AMERICAN PSYCHOLOGICAL ASSOCIATION

Emotion

ISSN: 1528-3542

2025, Vol. 25, No. 5, 1151–1168
https://doi.org/10.1037/emo0001483

Categorization Processes in Emotion Expression Recognition: The Roles of Language and Essentialism

Kaitlyn M. McMullen[1], Alexandra MacVittie[1], Jessica S. Leffers[2], Nicole Betz[3], and Jolie B. Wormwood[1]
[1] Department of Psychology, University of New Hampshire
[2] Department of Psychology, St. Mary's College of Maryland
[3] Department of Psychology, Yale University

We examined categorical processing biases in the perception and recognition of facial expressions of emotion across two studies. In both studies, participants first learned to discriminate between two ambiguous facial expressions of emotion selected from the middle of a continuous array of blended expressions (i.e., an array created from morphing images of two facial expressions together, with still images selected in equidistant increments). Participants were then asked to recognize the specific expressions they were trained to discriminate. In Study 1, target expressions labeled with emotion words (e.g., more angry face) during discrimination and recognition tasks were misremembered as more perceptually distinct from one another and therefore more perceptually similar to the stereotypical expression for their labeled emotion category than they were in reality. Critically, in Study 2, these recognition biases were significantly reduced or absent in conditions where the target expressions were not labeled with emotion words (e.g., Face A), demonstrating the role of emotion words in promoting categorical processing biases in emotion recognition. Moreover, in the absence of emotion labels, peoples' beliefs about the nature of emotion categories were related to the extent to which they employed categorical processing during emotion perception and recognition. Specifically, people with more essentialist beliefs about emotion categories—believing emotion categories are more innate, biologically-based, and immutable—exhibited more pronounced categorical processing biases during emotion recognition. Findings shed light on the critical role of language and cognition in constructing emotion and add to empirical findings on categorical processing in emotion perception.

Keywords: emotion perception, facial expressions, discrimination learning, perceptual memory, psychological essentialism

Supplemental materials: https://doi.org/10.1037/emo0001483.supp

Categorization is a primary function of our brains and a facilitator of our everyday lives (Harnad, 1987, 2006). In a broad sense, categories are sets of people or things grouped together based on real or perceived commonalities (Harnad, 2006). The categories we create and use to make sense of the world shape our perception of internal and external sensory information by emphasizing within-category commonalities and across-category distinctions, leading us to treat sensory input that varies continuously along some dimension as instead reflecting discrete classes of sensory input (Al-Rasheed, 2015; Harnad, 2006). For example, although there is a continuous spectrum of wavelengths of visible light, we perceive distinct categories of color (e.g., red, orange, blue; Gibson et al., 2017). Critical to the current investigation, it has also been theorized that the ways we experience and perceive emotions may be guided by categorization processes in a similar way (Barrett, 2006b, 2012, 2017a; Barrett et al., 2007; Calder et al., 1996). By this account, emotions are not naturally carved into discrete bins but constructed with the help of our proclivity to categorize. Here, we examine evidence for the role of categorization processes in the memory of ambiguous facial expressions of emotion with a focus on the contributing roles of language and individual beliefs about emotions as categories.

This article was published Online First January 20, 2025.
Erik C. Nook served as action editor.
Kaitlyn M. McMullen https://orcid.org/0009-0009-4244-4965
The authors report how they determined their sample size, all data exclusions (if any), all manipulations, and all measures in the study. Data from all studies in this article are publicly available in the Open Science Framework repository (https://osf.io/ydwf4/). The methods, analysis, and predicted results for Study 1 (conducted in 2021; https://osf.io/wmutc/) and Study 2 (conducted in 2022; https://osf.io/q6f8v/) were preregistered on the Open Science Framework.

Kaitlyn M. McMullen played a lead role in data curation, formal analysis, investigation, writing–original draft, and writing–review and editing and an equal role in software. Alexandra MacVittie played a supporting role in data curation, investigation, and writing–review and editing and an equal role in software. Jessica S. Leffers played a supporting role in writing–review and editing and an equal role in conceptualization. Nicole Betz played a supporting role in writing–review and editing and an equal role in conceptualization and methodology. Jolie B. Wormwood played a lead role in methodology, and supervision, a supporting role in data curation, formal analysis, and writing–original draft, and an equal role in conceptualization and writing–review and editing.

Correspondence concerning this article should be addressed to Kaitlyn M. McMullen, Department of Psychology, University of New Hampshire, 15 Academic Way, Durham, NH 03824, United States. Email: kaitlyn.mcmullen@unh.edu

Emotions as Categories

That emotion categories exist and are important in understanding both emotion experience and perception is generally accepted (see, e.g., Barrett, 2012, 2017a; Calder et al., 1996; Ekman, 1992; Ellsworth & Scherer, 2002; Etcoff & Magee, 1992; Russell & Bullock, 1986). However, there is ongoing debate among affective scientists about the precise nature of emotion categories (e.g., anger, happiness, fear). Specifically, researchers disagree about the extent of within- versus between-category variation in emotion experience and perception as well as whether the categories arise from innate biological distinctions or are more socially constructed (for a discussion, see, e.g., Barrett, 2006a, 2017b; Barrett et al., 2019).

According to what is frequently termed the classical view of emotions, discrete emotion categories are thought to represent distinct biologically based classes of experience such that the emotion categories are innate, existing from birth, and universal across cultures (Cordaro et al., 2020; Ekman, 1971, 1992; Ekman et al., 1987; Izard, 1971; Panksepp, 2004). From this theoretical perspective, emotions are also posited to each involve a distinct, fixed pattern of neural, physiological, and behavioral activity, including a discrete and universally recognizable facial expression (Ekman & Cordaro, 2011; Ekman et al., 1980). Thus, from the perspective of the classical view of emotions, emotion experience and perception are categorical in their very essence.

In contrast, according to constructionist theories of emotion, like the theory of constructed emotions (Barrett, 2006a, 2017b; Barrett et al., 2015; Lindquist, 2013; Lindquist & Barrett, 2008) emotions are not thought to represent biologically discrete phenomena but are instead thought to emerge from categorization processes. From this perspective, the experience of emotion is posited to result from a process of categorization whereby interoceptive sensations (i.e., sensations from the internal milieu) are combined with sensory information from the immediate surroundings using conceptual knowledge about emotion categories (Barrett, 2006a, 2006b; Frijda, 2009). Conceptual knowledge about emotion categories is thought to be based on one's language and culture, as well as one's experiences in similar situations in the past. In essence, an emotion is proposed to emerge when we use conceptual knowledge to categorize internal sensory information in a given context as an instance of a specific, labeled emotion category (e.g., anger).

According to constructionist theoretical perspectives, a similar categorization process involving conceptual knowledge about emotions is thought to be critical to how we perceive emotions in others as well (Satpute et al., 2016). Although people's faces and bodies can move in widely varied and dynamic ways, including during the experience of an emotion (Barrett, 2006a; Durán & Fernández-Dols, 2021; Hoemann et al., 2020; Siegel et al., 2018), we frequently attempt to categorize or label others' emotions as belonging to one of a set of discrete classes of emotion. That is, to perceive an instance of a particular emotion from another individual's face or body involves an act of categorizing nondiscrete motor movements into a discrete category labeled by an emotion word.

Essentialist Beliefs About Emotion

Conceptual disagreements about the nature of emotion are also reflected in everyday intuitions held by lay people, who differ in the extent to which they believe emotion categories are innate and discrete, an example of psychological essentialism. Psychological essentialism refers to the idea that categories have an underlying essence that defines category membership, such that categories have innate features that make them what they are, both at the surface level and in a metaphysical sense (i.e., unchangeable underlying property determining a category's identity; Haslam et al., 2000; Medin & Ortony, 1989; Prentice & Miller, 2007). Past research has demonstrated that people vary in the extent to which they hold essentialist beliefs about various social categories, including race (Mandalaywala et al., 2018; Roets & Van Hiel, 2011, for a review, see Roth et al., 2023), gender (Prentice & Miller, 2006; Skewes et al., 2018), and sexual orientation (Hoyt et al., 2019; Morandini et al., 2015, 2021; Morgenroth et al., 2021). This research also shows that holding essentialist beliefs can have important implications for how people perceive and reason about members of different social groups. For example, individuals who hold essentialist beliefs about race or gender tend to exhibit higher levels of prejudice toward racial and gender minority groups than those who hold less essentialist beliefs (e.g., race: Mandalaywala et al., 2018; gender: Skewes et al., 2018).

Of primary relevance to the present investigation, there is also evidence that individuals hold essentialist beliefs about emotion categories, believing each emotion has a biological cause or essence that defines it (Berent et al., 2020; Lindquist et al., 2013). For example, participants report that facial expressions of emotion currently normative in America would be recognized by hunter–gatherers with no contact with Western civilization, suggesting they believe emotions are both innate and embodied (i.e., evident in the face and body; Berent et al., 2020). Moreover, these beliefs persisted even when participants were explicitly informed that emotions are acquired categories (Berent et al., 2020). In another series of studies, participants rated emotions as high on a number of dimensions of essentialism, including that emotions are defined by a shared biological essence and have a casual mechanism in the body (vs. the mind; Lindquist et al., 2013). Critically, in one study, individuals who tended to have more essentialist beliefs about emotions in general also described their own emotional experiences as more highly differentiated, though this was not evident in self-report data on their emotions collected over time (Lindquist et al., 2013). These findings suggest that differences in the extent to which one essentializes emotion categories more generally may relate to differences in how they apply category labels when perceiving their own or others' emotional experiences.

Categorical Processing in Emotion Perception and Recognition

Categorial processing biases perception by simplifying reality into more discrete and comprehensible groupings. When a continuum is divided into an exhaustive set of classes, the divisions across classes or categories create biases by affecting discrimination of stimuli drawn from the continuum (Al-Rasheed, 2015). Discrimination is generally affected such that there is: (a) increased perceived dissimilarity across category boundaries and (b) increased perceived similarity within-category boundaries. For example, in general, individuals tend to be faster and more accurate at discriminating stimuli that cross categories than stimuli that both fall within the same category (Al-Rasheed, 2015; Harnad, 2006) and these categorical processing biases have been demonstrated across a wide variety of domains including the perception of social others (Anthony et al., 1992), color (Gibson et al., 2017), and speech (Harnad, 1987). Thus, observation of such biases are cues that categorical processing is occurring.

Given the ways people think and reason about emotion categories, it stands to reason that similar categorical processing biases might be present in emotion perception. Indeed, several prior studies have provided evidence of categorical processing biases for the perception of facial expressions of emotion (Calder et al., 1996; Etcoff & Magee, 1992; Fugate et al., 2010; Roberson et al., 2007, 2010; for a review, see Fugate, 2013). For example, early work with drawings of facial expressions (Etcoff & Magee, 1992) and morphed images of posed facial expressions (Calder et al., 1996) revealed that individuals perceive an expression continuum (e.g., from happiness to sadness) as having two distinct sections separated by a category boundary, and that discrimination for expressions along a given expression continuum was better for cross-boundary pairs than within boundary pairs. That is, participants were better able to discriminate between two faces that they had previously identified as belonging to different emotion categories (i.e., happiness vs. sadness) than two faces of equal physical difference that they had previously identified as belonging to the same emotion category. This pattern of results was also replicated with a larger variety of emotion expression continuums in a native Japanese sample using a similar set of identification and discrimination tasks (Fujimura et al., 2012). Moreover, research using a similar paradigm with chimpanzee facial expressions found evidence of categorical processing only when participants first learned to associate the faces with labels, suggesting that emotion words may play a causal role in producing categorical processing of facial expressions of emotion (Fugate et al., 2010).

