both distractor faces under these conditions. Emotion words also change participants' response bias (c) to include more "yes" responses. Thus, participants were more liberal in the use of the "yes" key when they saw an emotion word. Interesting, however, this shift in strategy is independent of the decrease to discriminate (sensitivity). Therefore, emotion words cause people to become less sensitive to the changes in distracting emotional content and also change their response bias to include more "yes" responses.

Experiment 3

In Experiment 3, we explored the possibility that emotion words have more of an effect on perceptual memory judgments when access to structural information in the face is limited, consistent with both the *language-as-context* hypothesis and the idea that words affect perceptual encoding. We also further addressed the possibility that the biases found in the previous two experiments reflected shifts in response selection rather than a perceptual phenomenon. To address both things, we made three methodological modifications in Experiment 3.

Table 3
Mean Differences for d', Confidence Intervals, Effects Sizes, and Statistical Power (Achieved Post Hoc) Between Emotion Words and Control Words to More Intense and Less Intense Distractor Faces in Experiment 2

Category	Difference, *M* (*SD*)	Confidence intervals	Effect size *dz*	Power
More intense				
Angry	−0.575 (1.45)	[−0.789, 0.187]	.398	.756
Happy	−0.318 (1.80)	[−0.927, 0.624]	.177	.274
Sad	−2.804 (2.68)*	[−3.709, −1.899]	1.05	.999
Calm	−0.274 (0.98)*	[−0.606, 0.574]	.280	.500
Less intense				
Angry	−0.302 (1.44)	[−0.790, 0.187]	.209	.340
Happy	0.124 (1.48)	[−0.377, 0.624]	.083	.124
Sad	−1.226 (1.88)*	[−1.863, −0.589]	.652	.987
Calm	−0.258 (1.39)	[−0.727, 0.211]	.186	.291

* significantly different, $p < .05$, based on dependent *t* tests.

Table 4
Mean Differences for C (Bias), Confidence Intervals, Effect Sizes, and Statistical Power (Achieved Post Hoc) Between Emotion Words and Control Words to More Intense and Less Intense Distractor Faces in Experiment 2

Category	Difference, *M* (*SD*)	Confidence intervals	Effect size *dz*	Power
More intense				
Angry	−0.279 (0.802)*	[−.550,−.007]	.348	.656
Happy	−0.279 (0.772)*	[−.540, −.018]	.362	.685
Sad	−0.017 (0.683)	[−.248, .214]	.025	.068
Calm	−0.328 (0.560)*	[−.518,−.139]	.587	.965
Less intense				
Angry	−0.142 (0.690)	[−.376, .913]	.206	.333
Happy	−0.058 (0.660)	[−.281, .165]	.088	.130
Sad	−0.508 (1.04)*	[−.861,−.155]	.488	.890
Calm	−0.321 (0.576)*	[−.516, −.182]	.557	.948

* significantly different, $p < .05$, based on dependent *t* tests.

First, we varied the presentation time of the target face: participants saw a target (unmorphed) emotional face for either 50 ms or 100 ms. We predicted that at shorter time intervals, when the structural information is limited, participants should have larger biases toward the chosen emotion word. If the word did not affect the encoding of the face but merely response bias, we predicted that there should be no difference based on how long the face was seen.

Second, we manipulated the relative congruency (based on valence or arousal) between faces and emotion word labels. Instead of supplying emotion and control words as primes, we now had participants choose an emotion word that was either partially or fully congruent with the face, prior to completing the same perceptual memory task as in Experiment 1. In some blocks one emotion word was only partially congruent with the target face; in another blocks, the emotion word was fully congruent with the target face (see Figure 4). By manipulating the emotion words to be fully or partially congruent with the faces in some blocks, we could further rule out that the biases we previously found reflected changes in response selection rather than encoding. We predicted the smallest biases in the fully congruent blocks. Although we predicted that the partially congruent blocks would produce larger biases than the fully congruent block, we did not have an a priori prediction about which (arousal-congruent or valence-congruent) would produce larger perceptual biases. Thus, we could also assess whether emotion word labels that matched faces (normatively) on arousal or valence were more effective in creating perceptual memory biases. Third, we had participants actively choose the best emotion word after the presentation of the target face (rather than passively view them before the target face, as in Experiment 1 and 2). By having participants choose an emotion word as a label, we ensured that participants were attending to the relevancy of the word and applying the words to the face.

[5] Significant two-way interaction is as follows: *emotion category * face strength:* $F(3, 105) = 31.307$, $p < .001$, Power = 1.00, $\eta_p^2 = .472$.

