

# Emotions in “Black and White” or Shades of Gray? How We Think About Emotion Shapes Our Perception and Neural Representation of Emotion

Ajay B. Satpute<sup>1,2</sup>, Erik C. Nook<sup>3</sup>, Sandhya Narayanan<sup>4</sup>,  
Jocelyn Shu<sup>4</sup>, Jochen Weber<sup>4</sup>, and Kevin N. Ochsner<sup>4</sup>

<sup>1</sup>Department of Psychology, Pomona College; <sup>2</sup>Department of Neuroscience, Pomona College; <sup>3</sup>Department of Psychology, Harvard University; and

<sup>4</sup>Department of Psychology, Columbia University

Psychological Science  
1–15

© The Author(s) 2016

Reprints and permissions:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/0956797616661555

pss.sagepub.com



## Abstract

The demands of social life often require categorically judging whether someone’s continuously varying facial movements express “calm” or “fear,” or whether one’s fluctuating internal states mean one feels “good” or “bad.” In two studies, we asked whether this kind of categorical, “black and white,” thinking can shape the perception and neural representation of emotion. Using psychometric and neuroimaging methods, we found that (a) across participants, judging emotions using a categorical, “black and white” scale relative to judging emotions using a continuous, “shades of gray,” scale shifted subjective emotion perception thresholds; (b) these shifts corresponded with activity in brain regions previously associated with affective responding (i.e., the amygdala and ventral anterior insula); and (c) connectivity of these regions with the medial prefrontal cortex correlated with the magnitude of categorization-related shifts. These findings suggest that categorical thinking about emotions may actively shape the perception and neural representation of the emotions in question.

## Keywords

emotions, neuroimaging, social cognition, categorization, affect

Received 10/16/15; Revision accepted 7/7/16

The ability to perceive and understand emotional experiences is critical for functioning effectively in society (Salovey & Mayer, 1989). Often this ability is directed outward, toward other individuals. For example, perceiving that a supervisor is happy or upset may inform an employee whether or not to ask for a raise, and perceiving that a defendant acted out of fear or anger may inform how a judge or jury metes out justice. Other times, this ability is directed inward, toward the self. For example, perceiving that you are feeling pleasant or unpleasant may inform you about what is worth attending to, remembering, or pursuing (Ochsner, 2000). Whatever the circumstance, the emotions people perceive in a situation may dictate both how they act and communicate in the moment and how they remember these moments in the future (Ochsner, 2000; Salovey & Mayer, 1989).

The broad importance of emotion perception has generated a wealth of research on its mechanisms and brain

bases (Adolphs, 1999; Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012; Whalen & Phelps, 2009). However, an important aspect of emotion perception that has been overlooked concerns the difference between the continuous nature of the sensory inputs that people receive and the categorical nature of their thinking about emotion. In facial expressions, the contractions of various facial muscles can vary continuously to create gradations of movements (Jack & Schyns, 2015). But people typically talk about these expressions in categorical terms, calling them expressions of “fear” or “calm,” for instance

## Corresponding Authors:

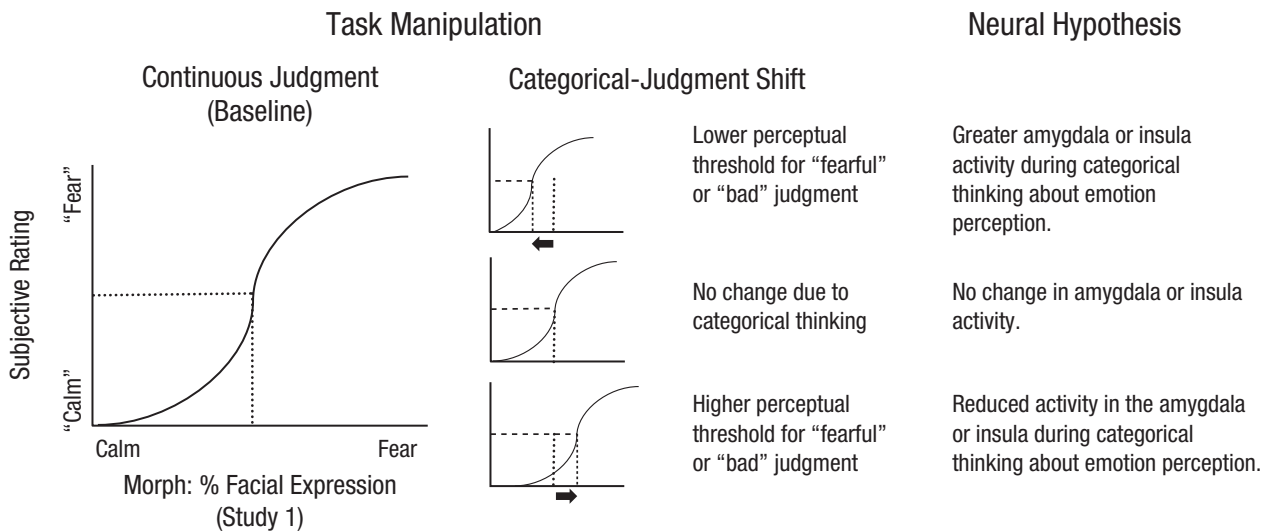
Ajay B. Satpute, 647 N. College Way, Claremont, CA 91711

E-mail: ajay.satpute@pomona.edu

Kevin N. Ochsner, 1190 Amsterdam Ave., 369 Schermerhorn Hall,

New York, NY 10027

E-mail: ochsner@psych.columbia.edu



**Fig. 1.** Schematic depiction of our test of whether categorical thinking may shift emotion perception, with corresponding neural hypotheses for the possible outcomes. In Study 1, participants judged faces on a morphed continuum as exhibiting fear or calm, using a continuous or categorical scale. As illustrated by the graphs, in each condition, we calculated the point of subjective equivalence (PSE; where the dotted lines intersect with the  $x$ -axes), the level of fear required for the participant’s judgment to be equally likely to fall into either of the two categories. In Study 2 (not depicted here), participants judged their affective response to images of natural scenes as bad or neutral, again using a continuous or categorical scale, and we calculated the PSEs in both conditions. In both experiments, we used the PSEs when participants made continuous judgments as a reference point to determine whether PSEs changed when participants made categorical judgments. There were three possible outcomes, as indicated by the middle column. Relative to making continuous judgments, making categorical judgments might shift the PSE to the left, so that a participant had a lower threshold for judging a face as “fearful” (rather than “calm”; or unpleasant affective feelings as “bad,” rather than “neutral,” in Study 2); in that case, we expected to find that activity in the amygdala or ventral anterior insula would be greater during categorical judgments than during baseline, continuous, judgments. Alternatively, there might be no change in the PSE due to categorical judgments, and no change in amygdala or insula activity. Finally, the PSE might shift to the right, so that a participant had a higher threshold for judging a face as “fearful” (or unpleasant affective feelings as “bad”); in that case, we expected to observe reduced activity in the amygdala or insula during the categorical-judgment condition, compared with the continuous-judgment condition.

(Barrett, 2006). Similarly, when people perceive their own emotions, their interoceptive signals may vary continuously, but they often talk about feeling “good” or “bad” (e.g., Satpute, Shu, Weber, Roy, & Ochsner, 2012).

Whether such categorical thinking matters for emotion perception remains unclear. Prior work suggests that it may, but this question has not yet been examined directly. One study found that orienting children to think in more categorical rather than continuous ways when perceiving another child’s emotional states influenced the children’s downstream judgments, leading to more extreme preferences about whether they wanted to play with that child (Master, Markman, & Dweck, 2012). Other studies have shown that emotion words may shape (Gendron, Lindquist, Barsalou, & Barrett, 2012; Pessoa, Japee, Sturman, & Ungerleider, 2006; Roberson & Davidoff, 2000), constitute (Barrett, 2006; Lindquist, Satpute, & Gendron, 2015; Nook, Lindquist, & Zaki, 2015), or disrupt processes that generate affect (Kassam & Mendes, 2013; Lieberman et al., 2007). This work has focused on the role of language in emotion (i.e., contrasting the presence vs. absence of emotion words in the task), but has not specifically addressed whether the categorical quality of the judgments also plays a role.

That said, research in cognitive psychology provides clues as to when such top-down, categorical thinking

about emotions could be more or less influential. The influence of such thinking is likely to be weakest at the ends of a perceptual continuum—where bottom-up inputs are least subject to interpretation—and greatest in the middle—where top-down knowledge may make meaning of perceptual signals that could be categorized in multiple ways (Harnad, 1990). This suggests that judging someone to be expressing “fear” or “calm” might make an observer “see” the judged emotion when the sensory input is somewhere in the middle. Similarly, judging one’s feelings categorically as “bad” or “neutral” could lead one to have stronger or weaker affective responses when they would otherwise be in the middle. Our hypothesis is that, in each case, such “black and white” thinking about continuously varying emotion cues changes the mental representation of those cues to be consistent with the chosen categorical term.

