A new look at emotion perception: Concepts speed and shape facial emotion recognition

Erik C. Nook[1], Kristen A. Lindquist[2] and Jamil Zaki[1]

[1]Department of Psychology – Stanford University
[2] Department of Psychology – University of North Carolina at Chapel Hill

Key words: Emotion perception, concepts, top-down processing, alexithymia

Address correspondence to:

Erik Nook or Jamil Zaki
Department of Psychology
450 Serra Mall
Stanford University
Stanford, CA 94305
enook@stanford.edu or jzaki@stanford.edu

Word count: 5,391
Abstract word count: 149

ABSTRACT

Decades ago, the "New Look" movement challenged how scientists think about vision by suggesting that conceptual processes shape visual perceptions. Currently, affective scientists are likewise debating the role of concepts in emotion perception. Here, we utilized a repetition-priming paradigm in conjunction with signal detection and individual difference analyses to examine how providing emotion labels—which correspond to discrete emotion concepts—affects emotion recognition. In Study 1, pairing emotional faces with emotion labels (e.g., "sad") increased individuals' speed and sensitivity in recognizing emotions. Additionally, individuals with alexithymia—who have difficulty labeling their own emotions—struggled to recognize emotions based on visual cues alone, but not when emotion labels were provided. Study 2 replicated these findings and further demonstrated that emotion concepts can shape perceptions of facial expressions. Together, these results suggest that emotion perception involves conceptual processing. We discuss the implications of these findings for affective, social and clinical psychology.

We quickly read others' emotions from their facial expressions, but how do we accomplish such a feat? This question has prompted long-standing debate between two theoretical models. The "basic emotion" model holds that emotional signals, including facial expressions, are evolutionarily prepared and automatically recognized as discrete types, requiring little conceptual processing on the part of the perceiver (Ekman, 1973; Ekman & Cordaro, 2011; Levenson, 2011). In this model, emotion concepts are epiphenomenal to facial emotion perception. Perceivers can reflexively decode a social target's emotions from purely visual information—such as the arrangement of their facial muscles—just as easily as they perceive light emanating from a lighthouse (Russell, Bachorowski, & Fernandez-Dols, 2003; Ekman et al., 2002).

However, studies from the "New Look" movement of the mid-twentieth century produced data suggesting that even "basic" visual perception is not always passive or automatic. Instead, scientists argued that people play active roles in constructing perceptions of the world around them based on their motivations, expectations, and prior experiences (Bruner, 1957; Bruner, Postman & Rodrigues, 1951). Constructivist theories apply this idea to *emotion* perception, asserting that people use emotion concepts (i.e., an individual's knowledge about the situations, feelings, facial expressions, and actions that tend to accompany a particular emotion, organized around a label for that emotion; Barrett & Lindquist, 2008; Niedenthal et al., 2009) to identify the emotion expressed by another person. In this theory, facial expressions do not unequivocally signal discrete emotion categories (Russell, 1991; Russell, 1994; Russell, Bachorowski, & Fernandez-Dols, 2003; Lindquist & Gendron, 2013; Lindquist, Gendron, Barrett, & Dickerson, 2014). Instead, expressions communicate categorically ambiguous information about positivity or negativity (valence) and activation or deactivation (arousal), which perceivers parse into discrete emotion categories using emotion concepts (Russell, 2003; Russell et al., 2003; Barrett, Lindquist & Gendron, 2007). Support for the notion that faces are "categorically ambiguous" (i.e., that there is not a one-to-one correspondence between an expression and an emotion type) is found in classic work on emotion perception (e.g., Russell, 1994) as well as recent demonstrations that certain facial movements can belong to multiple emotion categories (Jack, Garrod, & Schyns, 2014), that cultures differ in the facial cues they use to communicate emotion (Jack et al., 2012), and that the context of a face radically influences what emotion people believe it expresses (e.g., Aviezer et al., 2008; Hassin, Aviezer & Bentin, 2013).