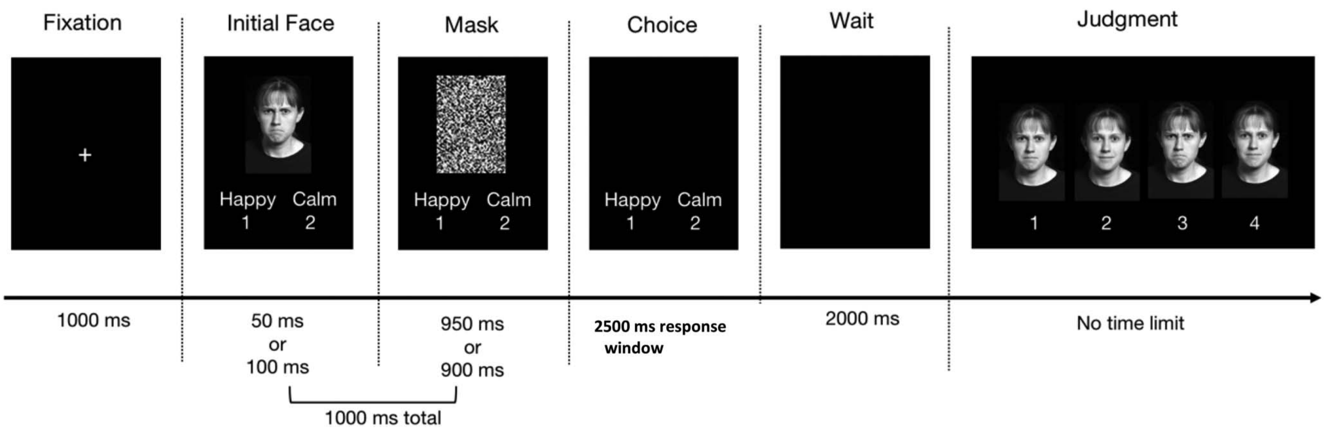

Figure 4. Procedure for Experiment 3

Method

Participants. Fifty participants recruited at Northeastern University (28 Male, 22 Female) between the ages of 18–45 years participated for one departmental research credit or $10 payment. Two participants' data were removed because they did not complete the experimental task. One additional participant's data was removed because they did not render any responses to the labeling task. Thirty-one participants identified as "White/Caucasian," 10 as "Asian," six as "African American/Black," and three as "more than one race." We did not link demographics to individual participant identifiers, so it was not possible to separate non-White participants' data from those of White participants in this Experiment.

Stimuli. The target faces were eight identities (4M and 4F) from the IASL face set used in Experiments 1 and 2. We used four emotional faces for each identity: *frowning, smiling, relaxing,* and *scowling.* In all blocks described next, one emotion word was fully incongruent to the target face on both arousal and valence. In the fully congruent emotion label block (CL), the other emotion word exactly matched the target face (normatively). The distractor faces were created from blending systemically (morphing) the target face with the face identity of the incongruent emotion word. For example, for a *smiling* target face in the CL block, the emotion words from which to choose were "happy" and "sad." The distractor faces were then incremental morphs of *smiling–frowning* faces. The same *smiling* target face in the partially congruent on arousal block (AL) would be paired with the emotion words "sad" and "anger" ("anger" shares high arousal with happiness depicted in a *smiling* face). The distractor faces were then incremental morphs of *smiling–scowling* faces. For the same target face in the partially congruent on valence block (VL), the emotion words were "sad" and "calm" ("calm" shares positive valence with happiness depicted in a *smiling* face). The distractor faces were then incremental morphs of *frowning–relaxing.* Table 5 shows the conditions in detail.[6] In total, we created six morphed sets for each identity's *smiling, frowning, scowling, and relaxing face* (e.g., *smiling–frowning, scowling–relaxing).* All faces were presented at 151 × 227 pixels on the center of the computer screen.