Although not directly testing categorical processing biases, research on the ways that language (i.e., the presence vs. absence of emotion words) can shape emotion perception is also relevant as it provides evidence of the ways category labels are utilized in the perception and memory of facial expressions of emotion (Barker et al., 2020; Champoux-Larsson & Nook, 2024; Lindquist et al., 2006, 2014; Lindquist & Gendron, 2013). For example, semantic dementia patients who have an inability to process and remember emotion words exhibit disrupted emotion perception (Lindquist et al., 2014). When asked to sort posed facial expressions of emotions into piles based on who felt the same, semantic dementia patients generally created valence-based piles (e.g., feeling good, feeling bad, feeling nothing) instead of separate categories for discrete emotions like anger, fear, and happiness. This pattern of emotion perception presumably emerged because of their inability to access or use discrete emotion words to guide categorization (Lindquist et al., 2014). Similarly, neurotypical individuals showed disrupted emotion perception following a semantic satiation paradigm in which emotion words were made temporarily inaccessible (Gendron et al., 2012; Lindquist et al., 2006); when an emotion label was made inaccessible to them, participants were slower and less able to match a posed stereotypical facial expression of emotion to its correct emotion category label.

Indeed, language appears to provide a unique context for emotion perception (Doyle et al., 2021; for a review, see Lindquist, 2021). For instance, participants are more accurate at perceptually matching facial expressions after being primed with emotion word labels versus a situational context (e.g., an image meant to evoke sadness; Doyle et al., 2021) and are faster to match facial expressions to emotion words than other facial expressions (Nook et al., 2015), even when the emotion words are from a second language (Champoux-Larsson & Nook, 2024). Interestingly, emotion words have also been demonstrated to impact neural representations of emotion (for a review, see Satpute & Lindquist, 2021). In a meta-analysis, Brooks et al. (2017) observed more frequent activation in brain regions related to semantic processing when emotion words (vs. general affect words, e.g., pleasant, unpleasant) were present in emotion perception and experience, suggesting that emotion words prompted semantic retrieval and application of relevant emotion concepts during both the experience and perception of emotion.

The presence of emotion category words also appears to impact the ability to accurately *remember* facial expressions of emotion at a later time. For example, J. B. Halberstadt and Niedenthal (2001) provided evidence that explaining the reason that an ambiguous facial expression was an example of a specific emotion category biased the memory of that facial expression toward the stereotypical expression for that category. In other words, providing an explanation as to why someone with an ambiguous facial expression was expressing happiness biased memory of the expression such that it was recalled as looking more stereotypically happy than it was in reality. Furthermore, this memory bias was shown to be driven by emotion category explanations drawing attention toward individual facial features at the expense of processing the whole facial configuration as a gestalt (J. Halberstadt, 2005). That is, participants who provided emotion category explanations for facial expressions showed less of a recall bias for individual features and configurally altered faces (e.g., inverted faces) but greater recall bias for whole, unaltered facial expressions. Taken together, these studies demonstrate that the accessibility of emotion categories and their verbal labels play a critical role in emotion perception and memory.

Present Studies

In the present studies, we conceptually replicate and extend prior work on the role of categories in emotion perception and memory, including work on categorical processing biases in emotion perception (e.g., Calder et al., 1996; Etcoff & Magee, 1992; Fugate et al., 2010) and the role of emotion category knowledge in recall of emotional expressions (e.g., J. Halberstadt, 2005; J. B. Halberstadt & Niedenthal, 2001). Specifically, we combine insights from these previous areas of research to focus on evidence of categorical processing biases in *remembering* facial expressions of emotion, while past work directly assessing categorical processing biases in emotion have focused on perception. In addition, unlike prior work on the role of category knowledge in memory for facial expressions of emotion, we do not examine *recall* of facial expressions (i.e., retrieval of related details from memory) but rather examine categorical processing biases in *recognition* of facial expressions (i.e., identifying a previously seen image), a memory process considered to require much less cognitive effort than recall.

To examine categorical processing in recognition of facial expressions of emotion, we use a paradigm novel to emotion research, called peak shift. Peak shift is an established paradigm for probing categorical processing from animal and vision science and refers to a form of stimulus generalization that occurs from discrimination learning, for example, with animals displaying a preference for more easily distinguishable stimuli which results in decreased risk of misclassification (Lynn, 2010; Lynn et al., 2005;

Nicholson & Gray, 1971). Here, we use a similar paradigm with human perceivers to see how categorical processing biases in perceptual memory emerge following discrimination training on two ambiguous facial expressions of emotion. Specifically, participants are trained to discriminate between two ambiguous facial expressions drawn from the middle of an expression continuum (i.e., a morphed array of images of facial expressions, each of equal physical distance), and each target ambiguous expression is labeled with a different emotion category label (e.g., more angry vs. more fearful) during training. Participants are then shown all morphed images from the expression continuum one at a time and asked to identify the two target expressions they were trained on. We predict that participants will demonstrate categorical processing biases such that they will: (a) recognize the target expressions labeled as belonging to different emotion categories as more perceptually distinct from one another than they were in reality; and thereby (b) recognize each target expression as more perceptually similar to the stereotypical expression for its labeled emotion category than it was in reality. This paradigm is unique in that it enables measurement of both types of categorical processing biases simultaneously (i.e., increased between-category dissimilarity and decreased within-category dissimilarity). Moreover, because participants purposefully learn to identify the target facial expressions through repeated exposure with feedback and then merely need to recognize the learned expressions, the paradigm tests memory for facial expressions of emotion under circumstances that should benefit accuracy, providing a particularly strong test of the role of categorical processing in memory for facial expressions of emotion.

In addition, given different theoretical and lay stances on the nature of emotion categories (i.e., as innate vs. constructed), we also examine the extent to which individuals' *beliefs* about emotion categories may be related to the extent to which they exhibit categorical processing biases for ambiguous emotion expressions. Specifically, we examine whether categorical processing biases are more pronounced among individuals who hold more essentialist beliefs about emotion—that is, the extent to which they believe emotion categories are inherently and biologically distinct from one another, each having vital/defining features that make it what it is (i.e., an "essence") as opposed to emotion categories being constructed through a process of categorization (Haslam et al., 2000; Medin & Ortony, 1989). As essentialism can be associated with the perception of categories as highly differentiated from one another (Prentice & Miller, 2007), we predicted that individuals with more essentialist beliefs about emotion categories would exhibit greater categorical processing biases for ambiguous facial expressions of emotion.

Study 1

In this first study, we created morphed images of facial expressions of emotion, ranging from a 100% anger stereotype expression to a 100% fear stereotype expression in 10% increments, such that increments represented physically equidistant changes from one expression to the other (see Figure 1). We trained participants to discriminate between two ambiguous expressions drawn from the center of the array: the 40% anger stereotype (60% fear stereotype) expression, labeled the "more fearful" expression, and the 60% anger stereotype (40% fear stereotype) expression, labeled the "more angry" expression. Finally, we had participants try to recognize the facial expressions they were trained on from among the full array of morphed expressions. We focused on the emotion categories of anger and fear because they share the same normative valence and arousal (i.e., both are normatively negative and high arousal emotions; Barrett & Bliss-Moreau, 2009). This allowed us to avoid confounding emotion category membership with differences in affect or affect category in Study 1.

As described above, we predicted that participants would demonstrate categorical processing biases, by (a) recognizing the target expressions labeled as belonging to different emotion categories as more perceptually distinct from one another than they were in reality; and (b) recognizing each target expression as more perceptually similar to the stereotypical expression for its labeled emotion category than it was in reality. Additionally, we assessed individual differences in essentialist beliefs about emotion and predicted that individuals with more essentialist beliefs about emotion categories (i.e., those who view emotion categories as more innate, distinct, and unmalleable) would exhibit greater categorical processing biases for ambiguous facial expressions of emotion.

Method

Transparency and Openness

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study. Data from all studies in this article are publicly available in the Open Science Framework repository (https://osf.io/ydwf4/). The methods,

Figure 1

Sample Stimuli From the Face Perception Task

Note. The gray and dotted boxes represent the more angry and more fearful target expressions, respectively. Explicit consent from the individual in the photo was obtained to create and use this morphed array. See the online article for the color version of this figure.

analysis, and predicted results for Study 1 (conducted in 2021) were preregistered on the Open Science Framework (https://osf.io/wmutc/). All studies received ethics approval and oversight from the University of New Hampshire IRB (IRB-FY2022-124).

Participants

A power analysis conducted in G*Power (V3.1) indicated we would have appropriate statistical power (80%) to detect a small effect size (Cohen's $f^2 = .02$) for a bivariate regression (two-tailed, $\alpha = .05$) with a sample size of $N = 123$. We preregistered that we would collect data from a sample of $N = 150$ participants, but decided to continue collecting data until the end of the semester in which data collection began; no analyses were conducted prior to the decision to extend data collection. Supplemental analyses with only the first $N = 150$ participants (to match the target N in our preregistration) revealed the same pattern of results and significance as analyses with the full sample (see Table S1).

Two-hundred thirty undergraduates from the University of New Hampshire consented to participate in an online study for course credit. Of the 230 participants who consented to participate, only 174 participants completed the study in full as the remainder either failed to link out to the Face Perception Task (run in InquisitWeb) or failed to return to Qualtrics following the Face Perception Task to complete the questionnaires. In addition, consistent with the preregistration, seven participants (4.0%) were removed from analyses for obtaining less than 50% accuracy in the training phase of the Face Perception Task; accuracy during the training phase was otherwise very high ($M = 0.90$, $SD = 0.01$).

This left a final analytic sample of 167 participants (24.6% male/man, 73.7% female/woman, and <2% nonbinary, gender fluid, or prefer not to answer). Participants were 19 years old on average ($M_{\text{age}} = 18.95$, $SD_{\text{age}} = 1.09$), all reported English as their primary language and were predominantly non-Hispanic White (88.6% non-Hispanic White, with <5% identifying as Black, East Asian, Middle Eastern, South Asian, Southeast Asian, Hispanic, or more than one race/ethnicity). Participants were also predominately right-handed (86.8% right-handed half the time or more, 6.6% equally right and left-handed, and 6.6% left-handed half the time or more).

Materials

Face stimuli were selected from a normed facial stimulus set, the Interdisciplinary Affective Science Laboratory Face Set (Interdisciplinary Affective Science Laboratory, 2006). Emotion expressions in this stimulus set were posed. Models were shown several sample photos and given instructions concerning which facial muscles to move for each stereotypical emotion expression. For example, for neutral expressions, models were told "let your face hang in a relaxed, natural position. Look straight into the camera." The specific models chosen for use in this study (three men, three women) were selected at random from this larger facial stimulus set. All six models appear to be White and roughly 18–25 years of age.

For the Face Perception Task, a series of morphed images were created for each model. Using Abrosoft Fantamorph software (https://www.fantamorph.com/), morphed images were created from two images of the same model displaying a posed, stereotypical fear expression and a posed, stereotypical anger expression. Still images were selected in 10% increments from 100% anger stereotype (0% fear stereotype) to 0% anger stereotype (100% fear stereotype) such that images represent physically equidistant increments from one expression to the other. Sample morphed facial stimuli can be seen in Figure 1.

Measures

Face Perception Task. The Face Perception Task involves two phases: a training phase (learning) and a testing phase (recognition). Participants completed the Face Perception Task once and were randomly assigned to complete both phases of the task with expressions from one of the six models. For each model, two faces drawn from the middle of the morph array were selected to be the target expressions: one labeled as the "more fearful" face (40% anger stereotype, 60% fear stereotype) and one labeled as the "more angry" face (60% anger stereotype, 40% fear stereotype).