The New Look notion that conceptual processes shape our experience of the world continues to influence research on visual perception (Lupyan, Thompson-Schill, & Swingley, 2010; Lupyan & Ward, 2013; Bar et al., 2006; Gilbert & Sigman, 2007). However, in the domain of *emotion* perception, the role of conceptual processes (i.e., the extent to which emotion concept knowledge is utilized when perceiving emotions) remains a matter of debate: emotion concepts are considered crucial to emotion perception in constructivist accounts and epiphenomenal in basic emotion accounts (see Brosch, Pourtois & Sander, 2010; Lindquist & Gendron, 2013). Here, we developed a paradigm that could adjudicate between these views by testing whether activating emotion concepts affects the speed and accuracy of emotion perception (Study 1) and whether these concepts can shape actual representations of facial expressions (Study 2).

Study 1

Study 1 utilized a repetition-priming paradigm in conjunction with signal detection and individual differences analyses to assess the role of concept use in emotion perception. On each trial, participants viewed a rapidly presented facial emotion expression (the *cue* stimulus) followed by a second emotional stimulus (the *target* stimulus). Participants indicated whether or not the emotions reflected in the cue and target stimuli matched. Critically, we manipulated (i) whether the target was a second facial expression (*face-face trials*) or an emotion category word (*face-word trials*) and (ii) whether or not the emotion category denoted by the target and cue in fact matched (i.e., were *congruent* or *incongruent*). This repetition-priming paradigm allowed us to test for concept use in emotion perception in four ways.

First, we assessed how categorizing the emotion of the cue face affected reaction times (RTs) to targets in each condition. A constructivist model posits that emotion perception draws on concept knowledge about emotion (Barrett et al. 2007; Barrett, 2011; Lindquist & Gendron, 2013). Hence, one instance of emotion perception should prime the concept for an emotion that speeds subsequent processing of stimuli congruent with that concept (for more on conceptual priming, see Neely, 1991; Collins & Loftus, 1975; Wheatley, Weisberg, Beauchamp, & Martin, 2005). Critically, perceiving an emotion in a face should prime congruent emotion labels—which are unambiguous representations of emotion concepts—more than congruent faces. Hence, priming should be stronger in face-word trials than face-face trials. However, according to a basic emotion approach, exposure to cue faces should most strongly prime congruent emotional expressions via purely visual priming. This leads to the opposite prediction: priming should be stronger when stimuli are visually similar (i.e., in face-face trials) rather than when they are cross-modal (i.e., in face-word trials).

Second, we used signal detection methods to compare participants' sensitivity (*d'*; Green & Swets, 1966) across face-face and face-word trials. Constructivist models would predict that emotion perception should be more difficult when conceptual information is unavailable to the perceiver than when it is available (Gendron et al. 2012; Lindquist et al. 2006; 2014). Hence, sensitivity should be lower for face-face trials than face-word trials in our task. Conversely, a basic emotion account predicts that facial muscle configurations are automatically recognized as discrete emotions regardless of the presence or absence of emotion language (Ekman & Cordaro, 2011). Thus sensitivity for face-face trials should be equal to or greater than sensitivity for face-word trials.

Third, we tested whether concept use is especially necessary when participants must differentiate stimuli that are visually similar. Stereotypical expressions of *anger* (a scowl), *disgust* (a wrinkled nose) and *sadness* (a frown) typically share a knitted brow, which tightens the eyes, whereas expressions of *fear* instead involve raised eyebrows and exposed sclera of the eyes (Ekman et al., 2002). In fact, extant research suggests that the widened eyes of fear expressions render them visually and psychologically dissimilar from expressions of other emotions (Jack, et al., 2014; Adolphs, et al., 2005; Whalen, et al., 2004; Somerville & Whalen, 2006; see supplemental materials for analyses confirming such differences in our stimuli). Hence, we compared RTs and *d'* scores for incongruent trials that involved pairs of two non-fear stimuli (*visually similar* trials) to those that involved both a fear and a non-fear stimulus (*visually dissimilar* trials). According to a basic emotion view, expressions are automatically understood as discrete kinds, so visual similarity between stimuli should not influence the speed or sensitivity of their perception (Tracy & Robins, 2008). However, a constructivist model would predict that concept use would be more evident when participants must categorize two visually similar expressions as distinct emotions than when these expressions are already visually dissimilar. If this is the case, visual similarity between stimuli should increase RT and decrease *d'*, especially when discrete concepts are not provided (i.e., in face-face trials).