We tested our hypothesis in two studies, using psychophysical and neuroimaging methods. In both studies, we assessed how using a categorical scale, rather than a continuously graded scale, to make emotion judgments shifted participants’ perceptual thresholds for judging that a given kind of emotion was present (see Fig. 1).

We then asked whether shifts in emotion perception when making categorical judgments were accompanied by

corresponding changes in the activity of neural regions involved in perceiving emotion, including the amygdala, given its importance for the perception of emotion in other individuals on the basis of exteroceptive cues, such as facial expressions (Adolphs, 1999; Anderson & Phelps, 2001; Pessoa et al., 2006), and the ventral anterior insula, given its importance in interoception (Craig, 2011; Critchley, Wiens, Rotshtein, Öhman, & Dolan, 2004; Zaki, Davis, & Ochsner, 2012). Study 1 examined how engaging in categorical thinking shaped participants' perception of other individuals' emotions depicted in pictures of faces. Study 2 examined participants' perception of their own affective responses to graphic images.

If categorical thinking does indeed shape activity in these brain regions during emotion perception, then it is important to ask what neural mechanisms exert this influence. Prior work suggests that the anterior medial prefrontal cortex (mPFC) might be particularly important given its role in the conceptualization of emotion (Barrett & Satpute, 2013; Ochsner & Gross, 2014; Roy, Shohamy, & Wager, 2012; Satpute et al., 2012), and particularly the finding that mPFC damage increases confusion about what facial expressions best depict various categories (Heberlein, Padon, Gillihan, Farah, & Fellows, 2008). Hence, we examined whether the mPFC is important for resolving continuously varying sensory inputs into discrete categories of perceived emotion. Specifically, we assessed whether connectivity of the amygdala and the ventral anterior insula with the mPFC was greater when categorical thinking had a stronger influence on emotion perception.

## Method

### Overview

In Study 1, we examined how categorical thinking shapes the perception of emotion from exteroceptive inputs by asking participants to judge photographs of facial expressions that were morphed along a continuum from calm to fearful (Fig. 2a). On *categorical* trials, participants were given only two response options: "fear" and "calm." On *continuous* trials, these terms were anchors on a continuously graded scale, and participants could provide a response anywhere along this scale. Each participant's point of subjective equivalence (PSE) for perceiving the morphs as calm or fearful was calculated, and we measured the degree to which categorical thinking shifted the threshold (percentage of fear in the morphed stimulus) for perceiving fear (or calm; Fig. 2b). We then asked whether the observed shifts correlated with participants' neural activity while viewing the faces.

In Study 2, we examined how categorical thinking shapes the perception of emotion from interoceptive sources by asking participants to judge their affective

response to photographs that varied in their normative induced affect, from neutral to highly negative (Fig. 3a). The design and analysis were similar to those of Study 1. Each participant's PSE for perceiving him- or herself to be feeling "bad" versus "neutral" was measured as a function of the normative negative-valence ratings of the pictures (Fig. 3b). As in Study 1, we first measured the degree to which categorical thinking shifted the threshold (degree of normative negative valence) for perceiving one's own feelings as bad (or neutral) and then asked whether the observed shifts correlated with participants' neural activity while viewing the picture stimuli. A different set of analyses involving portions of Study 2 has been published previously (Satpute et al., 2012).

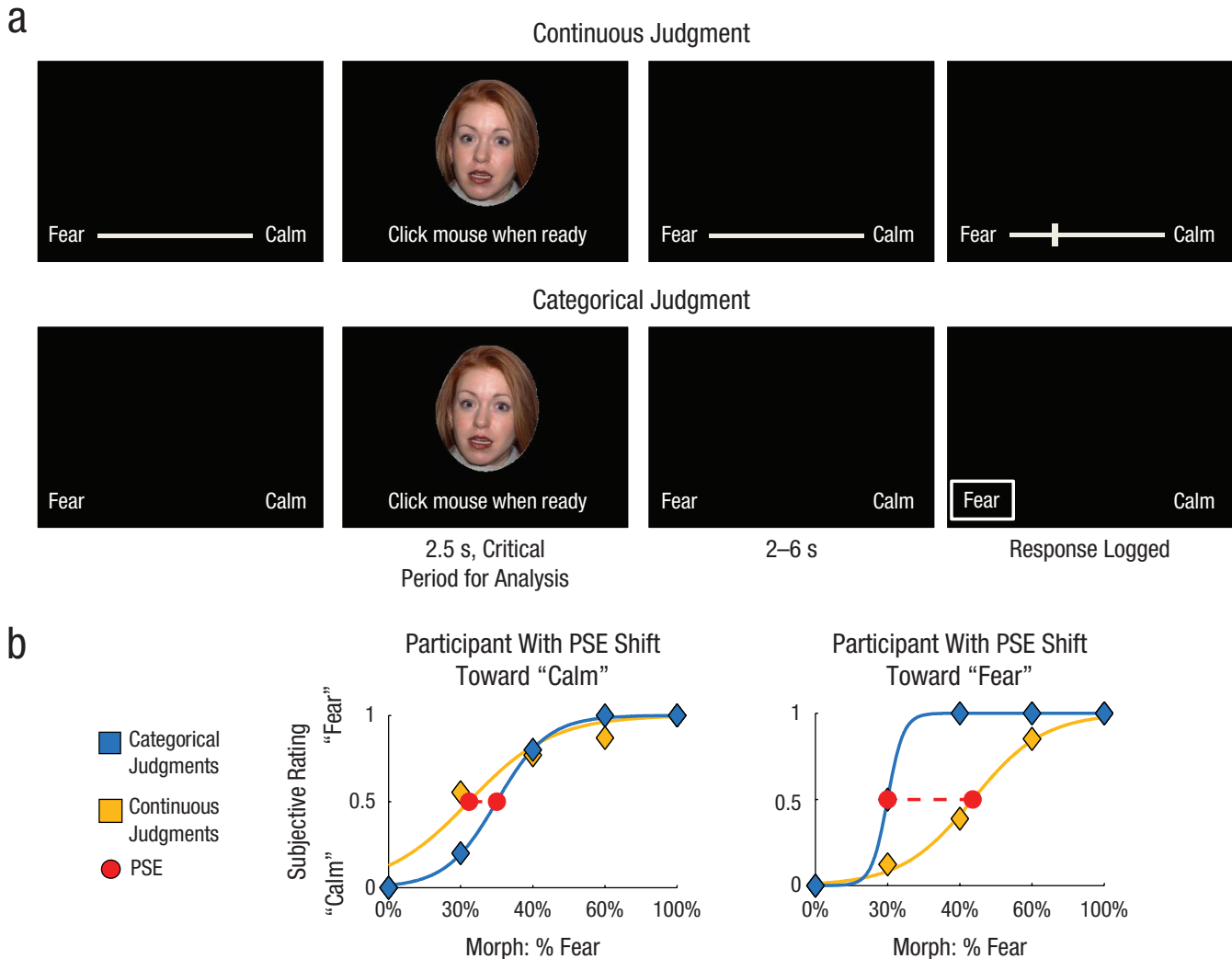
Notably, we restricted our experimental paradigm in both studies to examine the spectrum from neutral to negative affect. We did this to increase the number of trials within this spectrum and thus the statistical power of our design, and also to facilitate the interpretation of any observed findings involving the amygdala and ventral anterior insula. A parallel study examining the spectrum from neutral to positive affect would be of interest for future work.

### Participants

Participants were healthy, native English-speaking, right-handed adults who provided consent following the guidelines of Columbia University's institutional review board. Pregnant women, children, the elderly, and volunteers with MRI contraindications (i.e., nonremovable ferromagnetic metal implants or claustrophobia) were excluded. Participants received compensation of \$25 per hour. Study 1 included 20 participants (10 female, 10 male; age range = 19–36 years). Study 2 included a separate sample of 20 participants (14 female, 6 male; age range = 19–34 years). Prior to data collection, we decided on sample sizes of 20, to match the typical sample sizes of recent neuroimaging studies in affective neuroscience. A power calculation was not used because the primary manipulation (categorical vs. continuous judgment) had not been used in prior studies.