Before the training phase (i.e., learning phase), participants were shown the two target expressions for a given model with the labels "more fearful" and "more angry," and they were told they would be asked to identify these two faces throughout the task. Because the goal of the training phase was to learn to identify these two particular expressions using the provided labels, not to assess how participants would apply the labels organically, it was clarified that they should be identifying these two particular faces by their labels and not reporting on their perception of how angry or fearful any given expression appeared to them. During the training phase, the two target expressions were each shown five times, for a total of 10 trials. Participants were asked to indicate on each trial whether the face shown was the more fearful face or the more angry face, and they were given immediate feedback after each trial on whether their response was correct or incorrect. In this way, participants learned to discriminate between the two target expressions during the training phase and to identify them by their provided label.

Participants then completed the testing phase (i.e., recognition phase) of the Face Perception Task. In the testing phase, participants were shown the morphed expressions for the same model they saw during the training phase, ranging from 20% anger stereotype (80% fear stereotype) to 80% anger stereotype (20% fear stereotype) in 10% increments. Faces were unlabeled and shown one at a time, each appearing five times in a random order, for a total of 35 trials. On each trial of the testing phase (i.e., for each image shown), participants selected one of three response options: This is the more fearful face from the training phase; this is the more angry face from the training phase; or this is neither of the expressions from the training phase. Because the goal of the testing (recognition) phase of the task is to test how well participants can remember the specific expressions they learned to identify during training (and not how they might apply emotion category labels to other expressions), it was again clarified that participants should be trying to recognize the target expressions from the training phase and not reporting on their own perception of how angry or fearful any given expression appeared to them. Participants did not receive feedback on their answers during the testing phase.

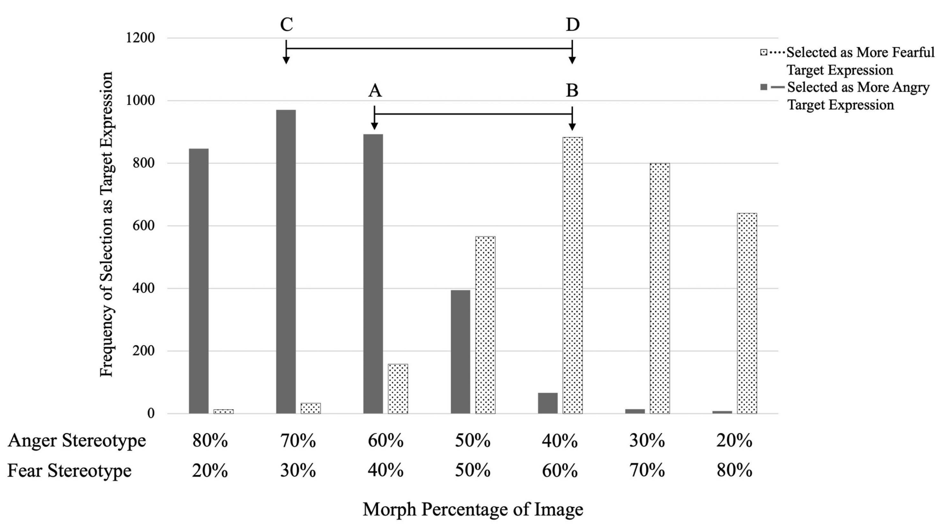
Dependent Measures. The primary dependent variable from this task is overall peak shift. Peak shift was calculated for each participant separately by first identifying two frequency distributions from their responses during the testing phase of the task: one giving the number of times they identified each morphed expression as the more angry face from training and one giving the number of times they identified each morphed expression as the more fearful face from training. Trials identified as neither of the target expressions during the testing phase were not assigned to either distribution. From these frequency distributions, the modal response for each target expression was identified (i.e., the morphed expressions identified most frequently as each of the target expressions by a given participant). In the event a participant chose more than one expression equally frequently, we took the average of all modes in the distribution. Next, the mode for the more fearful face response distribution was subtracted from the mode for the more angry face response distribution in anger morph percentage units to assess how far apart the peaks of the two response distributions were for a given participant. Finally, this distance between modes (i.e., distance between peaks) was compared to the actual morph percent difference between the target expressions (60% anger stereotype and 40% anger stereotype = 20) by subtracting this value (i.e., 20) from the distance between each participants' modal responses.

Although we did this at the participant level for analyses, we illustrate the process by providing raw frequency distributions across the full sample in Figure 2. In Figure 2, the target expressions are indicated by the Labels A and B, with the morph percentage difference between them given by the line between A and B in the figure. The modal responses for the two response distributions at the sample level are indicated by the Labels C and D in Figure 2, with the difference between them (i.e., distance between peaks of the two response distributions) given by the line between C and D. Peak shift is calculated by comparing the distance between C and D to the distance between A and B. Thus, peak shift is 0 when the recognized difference between the target expressions is accurate (i.e., when the distance between a participants' modal responses [C–D] equals the distance between the target expressions [B–A]). Larger numbers indicate the target expressions were misremembered as more perceptually distinct than they were in actuality (i.e., the distance between a participants' modal responses [C–D] is greater than the distance between the target expressions [B–A]). While Figure 2 illustrates this process at the sample level, peak shift was calculated per participant for analyses.

In addition to peak shift, we also calculated the extent to which the memory of each individual target expression was biased toward the stereotypical emotion expression indicated by its label. To do so, for each participant, we took the modal response for each of their two response distributions and calculated difference scores indicating the distance between each modal response and the relevant target expression (in morph percentage). Thus, when these measures are 0, that indicates the target expression was recognized accurately (i.e., A = C or B = D in Figure 2). Conversely, higher numbers indicate

Figure 2

Frequency of Each Expression Being Selected as the Target Expressions Across All Testing Phase Trials for All Participants in Study 1



Note. Bars represent the frequency with which each expression was recognized as each of the two target expressions across all testing (recognition) trials and participants in Study 1. A and B indicate the anticipated modal response if recognition is accurate/unbiased (i.e., the target expressions). C and D indicate the modal responses from the actual sample data (i.e., the faces most frequently identified as the more angry and the more fearful target expressions, respectively). Peak shift is calculated by subtracting the distance between A and B from the distance between C and D (peak shift = C–D–[A–B]) such that peak shift is 0 when recognition is unbiased and larger than 0 when the expressions most frequently identified as the target expressions are further apart than the target expressions were in reality. Biased perceptual memory for the individual target expressions were calculated by taking the difference between C and A for the more angry target expression (A–C), and by taking the difference between B and D for the more fearful target expression (B–D). Thus, these variables are 0 when recognition is unbiased and positive when the expression most frequently identified as a given target expression is closer to the ends of the distribution than the target expressions were in reality. This figure illustrates these values at the sample level, while all measures used in analyses were derived per participant.

the target expression was recognized as closer to the stereotypical (100%) emotion expression indicated by its label during the training phase (i.e., C is to the left of A or D is to the right of B in Figure 2). Again, while Figure 2 illustrates this calculation at the sample level, all variables were calculated per participant for analyses.

Questionnaires.

Emotion Essentialism. Emotion essentialism was measured via an unpublished 24-item self-report survey developed by one of the authors. This survey was developed to assess multiple facets of psychological essentialism that are commonly discussed when considering psychological essentialism for emotions (e.g., Berent et al., 2020; Lindquist et al., 2013) and social categories (Gelman, 2004; Haslam et al., 2000; Rangel & Keller, 2011), and item development for this questionnaire was guided by the content of existing survey materials from these sources. Forty items were originally developed for inclusion in the measure, but this was reduced to 24 items through feedback from experts on emotion theory and essentialism and informal pilot-testing that assessed the variability and reliability of ratings across items. The final 24-item survey measure captures eight facets of essentialism: Inherence (six items) measures the extent to which one believes emotional experiences form innate categories; Naturalness-Invented by People (two items) measures the extent to which one believes emotions were not invented by people; Naturalness-Natural Change (two items) measures the extent to which one believes emotions are brought on because of a natural change in the body; Naturalness-Subjective Interpretation (two items) measures the extent to which one believes emotions do not occur because of a subjective interpretation of a situation; Discreteness (two items) measures the extent to which one believes it is easy to perceive when someone is starting to feel a specific emotion; Informativeness (four items) measures the extent to which one believes they can readily predict someone's subjective experience by knowing what emotion word they label their experience with; Uniformity (four items) measures the extent to which one believes there is a commonality among experiences of the same emotion in different people; and finally, Necessary Features (two items) measures the extent to which one believes a specific neurological change is required to experience a given emotion and finally. The response scale for all items was a 5-point Likert scale ranging from *strongly agree* to *strongly disagree*. A copy of all questions in the Emotion Essentialism scale can be found in Appendix. After reverse scoring indicated items, responses to all items were averaged to create a single measure of emotion essentialism (Cronbach's $\alpha = .78$ in Study 1), such that higher values indicate more essentialist beliefs about emotion categories.

Demographic Information. Participants were asked to report on basic demographic information including gender, age, race/ethnicity, handedness, and if English was their primary language.

Procedure

The study was completed by participants in a single online session lasting approximately 20 min via Qualtrics and Inquisit Web. They first read an informed consent document in Qualtrics and indicated consent by clicking to the next page of the survey. Participants were then prompted to download Inquisit Web Player to complete the Face Perception Task online via Inquisit Web (Version: Inquisit 5). Following completion of the Face Perception Task, participants were instructed to return to Qualtrics to complete the questionnaires, after which they were debriefed and thanked for their participation.

Results

Peak Shift

A one-sample *t* test comparing the mean peak shift of the sample ($M = 13.85$, $SE = 0.98$) against 0 was significant, $t(166) = 14.11$, $p < .001$, $d = 1.01$, 95% CI [0.90, 1.28]; participants recognized the two target expressions to be, on average, 13.85 morph percentage points further apart than they were in reality. Exploratory analyses comparing peak shift for male and female models did not reveal any consistent differences in peak shift as a function of model gender (see Supplemental Figure S1).

Biased Perceptual Memory by Expression

Next, we examined whether overall peak shift was attributable to bias in the perceptual memory of one or both target expressions (i.e., the more angry expression or the more fearful expression). One-sample *t* tests confirmed that recognition was significantly biased for both the more angry expression ($M = 9.35$, $SE = 0.56$), $t(166) = 16.77$, $p < .001$, $d = 1.29$, 95% CI [1.09, 1.50], and the more fearful expression ($M = 4.50$, $SE = 0.77$), $t(166) = 5.89$, $p < .001$, $d = .46$, 95% CI [0.30, 0.61]; participants on average misidentified both the more fearful and the more angry target expressions as being closer to the stereotypical (100%) expression for the labeled emotion category than they were in reality. A paired-samples *t* test demonstrated this bias was significantly more pronounced for the more angry expression than the more fearful expression, $t(166) = 5.34$, $p < .001$, $d = .41$, 95% CI [0.25, 0.57]. Figure 3 depicts the average of participants' modal responses (i.e., expressions most frequently identified as each target expression by each participant) relative to each actual target expression.

Biased Perceptual Memory and Emotion Essentialism

Contrary to predictions, a zero-order Pearson's correlation failed to reveal any significant association between emotion essentialism and peak shift, $r(165) = -.11$, $p = .151$.

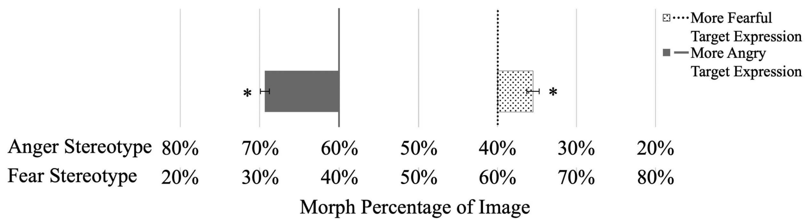
Discussion

Consistent with predictions, we observed evidence for categorical processing biases for ambiguous facial expressions of emotion in Study 1. Participants misremembered the two target expressions, which were labeled as belonging to two different emotion categories, as significantly more perceptually distinct from one another than they were in reality. Moreover, both target expressions (i.e., the more angry and more fearful target expressions) were recognized as being more perceptually similar to the stereotypical facial expressions for their labeled emotion category than they were in reality.