Finally, we investigated how emotion recognition varied as a function of individuals' levels of alexithymia, a trait characterized by difficulties in verbally describing one's own emotions (Grynberg, et al., 2012). Initially, one might predict that individuals high in alexithymia should exhibit the most severe emotion perception difficulties when they are forced to use language (i.e., in face-word trials). However, a constructivist model suggests a more counterintuitive prediction. To the extent that *all* discrete emotion perception requires access to and use of emotion concepts, individuals with alexithymia might fail to spontaneously access or use these concepts during emotion perception. Under this model, individuals high in alexithymia should actually perform worse in face-face trials (in which they are not given conceptual information) than in face-word trials (in which they are given such information). According to a basic emotion perspective, alexithymia would either not impair emotion perception ability or it would do so regardless of the presence or absence of emotion language.

Method

Participants. One hundred thirty participants provided informed consent and received course credit or payment for their participation, as approved by Stanford University's Institutional Review Board. Data from 4 participants were excluded because they were over 45 years old, a criterion age known to affect performance on measures of emotion perception (Ruffman, Henry, Livingstone, & Phillips, 2008) and reaction time (Der & Deary, 2006). An additional 3 participants were excluded because their mean reaction time (1 participant) or *d'* score (2 participants) for at least one condition was > 3 *SD*s from the sample mean.[1] Our final sample included 123 participants (78 females; mean age = 22.86 years, *SD* = 5.57).

Stimuli and Procedure. Participants viewed the cue and target sequentially and indicated as quickly and accurately as possible (i.e. within 1000ms after target onset) whether or not the stimuli conveyed the same emotion (see **Figure 1**). Two variables were manipulated within-subjects: whether the target was another face or an emotion category word (*target type*) and whether the emotions of the cue and target stimuli matched (*congruence*). Emotions included *fear*, *anger*, *disgust* and *sadness*, and faces were drawn from the NimStim (Tottenham et al., 2009) and IASLab (Gendron, Lindquist, & Barrett, unpublished data) stimuli sets. After the task, participants completed the Toronto Alexithymia Scale (TAS-26; Wise, Simpson & Sheridan, 2000). See supplemental materials for more details on the methods of Study 1.

We analyzed mean RTs for correctly answered trials in each condition using a 2 [target type: face vs. word] x 2 [congruence: congruent vs. incongruent] repeated measures ANOVA.

We calculated *d'* scores for face-face and face-word trials according to traditional signal detection techniques [*d'* = z(hit rate) – z(false alarm rate); Macmillan & Creelman, 2004; Stanislaw & Todorov, 1999] and compared these scores using a paired sample t-test.

We assessed whether *visual similarity* of stimuli affected emotion perception speed or sensitivity by coding incongruent trials including combinations of *anger*, *sadness* and *disgust* as *visually similar*, and incongruent trials including *fear* as *visually dissimilar*. We created these categories for both face-face and face-word trials. We then analyzed RTs for these subtypes of incongruent trials using a 2 [target type] x 2 [visual similarity] repeated measures ANOVA, and

[1] We scanned for trial-level RT outliers using the same criteria as our group-level analysis. Although we found a small number of trials that qualified as within-participant outliers (*M* = 0.15 across all participants), excluding these trials did not affect the significance of any of our effects. Hence, we included them in our analyses.

we analyzed *d'* for all trials using another 2 [target type] x 2 [visual similarity] repeated measures ANOVA.

Finally, we calculated alexithymia scores for each participant by summing items from the "identifying emotions" and "describing emotions" subscales of the TAS-26. These scales most closely measure one's difficulty translating affect into discrete emotion categories. To assess how alexithymia relates to one's ability to identify others' emotions both with and without a linguistic context, we correlated participants' alexithymia scores with their mean *d'* for both face-face and face-word trials. We then used the Steiger (1980) method to test for significant differences between these correlations.