### Stimuli

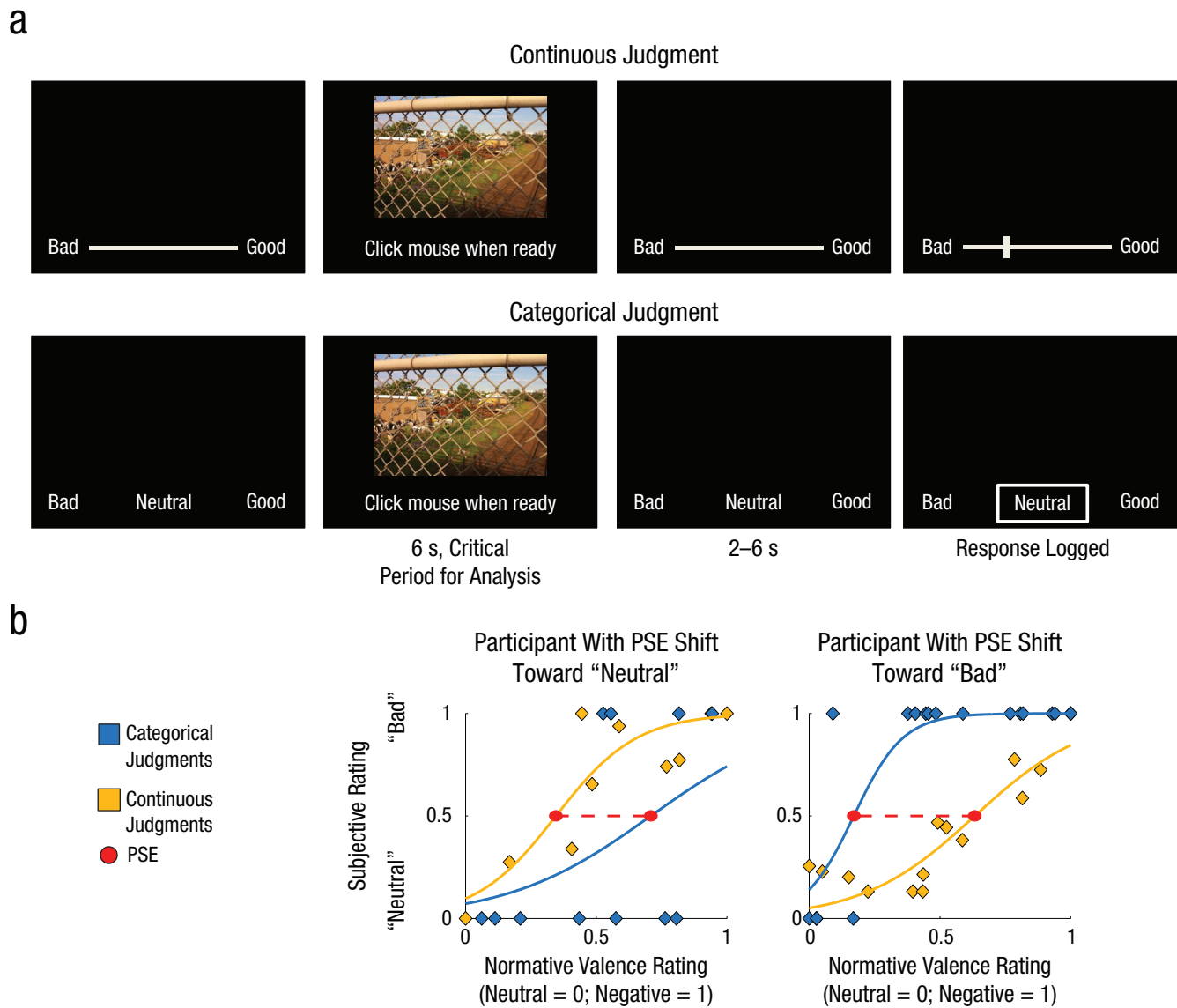
**Study 1.** Photographs of 20 Caucasian actors (10 male, 10 female) portraying fearful and calm facial expressions were selected from the NimStim set (Tottenham et al., 2009) and morphed using Morpheus Photo Morpher software (Morpheus, <http://www.morpheussoftware.net>) to produce faces along a continuum from calm to fear in 10% increments. Although calm has a positive valence connotation, calm faces are effectively neutral in this stimulus set (Tottenham et al., 2009). Open-mouthed calm



**Fig. 2.** Study 1: task design and examples of categorization-related shifts in the perception of emotion in other individuals. Each participant judged facial expressions morphed along a continuum from calm to fearful using a continuous scale in some blocks and a categorical scale in others (a). Each face was presented for 2.5 s, during which time participants were asked to indicate when they had made a decision concerning the emotion displayed. Next, following a 2- to 6-s interval, participants recorded their judgment. For each condition, the average subjective response (“calm” = 0, “fear” = 1) for each morph level was then calculated, and the point of subjective equivalence (PSE) was identified. The difference in the PSE between the conditions indexed the participant’s categorization-related shift in the PSE. Note that because of copyright concerns, the face image shown here is not one that was used in the experiment. In (b), the plot on the left shows results for a participant with a categorization-related shift (dashed red line) toward “calm,” and the plot on the right shows results for a participant with a categorization-related shift toward “fear.”

and fearful expressions were used to aid smooth morphing. Manual corrections were made in Adobe Photoshop to ensure that the morphs appeared natural. On the basis of pilot testing with a separate sample, we identified a subset of stimuli for this study. We selected the morphs at the level of fear (40%) that was closest to the average PSE across stimuli, the morphs at two adjacent levels (30% and 60% fear; the morphs with 50% fear were not perceptually different from the 40% morphs, according to the pilot study, and thus using them would not improve upon the fits of the psychometric curves), and the nonmorphed faces (100% calm and 100% fear), to produce a set of

stimuli that allowed estimation of psychometric curves for which PSEs varied across participants. The photographs were also tilted slightly leftward, to varying degrees (0–10°), so that we could compare neural activity when participants judged emotion and when they judged a non-affective dimension of the stimuli (i.e., tilt); we plan to present the results of those analyses in a separate report. The 100 selected photographs were divided into four balanced lists of 25 photographs each; each list included 5 photos from each morph level. One list was assigned to the categorical emotion-judgment task, and another was assigned to the continuous emotion-judgment task. The



**Fig. 3.** Study 2: task design and examples of categorization-related shifts in the perception of emotion in the self. Each participant judged his or her own affective experience in response to graphic images using a continuous scale in some blocks and a categorical scale in others (a). Each image was presented for 6 s, during which time participants were asked to indicate when they had made a decision concerning the affect they experienced. Next, following a 2- to 6-s interval, participants recorded their judgment. For each condition, the average subjective response ("neutral" = 0, "bad" = 1) for each normative rating level was then calculated, and the point of subjective equivalence (PSE) was identified. The difference in the PSE between the conditions indexed the participant's categorization-related shift. In (b), the plot on the left shows results for a participant with a categorization-related shift toward "neutral," and the plot on the right shows results for a participant with a categorization-related shift toward "bad."

other two lists were assigned to the tilt-judgment task. The lists were counterbalanced across experimental conditions across participants.

**Study 2.** Varying degrees of negative affective experience were induced by showing participants images from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008). Using the published IAPS normative ratings, we created three image sets that varied

in average negative affective intensity: neutral (valence:  $M = 4.98$ ; arousal:  $M = 3.64$ ), mildly aversive (valence:  $M = 3.54$ ; arousal:  $M = 5.56$ ), and highly aversive (valence:  $M = 1.97$ ; arousal:  $M = 6.18$ ). These image sets were divided into four lists balanced for average levels of valence and arousal; the lists were counterbalanced across experimental conditions and participants. One list was assigned to the categorical affect-judgment task, and another was assigned to the continuous affect-judgment

task. The other two lists were assigned to a line-judgment task, which involved judging whether the lines in each image were overall more curvy or straight; this task was not relevant to the present analyses (see Satpute et al., 2012).

## Tasks

**Study 1.** This experiment had a 2 (kind of judgment: categorical or continuous)  $\times$  2 (content of judgment: emotion or tilt) factorial design. Participants were provided with instructions and practice using a separate set of stimuli prior to the experimental trials in the MRI scanner. In the scanner, they made categorical and continuous judgments of the facial-expression photographs. Each of the four tasks (categorical emotion judgments, continuous emotion judgments, categorical tilt judgments, continuous tilt judgments) was presented in separate blocks to minimize costs of switching between conditions. Each block consisted of five trials, and a block of each task condition was included within each of five functional runs; the order of the blocks was randomized within each run.

In the categorical emotion-judgment task, participants decided whether each depicted person was experiencing calm or fear. We chose the label “calm” because it emphasizes low arousal, because this label is commonly used to describe facial expressions, and because calm faces are not discernibly different from neutral faces in the NimStim set (Tottenham et al., 2009). Participants were first shown a 3-s instruction indicating that they were to make categorical emotion judgments. A block of five trials including one stimulus from each morph level was then presented. As illustrated in Figure 2a, each image was presented for 2.5 s. It was initially accompanied by a prompt asking participants to click the right-side mouse button when they came to their decision. The prompt (but not the face image) disappeared when this button was pressed, and the reaction time for this decision, *time to decision*, was recorded. Following a 2- to 6-s jittered interval, a box appeared, and participants used a trackball to move the box to select an emotion category; they recorded their judgment by pressing the right-side mouse button. The box shifted discretely between the categories, which were presented at the bottom of the screen during the entire block except for the periods during which the images were presented. A total of 25 categorical emotion-judgment trials were presented across five blocks.

In the continuous emotion-judgment task, participants judged the relative degree to which each depicted person was experiencing calm or fear. The procedure was the same as for the categorical emotion-judgment task except that (a) the 3-s task instruction indicated that participants

were to make continuous emotion judgments, (b) the response scale was a continuous line, and (c) following the jittered interval, a vertical bar (instead of a box) appeared, and participants slid the bar to any point along the line to record their response (see Fig. 2a).