Contrary to predictions, we failed to find a significant association between peak shift and emotion essentialism in Study 1. Of note, however, the Face Perception Task in Study 1 involved the explicit use of emotion category labels during the training and testing phases. One possibility is that the application of these labels strongly promotes categorical processing biases among all perceivers because they explicitly reference categorical knowledge about emotions. If so, the use of emotion category labels during the task may be creating something like a ceiling effect such that other individual differences that might otherwise contribute to the extent of categorical processing

Figure 3

Expressions Most Frequently Recognized as the Target Expressions on Average in Study 1



Note. The expression most frequently recognized as each target expression by each participant was identified. Bars represent the average modal response across participants for each target expression. The gray and dotted lines represent the more angry and more fearful target expressions, respectively. Error bars represent ±1 *SE*. *SE* = standard error.
* indicates significance at the $p < .05$ level.

biases are overshadowed. Study 2 will examine the extent to which individuals exhibit categorical processing biases for ambiguous facial expressions of emotion when the ambiguous expressions are not labeled with emotion category words, and whether individual differences in the ways people reason about and experience emotions play a larger role in the extent to which people exhibit categorical processing biases for emotion expressions in such contexts.

Study 2

Previous research suggests that the presence of emotion labels alone may be sufficient to influence perceptual memory of ambiguous emotion expressions. In a set of studies, J. B. Halberstadt and Niedenthal (2001) found that participants misremembered an ambiguous facial expression as looking happier than it was in reality if they had provided an explanation for why the original expression was an expression of happiness (vs. an explanation of why it was an expression of anger). That is, they were asked during encoding of the expression to envision a scenario that would cause an experience of a given emotion and result in the presented expression. More important to the present investigation, however, they also observed biased perceptual memory of expressions when they merely provided emotion category labels for the original expressions without asking participants to generate an explanation for the expression. That is, simply labeling an ambiguous face as angry (vs. happy) during encoding led participants to misremember the face as looking more angry than it was in reality. However, the effect of mere labeling was less consistent than the effect of generating an explanation across studies and stimuli and was not always significant. Nevertheless, findings suggest that the mere availability of an emotion concept at encoding may be sufficient to produce memory biases. Previous work on categorical processing biases for faces have also demonstrated that the mere inclusion of category labels during initial exposure to stimuli produced categorical processing biases that otherwise were not evident for chimpanzee expressions (Fugate et al., 2010) and for unfamiliar faces (Kikutani et al., 2008). Taken together, these findings suggest that merely labeling the target expressions during the Face Perception Task as more angry and more fearful may be sufficient to produce categorical processing biases in recognition like we observed in Study 1.

To test this, in Study 2, we used the Face Perception Task as described in Study 1 with the addition of a label condition in which participants were randomly assigned to complete the task in a labeled (replicating Study 1) or an unlabeled condition (where faces were identified as “Face A” and “Face B”). We predicted that findings from Study 1 would be replicated, in that we would find evidence of significant categorical processing biases in the recognition of ambiguous emotion expressions in labeled conditions. We also predicted that we would observe significantly less biased recognition of target expressions in the unlabeled conditions than the labeled conditions, suggesting that the presence of emotion category words during the perception and recognition of the ambiguous expressions influences the extent to which they are misremembered as more like stereotypical emotion expressions (and less like one another).

Moreover, although we failed to find evidence that individual differences in emotion essentialism were related to categorical processing biases for emotion expressions in Study 1, it is possible that how an individual reasons about emotion categories will be a more powerful predictor of categorical processing biases in contexts where all participants are not explicitly directed to use category labels to guide discrimination learning. We explore this possibility in Study 2, predicting that emotion essentialism would be a stronger predicter of peak shift in the unlabeled conditions than in the labeled conditions of Study 2.

In addition to emotion essentialism, in Study 2, we also added two additional individual difference measures related to the ways a person identifies and labels their emotions: emotional clarity, an individual difference in how unambiguously an individual labels, identifies and describes their own emotional experiences (Boden & Thompson, 2017), and alexithymia, the extent to which an individual has difficulty identifying or describing their own emotional experiences (Grynberg et al., 2012; Nemiah et al., 1976). Alexithymia is elevated in a range of psychiatric and neurodevelopmental conditions associated with affective psychopathology, including depression (Li et al., 2015), anxiety (D. A. Preece et al., 2024), and autism spectrum disorder (Cook et al., 2013). Elevated alexithymia has also consistently been related to deficits in emotional processing, including difficulty in emotion regulation and the recognition of others’ facial expressions (Di Tella et al., 2020; Grynberg et al., 2012); specifically decreased abilities to detect and rate the intensity of facial expressions (Prkachin et al., 2009). Of interest to the present investigation, the development of certain facets of alexithymia have been linked to language deficits in childhood (Lee et al., 2024), suggesting that language-related difficulties in learning and applying emotion category words may be particularly relevant to its etiology.

Given that both emotional clarity and alexithymia represent individual differences in the ways in which individuals apply emotion category knowledge to make sense of their own and others' emotions (Boden & Thompson, 2017; Grynberg et al., 2012), we wanted to explore whether these traits were associated with differences in the extent to which individuals exhibited categorical processing biases in emotion recognition as well. In particular, given that alexithymia is consistently associated with deficits in emotion perception and processing (e.g., Cook et al., 2013; Di Tella et al., 2020; D. A. Preece et al., 2024), it is possible that greater alexithymia may be related to more pronounced biases in the memory of facial expressions of emotion. Conversely, to the extent that a lack of knowledge or use of emotion categories underlies difficulty labeling emotional experiences and expressions in alexithymia, we would anticipate decreased categorical processing biases in our paradigm due to a reduced reliance on categories to guide learning and recognition. Critically, a previous study found that alexithymia was related to poorer performance in an emotion matching task when matching two facial expressions, but not when matching a facial expression to an emotion word (Nook et al., 2015), suggesting alexithymia may be related to deficits in spontaneous application of emotion categories (as in our unlabeled conditions) while sparing use of emotion category information when it is made available (as in our labeled conditions).

Method

Transparency and Openness

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study. Data from all studies in this article are publicly available in the Open Science Framework repository (https://osf.io/ydwf4/). The methods, analysis, and predicted results for Study 2 (conducted in 2022) were preregistered on the Open Science Framework (https://osf.io/q6f8v/).

Participants

Power analyses conducted in G*Power (V3.1) determined that a sample size of $N = 395$ would afford appropriate statistical power (80%) to detect a small effect size (Cohen's $f^2 = .02$) at $\alpha = .05$ for multiple regression analyses assessing an interaction term between a continuous individual difference measure and label condition. We aimed to collect participants until $N = 400$ or until we reached the end of the academic semester in which data collection began, whichever occurred second. Supplemental analyses with only the first $N = 400$ participants (to match the target N in our preregistration) revealed the same pattern of results and significance as analyses with the full sample (see Supplemental Tables S2–6).

Five hundred ninety students from the University of New Hampshire consented to participate in an online study for course credit. Of the 590 participants who consented, only 524 participants completed the study in full. Of the 524 participants who completed the study, 67 participants (12.8%) were removed from analyses for obtaining less than 50% accuracy in the training portion of the Face Perception Task, consistent with our preregistration; accuracy during the training phase was otherwise very high ($M = 0.88$, $SE = 0.01$). In addition, in response to increasing concerns about online data quality, nine participants (1.7%) were removed from analyses a priori for having more than 95% of their responses be identical on the questionnaires, and four participants (<1%) were removed as outliers based on their performance in the Face Perception task (i.e., peak shift more than 3 SD from mean).

This left a final analytic sample of 444 participants (242 participants [54.5%] in labeled condition; 78.4% female/woman, 18.2% male/man, with <2.5% identifying as transgender, nonbinary, or prefer not to answer). Participants had a mean age of 19 years ($M_{\text{age}} = 19.04$, $SD_{\text{age}} = 2.95$), 99% identified English as their primary language, and they were predominantly non-Hispanic White (86.0% non-Hispanic White, 5.4% more than one race/ethnicity, <2.5% identifying as Black, Hispanic, South Asian, Middle Eastern, Indigenous, East Asian, or prefer not to answer). Participants were also predominantly right-handed (90.5% right-handed half the time or more).

Materials and Procedure

The materials and procedure for Study 2 utilized the same Face Perception Task as described in Study 1, with the following exceptions. First, in Study 2, participants were randomly assigned to complete the Face Perception Task in one of two label conditions. In the labeled condition, the two target expressions (40/60% blends of anger and fear stereotype expressions) were labeled as the "more fearful" face and the "more angry" face, as in Study 1. In the unlabeled condition, the same target expressions were instead labeled as "Face A" and "Face B." Second, participants completed the same emotional essentialism (Cronbach's $\alpha = .74$), and demographic questionnaires as used in Study 1 as well as measures of emotional clarity (Salovey et al., 1995) and alexithymia (Bagby et al., 1994). They also completed a questionnaire probing their experience of the Face Perception Task.

Face Perception Task Questionnaire. Immediately following the Face Perception Task, we asked participants specific questions regarding their personal process for completing the task. These questions were meant to shed light on potential mechanisms underlying differences in performance on the Face Perception Task. Participants were asked to identify if they used verbal labels of any kind (e.g., words or phrases) to help them learn or recognize each target expression, and if so, they were asked to provide the verbal labels they used

> Did you use a verbal label of any kind (e.g., word or phrase) to help you learn and/or recall the target expression provided to you as the More Angry ["More Fearful"/"Face A"/"Face B"] face? If so, please provide the label you used in the space below. If not, please leave blank.

Participants received this question twice, once for the more angry expression (labeled "more angry" or "Face A") and once for the more fearful expression (labeled "more fearful" or "Face B").

In addition, participants were shown each target expression individually and asked, "What emotion word would you use to label this expression?" They freely generated a word for this open-ended question. Participants then selected an emotion word from a set of provided words that they believed best matched each target expression ("Which of these emotion words would you use to label this expression?," response options: happy, sad, angry, disgust, and fearful). These items captured the extent to which participants are naturally using emotion category knowledge when trying to learn and remember ambiguous expressions of emotion, as well as whether participants are perceiving the ambiguous expressions as belonging to one of the discrete emotion categories used in their creation (e.g., whether an angry-fear morph is perceived as angry, fearful, or some other emotion). Note that the order of items was consistent for all participants in the Face Perception Task Questionnaire due to a failure to randomize: for each subset of questions, participants were always asked about the more angry expression first followed by the more fearful expression. Any differences in the pattern of results across expressions on this questionnaire should be interpreted cautiously in light of potential order effects.

Emotional Clarity. Emotional clarity was measured via the 11-item self-report subscale, Clarity of Feelings, from the Trait Meta-Mood Scale (Salovey et al., 1995). Emotional clarity measures the extent to which individuals unambiguously label, identify and define their own emotional experiences, and was shown to have good reliability in our study (Cronbach's $\alpha = .85$; example item: "I am rarely confused about how I feel"). The response scale for all items is a 5-point Likert scale ranging from *strongly agree* to *strongly disagree*.