Results and Discussion

Repetition-Priming Task: Reaction Times. Participants responded faster to congruent (M = 0.69s) than incongruent (M = 0.72s) target-cue pairs [$F(1,122) = 118.15$, $p < .001$, $\eta_p^2 = .49$]. Participants also responded faster to face-word trials (M = 0.67) than face-face trials [M = 0.74s, $F(1,122) = 253.71$, $p < .001$, $\eta_p^2 = .67$]. A significant interaction between congruence and target type [$F(1,122) = 13.08$, $p < .001$, $\eta_p^2 = .10$] further revealed that priming was greater for face-word than face-face trials (**Figure 2a**). These data (which can be downloaded from www.ssnl.stanford.edu/publications) are consistent with a constructivist model of emotion perception, under which perceiving emotions in facial expressions activates emotion concepts that facilitate subsequent processing of stimuli congruent with that concept. In fact, recognizing an emotion in a face primed subsequent congruent *words* more than congruent *faces*, further supporting the notion that emotion perception primes concepts more than visual facial information.

Sensitivity. Participants' *d'* scores were significantly lower for face-face trials ($M = 0.96$) than face-word trials ($M = 1.26$), [$t(122) = -6.33$, $p < .001$, Cohen's $d = 0.55$], supporting the constructivist model's prediction. Facial expressions do not unambiguously convey discrete emotion categories. Instead, emotion labels are clearer signals of emotion categories than faces, allowing participants to make a more sensitive assessment of whether a word matches a face than whether two faces express the same emotion. These results also suggest that the "label superiority effect" (Russell & Widen, 2002)—under which children are significantly more accurate when matching faces with emotion words than with other faces—holds into adulthood.

Visual Similarity. When we analyzed RTs for incongruent targets with visual similarity as a factor, the main effect of target type remained [$M = 0.69$s; $F(1,121) = 191.29$, $p < .001$, $\eta_p^2 = .61$]. However, a main effect of visual similarity [$M = .72$; $F(1,121) = 5.76$, $p = .018$, $\eta_p^2 = .05$] and a significant interaction [$F(1,121) = 20.22$, $p < .001$, $\eta_p^2 = .14$] also emerged. Follow-up paired samples t-tests revealed that visual similarity slowed responses to incongruent face-face pairs [$t(121) = 4.50$, $p < .001$, $d = 0.30$] but had no effect on RTs for incongruent face-word trials [$t(122) = 1.70$, $p = .09$, $d = 0.09$]. Hence, participants were slower to accurately distinguish two visually similar faces (e.g., *anger* and *disgust*) than two visually dissimilar faces (e.g., *anger* and *fear*). However, when targets were words, this relationship disappeared (e.g., participants responded just as quickly to trials involving anger faces paired with the word "disgust" as they did to trials paired with the word "fear"). We found a similar pattern in our analysis of *d'* (**Figure 2b**). The main effect of target type held [$F(1,122) = 38.24$, $p < .001$, $\eta_p^2 = .24$], but a main effect of visual similarity [$F(1,122) = 82.69$, $p < .001$, $\eta_p^2 = .40$] and a target type x visual similarity interaction [$F(1,122) = 20.05$, $p < .001$, $\eta_p^2 = .14$] also emerged. Follow-up t-tests indicated that visual similarity reduced sensitivity for both target types, but it had a much larger effect for face-face trials [$t(122) = -9.98$, $p < .001$, $d = 0.70$] than face-word trials [$t(122) = -3.44$, $p = .001$, $d = 0.20$, see **Supplemental Table 1** for all signal detection measures from both studies]. As the visual similarity of stimuli pairs increased, participants' sensitivity for judging whether one face matched another decreased. However, matching a face to a label mitigated this effect. These parallel results suggest that: (i) emotion concept use is most pronounced when participants must distinguish visually similar facial stimuli, and (ii) emotion labels contribute to faster and more sensitive emotion perception because they activate concepts that help participants to disambiguate affective information communicated by a face.

Alexithymia and Sensitivity. The sum of the "identifying emotions" and "describing emotions" subscales of the TAS-26 correlated negatively with sensitivity for face-face trials [$r(121) = -0.27$, $p = .003$]. However, alexithymia did not correlate with sensitivity in face-word trials [$r(121) = -0.031$, $p = .734$]. A Steiger (1980) comparison indicated that these correlation coefficients differed significantly ($z = -2.87$, $p = .004$, **Figure 2c**). Intriguingly, these findings suggest that alexithymic individuals fare worse at perceiving emotions when they are not given emotion concepts to make meaning of faces. Hence, alexithymia might not be due to a misunderstanding of emotion terms, but rather an inability to spontaneously apply such conceptual knowledge to parse others' (and perhaps one's own) emotional states.