The categorical and continuous tilt-judgment tasks, performed in separate blocks, were similar to the emotion-judgment tasks except that participants judged how much each face tilted leftward. After each task block, participants completed 18 s of a standard odd-even baseline task (Satpute et al., 2012), in which they were shown single-digit numbers and used mouse clicks to report whether the numbers were odd or even. The baseline task provided additional separation between the task conditions to minimize carry over.

Note that in all task conditions, participants were prompted to respond twice on each trial: first, while the image was on the screen, to signal (with a simple mouse click) that they had made their judgment and then, after a jittered delay, to log their response on the scale. Separating these responses allowed us to distinguish brain activity during the image-viewing period—when participants were making their decisions—from brain activity during the motor response period.

**Study 2.** The design for Study 2 was very similar to that for Study 1. Again, a 2 (kind of judgment: categorical or continuous)  $\times$  2 (content of judgment: affect or line curvature) design was used. As in Study 1, the four task conditions were blocked to minimize switching costs. Participants were provided with instructions and practice using a separate set of stimuli prior to the experimental trials in the MRI scanner.

In the scanner, participants made categorical and continuous judgments while viewing IAPS images. For both the categorical and the continuous affect-judgment tasks, participants were instructed to judge how viewing the photographs made them feel. In the categorical-judgment condition, participants decided whether their own affective experience was best categorized as “bad,” “neutral,” or “good.” The trial structure was otherwise very similar to that of the categorical-judgment task in Study 1, except that each photograph was presented for 6 s (see Fig. 3a).

In the continuous affect-judgment condition, participants judged the relative degree to which they were experiencing affect on a continuous scale ranging from “bad” to “good” (see Fig. 3a). The procedure was the same as for the categorical affect-judgment condition except for the same key differences that distinguished the continuous and categorical emotion-judgment tasks in Study 1.

The categorical and continuous curvature-judgment tasks were similar to the emotion-judgment tasks except that participants judged how overall curvy or straight the lines in each image were. After each task block,

participants completed 18 s of the odd-even baseline task. Each block type was included once in each of three functional runs, and the order of the blocks was randomized within each run. Each block contained two photographs from each level of affective intensity (neutral, mildly aversive, and highly aversive). Note that Study 2 had 6 trials per block (total of 18 trials per task condition) so that the three levels of affective intensity were equally represented in each block, whereas Study 1 had 5 trials per block so that the five emotion levels would be equally represented.

Study 1 examined perception of facial expressions that ranged from expressions of calm to expressions of fear. Similarly, Study 2 focused on affective responses in the neutral to aversive range; no normatively pleasant photographs were included as stimuli. However, as already described, both the categorical and the continuous scales included "good" as an endpoint, so that we could identify any trials in which participants had a pleasant affective response. In the categorical-judgment task, participants selected "good" on a few trials ( $M = 6.22\%$  of trials,  $SD = 6.04\%$ ). These trials were excluded from further analysis, as were a similar number of continuous-judgment trials in which responses fell within the positive range of the judgment scale.

## Apparatus

Scanning was conducted on a GE TwinSpeed 1.5-T scanner equipped with an eight-channel head coil. Functional spiral echo in/out pulse-sequence scanning parameters were as follows: repetition time (TR) = 2,000 ms, slice thickness = 4.5 mm, gap = 0 mm, flip angle = 84°, field of view (FOV) = 22.4 cm, matrix size = 64 × 64, 23 slices, axial orientation, voxel size = 3.5 × 3.5 × 4.5 mm, interleaved bottom-to-top acquisition. Structural-spoiled-gradient scan parameters were as follows: TR = 19 s, echo time = 5 ms, slice thickness = 1 mm, gap = 0 mm, flip angle = 20°, FOV = 25.6 cm, sagittal orientation, matrix size = 256 × 256 × 180, voxel size = 1 × 1 × 1 mm. Stimuli were projected onto a screen and deflected by a mirror for viewing, and responses were made with a scanner-compatible trackball with two buttons. Stimuli were presented and behavioral data were collected using a desktop PC and MATLAB software with the Psychophysics Toolbox (Pelli, 1997).

## Behavioral analysis

**Study 1.** In our behavioral analysis, we examined whether each participant's PSE—the point at which the participant was as likely to judge a depicted person's emotional experience as being calm as to judge it as

being fearful—varied according to whether he or she was making categorical judgments or continuous judgments (Fig. 2b). To perform this analysis, we scaled each participant's judgments from 0 to 1 within each task condition and then used logit functions to regress these judgments against the five levels of morphing. More specifically, we estimated the parameters  $a$  and  $b$  in the following equation:  $y = 1/(1 + \exp((-1) * a * (x - b)))$ , where  $x$  refers to the percentage of morphed fear in the stimulus, and  $y$  is the averaged subjective report for stimuli with that level of morphed fear. These logit functions were calculated for each participant and each task condition separately (see Fig. 2b). We identified each participant's PSE as the estimated morphed-stimulus intensity at which there was a 50% probability that he or she would judge the depicted person to be experiencing fear. We then subtracted each participant's continuous-judgment PSE from his or her categorical-judgment PSE to quantify the participant's categorization-related shift in PSE. We then multiplied this value by  $-1$  for easier interpretation. Thus, higher (positive) values indicated that participants were more likely to select the "fear" category when judging depicted emotions categorically than when judging them using the continuous scale (for a given morph level; Fig. 2b, right). Lower (negative) values indicated that participants were more likely to select the "calm" category when making that judgment categorically than when using the continuous scale (Fig. 2b, left).

**Study 2.** A similar behavioral analysis was conducted for Study 2. Again, PSEs were calculated separately for the categorical and continuous-judgment conditions, separately for each participant. However, whereas in Study 1 the morph levels were used as the reference points, in Study 2 we used the normative valence ratings of the IAPS images (Lang et al., 2008; Fig. 3b). As in Study 1, we first scaled each participant's judgments from 0 to 1 within each task condition and then used logit functions to regress these values against the normative valence ratings (see the equation in the previous paragraph). Data from 2 participants were excluded because their subjective ratings did not correspond with the normative valence ratings and thus could not be modeled well (goodness of fit was .0008 for one participant and indeterminate for the other). We identified each participant's PSE as the normative valence rating at which the participant was equally likely to judge his or her affective experience as being more neutral as to judge it as being more bad. We then subtracted participants' continuous-judgment PSEs from their categorical-judgment PSEs to measure their categorization-related shifts in PSEs, and multiplied these scores by  $-1$ . Hence, higher (positive) values indicated that participants were more likely to select the "bad" category when making their judgment categorically than when

using the continuous scale (for a given level of normative valence; Fig. 3b, right). Similarly, lower (negative) values indicated that participants were more likely to select the “neutral” category when making a categorical judgment than when using the continuous scale (Fig. 3b, left).

### Image analysis

For Studies 1 and 2, images were realigned, normalized to the Montreal Neurological Institute (MNI) 152 template space, and smoothed (6-mm kernel) using SPM5 software (Wellcome Trust Centre for Neuroimaging, <http://www.fil.ion.ucl.ac.uk/spm>). Statistical models were implemented in NeuroElf v0.9d ([www.neuroelf.com](http://www.neuroelf.com)). First-level regressors modeled the image-viewing periods for trials within each task condition as epochs (durations: 2.5 s in Study 1, 6 s in Study 2) convolved with the canonical hemodynamic response function. The models also included nuisance regressors controlling for motor responses in the response-logging phase (from onset of the decision screen to time of response in all conditions) and a high-pass filter (discrete cosine transform, 128-s cutoff). For 2 participants in Study 2, one run (of three) was lost because of scanner failure; the available two runs were included in analyses.