Procedure. On each trial, participants first saw a central fixation cross presented on the computer screen for 1,000 ms. We next presented participants with an original, unmorphed emotional face (*scowling, frowning, relaxing,* and *happy*) for 50 ms and 100 ms, in separate trials. Immediately after the target face, we presented a mask so that the total time between the target and presentation of emotion word labels was always 1,000 ms. Concurrent with the face, two emotion words appeared on either side of a monitor and participants were asked to choose the best emotion word to label the face by pressing one of two keys. Participants could respond any time after the face initially appeared on screen until 2,500 ms total had elapsed (see Figure 4). We instructed participants to consider both words but to indicate their decision as fast as possible. We randomized which word appeared on the right and the left of the screen across trials. Following the response period, we presented a blank screen for 2,000 ms. Finally, we presented the perceptual array of four faces: the target face and the three distractor faces (identical to the procedure in Experiment 1 except that we used three distractor faces and the target face in the array). The four faces' positions were randomized across trials. Each face appeared with a number underneath it (1, 2, 3, 4). Participants pressed the number that indicated which face was the target face. There was no time limit to indicate which face was the target face. There were 96 experimental trials in four blocks described above. All trials of a given type were presented within a block, and the blocks were presented in a randomized order (in a single session) across participants. We collected response choice and RT data for both the emotion word that was chosen as the label and the actual perceptual judgment on each trial.

Results

Calculation of bias scores. We removed participants' responses to the perceptual matching judgment that were less than 250 ms and greater than 10,000 ms (decreasing the number of trials submitted for analysis by 3.53%). As in Experiment 1, we calculated a bias score for each judgment based on whether participants

[6] In a fourth trial type, participants completed the same task without assigning an emotion word after the presentation of the target face. These trials were not included in analysis for this paper, as they did not involve the presence of emotion words.

Table 5
Combination of Emotion Words for Each Emotion

	Angry		Happy		Sad		Calm	
Trial choice type	Congruent	Incongruent	Congruent	Incongruent	Congruent	Incongruent	Congruent	Incongruent
Category congruent (CL)	Angry		Happy		Sad		Calm	
Arousal congruent (AL)	Happy	Calm	Angry	Sad	Calm	Happy	Sad	Angry
Valence congruent (VL)	Sad		Calm		Angry		Happy	

selected the correct target face or one of the three morphed distractors. If a participant selected the correct face (target), we scored their answer as a 0 (no bias). If the participant selected the distractor face with the least emotional content of the incongruent emotion word (*70% target emotion-30% incongruent emotion*), we scored their answer as 1 (i.e., a weak bias toward the congruent/partially congruent emotion word). If participants selected the distractor face with the most emotional content of the incongruent emotion word (*30% target emotion −70% incongruent emotion*) we scored their answer as 3 (i.e., a maximal bias toward the congruent/partially congruent emotion word). Finally, when participants selected the 50%-50% morph, we scored their answer was scored as a 2 (i.e., moderate bias).

Perceptual memory bias on "correctly" labeled faces. We analyzed only the trials in which participants chose the congruent (CL) or partially congruent matching word (AL or VL). Participants rarely chose the fully incongruent label for CL and VL trials. First, we performed a two-factor rmANOVA using choice type (AL, VL, CL), and presentation time of target face (50 vs. 100 ms) as the within subject factors. We found main effects for choice type: $F(2, 78) = 29.065$, $p < .001$; Power = 1.00; $\eta_p^2 = .427$, presentation time of the target face: $F(1, 39) = 7.783$, $p < .001$; Power = 0.973; $\eta_p^2 = 0.166$. There was no significant interaction. Participants had larger biases at the shorter stimulus duration (M = 1.218, SE = 0.035) compared with the longer durations (M = 1.104, SE = 0.050). Participants had larger biases to the valence-congruent blocks (M = 1.347, SE = 0.041) compared with the arousal congruent blocks (M = 1.250, SE = 0.072) and the fully congruent blocks (M = 0.888, SE = 0.039). Pairwise comparisons showed that VL and AL trials produced larger biases than the CL condition, consistent with predictions, $p < .05$ for both (see Figure 5). We also performed a one-factor rmANOVA using face type (*relaxed, smiling, frowning and scowling*) as the within subjects factor. This analysis was conducted separately due to missing data (i.e., insufficient endorsement of the arousal congruent label) for the AL condition across face types. This analysis revealed a main effect of face type: $F(3, 135) = 13.703$, $p < .001$; Power = .999; $\eta_p^2 = 0.233$. Additionally, participants had the strongest bias to *relaxed* (M = 1.260, SD = 0.044), followed by *frowning* (M = 1.051, SD = 0.050), *scowling* (M = 0.971, SD = 0.055), and finally *smiling* faces (M = 0.964, SD = 0.050). Pairwise comparisons showed that *relaxed* faces differed from all other faces, $p < .01$. Overall, the results suggest that when exposure to the target is short, participants are more biased by the emotion words applied. Specifically, when participants are asked to choose among emotion words which are congruent on valence (VL) and arousal (AL), they are more biased by the emotion word. Finally, some categories of emotion (namely, *relaxed* faces portraying calm) produce larger biases than others, as was the case in all three experiments.