Alexithymia. Alexithymia was measured using two of the subscales from the Toronto Alexithymia Scale-20 (Bagby et al., 1994): the Difficulty Identifying Feelings subscale (seven items), which measures the extent to which individuals have the capacity to identify emotions and distinguish between feelings (example item: "I often don't know why I am angry"); and the Difficulty Describing Feelings subscale (five items), which measures the extent to which individuals are able to communicate about their own emotional experiences (example item: "People tell me to describe my feelings more"). The response scale for all items is a 5-point Likert scale ranging from *strongly agree* to *strongly disagree*. We averaged ratings across these 12 items to create a single index of alexithymia, and this index variable had good reliability in our study (Cronbach's $\alpha = .89$). The Externally Oriented Thinking subscale of the Toronto Alexithymia Scale-20 (eight items) was not included owing to it having poor internal consistency in prior studies (D. Preece et al., 2018).

Results

Figure 4 depicts the average of participants' modal responses for each target expression (i.e., the expressions most frequently identified as each target expression by each participant) relative to the actual target expressions participants were trying to recognize.

Peak Shift

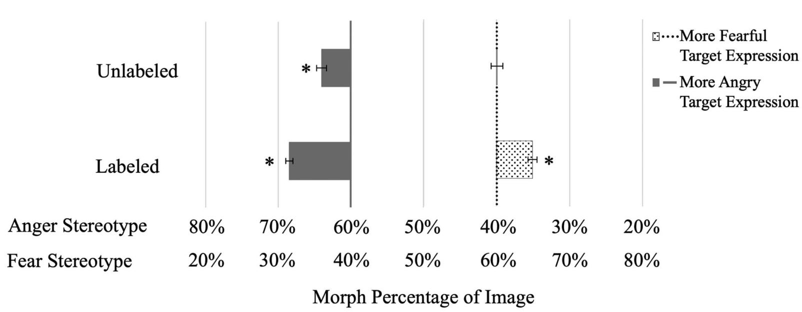
A one-sample t test comparing the mean peak shift of the sample across labeled and unlabeled conditions ($M = 9.17$, $SE = 0.68$) against 0 was significant, $t(443) = 13.47$, $p < .001$, $d = .639$, 95% CI [0.54, 0.74]. As predicted, an independent samples t test demonstrated a greater peak shift for the label condition ($M = 13.41$, $SE = 0.82$) than for the unlabeled condition ($M = 4.08$, $SE = 1.02$), $t(442) = 7.21$, $p < .001$, $d = .69$, 95%, CI [0.49, 0.88]. Separate one-sample t tests indicated that there was significant peak shift in both the labeled, $t(241) = 16.29$, $p < .001$, $d = 1.05$, 95%, CI [0.89, 1.20] and unlabeled conditions, $t(201) = 4.02$, $p < .001$, $d = .28$, 95% CI [0.14, 0.42], individually. See Figure 4. Exploratory analyses comparing peak shift for male and female models did not reveal any consistent differences in peak shift as a function of model gender across label conditions (see Supplemental Figure S2).

Biased Perceptual Memory by Expression

We next assessed if the overall peak shift effect was attributable to biases in the perceptual memory for each of the individual target expressions. A 2 (Label Condition: Labeled vs. Unlabeled) × 2 (Emotion Expression: More Angry and More Fearful) mixed model analysis of variance, with label condition as a between-subject factor and emotion expression as a within-subject factor, demonstrated a significant main effect of emotion expression, $F(1, 442) = 33.92$, $p < .001$, $\eta^2 = .071$, such that recognition was significantly more biased for the more angry expression ($M = 6.46$, $SE = 0.43$) than the

Figure 4

Expressions Most Frequently Recognized as the Target Expressions on Average in Study 2



Note. The expression most frequently recognized as each target expression by each participant was identified. Bars represent the average modal response across participants for each target expression. The gray and dotted lines represent the more angry and more fearful target expressions, respectively. Error bars represent ± 1 *SE*. *SE* = standard error.
* indicates significance at the $p < .05$ level.

more fearful expression (M = 2.70, SE = 0.51). In addition, there was a significant main effect for label condition, $F(1, 442) = 51.92$, $p < .001$, $\eta^2 = .105$, in that recognition was significantly more biased in the labeled condition (more angry expression: M = 8.50, SE = 0.50; more fearful expression: M = 4.91, SE = 0.61), compared to in the unlabeled condition, more angry expression (Face A): M = 4.03, SE = 0.69; more fearful expression (Face B): M = 0.06, SE = 0.80. See Figure 4. However, there was no significant interaction between label condition and emotion expression, $F(1, 442) = .086$, $p = .769$, $\eta^2 < .001$, suggesting that the presence (vs. absence) of emotion word labels impacted recognition of both expressions to a similar extent.

To examine whether recognition for each target expression was significantly biased (i.e., differed from zero), a series of one-sample t tests were conducted. These analyses revealed that recognition for each target expression was significantly biased toward the stereotyped face in all conditions (ps < .001; ds > .41), except for the more fearful expression (Face B) in the unlabeled condition, $t(201) = .073$, $p = .471$, $d = .005$, 95%, CI [−0.133, 0.143].

Individual Differences in Emotion Beliefs and Knowledge

As predicted, Pearson's correlations within each label condition separately revealed a significant positive association between peak shift and emotion essentialism in the unlabeled condition, $r(200) = .218$, $p = .002$, but no association in the labeled condition, $r(240) = -0.024$, $p = .713$. A multiple regression analysis predicting peak shift from emotion essentialism, label condition, and their interaction revealed a significant interaction effect, confirming the association between emotion essentialism and peak shift differed significantly across conditions, $b = -10.62$, $SE_b = 3.72$, $\beta = -1.28$, $t(440) = -2.86$, $p = .004$, $sr^2 = .016$; sr^2 is an abbreviation for the squared semipartial correlation, an effect size estimate interpretable as the proportion of variance in the outcome variable uniquely attributable to a given predictor variable when controlling for all other variables in a multivariable regression analysis. See Figure 5 for a simple slopes depiction of this interaction.

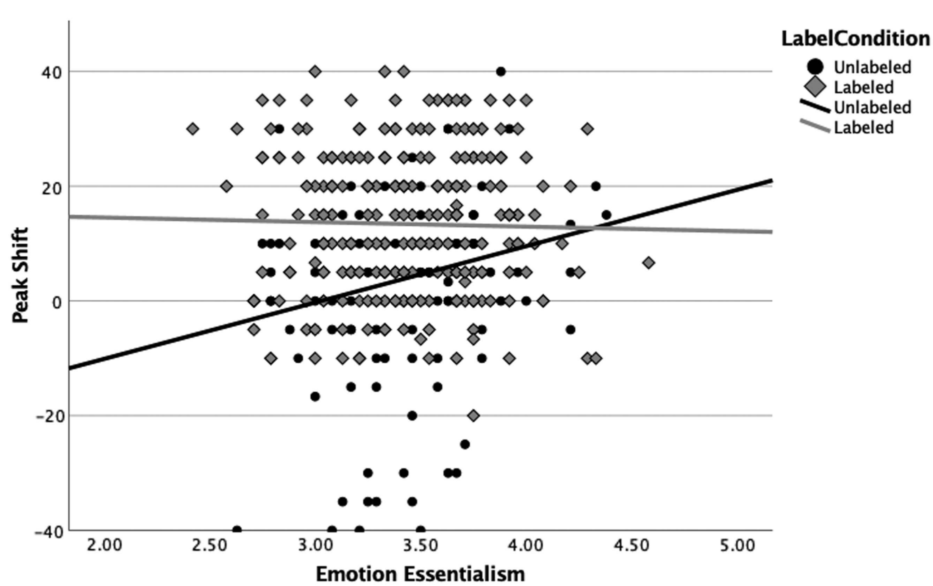
There were no significant associations between peak shift and alexithymia (labeled condition: $r(240) = .093$, $p = .150$; unlabeled condition: $r(200) = -.026$, $p = .715$) or between peak shift and emotional clarity (labeled condition: $r(240) = -.004$, $p = .946$; unlabeled condition: $r(200) = .035$, $p = .618$), and no significant interactions with label condition in multivariable regression analyses (alexithymia: $b = .16$, $SE_b = .14$, $\beta = .208$, $t(440) = 1.20$, $p = .233$, $sr^2 = .003$; emotional clarity: $b = -.08$, $SE_b = .18$, $\beta = -.101$, $t(440) = -.43$, $p = .668$, $sr^2 = .001$).

Face Perception Task Questionnaire

Reported Use of a Verbal Label. Of the 444 total participants in analyses, only 34.5% (153 participants) reported using a verbal label to help them discriminate and/or recognize the target expressions during the Face Perception Task. Moreover, of the 153 participants to report utilizing a verbal label during the Face Perception Task, only one third (51 participants) reported using a label that contained an emotion word (e.g., angry, confused, sad). The majority of participants who reported using a label (66.7%, 102 participants) reported labels that used terms more associated with facial features or characteristics of the model (e.g., eyebrows, wrinkles, attitude). Thus, the reported conscious use of labels, particularly ones containing emotion words, were rare overall.

Interestingly, a chi-square test indicated that a significantly larger percentage of participants in the unlabeled condition (45.0%; 91/202) reported using a verbal label than did participants in the labeled condition (25.6%; 62/242), $\chi^2(1, 444) = 18.40$, $p < .001$. We believe this is because labels were already provided to participants in the labeled condition and so participants did not report using labels of their own.

Figure 5
Simple Slopes Depiction of Interaction Between Emotion Essentialism and Peak Shift in Study 2



Peak Shift by Reported Verbal Label. As an exploratory analysis, an independent t test was conducted to examine if peak shift differed across participants who did ($N = 91$) versus did not ($N = 111$) report using labels in the unlabeled condition ($N = 202$). Individuals who reported using a verbal label exhibited more biased recognition ($M = 6.17$, $SE = 1.54$) than those who did not report using a verbal label ($M = 2.37$, $SE = 1.34$) in the unlabeled condition, but this difference failed to reach significance, $t(200) = 1.87$, $p = .063$, $d = .27$, 95% CI [−0.01, 0.54].

Identifying the Emotion of Target Expressions.

Multiple Choice. For the more fearful target expression, 74.8% (332 participants) identified the expression to be "fear," and a chi-square test found that a significantly larger percentage of participants in the labeled condition (81.8%; 198/242) identified it as fear than did participants in the unlabeled condition (66.3%; 134/202), χ^2 (1, 444) = 17.60, $p < .001$ (see Table 1). For the more angry target expression, 60.1% (267 participants) identified the expression to be "angry," and a chi-square test found no significant difference in the percentage of participants in the labeled condition (58.7%; 142/242) versus the unlabeled condition (61.9%; 125/202) who identified it as angry, χ^2 (1, 444) = 5.57, $p = .234$. For a full breakdown of response proportions for each target expression, see Supplemental Table S7.

Open Ended Responses. Results for the open-ended emotion perception questions largely paralleled those for the multiple-choice emotion perception questions. For the questions asking participants to free label the emotion depicted by each target expression, we coded responses into categories based on semantic similarity (e.g., a fear category comprised of fearful, afraid, scared, frightened, etc.). For a list of all words coded into each category see Supplemental Table S8. For the more angry target expression, 55.4% (246 participants) identified the expression to be anger, and a chi-square test found that a significantly larger percentage of participants in the labeled condition (63.6%; 154/242) identified it as anger than did participants in the unlabeled condition (45.5%; 92/202), χ^2 (1, 444) = 14.59, $p < .001$. For the more fearful target expression, 53.4% (237 participants) identified the expression to be fear, and a chi-square test found that a significantly larger percentage of participants in the labeled condition (66.5%; 161/242) identified it as fear than did participants in the unlabeled condition (37.6%; 76/202), χ^2 (1, 444) = 36.96, $p < .001$ (see Table 1). For a full breakdown of response proportions for each target expression, see Supplemental Table S9.