In both studies, our analysis focused on the amygdala and ventral anterior insula as regions of interest (ROIs), given their critical role in affective responding (Lindquist, Satpute, Wager, Weber, & Barrett, 2016). The amygdala has extensive connections with exteroceptive inputs (Whalen & Phelps, 2009), and patients with damage to the amygdala show deficits in perceiving and experiencing affect and emotions (Feinstein, 2013). The anterior insula plays an important role in processing interoceptive information (Critchley et al., 2004), a conclusion that is supported by its anatomical connectivity with posterior insular regions receiving inputs from the viscera (Craig, 2011) and by its functional activity observed during neuroimaging studies of affective experience (Lindquist et al., 2016). To test our hypothesis, we first selected voxels using NeuroSynth (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011), which provides term-based meta-analytic maps derived from a database of more than 5,000 neuroimaging studies. Specifically, we used the reverse-inference maps for the term “emotion” and used clusters of voxels appearing in the amygdala and ventral anterior insula bilaterally to make masks defining our ROIs. Independently defined ROIs provide unbiased estimates of effect size, but at the expense of potential partial-volume effects. Hence, we used one-tailed  $p$  values to assess significance given our directional hypotheses and use of ROIs based on prior behavioral work and theory in categorization (Harnad, 1990) and neuroscience. However, we also performed analyses using whole-brain

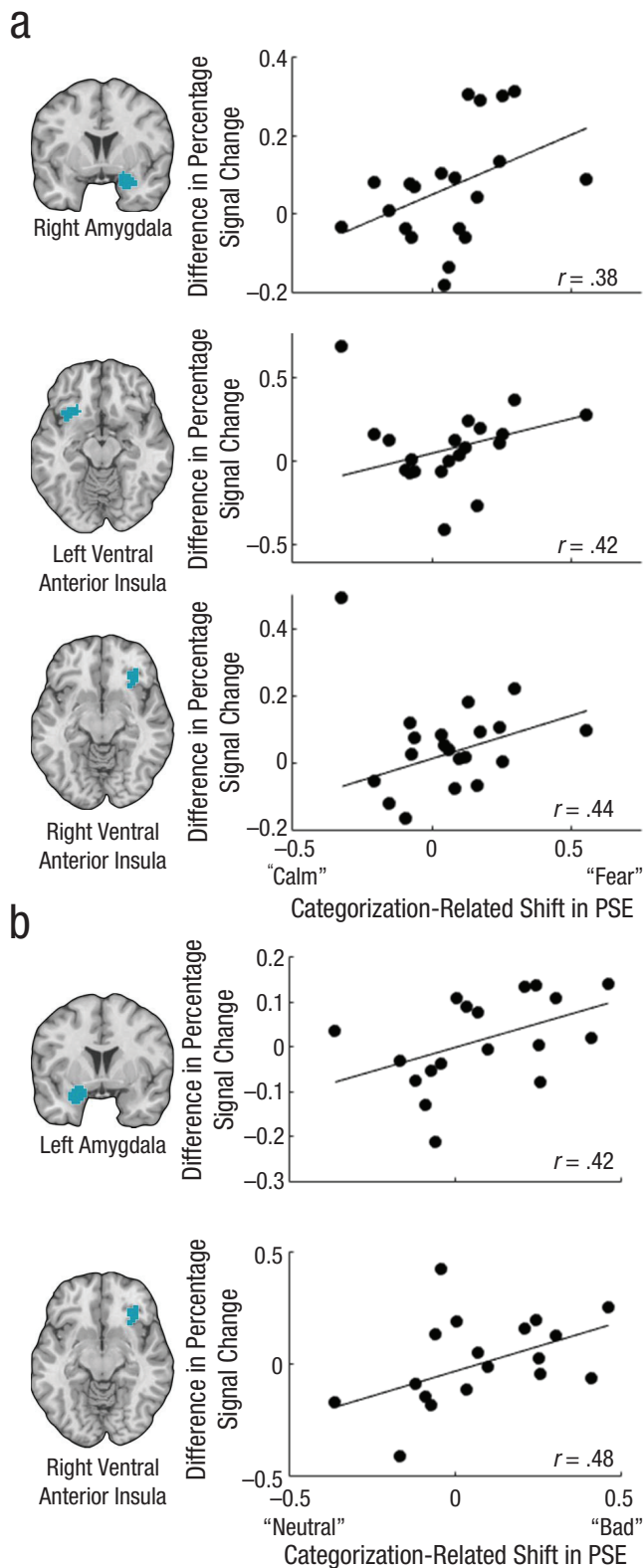
corrections ( $p < .05$ , corrected; smoothness estimated from the data). To limit the influence of outliers, we used robust regressions (Wager, Keller, Lacey, & Jonides, 2005), which reduce the probability of false positives due to values with inordinately high leverage.

For significant correlations from independent-ROI analyses, bootstrapping procedures (MATLAB software, function *bootci.m*; The MathWorks, Natick, MA) were used to obtain 95% confidence intervals (CIs). CIs were constructed using 5,000 bootstrap samples, Pearson correlations, and the default, bias-corrected, and accelerated percentile method. Outliers were identified as values having Studentized residuals (from the robust regressions) greater than 2; one outlier for the left anterior insula in Study 1 (Fig. 4a, second graph, top left quadrant), one outlier for the right anterior insula in Study 1 (Fig. 4a, third graph, top left quadrant) and one outlier for the right anterior insula in Study 2 (Fig. 4b, second graph, data point with the maximal  $y$  value) were removed. Given the resemblance of the two studies, we also combined the two samples (after standardizing scores within each sample) so as to obtain more robust estimates of the CIs.

**Study 1.** We tested whether changes in participants' PSEs, toward “fearful” or “calm,” when they judged emotional experiences in other individuals using categorical terms (as compared with the continuous scale) correlated with neural activity in our ROIs. Parameter estimates of activity from voxels in the ROIs during the image-viewing period were extracted for both the categorical-judgment and the continuous-judgment conditions. We then averaged these values across voxels within ROIs to produce two values per ROI, corresponding to activity during the categorical-judgment condition and activity during the continuous-judgment condition. The latter value was subtracted from the former to produce a measure of the extent to which categorical judgments were associated with increased or decreased neural activity during emotion perception. We then regressed this difference in activity against participants' categorization-related PSE shifts to determine whether categorical-thinking-induced shifts in emotion perception toward “fear” or toward “calm” were associated with corresponding shifts in neural activity.

If categorical judgments about emotion shape the underlying neural representation of emotion, then to the extent that making categorical judgments increases the perception of fear (as measured by the difference in PSE between the categorical- and continuous-judgment conditions), brain regions supporting affect should show greater activity in the categorical-judgment condition than in the continuous-judgment condition. Likewise, to the extent that making categorical judgments increases





**Fig. 4.** Brain images showing regions in which activity correlated with the categorization-related shift in the point of subjective equality (PSE) and scatterplots (with best-fitting regression lines) illustrating those correlations. From top to bottom, the figure shows (a) results for the right amygdala, left ventral anterior insula, and right ventral anterior insula in Study 1 and (b) results for the left amygdala and right ventral anterior insula in Study 2.

the perception of “calm,” these areas should show reduced response in the categorical-judgment condition. We note that an alternative to this between-participants means of testing our hypothesis would be to sort each participant’s trials on an ad hoc basis, according to his or her responses, and then correlate these response categories with brain activity across trials. We chose not to conduct this type of within-participants analysis, however, because it would have two drawbacks. First, ad hoc sorting might disrupt stimulus counterbalancing, making it unclear whether findings were due to an unbalanced selection of stimuli from different presentation blocks rather than to the top-down effects of making categorical judgments. Second, the ad hoc approach would require several more trials than was permissible given the task design and scanning duration.

We tested whether those ROIs whose activity correlated with categorization-related shifts in emotion perception showed greater connectivity with the mPFC when categorical thinking held more sway over emotion perception. This required two steps. First, we used separate psychophysiological interaction (PPI) models to measure differences in moment-to-moment connectivity of the amygdala and insula seeds with mPFC brain regions. Each PPI model included two regressors in addition to those included in the initial model described earlier: activity in a seed region and the interaction of the seed region’s activity with the difference between the image-viewing regressors for the categorical- and continuous-judgment conditions. The latter (PPI) regressor assessed the degree to which moment-to-moment mPFC connectivity with activity in the seed region was greater during the categorical-judgment condition than during the continuous-judgment condition. Second, we correlated amygdala-PPI and insula-PPI connectivity with the absolute value of the categorization-related shift in the PSE. These absolute values were not normally distributed (skewness = 1.796,  $SE = 0.512$ ), so we applied a square-root transform (skewness = 0.163,  $SE = 0.512$ ). Given our a priori interest in the mPFC because of its well-known role in regulating the amygdala, but also given that the mPFC involves several subregions, we used a small-volume correction. Analysis was restricted to the mPFC extending from the central sulcus to the medial orbitofrontal cortex (see Fig. S3 in the Supplemental Material available online). We examined whether portions of the mPFC had greater PPI connectivity with the amygdala seed as a function of greater categorization-related shifts in PSEs.

**Study 2.** We conducted a similar set of analyses for Study 2. Categorization-related changes in emotion perception when participants judged their own affective experience, as measured by the difference in PSEs during categorical versus continuous judgments, were correlated with activity in our ROIs. The PPI analysis was

also conducted similarly to the one for Study 1. We correlated connectivity between the mPFC and the amygdala or ventral anterior insula with the absolute value of the difference between the PSEs in the categorical- and continuous-judgment conditions (no skewness correction was needed; skewness = 0.646,  $SE = 0.536$ ), to search for mPFC regions that had greater PPI connectivity with the ROIs as a function of greater categorization-related shifts in PSEs.