Analysis summary. Our results indicate participants are more affected by emotion words when encoding of the face is limited by time constraints. This suggests that the less time the perceiver has to access the structural information (prior to the onset of the choices), the more effect words have on perceptual memory judgments, consistent with the *language-as-context* hypothesis. In addition, our results indicate that perceptual memory judgments are most affected when the emotion words are congruent on valence (rather than fully congruent, for which we would not expect labels

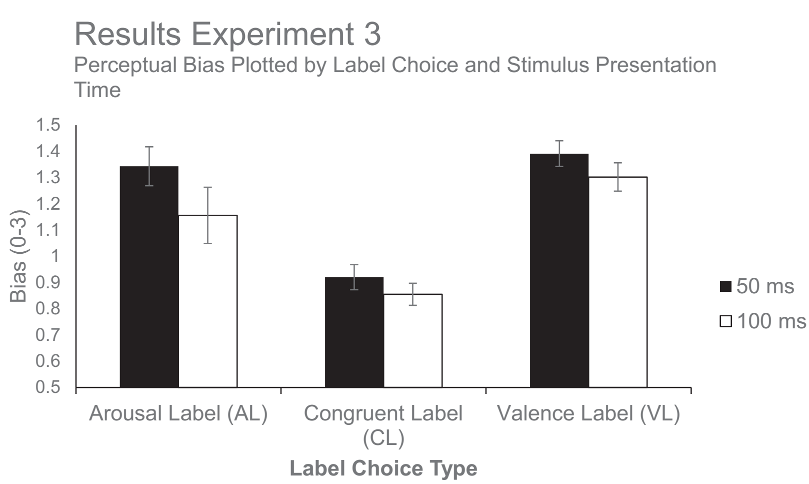

Figure 5. Bias results for Experiment 3

to produce a biasing result). Together, our results support our central hypothesis that when encoding of the structural information in the face is limited, emotion words have more effect on perceptual memory judgments. The findings are also consistent with the idea that emotion words can affect the perceptual encoding of the stimulus rather than only response selection.

Reaction times on perceptual memory judgments. As in previous experiments, there was no significant main effect of choice type, but a marginal main effect of timing: $F(1, 39) = 3.413$, $p = .072$; Power $= .730$; $\eta_p^2 = 0.080$. There was was no interaction between the two. These results suggest that RTs increased only marginally to complete the task when the presentation timing of the initial stimulus was shorter (50 ms). Given that this effect was only marginal, the biases observed were unlikely to be due to participants performing the task more or less quickly in some conditions.

General Discussion

For many years, psychology has assumed that emotion perception is a largely bottom-up process that is driven by the information provided in the face of a target person. Structural views suggest that the information in another person's face supplies all the necessary structural information to arrive at a discrete emotion category judgment (e.g., Ekman, 1972, 1992; Ekman & Cordaro, 2011; Izard, 1971, 1994; Matsumoto et al., 2008; Tomkins, 1962, 1963). According to these views, language is either independent to emotion perception (e.g., Brosch et al., 2010; Sauter et al., 2011) or serves to simply label preexisting categories (i.e., innate categories sediment out into language; Scherer, 2000). More recently, evidence has begun to accumulate that there are significant perceiver-based influences that affect perceptual judgments. Specifically the *language-as-context* hypothesis suggests that emotion perception should be more difficult when conceptual information is unavailable, or limited, than when it is available (see reviews by Barrett et al., 2007; Fugate & Barrett, 2014; Lindquist & Gendron, 2013; Lindquist et al., 2015). This view is also consistent with those which suggest that words serve to constrain and highlight possible interpretations of a visual stimulus (Lupyan, 2012; Lupyan & Clark, 2015; Lupyan & Thompson-Schill, 2012; Lupyan & Ward, 2013; Bar, 2007; Kveraga et al., 2007). The studies presented here explore how emotion words serve as one of these influences.

In two different types of perceptual memory tasks, we showed that emotion words (compared with control words) affected participants' abilities to match and determine which emotional face was recently seen. Specifically, participants primed with an emotion word (compared with a control word) selected (Experiment 1) and accepted (Experiment 2) a *more intense* emotional distractor face as the target. Specifically, participants primed with emotion words (compared with control words) showed decreased sensitivity (d') and more of a bias toward the "yes" key (c) for both distractor faces. Therefore, in the presence of emotion words, participants had difficultly discriminating between targets and distractors because emotion words reduced the perceptual distance between the target category and the distractor category. In addition, participants adjusted their responding in the presence of emotion words by using the "yes" key more, consistent with the fact that they made more false alarms.