Discussion

Consistent with predictions, we observed categorical processing biases for ambiguous facial expressions of emotion, and these biases were significantly more pronounced in the labeled condition (i.e., emotion word labels provided) than the unlabeled condition (i.e., absence of emotion word labels). Participants misremembered the two target expressions as significantly more perceptually distinct from one another than they were in reality, and this effect was significantly more pronounced in the presence of emotion category words than in their absence. Though it is important to note that significant peak shift was still observed in the unlabeled condition, suggesting that the presence of emotion word labels increased reliance on categorical processing during recognition of ambiguous facial expressions of emotion but was not necessary to produce it. Furthermore, in the labeled condition, both target expressions (i.e., the more angry and more fearful target expressions) were misremembered as more perceptually similar to the stereotypical facial expression for their labeled category than they were in reality. However, in the unlabeled condition, this effect was only significant for the more angry expression ("Face A") and not for the more fearful expression ("Face B"), for which participants exhibited more accurate recognition performance. This pattern of results demonstrates that the presence of emotion category words during the perception and recognition of ambiguous expressions of emotion increases the extent to which participants demonstrate categorical processing biases in recognition.

To provide insight into potential mechanisms underlying recognition performance, participants reported on their experiences in the Face Perception Task and their perception of the target expressions. Participants were generally not perceiving the ambiguous target expressions as consistently belonging to a discrete emotion category (i.e., consensus on emotion category labels for the target expressions were between 55% and 63% for freely generated responses and were between 60% and 75% for multiple choice); this is consistent with the expectation that they are ambiguous expressions and do not represent stereotypical expressions for any given emotion category. However, participants in the labeled condition, who were presented with emotion category labels for the expressions during the learning and recognition phases of the task, were significantly more likely to identify the ambiguous target expressions as expressing the emotion given by their label than were participants in the unlabeled condition. This is consistent with the idea that providing emotion labels increases reliance on categorical knowledge about emotions during perception and/or recognition. As a result, participants are more likely to view an ambiguous expression labeled with an emotion word as an exemplar of that emotion category, and thus to misremember it as looking more like the stereotype for that category.

Critically, and consistent with predictions, results also indicated a significant positive association between peak shift and emotion essentialism in the unlabeled condition. Individual differences in emotion essentialism were again not related to categorical processing biases in the recognition of ambiguous facial expressions of emotion when these expressions were labeled with emotion category words and categorical processing biases were fairly robust. However, in the absence of explicit category labels, where the extent of categorical

Table 1
Percent Label Agreement for Ambiguous Target Expressions in Study 2

	More angry target expression		More fearful target expression	
Question style	Labeled ($N = 242$)	Unlabeled ($N = 202$)	Labeled ($N = 242$)	Unlabeled ($N = 202$)
Multiple choice	58.7% (142)	61.9% (125)	81.8% (198)	66.3% (134)
Open-ended	63.6% (154)	45.5% (92)	66.5% (161)	37.6% (76)

processing biases was less pronounced and more varied, individuals' beliefs about the nature of emotion categories was related to the extent to which they exhibited such biases in emotion expression recognition. Specifically, in the absence of explicit emotion labels, individuals with more essentialist beliefs about emotion categories—perceiving them as more discrete, biologically defined, and immutable—exhibited more pronounced categorical processing biases for ambiguous emotion expressions. Speculatively, this is because believing emotions are discrete, uniform, essentialized entities aligns with the belief in stereotypical emotional experiences, prompting people to automatically employ biased categorization processes during emotion perception and recognition. This aligns with previous findings surrounding psychological essentialism and stereotypes that has largely been considered through the lens of social categories (e.g., Bastain & Haslam, 2006; Morgenroth et al., 2021; Roth et al., 2023). Whether the essentialized categories are social groups or emotions, believing that categories are innate, natural, and uniform may impact how individuals process exemplars of those categories.

In addition, we failed to find any significant associations between peak shift and participants' self-reported beliefs about the ways they use emotion categories to label and define their own emotional experiences (i.e., alexithymia and emotional clarity). This pattern of results suggests that beliefs about the *nature* of emotion categories may be more strongly related to the use of categorical processing in emotion recognition than beliefs about how one *uses* emotion category labels in identifying their own emotions. That is, emotional clarity and alexithymia focus on self-perceived ability to apply category labels to one's own and others' emotions, while emotion essentialism is more concerned with the nature of the categories themselves (i.e., whether they are more discrete or fuzzy).

General Discussion

Across two studies, we found evidence of categorical processing biases in the recognition of ambiguous facial expressions of emotion. Participants misremembered two ambiguous target expressions (drawn from the middle of an array of morphed emotion expressions) as significantly more perceptually distinct from one another than they were in reality, and this bias was more pronounced when the target expressions were accompanied by emotion category labels. In addition, ambiguous target expressions labeled with an emotion word (e.g., more angry and more fearful target expressions) were recognized as being more perceptually similar to the stereotypical facial expressions for their labeled emotion category than they were in reality. These findings highlight the role of categorical processing in the recognition of ambiguous facial expressions of emotion.

Findings from Study 2, where we manipulated the presence (or absence) of emotion category labels during the Face Perception Task, demonstrated that the presence of emotion category labels significantly increased categorical processing biases. Given the vast literature on the role of language in emotion perception (e.g., Barker et al., 2020; Champoux-Larsson & Nook, 2024; Lindquist et al., 2006, 2014; Lindquist & Gendron, 2013) as well as work on categorical processing biases in emotion perception (Calder et al., 1996; Etcoff & Magee, 1992; Fugate et al., 2010; Roberson et al., 2007, 2010), we posit that this bias may be being introduced during the learning phase of the task, at stimulus encoding. That is, when an ambiguous facial expression is paired with an emotion category label, individuals may rely on the category label to guide perception or mental inference regarding what is being conveyed by the ambiguous expression such that they are biased toward the stereotype for the provided emotion category. Conversely, when an emotion label is not provided, individuals may more freely interpret the ambiguous expressions and so are less consistently biased toward a specific emotion category stereotype. For example, when unlabeled, some individuals may apply emotion category knowledge to aid in perception while others may not, and individuals who do apply emotion category knowledge may interpret the expression as an exemplar of different emotion categories (e.g., the same ambiguous expression may be interpreted using anger conceptual knowledge for one perceiver but disgust conceptual knowledge for another).

Critically, however, because we used emotion category labels in both the learning and recognition phases of our studies, we are unable to determine whether the emotion labels are increasing recognition biases during the initial perception of the ambiguous expressions or when participants are trying to remember the expressions later on. That is, it is possible participants are forming accurate representations of the two target expressions during the learning phase of the task, regardless of whether labels are present or not, but then the memory or reconstruction of the target expressions are being biased by the presence of the emotion category labels during recognition. Future studies should attempt to disentangle these possible sources of bias by manipulating whether emotion category labels are provided in only one, both, or neither of the two phases of the face perception task. For instance, ambiguous target expressions could be unlabeled in the learning/perception phase of the task, and then labeled with emotion category words in the recognition phase of the task. To the extent that increased recognition biases occurred in such a condition, this would suggest that the presence of emotion category labels during the memory stage of the task alone is sufficient to induce categorical processing biases in recognition.

Our results also showed that, in the less constrained context of the unlabeled condition, individuals' beliefs about the nature of emotion categories predicted the extent to which they would exhibit categorical processing biases. Specifically, individuals who reported believing that emotion categories are more distinct, innate, and immutable (i.e., more essentialist beliefs) exhibited more pronounced categorical processing biases during the recognition of ambiguous emotion expressions, though only when the expressions were not labeled with emotion category words. This finding builds on prior work which found that individuals with more essentialist beliefs about emotion categories retrospectively describe their own emotions as more differentiated (Lindquist et al., 2013). Specifically, our results suggest that individuals with more essentialist beliefs about emotion also have more differentiated memory for the emotional expressions of others.

Interestingly, no similar relationships were observed for individual differences in emotional clarity or alexithymia, traits related to one's perceived ability to identify and label their own emotional experiences. These results suggest that beliefs about the nature of emotion categories (e.g., whether they have clear-cut or fuzzy boundaries) may be more predictive of reliance on these categorical boundaries during emotion perception and recognition than one's perceived ability to use category labels during emotion experience. These null associations are especially interesting for alexithymia given that it has consistently been demonstrated to be associated with deficits in the processing of facial expressions of emotion (Grynberg et al., 2012; Hobson et al., 2019), particularly in the absence of emotion labels (Nook et al., 2015). However, our results suggest that these deficits do not extend to the use of categorical processing when remembering facial expressions of emotion. This may suggest that alexithymia selectively impacts the perception or encoding of facial expressions of emotion (Grynberg et al., 2012) but spares processes related to remembering/reconstructing observed expressions. However, caution is warranted in generalizing findings from a single study, particularly null effects; future research should seek to replicate this pattern of results and further probe how individual differences in emotion knowledge, expertise, and experience may relate to categorical processing in emotion recognition.

Interestingly, and unexpectedly, we observed that recognition biases were more pronounced for the more angry target expression than the more fearful target expression in both studies. Future work should continue to explore whether differences in the extent of recognition biases persist across different emotion categories as well as possible mechanisms underlying any such differences. For example, although both anger and fear expressions are normatively high arousal and negatively valenced and are frequently posited to signify threat to perceivers, past work has demonstrated that fearful expressions may actually facilitate approach-related behavior and thus may represent an affiliative stimulus (Marsh et al., 2005). This suggests participants may exhibit differences in motivational stance toward exemplars of these two categories throughout the task. In addition, past work suggests that angry expressions grab and hold people's attention (Becker et al., 2019), so it is possible that attentional differences toward certain expressions may contribute to differences in the extent to which categorical processing is applied during the perception and recognition of facial expressions. Furthermore, past work has demonstrated that anger expressions are encountered more frequently in daily life than fear expressions, and that increased exposure to anger expressions was related to faster and more accurate identification of anger (Calvo et al., 2014). As such, participants may have a more accessible or salient representation of a prototypical facial expression for anger than for fear. Finally, it is possible that participants have differing biases when it comes to labeling someone as having certain emotions. For example, people may not want to call someone angry unless they are fairly certain they are in fact angry, which may lead them to remember facial expressions labeled as angry as more extreme than they were in reality.

Our results are consistent with constructionist theories of emotion and add to the literature on the role of language and conceptual knowledge in emotion perception and experience (see, e.g., Betz et al., 2019; Doyle et al., 2021; Gendron et al., 2012; J. B. Halberstadt & Niedenthal, 2001; Lindquist et al., 2014; Lindquist & Gendron, 2013; Nook et al., 2015), but extend this work by focusing on the role of language in promoting categorical processing biases and in manifestations of these biases in memory for facial expressions of emotion more specifically. We found that the mere presence of emotion words provides an important context that promotes categorical processing in the recognition of previously learned ambiguous emotion expressions. Moreover, beliefs about the nature of emotion categories, a type of conceptual knowledge about emotions, were related to the extent to which one exhibits categorical processing biases when recognizing ambiguous emotion expressions, at least in the absence of explicit category labels.

Results also conceptually replicate past work demonstrating categorical processing in emotion perception (Calder et al., 1996; Etcoff & Magee, 1992; Fugate et al., 2010; Roberson et al., 2007, 2010), extending it to memory for emotion expressions. Taken together, our study and this body of literature suggest that the ways in which people perceive and remember facial expressions is not specific to emotion; rather, perception and memory for facial expressions of emotion appear subject to the same categorical processing biases that are evident across a wide range of domains and stimuli beyond emotion (e.g., social categories, Anthony et al., 1992; color categories, Gibson et al., 2017; and speech-related categories, Harnad, 1987).