## Results

### ***Study 1: categorization-related differences in PSEs during the perception of other individuals' emotions***

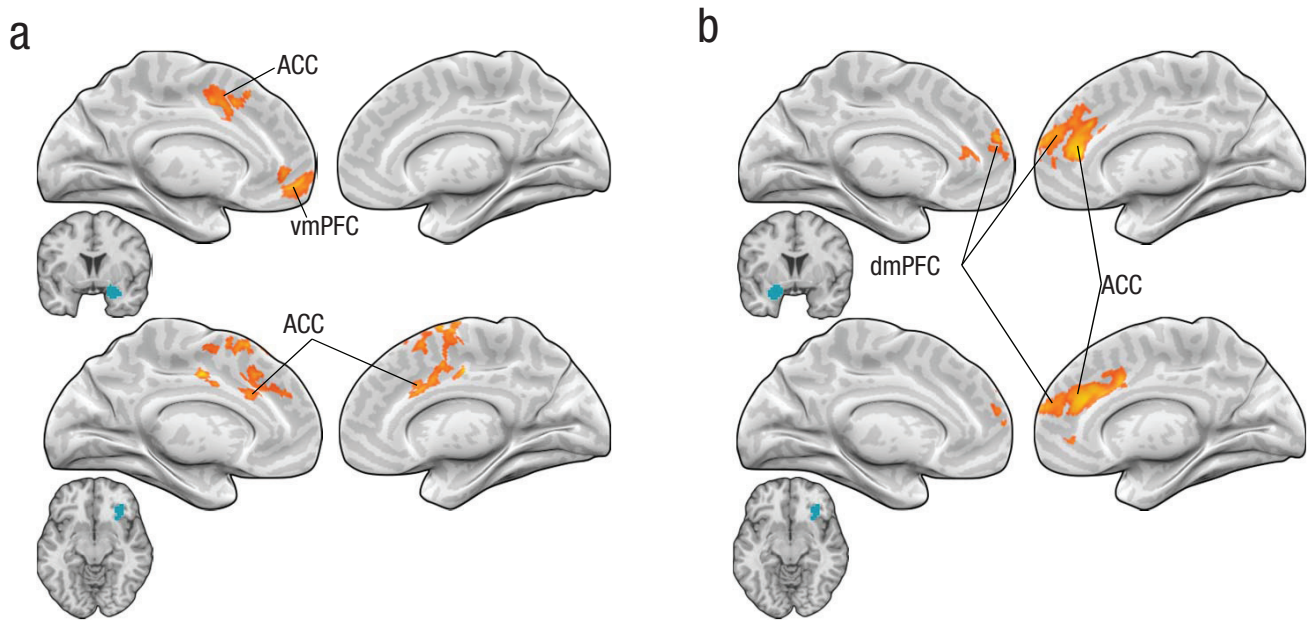
**Behavioral results.** We first used logit functions to examine participants' subjective reports as a function of the percentage of fear in the morphed faces. The model fits were high for both categorical judgments (mean  $R^2 = .91$ ) and continuous judgments (mean  $R^2 = .92$ ), and the difference in fit between the two task conditions was not significant,  $t(19) = 0.06$ ,  $p = .95$ . The PSEs in the two conditions were identified, and the difference between them was calculated for use in neuroimaging analyses. Mean time to decision was faster in the categorical-judgment condition ( $M = 1,492$  ms,  $SE = 77$ ) than in the continuous-judgment condition ( $M = 1,638$  ms,  $SE = 88$ ),  $t(19) = 3.64$ ,  $p < .002$ , which suggests that, on average, participants found it easier to make categorical than continuous judgments (also see the Supplemental Material). We performed subsequent analyses both including and excluding time to decision as a covariate, and found that including this covariate did not meaningfully change the pattern of results.

**Categorization-related shifts toward "fear" correlate with greater activity in the amygdala and ventral anterior insula.** As predicted, participants who had a categorization-related shift in their PSEs toward "fear" also had greater activity in the independently defined ROI of the right amygdala during categorical judgments than during continuous judgments (Fig. 4a),  $r(18) = .383$ ,  $p = .0438$  (one-tailed), 95% CI = [.176, .684]; controlling for mean time to decision reduced the coefficient, semipartial  $r(17) = .367$ ,  $p = .052$  (one-tailed). The same analysis conducted across the whole brain also showed that greater activity in the right amygdala correlated with categorization-related shifts in the PSE toward "fear" (peak MNI coordinates = [33, -3, -21],  $p < .05$ , family-wise-error-rate, or FWER, corrected). Corresponding between-condition differences in activity in the left amygdala ROI showed a numerically positive association with differences in PSEs, but this association was not

significant,  $r(18) = .243$ ,  $p = .145$  (one-tailed), and remained nonsignificant when we controlled for mean time-to-decision differences, semipartial  $r(17) = .246$ ,  $p = .144$  (one-tailed). Categorization-related PSE shifts toward "fear" also were correlated with increased activity in the independently defined ROI of the right ventral anterior insula (Fig. 4a),  $r(18) = .445$ ,  $p = .022$  (one-tailed), 95% CI = [.107, .719]; the correlation remained significant when we controlled for mean time-to-decision differences, semipartial  $r(17) = .436$ ,  $p = .0248$  (one-tailed), and was also significant at whole-brain-corrected levels (peak MNI coordinates = [36, 12, -15],  $p < .05$ , FWER corrected). Left ventral anterior insula activity showed a positive correlation with categorization-related shifts in emotion perception, but the bootstrapped CI included zero (Fig. 4a),  $r(18) = .425$ ,  $p = .0276$  (one-tailed), 95% CI = [-.164, .698]; the correlation remained significant when we controlled for mean time-to-decision differences, semipartial  $r(17) = .398$ ,  $p = .0378$  (one-tailed), but was not significant at whole-brain-corrected levels.

Overall, these findings indicate that, across participants, a tendency to perceive emotion as being more affectively intense ("fear") when judgments were categorical rather than continuous covaried with greater activity in the right amygdala and right ventral anterior insula during categorical judgments. Likewise, a tendency to perceive emotion as being less affectively intense ("calm") when judgments were categorical rather than continuous covaried with reduced activity in these areas during categorical judgments.

**Connectivity of the amygdala and ventral anterior insula with mPFC areas covaries with categorization-related PSE shifts.** Next, we conducted PPI analyses to search for brain regions that may regulate activity in the right amygdala and right insula, particularly when categorical thinking has a stronger influence over emotion perception. Specifically, we examined whether the mPFC had greater moment-to-moment PPI connectivity with the amygdala when the PSE difference between the categorical- and continuous-judgment conditions was greater by using the absolute value of the categorization-related PSE shift (corrected for skewness). As shown in Figure 5a, across participants, greater categorization-related shifts in PSEs were associated with greater connectivity of the right amygdala with portions of the mPFC (556 resels), including the anterior cingulate cortex (peak MNI coordinates = [0, 45, -15],  $p < .05$ , small-volume corrected) and ventromedial prefrontal cortex (peak MNI coordinates = [-12, 18, 42],  $p < .05$ , small-volume corrected). PPI connectivity of the right ventral anterior insula with the anterior cingulate cortex was also greater when the PSE shift was larger (558 resels; peak MNI coordinates = [9, 15, 33],  $p < .05$ , small-volume corrected;



**Fig. 5.** Results of the psychophysiological interaction analyses. The yellow shading shows the location of regions of medial prefrontal cortex for which connectivity with the seed regions was greater when the categorization-related shift in the point of subjective equivalence was larger. Results are presented separately for (a) Experiment 1 and (b) Experiment 2. In each panel, the top row shows results for the seed region of the amygdala, and the bottom row shows results for the seed region of the ventral anterior insula (seed regions indicated in blue). ACC = anterior cingulate cortex; dmPFC = dorsomedial prefrontal cortex; vmPFC = ventromedial prefrontal cortex.

Fig. 5a). PPI connectivity of the left ventral anterior insula with the mPFC was not correlated with categorization-related PSE shifts. These results are consistent with the view that during perception of emotion in other individuals, mPFC regions may regulate activity in the right amygdala and right ventral anterior insula—and that this role increases with greater demand to resolve bottom-up affective responses with top-down, categorical choices.

**Study 2: categorization-related differences in PSEs during the perception of one’s own affect**

**Behavioral results.** Using the same approach as in Study 1, we first used logit functions to examine participants’ subjective reports as a function of the normative image ratings of the images. The model fits were reasonable for both task conditions, although the fit was better for the continuous-judgment condition (mean  $R^2 = .62$ ) than for the categorical-judgment condition (mean  $R^2 = .47$ ),  $t(17) = 3.09$ ,  $p < .01$ . The PSEs were identified for each participant and for each judgment condition separately, and the PSE differences between the two conditions were calculated for use in neuroimaging analyses. Mean time to decision did not differ significantly between the categorical-judgment condition ( $M = 2,598$  ms,  $SE = 18.3$ ) and the continuous-judgment condition ( $M = 2,605$  ms,  $SE = 18.1$ ),  $t(17) = -0.06$ ,  $p = .96$ .