In Experiment 3, we varied the timing of target face presentations while also requiring that participants label the face with an emotion word (rather than having emotion words serve as a prime). In some of the blocks, the emotion words were only partially congruent with the target face. Participants in Experiment 3 then performed the same perceptual memory task as in Experiment 1. Participants showed larger biases (i.e., choosing a *more intense* distractor face which matched the emotion word) when the target face was shown for the shorter of two times and when the emotion word was partially congruent with the target face. This suggests that when the structural information in the face is limited, people rely more on conceptual information to "fill in" the information, consistent with the *language-as-context* hypothesis. Said another way, the less structural information is available, the more language as a top-down source of information is important. Moreover, emotion words which were congruent on valence and arousal were more effective in changing perceptual judgments than those which were fully congruent. That is, participants "note" the relevancy of the word to the face in creating the perceptual biases. These results cast doubt (although cannot completely eliminate) that in the early experiments here (as well as in the literature, more broadly), emotion words are impacting only response selection, because our results were sensitive to the type of word and the degree of stimulus-driven information. In addition, the change in sensitivity (in addition to response bias) in Experiment 2 suggests that our effects are more than just differences in response selection.

Distinguishing effects of stimuli at the time of encoding from those that happen later in processing (i.e., post-processing, memory biases) is controversial and difficult to discern empirically (see Schyns, Goldstone, & Thibaut, 1998). In fact, the El Greco debate has recently recaptured this controversy (Firestone & Scholl, 2014, 2015). Specifically, how much of reported effects of top-down knowledge (including language) are truly effects on perception, or are they actually memory biases or response judgments? The El Greco effect comes from the idea that if these purported effects are perceptual in nature then they should be replicated in judgment (i.e., they would cancel each other out). Effects which remain at judgment would have to be post-processing effects. Interpreting any such effects as anything but after-the-fact has therefore become known as the "El Greco Fallacy" (see Firestone, 2013; Firestone & Scholl, 2014, 2015). The fact that several well-known studies which argue for top-down effects at encoding only occur when participants suspect the manipulation suggests that the effects must be due to response biases (Firestone & Scholl, 2014; see also Baker & Levin, 2016). Although our data do not address this debate head-on, we do show that when participants' exposure to a face is limited (Experiment 3), their bias is affected by the relationship of the word to the face. Therefore, the results are consistent with the idea that emotion words are affecting encoding processes rather than a simple response-bias.

Mechanisms of Language-as-Context

Clearly, from the results of the studies presented herein and others, emotion words are contributing to emotion perception at many levels. Studies of object perception (outside the emotion domain) can shed more light on the ways that perceiver-based influences might affect perceptual processing of stimuli. One model, proposed by Bar and colleagues, suggests that the orbitofrontal cortex (OFC), a multimodal brain region in the prefrontal cortex, helps integrate low spatial frequency information from the dorsal stream with top-down conceptual knowledge coming from other parts of cortex (Bar, 2003; Kveraga et al., 2007). The OFC is connected to many parts of the prefrontal cortex, including the inferior frontal gyrus which is important for semantic processing (e.g., Gitelman, Nobre, Sonty, Parrish, & Mesulam, 2005; for a review, see Binder, Desai, Graves, & Conant, 2009). Through such connections, language might affect how an object is seen or even whether an object is seen at all. A recent electroencephalography study suggests that one way in which language can affect visual perception of faces is by altering the N170 in posterior regions of the left hemisphere, likely through the ventral visual stream (Landau, Aziz-Zadeh, & Ivry, 2010). In addition, the P1 component is affected by learning a label for a category, in which the label is enough to induce CP for novel visual stimuli (Maier, Glage, Hohlfeld, & Abdel Rahman, 2014). In that study, adding more rich semantic content did not additionally contribute to the P1 effects of the word but did have separate effects at later stages of processing, including P2. The results are also consistent with the results of other ERP studies which show that the P1 component represents sustained visual activation that already includes several iterations of feedback from other cortical areas, as opposed to just feedforward processing (Foxe & Simpson, 2002).