In addition, our results may help to inform future work on the role of emotion naming in emotion regulation. Nook et al. (2021) demonstrated that actively categorizing emotional experiences with emotion labels impacts individuals' ability to effectively regulate their feelings in the moment. Emotion categories seemingly "crystalize" an individual's affective state, making it more difficult to modify or regulate their feelings. Our results could be interpreted in a similar fashion, such that providing emotion labels "crystalized" an expected or prototypical facial expression in participants' minds, biasing memory toward this crystalized mental representation. Future work could examine whether actively categorizing an emotional experience leads not only to poorer emotion regulation in the moment, but also more biased memory for the intensity of the experience later on.

Constraints on Generality

Our findings should be interpreted in the context of several critical limitations. First, the models used to create the series of morphed images are fairly homogenous in appearance: All appear to be White, between 18 and 25 years of age, and of binary gender presentation. These characteristics, while fairly representative of the participants in our sample, do not generalize well to the broader population. As such, they fail to account for prevalent stereotypes surrounding emotion perception and expression concerning how individuals of certain ages (Freudenberg et al., 2020), races/ethnicities (Bijlstra et al., 2014), or genders (Plant et al., 2000) are expected to experience and express their emotions. Future studies should expand on our findings using a more diverse sample as well as more diverse models posing the facial expressions of emotion. Such work could also examine how characteristics of the perceiver interact with characteristics of the emotion expressor to shape the perception and recognition of ambiguous facial expressions of emotion.

References

Al-Rasheed, A. S. (2015). Categorical perception of color: Evidence from secondary category boundary. *Psychology Research and Behavior Management*, *8*, 273–285. https://doi.org/10.2147/PRBM.S78348

Anthony, T., Copper, C., & Mullen, B. (1992). Cross-racial facial identification: A social cognitive integration. *Personality and Social Psychology Bulletin*, *18*(3), 296–301. https://doi.org/10.1177/0146167292183005

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*(1), 23–32. https://doi.org/10.1016/0022-3999(94)90005-1

Barker, M. S., Bidstrup, E. M., Robinson, G. A., & Nelson, N. L. (2020). "Grumpy" or "furious"? arousal of emotion labels influences judgments of facial expressions. *PLOS ONE*, *15*(7), Article e0235390. https://doi.org/10.1371/journal.pone.0235390

Barrett, L. F. (2006a). Emotions as natural kinds? *Perspectives on Psychological Science*, *1*(1), 28–58. https://doi.org/10.1111/j.1745-6916.2006.00003.x

Barrett, L. F. (2006b). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, *10*(1), 20–46. https://doi.org/10.1207/s15327957pspr1001_2

Barrett, L. F. (2012). Emotions are real. *Emotion*, *12*(3), 413–429. https://doi.org/10.1037/a0027555

Barrett, L. F. (2017a). Categories and their role in the science of emotion. *Psychological Inquiry*, *28*(1), 20–26. https://doi.org/10.1080/1047840X.2017.1261581

Barrett, L. F. (2017b). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(11), 1–23. https://doi.org/10.1093/scan/nsx060

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1–68. https://doi.org/10.1177/1529100619832930

Barrett, L. F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology*, *41*, 167–218. https://doi.org/10.1016/S0065-2601(08)00404-8

Barrett, L. F., Lindquist, K. A., & Gendron, M. (2007). Language as context for the perception of emotion. *Trends in Cognitive Sciences*, *11*(8), 327–332. https://doi.org/10.1016/j.tics.2007.06.003

Barrett, L. F., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). The conceptual act theory: A roadmap. In L. F. Barrett & J. A. Russell (Eds.), *The psychological construction of emotion* (pp. 83–110). Guilford Press.

Bastain, B., & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology*, *42*(2), 228–235. https://doi.org/10.1016/j.jesp.2005.03.003

Becker, D. V., Rheem, H., Pick, C. M., Ko, A., & Lafko, S. R. (2019). Angry faces hold attention: Evidence of attentional adhesion in two paradigms. *Progress in Brain Research*, *247*, 89–110. https://doi.org/10.1016/bs.pbr.2019.03.033

Berent, I., Feldman Barrett, L., & Platt, M. (2020). Essentialist biases in reasoning about emotions. *Frontiers in Psychology*, *11*, Article 562666. https://doi.org/10.3389/fpsyg.2020.562666

Betz, N., Hoemann, K., & Barrett, L. F. (2019). Words are a context for mental inference. *Emotion*, *19*(8), 1463–1477. https://doi.org/10.1037/emo0000510

Bijlstra, G., Holland, R. W., Dotsch, R., Hugenberg, K., & Wigboldus, D. H. (2014). Stereotype associations and emotion recognition. *Personality and Social Psychology Bulletin*, *40*(5), 567–577. https://doi.org/10.1177/0146167213520458

Boden, M. T., & Thompson, R. J. (2017). Meta-analysis of the association between emotional clarity and attention to emotions. *Emotion Review*, *9*(1), 79–85. https://doi.org/10.1177/1754073915610640

Brooks, J. A., Shablack, H., Gendron, M., Satpute, A. B., Parrish, M. H., & Lindquist, K. A. (2017). The role of language in the experience and perception of emotion: A neuroimaging meta-analysis. *Social Cognitive and Affective Neuroscience*, *12*(2), 169–183. https://doi.org/10.1093/scan/nsw121

Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996). Categorical perception of facial expressions. *Cognition*, *3*(2), 81–117. https://doi.org/10.1080/713756735

Calvo, M. G., Gutiérrez-García, A., Fernández-Martín, A., & Nummenmaa, L. (2014). Recognition of facial expressions of emotion is related to their frequency in everyday life. *Journal of Nonverbal Behavior*, *38*(4), 549–567. https://doi.org/10.1007/s10919-014-0191-3

Champoux-Larsson, M.-F., & Nook, E. C. (2024). How first- and second-language emotion words influence emotion perception in Swedish–English bilinguals. *Bilingualism: Language and Cognition*. Advance online publication. https://doi.org/10.1017/S1366728923000998

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychological Science*, *24*(5), 723–732. https://doi.org/10.1177/0956797612463582

Cordaro, D. T., Sun, R., Kamble, S., Hodder, N., Monroy, M., Cowen, A., Bai, Y., & Keltner, D. (2020). The recognition of 18 facial-bodily expressions across nine cultures. *Emotion*, *20*(7), 1292–1300. https://doi.org/10.1037/emo0000576

Di Tella, M., Adenzato, M., Catmur, C., Miti, F., Castelli, L., & Ardito, R. B. (2020). The role of alexithymia in social cognition: Evidence from a non-clinical population. *Journal of Affective Disorders*, *273*, 482–492. https://doi.org/10.1016/j.jad.2020.05.012

Doyle, C. M., Gendron, M., & Lindquist, K. A. (2021). Language is a unique context for emotion perception. *Affective Science*, *2*(2), 171–177. https://doi.org/10.1007/s42761-020-00025-7

Durán, J. I., & Fernández-Dols, J.-M. (2021). Do emotions result in their predicted facial expressions? A meta-analysis of studies on the co-occurrence of expression and emotion. *Emotion*, *21*(7), 1550–1569. https://doi.org/10.1037/emo0001015

Ekman, P. (1971). Universals and cultural differences in facial expressions of emotion. *Nebraska Symposium on Motivation*, *19*, 207–283. https://www.paulekman.com/wp-content/uploads/2013/07/Universals-And-Cultural-Differences-In-Facial-Expressions-Of.pdf

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, *3*(4), 364–370. https://doi.org/10.1177/1754073911410740

Ekman, P., Frank, M. G., & Ancoli, S. (1980). Facial signs of emotion experience. *Journal of Personality and Social Psychology*, *39*(6), 1125–1134. https://doi.org/10.1037/h0077722

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W. A., Pitcairn, T., Ricci-Bitti, P. E., Scherer, K., Tomita, M., & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, *53*(4), 712–717. https://doi.org/10.1037/0022-3514.53.4.712

Ellsworth, P. C., & Scherer, K. R. (2002). Appraisal processes in emotion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 572–595). Oxford University Press New York. https://doi.org/10.1093/oso/9780195126013.003.0029

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition*, *44*(3), 227–240. https://doi.org/10.1016/0010-0277(92)90002-Y

Freudenberg, M., Albohn, D. N., Kleck, R. E., Adams, R. B., Jr., & Hess, U. (2020). Emotional stereotypes on trial: Implicit emotion associations for young and old adults. *Emotion*, *20*(7), 1244–1254. https://doi.org/10.1037/emo0000626

Frijda, N. H. (2009). Emotion experience and its varieties. *Emotion Review*, *1*(3), 264–271. https://doi.org/10.1177/1754073909103595

Fugate, J. M. B. (2013). Categorical perception for emotional faces. *Emotion Review*, *5*(1), 84–89. https://doi.org/10.1177/1754073912451350

Fugate, J. M. B., Gouzoules, H., & Barrett, L. F. (2010). Reading chimpanzee faces: Evidence for the role of verbal labels in categorical perception of emotion. *Emotion*, *10*(4), 544–554. https://doi.org/10.1037/a0019017

Fujimura, T., Matsuda, Y.-T., Katahira, K., Okada, M., & Okanoya, K. (2012). Categorical and dimensional perceptions in decoding emotional facial expressions. *Cognition and Emotion*, *26*(4), 587–601. https://doi.org/10.1080/02699931.2011.595391

Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, *8*(9), 404–409. https://doi.org/10.1016/j.tics.2004.07.001

Gendron, M., Lindquist, K. A., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion*, *12*(2), 314–325. https://doi.org/10.1037/a0026007

Gibson, E., Futrell, R., Jara-Ettinger, J., Mahowald, K., Bergen, L., Ratnasingam, S., Gibson, M., Piantadosi, S. T., & Conway, B. R. (2017). Color naming across languages reflects color use. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(40), 10785–10790. https://doi.org/10.1073/pnas.1619666114

Grynberg, D., Chang, B., Corneille, O., Maurage, P., Vermeulen, N., Berthoz, S., & Luminet, O. (2012). Alexithymia and the processing of emotional facial expressions (EFEs): Systematic review, unanswered questions and further perspectives. *PLOS ONE*, *7*(8), Article e42429. https://doi.org/10.1371/journal.pone.0042429

Halberstadt, J. (2005). Featural shift in explanation-biased memory for emotional faces. *Journal of Personality and Social Psychology*, *88*(1), 38–49. https://doi.org/10.1037/0022-3514.88.1.38

Halberstadt, J. B., & Niedenthal, P. M. (2001). Effects of emotion concepts on perceptual memory for emotional expressions. *Journal of Personality and Social Psychology*, *81*(4), 587–598. https://doi.org/10.1037/0022-3514.81.4.587

Harnad, S. (1987). Psychophysical and cognitive aspects of categorical perception: A critical overview. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 1–25). Cambridge University Press.