**Categorization-related shifts toward “bad” correlate with greater activity in the amygdala and ventral anterior insula.** Participants whose PSEs had a categorization-related shift toward “bad” also had greater activity in the left amygdala during the image-viewing period when making categorical judgments than when making continuous judgments (Fig. 4b),  $r(16) = .424$ ,  $p = .0357$  (one-tailed), 95% CI = [.067, .721]; this correlation remained significant when we controlled for time to decision,  $r = .437$ ,  $p = .0317$  (one-tailed). The same analysis conducted across the whole brain did not show a correlation between the categorization-related PSE shift and activity in the left amygdala. A corresponding analysis for activity in the right amygdala ROI did not reveal a significant correlation,  $r(16) = .264$ ,  $p = .138$ , although a correlation was observed at whole-brain-corrected levels at two local maxima,  $r = .645$  (MNI coordinates = [18, 0, -27]) and  $r = .729$  (MNI coordinates = [24, 9, -27]),  $ps < .05$ , FWER corrected. As predicted, participants’ PSE shifts also correlated with activity in our independently defined ROI of the right ventral anterior insula (Fig. 4b),  $r(16) = .482$ ,  $p = .019$  (one-tailed); the correlation remained significant when we controlled for time to decision, semi-partial  $r(15) = .432$ ,  $p = .033$  (one-tailed), 95% CI = [.128, .819], and was significant at whole-brain-corrected levels (peak MNI coordinates = [36, 18, -18],  $p < .05$ , FWER corrected). The correlation between the PSE shift and the between-condition difference in activity in the left

ventral anterior insula ROI was not significant ( $p > .25$ , one-tailed).

These findings indicate that, across participants, a tendency to perceive affective experience as more intense (“bad”) when judgments were categorical rather than continuous resulted in greater activity in the right ventral anterior insula and left amygdala during categorical judgments. Likewise, a tendency to perceive affective experience as less intense when judgments were categorical rather than continuous was associated with reduced activity in these regions during categorical judgments.

**Connectivity of the amygdala and ventral anterior insula with mPFC areas covaries with categorization-related PSE shifts.** Finally, we conducted a PPI analysis similar to the one in Study 1, to search for mPFC brain regions that may up- or down-regulate activity in the left amygdala and the right ventral anterior insula, particularly when categorical thinking has a stronger influence on emotion perception (see Fig. 5b). Greater categorization-related PSE shifts across participants were associated with greater connectivity of the left amygdala with portions of the mPFC (508 resels), including the anterior cingulate cortex (peak MNI coordinates = [9, 42, 30],  $p < .05$ , small-volume corrected), dorsomedial prefrontal cortex (peak MNI coordinates = [3, 54, 12],  $p < .05$ , small-volume corrected), and ventromedial prefrontal cortex (peak MNI coordinates = [15, 57, -3],  $p < .05$ , small-volume corrected). PPI connectivity of the mPFC (478 resels) with the right ventral anterior insula was also greater in the anterior cingulate cortex (peak MNI coordinates = [18, 33, 30],  $p < .05$ , small-volume corrected) and dorsomedial prefrontal cortex (peak MNI coordinates = [0, 54, 21],  $p < .05$ , small-volume corrected). Overall, these findings are consistent with the view that during perception of emotion in the self, mPFC regions may regulate activity in the amygdala and the ventral anterior insula—and that this role increases with greater demand to resolve bottom-up affective responses with top-down, categorical choices about one’s own affective experience.

### **ROI analyses combining the two studies**

Given the modest sample sizes of our studies, we fully acknowledge that replication in future studies on this topic will be required to further establish the statistical reliability of the findings. In light of this concern, and to obtain a better estimate of the CI for ROI analyses given the overall similarity of the studies, we combined their data to achieve a larger sample size. Participants who had a categorization-related shift in their PSEs toward the more affectively potent category (“fearful” or “bad”) also had greater activity in three of the ROIs during the categorization-judgment task than during the continuous-judgment task: left amygdala (95% CI for  $r = [.163, .548]$ ),

right amygdala (95% CI = [.026, .541]), and right ventral anterior insula (95% CI = [.28, .719]). The correlation for the left ventral anterior insula narrowly missed the required threshold for significance, 95% CI = [-.009, .563].

### **Task variables**

It also may be possible that having to make categorical judgments is more aversive, relative to making continuous judgments, for some participants than for others. Such an aversion could incidentally shift participants’ PSEs toward more arousing emotion categories and also be responsible for the correlation we observed between the categorization-related PSE shift and activity in our amygdala and insula ROIs. In a posttask questionnaire, we measured whether participants found the categorical-judgment task to be easier, more natural, or more emotional than the continuous-judgment task. Contrary to this alternative account, participants’ task preferences were not associated with the categorization-related shift in PSE, or with activity in the ROIs (see the Supplemental Material for details). In a prior study (Satpute et al., 2012), we investigated which brain regions were more responsive during categorical- relative to continuous-judgment conditions in general, that is, regardless of shifts in emotion perception thresholds (i.e., PSEs). We found that the ventrolateral prefrontal cortex—a region commonly associated with semantic (including social) judgments (Binder, Desai, Graves, & Conant, 2009; Lindquist et al., 2015; Satpute, Badre, & Ochsner, 2014)—was significantly more activated during categorical-judgment conditions, whereas activity in the amygdala and the ventral anterior insula was not. In combination with the present findings, these results are consistent with the view that mere instruction to think about emotion in categorical or continuous ways does not influence activity in the amygdala or the ventral anterior insula. Instead, activity in those areas increases when categorical thinking shifts emotion perception thresholds toward high-arousal emotion and affect categories, such as “fear” or “bad,” and decreases when categorical thinking shifts emotion perception thresholds toward low-arousal emotion and affect categories, such as “calm” or “neutral.”

### **Discussion**

People often act on the basis of which emotions they perceive to be present in others or in themselves. At times, this may involve thinking about emotions in “black and white” terms rather than in “shades of gray” terms. Here, we asked whether judging emotions in categorical or continuous ways shapes how those emotions are perceived and neurally represented. Employing a combination of psychophysical and functional MRI methods, we observed three key findings. First, participants’ thresholds for

perceiving other people's or their own emotions shifted when they made categorical, relative to continuous, judgments. Second, activation in systems that process the affective qualities of stimuli (i.e., the amygdala and ventral anterior insula) shifted in accordance with this PSE shift. Third, activity in these regions was more strongly coupled with activity in mPFC regions when categorical judgments induced a greater shift in emotion perception.

Our findings suggest that categorical judgments of emotion are not just “readouts” of stimulus features, but rather contribute to emotion perception and its underlying neural representation—particularly when sensory inputs are uncertain or ambiguous. In such situations, using categories appears to shift emotion perception to be more consistent with the category selected. Top-down categorical knowledge, when used to interpret gradations of sensory inputs, be they other people's facial movements or interoceptive signals, may actively shape the perception of emotion.

### ***Implications for research on emotion perception***

Our findings contribute to emotion research in three ways. First, although emotion theory has long accounted for top-down contributions to emotion perception (Barrett, 2006; Ochsner et al., 2004), few studies have examined how the categories people use to describe emotion shape emotions themselves (e.g., Nook et al., 2015), and fewer still have asked what neural mechanisms underlie such effects. Previous studies of emotion perception have focused on the presence versus absence of emotion language when participants judge a stimulus rather than on the categorical or continuous nature of that judgment per se (Gendron et al., 2012; Kassam & Mendes, 2013; Lieberman et al., 2007; Lindquist et al., 2015; Nook et al., 2015; Roberson & Davidoff, 2000). Our design differed slightly from prior designs in that linguistic terms were provided in both conditions and what we varied was the categorical versus continuous way in which they guided judgments. With linguistic terms present in both conditions, we found that thinking about emotions in categorical ways shaped emotion perception. Thus, our findings support prior work showing that emotion judgments shape emotion perception, but also identify a unique role for categorical thinking. An issue of interest is the stage at which categorical thinking plays a role, be it during encoding of stimulus features or later, at response selection. This may be addressed in future work using techniques that provide greater temporal resolution.

Second, our findings inform current understanding of the neural mechanisms underlying emotion perception. We found that the mPFC plays an important role at the interface between the top-down use of interpretive categories and sensory inputs (Barrett & Satpute, 2013; Ochsner & Gross, 2014). Our findings provide empirical

support for the view that the mPFC plays a role in meaning making or conceptualization. This idea has been supported by meta-analyses and literature reviews (Barrett & Satpute, 2013; Ochsner & Gross, 2014; Roy et al., 2012), but not examined directly. Our studies suggest that the mPFC may be involved in conceptualizing emotion by resolving continuous sensory inputs with bounded concepts (cf. Grinband, Hirsch, & Ferrera, 2006).

And third, this is the first study to examine how categorical thinking plays a role in people's perception of their own emotions. We found that parallel systems involving the mPFC and the insula or amygdala supported categorical influences on participants' perception of other individuals' and their own emotions. Although we expected the insula to be involved when participants perceived their own emotions, its involvement when they judged emotions in others is consistent with findings from empathy studies suggesting that interoceptive information also informs emotion perception in others (Zaki et al., 2012). But more pertinently, thinking about emotion in categorical terms appears to contribute to both how one perceives others and how one perceives one's own affective reactions to the world.