These findings are in agreement with several studies that used different methods to demonstrate the impact of emotion words as a form of top-down context, adding to information given by the structural aspects of an emotional face (e.g., Barrett et al., 2007; Fugate et al., 2010; Gendron et al., 2012; Lindquist et al., 2006). Our findings are consistent with "predictive coding" accounts as well as embodied accounts in which concepts (anchored by emotion labels) are thought to be *instantiated* as activity in sensory cortices (e.g., Clark, 2013; Chanes & Barrett, 2016; Friston, 2010; Hohwy, 2013; Lupyan & Clark, 2015). The predictive coding perspective flips the traditional structure of neural organization, by suggesting that the brain operates in a largely predictive (top-down) manner, with bottom-up activity reflecting the errors of this predictive architecture. In this viewpoint, language is an effective means of propagating activity across the brain's predictive hierarchy because it can activate a distributed representation that allows for flexible predictions (Lupyan & Clark, 2015). This viewpoint suggests that the impact of language on emotion judgments is not a special case, but rather a natural consequence of neural organization.

Broader Implications for the Study of Emotion

Our results have important implications for studies of emotion. Many previous studies involving category judgments of emotion embed words within the paradigm. Here we have shown that emotion words are a powerful source of context, shaping how we perceive the emotional world. Thus, whether anticipated by the experimenter, the presence of words creates a context that can influence the results of an experiment. Hence, what the majority of published emotion judgment experiments demonstrate is how a person perceives emotion in the context of words (cf. Russell, 1994; see also Barrett et al., 2011). As a result, efforts to understand how emotion judgments occur in vivo should seriously consider how including emotion words in task paradigms can alter the phenomenon being investigated.

To conclude, this present research adds to prior findings (e.g., Gendron et al., 2012; Fugate et al., 2010; Lindquist et al., 2006; Roberson & Davidoff, 2000; for a review, see Barrett et al., 2007; Lindquist & Gendron, 2013) to help illuminate how emotion perception "goes beyond" the structural information provided by a stimulus, to include the words a perceiver knows.

References

Adams, R. B., Jr., & Kleck, R. E. (2003). Perceived gaze direction and the processing of facial displays of emotion. *Psychological Science, 14,* 644–647. http://dx.doi.org/10.1046/j.0956-7976.2003.psci_1479.x

Baker, L. J., & Levin, D. T. (2016). The face-race lightness illusion is not driven by low-level stimulus properties: An empirical reply to Firestone and Scholl (2014). *Psychonomic Bulletin & Review, 23,* 1989–1995. http://dx.doi.org/10.3758/s13423-016-1048-z

Bar, M. (2003). A cortical mechanism for triggering top–down facilitation in visual object recognition. *Journal of Cognitive Neuroscience, 15,* 600–609. http://dx.doi.org/10.1162/089892903321662976

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11,* 280–289. http://dx.doi.org/10.1016/j.tics.2007.05.005

Barrett, L. F. (2006a). Are emotions natural kinds? *Perspectives on Psychological Science, 1,* 28–58.

Barrett, L. F. (2006b). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review, 10,* 20–46. http://dx.doi.org/10.1207/s15327957pspr1001_2

Barrett, L. F. (2009). The future of psychology: Connecting mind to brain. *Perspectives on Psychological Science, 4,* 326–339. http://dx.doi.org/10.1111/j.1745-6924.2009.01134.x

Barrett, L. F. (2011). Constructing emotion. *Psychological Topics, 3,* 359–380.

Barrett, L. F. (2012). Emotions are real. *Emotion, 12,* 413–429. http://dx.doi.org/10.1037/a0027555

Barrett, L. F. (2013). Psychological construction: A Darwinian approach to the science of emotion. *Emotion Review, 5,* 379–389. http://dx.doi.org/10.1177/1754073913489753

Barrett, L. F. (2014). The conceptual act theory: A précis. *Emotion Review, 6,* 292–297. http://dx.doi.org/10.1177/1754073914534479

Barrett, L. F., Lindquist, K. A., & Gendron, M. (2007). Language as context for the perception of emotion. *Trends in Cognitive Sciences, 11,* 327–332. http://dx.doi.org/10.1016/j.tics.2007.06.003

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20,* 286–290. http://dx.doi.org/10.1177/0963721411422522

Barrett, L. F., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2014). The conceptual act theory: A road map. In L. F. Barrett & J. A. Russell (Eds.), *The Psychological construction of emotion* (pp. 83–110). New York, NY: Guilford Press.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19,* 2767–2796. http://dx.doi.org/10.1093/cercor/bhp055

Boucher, J. D., & Carlson, G. E. (1980). Recognition of facial expression in three cultures. *Journal of Cross-Cultural Psychology, 11,* 263–280. http://dx.doi.org/10.1177/0022022180113003

Brosch, T., Pourtois, G., & Sander, D. (2010). The perception and categorization of emotional stimuli: A review. *Cognition and Emotion, 24,* 377–400. http://dx.doi.org/10.1080/02699930902975754

Carroll, N. C., & Young, A. W. (2005). Priming of emotion recognition. *The Quarterly Journal of Experimental Psychology: Human Experimen-*

tal Psychology, 58A, 1173–1197. http://dx.doi.org/10.1080/02724980443000539

Chanes, L., & Barrett, L. F. (2016). Redefining the role of limbic areas in cortical processing. *Trends in Cognitive Sciences, 20,* 96–106. http://dx.doi.org/10.1016/j.tics.2015.11.005

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36,* 181–204. http://dx.doi.org/10.1017/S0140525X12000477

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska Symposium on Motivation* (pp. 207–282). Lincoln, NE: University of Nebraska Press.