Harnad, S. (2006). Categorical perception. *Encyclopedia of cognitive science*. Nature Publishing Group. https://doi.org/10.1002/0470018860.s00490

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology*, *39*(1), 113–127. https://doi.org/10.1348/014466600164363

Hobson, H., Brewer, R., Catmur, C., & Bird, G. (2019). The role of language in Alexithymia: Moving towards a multiroute model of Alexithymia. *Emotion Review*, *11*(3), 247–261. https://doi.org/10.1177/1754073919838528

Hoemann, K., Khan, Z., Feldman, M. J., Nielson, C., Devlin, M., Dy, J., Barrett, L. F., Wormwood, J. B., & Quigley, K. S. (2020). Context-aware experience sampling reveals the scale of variation in affective experience. *Scientific Reports*, *10*(1), Article 12459. https://doi.org/10.1038/s41598-020-69180-y

Hoyt, C. L., Morgenroth, T., & Burnette, J. L. (2019). Understanding sexual prejudice: The role of political ideology and strategic essentialism. *Journal of Applied Social Psychology*, *49*(1), 3–14. https://doi.org/10.1111/jasp.12560

Interdisciplinary Affective Science Laboratory. (2006). *Interdisciplinary Affective Science Laboratory face set*. https://www.affective-science.org/face-set.shtml

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

Kikutani, M., Roberson, D., & Hanley, J. R. (2008). What's in the name? Categorical perception for unfamiliar faces can occur through labeling. *Psychonomic Bulletin & Review*, *15*(4), 787–794. https://doi.org/10.3758/PBR.15.4.787

Lee, K. S., Catmur, C., & Bird, G. (2024). Childhood language development and alexithymia in adolescence: An eight-year longitudinal study. *Development and Psychopathology*. Advance online publication. https://doi.org/10.1017/S0954579424001007

Li, S., Zhang, B., Guo, Y., & Zhang, J. (2015). The association between Alexithymia as assessed by the 20-item Toronto Alexithymia Scale and depression: A meta-analysis. *Psychiatry Research*, *227*(1), 1–9. https://doi.org/10.1016/j.psychres.2015.02.006

Lindquist, K. A. (2013). Emotions emerge from more basic psychological ingredients: A modern psychological constructionist approach. *Emotion Review*, *5*(4), 356–368. https://doi.org/10.1177/1754073913489750

Lindquist, K. A. (2021). Language and emotion: Introduction to the special issue. *Affective Science*, *2*(2), 91–98. https://doi.org/10.1007/s42761-021-00049-7

Lindquist, K. A., & Barrett, L. F. (2008). Constructing emotion: The experience of fear as a conceptual act. *Psychological Science*, *19*(9), 898–903. https://doi.org/10.1111/j.1467-9280.2008.02174.x

Lindquist, K. A., Barrett, L. F., Bliss-Moreau, E., & Russell, J. A. (2006). Language and the perception of emotion. *Emotion*, *6*(1), 125–138. https://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*(1), 66–71. https://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*(2), 375–387. https://doi.org/10.1037/a0035293

Lindquist, K. A., Gendron, M., Oosterwijk, S., & Barrett, L. F. (2013). Do people essentialize emotions? Individual differences in emotion essentialism and emotional experience. *Emotion*, *13*(4), 629–644. https://doi.org/10.1037/a0032283

Lynn, S. K. (2010). Decision-making and learning: The peak shift behavioral response. In M. Breed & J. Moore (Eds.), *Encyclopedia of animal behavior* (pp. 470–475). Academic Press. https://doi.org/10.1016/B978-0-08-045337-8.00146-7

Lynn, S. K., Cnaani, J., & Papaj, D. R. (2005). Peak shift discrimination learning as a mechanism of signal evolution. *Evolution*, *59*(6), 1300–1305. https://doi.org/10.1111/j.0014-3820.2005.tb01780.x

Mandalaywala, T. M., Amodio, D. M., & Rhodes, M. (2018). Essentialism promotes racial prejudice by increasing endorsement of social hierarchies. *Social Psychological & Personality Science*, *9*(4), 461–469. https://doi.org/10.1177/1948550617707020

Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Emotion*, *5*(1), 119–124. https://doi.org/10.1037/1528-3542.5.1.119

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). Cambridge University Press. https://doi.org/10.1017/CBO9780511529863.009

Morandini, J. S., Blaszczynski, A., Ross, M. W., Costa, D. S. J., & Dar-Nimrod, I. (2015). Essentialist beliefs, sexual identity uncertainty, internalized homonegativity and psychological wellbeing in gay men. *Journal of Counseling Psychology*, *62*(3), 413–424. https://doi.org/10.1037/cou0000072

Morandini, J. S., Dacosta, L., & Dar-Nimrod, I. (2021). Exposure to continuous or fluid theories of sexual orientation leads some heterosexuals to embrace less-exclusive heterosexual orientations. *Scientific Reports*, *11*(1), Article 16546. https://doi.org/10.1038/s41598-021-94479-9

Morgenroth, T., Kirby, T. A., Gee, I. A., & Ovett, T. A. (2021). Born this way- or not? The relationship between essentialism and sexual minorities' LGBTQ+ identification and belonging. *Archives of Sexual Behavior*, *50*(8), 3447–3458. https://doi.org/10.1007/s10508-021-02145-y

Nemiah, J. C., Freyberger, H., & Sifneos, P. E. (1976). Alexithymia: A view of the psychosomatic process. In O. W. Hill (Ed.), *Modern trends in psychosomatic medicine* (Vol. 3, pp. 430–439). Butterworths.

Nicholson, J. N., & Gray, J. A. (1971). Behavioural contrast and peak shift in children. *British Journal of Psychology*, *62*(3), 367–373. https://doi.org/10.1111/j.2044-8295.1971.tb02047.x

Nook, E. C., Lindquist, K. A., & Zaki, J. (2015). A new look at emotion perception: Concepts speed and shape facial emotion recognition. *Emotion*, *15*(5), 569–578. https://doi.org/10.1037/a0039166

Nook, E. C., Satpute, A. B., & Ochsner, K. N. (2021). Emotion naming impedes both cognitive reappraisal and mindful acceptance strategies of emotion regulation. *Affective Science*, *2*(2), 187–198. https://doi.org/10.1007/s42761-021-00036-y

Panksepp, J. (2004). *Affective neuroscience: The foundations of human and animal emotions*. Oxford University Press.

Plant, E. A., Hyde, J. S., Keltner, D., & Devine, P. G. (2000). The gender stereotyping of emotions. *Psychology of Women Quarterly*, *24*(1), 81–92. https://doi.org/10.1111/j.1471-6402.2000.tb01024.x

Preece, D., Becerra, R., Robinson, K., & Dandy, J. (2018). Assessing Alexithymia: Psychometric properties and factorial invariance of the

20-item Toronto Alexithymia Scale in nonclinical and psychiatric samples. *Journal of Psychopathology and Behavioral Assessment*, *40*(2), 276–287. https://doi.org/10.1007/s10862-017-9634-6

Preece, D. A., Mehta, A., Petrova, K., Sikka, P., Pemberton, E., & Gross, J. J. (2024). Alexithymia profiles and depression, anxiety, and stress. *Journal of Affective Disorders*, *357*, 116–125. https://doi.org/10.1016/j.jad.2024.02.071

Prentice, D. A., & Miller, D. T. (2006). Essentializing differences between women and men. *Psychological Science*, *17*(2), 129–135. https://doi.org/10.1111/j.1467-9280.2006.01675.x

Prentice, D. A., & Miller, D. T. (2007). Psychological essentialism of human categories. *Current Directions in Psychological Science*, *16*(4), 202–206. https://doi.org/10.1111/j.1467-8721.2007.00504.x

Prkachin, G. C., Casey, C., & Prkachin, K. M. (2009). Alexithymia and perception of facial expressions of emotion. *Personality and Individual Differences*, *46*(4), 412–417. https://doi.org/10.1016/j.paid.2008.11.010

Rangel, U., & Keller, J. (2011). Essentialism goes social: Belief in social determinism as a component of psychological essentialism. *Journal of Personality and Social Psychology*, *100*(6), 1056–1078. https://doi.org/10.1037/a0022401

Roberson, D., Damjanovic, L., & Kikutani, M. (2010). Show and tell: The role of language in categorizing facial expression of emotion. *Emotion Review*, *2*(3), 255–260. https://doi.org/10.1177/1754073910361979

Roberson, D., Damjanovic, L., & Pilling, M. (2007). Categorical perception of facial expressions: Evidence for a "category adjustment" model. *Memory & Cognition*, *35*(7), 1814–1829. https://doi.org/10.3758/BF03193512

Roets, A., & Van Hiel, A. (2011). The role of need for closure in essentialist entitativity beliefs and prejudice: An epistemic needs approach to racial categorization. *British Journal of Social Psychology*, *50*(1), 52–73. https://doi.org/10.1348/014466610X491567

Roth, W. D., Van Stee, E. G., & Regla-Vargas, A. (2023). Conceptualizations of race: essentialism and constructivism. *Annual Review of Sociology*, *49*(1), 39–58. https://doi.org/10.1146/annurev-soc-031021-034017

Russell, J. A., & Bullock, M. (1986). Fuzzy concepts and the perception of emotion in facial expressions. *Social Cognition*, *4*(3), 309–341. https://doi.org/10.1521/soco.1986.4.3.309

Salovey, P., Mayer, J. D., Goldman, S. L., Turvey, C., & Palfai, T. P. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the Trait Meta-Mood Scale. In J. W. Pennebaker (Ed.), *Emotion, disclosure, & health* (pp. 125–154). American Psychological Association. https://doi.org/10.1037/10182-006

Satpute, A. B., & Lindquist, K. A. (2021). At the neural intersection between language and emotion. *Affective Science*, *2*(2), 207–220. https://doi.org/10.1007/s42761-021-00032-2

Satpute, A. B., Nook, E. C., Narayanan, S., Shu, J., Weber, J., & Ochsner, K. N. (2016). Emotions in "black and white" or shades of gray? how we think about emotion shapes our perception and neural representation of emotion. *Psychological Science*, *27*(11), 1428–1442. https://doi.org/10.1177/0956797616661555

Siegel, E. H., Sands, M. K., Van den Noortgate, W., Condon, P., Chang, Y., Dy, J., Quigley, K. S., & Barrett, L. F. (2018). Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. *Psychological Bulletin*, *144*(4), 343–393. https://doi.org/10.1037/bul0000128

Skewes, L., Fine, C., & Haslam, N. (2018). Beyond Mars and Venus: The role of gender essentialism in support for gender inequality and backlash. *PLOS ONE*, *13*(7), Article e0200921. https://doi.org/10.1371/journal.pone.0200921

(*Appendix follows*)

Appendix

Emotion Essentialism Questionnaire

1. With enough thought, I would be able to come up with an explanation of what makes people's experiences of anger members of the category "anger." (R)
2. With enough thought and knowledge, experts on emotions would be able to come up with an explanation of what makes people's experiences of anger members of the category "anger." (R)
3. With enough thought and knowledge, experts on emotions would be able to come up with an explanation of what makes people's experiences of love members of the category "love." (R)
4. With enough thought, I would be able to come up with an explanation of what makes people's experiences of love members of the category "love." (R)
5. The category "fear" was invented by people.
6. The category "amusement" was invented by people.
7. Episodes of embarrassment start because of a natural change in one's body. (R)
8. Episodes of surprise start because of a natural change in one's body. (R)
9. Episodes of surprise start because of a subjective interpretation of a situation.
10. Episodes of embarrassment start because of a subjective interpretation of a situation.
11. Angry people act in very similar ways. (R)
12. Excited people act in very similar ways. (R)
13. Whether or not I can put my finger on it, there is something common across all people's experiences of love. (R)
14. Whether or not I can put my finger on it, there is something common across all people experiences of sadness. (R)
15. There is one particular neurological change that must occur of an experience to be an episode of disgust. (R)
16. There is one particular neurological change that must occur of an experience to be an episode of pride. (R)
17. It is easy for me to tell when someone is starting to feel happy. (R)
18. It is easy for me to tell when someone is starting to feel worried. (R)
19. If someone were to tell me "I'm disgusted," I would know a lot about how they are feeling in that moment. (R)
20. If someone were to tell me "I'm amused," I would know a lot about how they are feeling in that moment. (R)
21. It makes sense to use my previous experiences of fear to predict what someone else's experience of fear will be like. (R)
22. It makes sense to use my previous experiences of pride to predict what someone else's experience of pride will be like. (R)
23. There is an essence to what sadness is. (R)
24. There is an essence to what love is. (R)

Received March 7, 2024
Revision received October 30, 2024
Accepted October 31, 2024 ■