### ***Implications for decision making, categorization, and social cognition***

Our studies have implications for several research areas in which forced-choice judgments are used to assess behavior and cognition. Decision-making studies often focus on choices between two stimuli under conditions of uncertainty. But this work rarely considers whether the categorical nature of such judgments influences the value representation of the stimuli. Categorization research has examined how continuously varying stimuli are perceived categorically (Calder, Young, Perrett, Etcoff, & Rowland, 1996; Harnad, 1990), but this work has focused on whether perception is categorical regardless of the judgment type and not as a consequence of it (cf. Fugate, 2013; for a notable exception, see Kay & Kempton, 1984). In social-cognition tasks, people often judge a conspecific as “Black” or “White,” or as “trustworthy” or “untrustworthy,” or a behavior as “good” or “evil.” It is often assumed (if only implicitly) that these forced-choice judgments are inert with respect to the stereotypes, impressions, or morals they measure. Our findings suggest that categorical judgments—especially when made about people, behaviors, or options that fall in the gray zone—may change perception and mental representation of targets to be consistent with the category selected.

Our study dovetails with research on object perception, which has found that changing the content of judgments about objects (e.g., judgments about color vs. movement) influences the neural representations of those objects (Harel, Kravitz, & Baker, 2014). Extending those

findings, we found that varying whether judgments within a single content dimension—in this case, emotion—were categorical or continuous shifted participants' perceptions of emotion and neural representations. Taken together, these findings suggest that both the content of a judgment and its categorical quality may play roles in shaping perception. The resemblance of our findings to recent findings regarding object perception also suggests that common prefrontal mechanisms may underlie the influence of judgments on perception across content domains (cf. Grinband et al., 2006), a possibility that may be tested in future work.

### ***Implications for individual differences in healthy and clinical populations***

People vary in how they perceive emotions in others' expressions, and these differences are associated with demographic characteristics (e.g., Thomas, De Bellis, Graham, & LaBar, 2007) and clinical conditions, including social phobia, depression, schizophrenia, and psychopathy (e.g., Joormann & Gotlib, 2006). Past studies have relied on categorical judgments of emotion to examine differences in perception. Our results suggest that how people think about their emotions changes their perceptions, which highlights the need to understand in what situations—and for what individuals—categorical and continuous thinking are most likely to be evoked.

#### **Action Editor**

Ralph Adolphs served as action editor for this article.

#### **Author Contributions**

A. B. Satpute and K. N. Ochsner developed the study concept; A. B. Satpute, E. C. Nook, J. Shu, and S. Narayanan further developed the study content and devised the tasks. E. C. Nook and S. Narayanan developed the stimuli for Study 1, and E. C. Nook programmed Study 1. J. Shu organized the stimuli for and programmed Study 2. Data collection was performed by E. C. Nook, S. Narayanan, J. Shu, and A. B. Satpute. A. B. Satpute designed and performed the analyses with support from J. Weber and K. N. Ochsner. A. B. Satpute and K. N. Ochsner drafted the manuscript, and all the authors provided revisions. All the authors approved the final version of the manuscript for submission.

#### **Acknowledgments**

We thank Andrew Kogan for MRI assistance and Managing Editor Michele Nathan for much support in manuscript revisions.

#### **Declaration of Conflicting Interests**

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

#### **Funding**

This research was supported by Grant MH076137 from the National Institute of Mental Health, Grant AG043463 from the National Institute on Aging, and Grant HD069178 from the National Institute of Child Health and Human Development (all awarded to K. N. Ochsner).

#### **Supplemental Material**

Additional supporting information can be found at <http://pss.sagepub.com/content/by/supplemental-data>

#### **References**

- Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences*, 3, 469–479.
- Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*, 411, 305–309.
- 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., & Satpute, A. B. (2013). Large-scale brain networks in affective and social neuroscience: Towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology*, 23, 361–372.
- 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.
- Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996). Categorical perception of morphed facial expressions. *Visual Cognition*, 3, 81–117.
- Craig, A. D. (2011). Significance of the insula for the evolution of human awareness of feelings from the body. *Annals of the New York Academy of Sciences*, 1225, 72–82.
- Critchley, H. D., Wiens, S., Rotshtein, P., Öhman, A., & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience*, 7, 189–195.
- Feinstein, J. S. (2013). Lesion studies of human emotion and feeling. *Current Opinion in Neurobiology*, 23, 304–309.
- Fugate, J. M. (2013). Categorical perception for emotional faces. *Emotion Review*, 5, 84–89.
- Gendron, M., Lindquist, K. A., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion*, 12, 314–325.
- Grinband, J., Hirsch, J., & Ferrera, V. P. (2006). A neural representation of categorization uncertainty in the human brain. *Neuron*, 49, 757–763.
- Harel, A., Kravitz, D. J., & Baker, C. I. (2014). Task context impacts visual object processing differentially across the cortex. *Proceedings of the National Academy of Sciences, USA*, 111, E962–E971.
- Harnad, S. R. (Ed.). (1990). *Categorical perception: The groundwork of cognition*. Cambridge, England: Cambridge University Press.
- Heberlein, A. S., Padon, A. A., Gillihan, S. J., Farah, M. J., & Fellows, L. K. (2008). Ventromedial frontal lobe plays a critical role in facial emotion recognition. *Journal of Cognitive Neuroscience*, 20, 721–733.

- Jack, R. E., & Schyns, P. G. (2015). The human face as a dynamic tool for social communication. *Current Biology*, *25*, R621–R634.
- Joormann, J., & Gotlib, I. H. (2006). Is this happiness I see? Biases in the identification of emotional facial expressions in depression and social phobia. *Journal of Abnormal Psychology*, *115*, 705–714.
- Kassam, K. S., & Mendes, W. B. (2013). The effects of measuring emotion: Physiological reactions to emotional situations depend on whether someone is asking. *PLoS ONE*, *8*(6), Article e64959. doi:10.1371/journal.pone.0064959
- Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologist*, *86*, 65–79.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual* (Technical Report No. A-8). Gainesville: University of Florida.
- Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H., & Way, B. M. (2007). Putting feelings into words: Affect labeling disrupts amygdala activity in response to affective stimuli. *Psychological Science*, *18*, 421–428.
- Lindquist, K. A., Satpute, A. B., & Gendron, M. (2015). Does language do more than communicate emotion? *Current Directions in Psychological Science*, *24*, 99–108.
- Lindquist, K. A., Satpute, A. B., Wager, T. D., Weber, J., & Barrett, L. F. (2016). The brain basis of positive and negative affect: Evidence from a meta-analysis of the human neuroimaging literature. *Cerebral Cortex*, *26*, 1910–1922.
- Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral & Brain Sciences*, *35*, 121–143.
- Master, A., Markman, E. M., & Dweck, C. S. (2012). Thinking in categories or along a continuum: Consequences for children's social judgments. *Child Development*, *83*, 1145–1163.
- 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.
- Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology: General*, *129*, 242–261.
- Ochsner, K. N., & Gross, J. J. (2014). The neural bases of emotion and emotion regulation: A valuation perspective. In J. J. Gross (Ed.), *Handbook of emotional regulation* (2nd ed., pp. 23–41). New York, NY: Guilford Press.
- Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D. E., & Gross, J. J. (2004). For better or for worse: Neural systems supporting the cognitive down- and up-regulation of negative emotion. *NeuroImage*, *23*, 483–499.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Pessoa, L., Japee, S., Sturman, D., & Ungerleider, L. G. (2006). Target visibility and visual awareness modulate amygdala responses to fearful faces. *Cerebral Cortex*, *16*, 366–375.
- Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & Cognition*, *28*, 977–986.
- Roy, M., Shohamy, D., & Wager, T. D. (2012). Ventromedial prefrontal-subcortical systems and the generation of affective meaning. *Trends in Cognitive Sciences*, *16*, 147–156.
- Salovey, P., & Mayer, J. D. (1989). Emotional intelligence. *Imagination, Cognition and Personality*, *9*, 185–211.
- Satpute, A. B., Badre, D., & Ochsner, K. N. (2014). Distinct regions of prefrontal cortex are associated with the controlled retrieval and selection of social information. *Cerebral Cortex*, *24*, 1269–1277.
- Satpute, A. B., Shu, J., Weber, J., Roy, M., & Ochsner, K. N. (2012). The functional neural architecture of self-reports of affective experience. *Biological Psychiatry*, *73*, 631–638.
- Thomas, L. A., De Bellis, M. D., Graham, R., & LaBar, K. S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental Science*, *10*, 547–558.
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*, 242–249.
- Wager, T. D., Keller, M. C., Lacey, S. C., & Jonides, J. (2005). Increased sensitivity in neuroimaging analyses using robust regression. *NeuroImage*, *26*, 99–113.
- Whalen, P. J., & Phelps, E. A. (2009). *The human amygdala*. New York, NY: Guilford Press.
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, *8*, 665–670.
- Zaki, J., Davis, J. I., & Ochsner, K. N. (2012). Overlapping activity in anterior insula during interoception and emotional experience. *NeuroImage*, *62*, 493–499.