Ekman, P. (1992). Are there basic emotions? *Psychological Review, 99,* 550–553. http://dx.doi.org/10.1037/0033-295X.99.3.550

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review, 3,* 364–370. http://dx.doi.org/10.1177/1754073911410740

Firestone, C. (2013). On the origin and status of the "El Greco fallacy." *Perception, 42,* 672–674. http://dx.doi.org/10.1068/p7488

Firestone, C., & Scholl, B. J. (2014). "Top–down" effects where none should be found: The El Greco fallacy in perception research. *Psychological Science, 25,* 38–46. http://dx.doi.org/10.1177/0956797613485092

Firestone, C., & Scholl, B. J. (2015). Can you experience 'top-down' effects on perception? The case of race categories and perceived lightness. *Psychonomic Bulletin & Review, 22,* 694–700. http://dx.doi.org/10.3758/s13423-014-0711-5

Foxe, J. J., & Simpson, G. V. (2002). Flow of activation from V1 to frontal cortex in humans. A framework for defining "early" visual processing. *Experimental Brain Research, 142,* 139–150. http://dx.doi.org/10.1007/s00221-001-0906-7

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11,* 127–138. http://dx.doi.org/10.1038/nrn2787

Fugate, J. M. B. (2013). Categorical perception of emotional faces. *Emotion Review, 5,* 84–89. http://dx.doi.org/10.1177/1754073912451350

Fugate, J. M., & Barrett, L. F. (2014). The role of language on the perception and experience of emotion. In T. Holtgraves (Ed.), *The Oxford handbook of language and social psychology* (pp. 280–298). New York, NY: Oxford.

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. http://dx.doi.org/10.1037/a0019017

Fugate, J. M. B., O'Hare, A. J., & Emmanuel, W. J. (in press). Emotion words: *Facing* change.

Gendron, M., Lindquist, K. A., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion, 12,* 314–325. http://dx.doi.org/10.1037/a0026007

Gitelman, D. R., Nobre, A. C., Sonty, S., Parrish, T. B., & Mesulam, M. M. (2005). Language network specializations: An analysis with parallel task designs and functional magnetic resonance imaging. *NeuroImage, 26,* 975–985. http://dx.doi.org/10.1016/j.neuroimage.2005.03.014

Halberstadt, J. (2005). Featural shift in explanation-biased memory for emotional faces. *Journal of Personality and Social Psychology, 88,* 38–49. http://dx.doi.org/10.1037/0022-3514.88.1.38

Halberstadt, J. I. (2003). The paradox of emotion attribution explanation biases perceptual memory for emotional expressions. *Current Directions in Psychological Science, 12,* 197–201. http://dx.doi.org/10.1046/j.0963-7214.2003.01261.x

Halberstadt, J. B., & Niedenthal, P. M. (2001). Effects of emotion concepts on perceptual memory for emotional expressions. *Journal of Personality and Social Psychology, 81,* 587–598. http://dx.doi.org/10.1037/0022-3514.81.4.587

Halberstadt, J., Winkielman, P., Niedenthal, P. M., & Dalle, N. (2009). Emotional conception: How embodied emotion concepts guide perception and facial action. *Psychological Science, 20,* 1254–1261. http://dx.doi.org/10.1111/j.1467-9280.2009.02432.x

Hohwy, J. (2013). *The predictive mind*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199682737.001.0001

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

Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin, 115,* 288–299. http://dx.doi.org/10.1037/0033-2909.115.2.288

Kveraga, K., Ghuman, A. S., & Bar, M. (2007). Top–down predictions in the cognitive brain. *Brain and Cognition, 65,* 145–168. http://dx.doi.org/10.1016/j.bandc.2007.06.007

Landau, A. N., Aziz-Zadeh, L., & Ivry, R. B. (2010). The influence of language on perception: Listening to sentences about faces affects the perception of faces. *The Journal of Neuroscience, 30,* 15254–15261. http://dx.doi.org/10.1523/JNEUROSCI.2046-10.2010

Lindquist, K. A., Barrett, L. F., Bliss-Moreau, E., & Russell, J. A. (2006). Language and the perception of emotion. *Emotion, 6,* 125–138. http://dx.doi.org/10.1037/1528-3542.6.1.125

Lindquist, K. A., & Gendron, M. (2013). What's in a word? Language constructs emotion perception. *Emotion Review, 5,* 66–71. http://dx.doi.org/10.1177/1754073912451351

Lindquist, K. A., Gendron, M., Barrett, L. F., & Dickerson, B. C. (2014). Emotion perception, but not affect perception, is impaired with semantic memory loss. *Emotion, 14,* 375–387. http://dx.doi.org/10.1037/a0035293

Lindquist, K. A., Satpute, A. B., & Gendron, M. (2015). Does language do more than communicate emotion? *Current Directions in Psychological Science, 24,* 99–108. http://dx.doi.org/10.1177/0963721414553440

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology, 3*. Article, 54, 114–126.

Lupyan, G., & Clark, A. (2015). Words and the world: Predictive coding and the language–perception–cognition interface. *Current Directions in Psychological Science, 24,* 279–284. http://dx.doi.org/10.1177/0963721415570732

Lupyan, G., & Thompson-Schill, S. L. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General, 141,* 170–186. http://dx.doi.org/10.1037/a0024904

Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences of the United States of America, 110,* 14196–14201. http://dx.doi.org/10.1073/pnas.1303312110

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and" nonparametric" indexes. *Psychological Bulletin, 107,* 401–413. http://dx.doi.org/10.1037/0033-2909.107.3.401

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York, NY: Cambridge University Press.

Maier, M., Glage, P., Hohlfeld, A., & Abdel Rahman, R. (2014). Does the semantic content of verbal categories influence categorical perception? An ERP study. *Brain and Cognition, 91,* 1–10. http://dx.doi.org/10.1016/j.bandc.2014.07.008

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

Niedenthal, P. M. (2007, May 18). Embodying emotion. *Science, 316,* 1002–1005. http://dx.doi.org/10.1126/science.1136930

Niedenthal, P. M. (2008). *Emotion concepts. Handbook of emotions* (3rd ed., pp. 587–600). New York, NY: Guilford Press.

Nook, E. C., Lindquist, K. A., & Zaki, J. (2015). A new look at emotion perception: Concepts speed and shape facial emotion recognition. *Emotion, 15,* 569–578. http://dx.doi.org/10.1037/a0039166

Richardson, J. T. E. (1994). Continuous recognition memory tests: Are the assumptions of the theory of signal detection really met? *Journal of Clinical & Experimental Neuropsychology, 16,* 482–486.

Roberson, D., Damjanovic, L., & Pilling, M. (2007). Categorical perception of facial expressions: Evidence for a "category adjustment" model. *Memory & Cognition, 35,* 1814–1829. http://dx.doi.org/10.3758/BF03193512

Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & Cognition, 28,* 977–986. http://dx.doi.org/10.3758/BF03209345

Russell, J. A. (1980). A cirumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178. http://dx.doi.org/10.1037/h0077714

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.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110,* 145–172. http://dx.doi.org/10.1037/0033-295X.110.1.145

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76,* 805–819. http://dx.doi.org/10.1037/0022-3514.76.5.805

Sauter, D. A., LeGuen, O., & Haun, D. B. (2011). Categorical perception of emotional facial expressions does not require lexical categories. *Emotion, 11,* 1479–1483. http://dx.doi.org/10.1037/a0025336

Scherer, K. R. (2000). Psychological models of emotion. In J. Borod (Ed.), *The Neuropsychology of emotion* (pp. 137–162). New York, NY: Oxford.

Schyns, P. G., Goldstone, R. L., & Thibaut, J. P. (1998). The development of features in object concepts. *Behavioral and Brain Sciences, 21,* 1–17. http://dx.doi.org/10.1017/S0140525X98000107

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31,* 137–149. http://dx.doi.org/10.3758/BF03207704

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

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

Widen, S. C. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review, 5,* 72–77. http://dx.doi.org/10.1177/1754073912451492

Widen, S. C., & Russell, J. A. (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology, 39,* 114–128. http://dx.doi.org/10.1037/0012-1649.39.1.114

Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., & Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia, 49,* 1105–1127. http://dx.doi.org/10.1016/j.neuropsychologia.2010.12.032

Received April 10, 2015
Revision received March 30, 2017
Accepted April 17, 2017 ■