Study 2

Study 1 provided four lines of converging evidence suggesting that conceptual processes are involved in emotion perception. In Study 2, we drew upon another idea inspired by the New Look movement, namely that conceptual processes influence the formation of visual percepts (Bruner, 1957) and memories (Bartlett, 1932; Carmichael, Hogan, & Walter, 1932). Specifically, we assessed whether emotion concepts can shape emotional percepts as they unfold over time (cf. Neisser, 1967). Just as priming the concept of a tomato can make an ambiguous color swatch seem more red (Bruner et al., 1951), so too might emotion concepts bias processing of a facial expression, shifting its representation to be congruent with that concept. In our paradigm, this corresponds to testing whether emotion concepts activated by a target shape participants' evolving representation of a prior emotion cue.

To provide such a test, we modified the paradigm of Study 1 by including a *cue perception* task at the end of each trial. After viewing both the cue and target, participants used a sliding scale to browse a set of facial morphs and select one they believed matched the initial cue face for that trial (see Halberstadt & Niedenthal, 2001). For each trial, we measured *cue shift*, or the difference between the face participants believed was the cue and the actual cue. Hence, higher *cue shift* scores indicate a shift in participants' representation of the cue away from its actual appearance.

A constructivist model makes two predictions for cue shift scores in Study 2. First, incongruent targets should activate concepts that bias conceptual processing of the cue, shifting its representation to be less like the actual cue face and more like an expression of the target's emotion. Conversely, congruent targets should activate concepts that reinforce—or even exaggerate (Halberstadt and Niedenthal, 2001)—the cue's emotion, leading to the prediction that cue shift should be greater following incongruent targets than congruent targets.

Second, providing emotion concepts in face-word trials should "sharpen" categorical representations of the cue (i.e. with reduced cue shift and increased sensitivity), whereas the relative dearth of emotion concepts in face-face trials should produce relatively "blurred" representations of the cue (i.e., with increased cue shift and reduced sensitivity). Such a finding may help explain participants' greater sensitivity for face-word trials in Study 1. Hence, a constructivist model predicts that cue shift should be lower for face-word than face-face trials and that this difference in cue shift could contribute to differences in emotion perception sensitivity across these conditions. By contrast, a basic emotions model would not make strong predictions with regards to the cue perception task. If conceptual processes are not involved in emotion perception, then targets should have little (or no) influence on representations of the cue. This would result in no significant differences in cue shift across conditions.

Method

Participants. Eighty-two participants completed Study 2. Data from 11 participants were excluded because they were older than 45 (1 participant), failed to follow experimental instructions (3 participants), had completed Study 1 (1 participant), or produced accuracy or RT scores > 3 *SD*s from the mean (6 participants).[2] Our final sample included 71 participants (50 females, 1 undisclosed gender; mean age = 22.9 years, *SD* = 5.72).

Stimuli and Procedure. Four modifications were made to Study 1's paradigm. First, we manipulated target type between-subjects rather than within-subjects. Second, emotion categories included only *anger*, *disgust* and *sadness*, restricting our facial stimuli to those that had high visual similarity. Given our results in Study 1, we believed these two alterations would allow for a more focused test of the role of conceptual processing in emotion perception. Specifically, this design allows us to compare how participants parse visually similar faces into discrete emotion types either when discrete emotion concepts are not provided at all (face-face condition, $n = 37$) or when these concepts are provided in all trials (face-word condition, $n = 34$).

[2] Again, very few trial-level RT outliers emerged ($M = 0.15$). Excluding these trials did not change the significance of any results presented here, so we included them in our analyses.

Third, to measure alexithymia, the TAS-26 was replaced with the TAS-20 (Bagby, Parker & Taylor, 1994), a shorter and more commonly used version of the scale (Wise et al., 2000). Finally, we added the *cue perception* task described below.

Trials proceeded exactly as in Study 1 until the end of the target stimulus presentation (see **Figure 3**). After this, a visual mask appeared for 200ms, followed by a "slider" beneath a facial expression of emotion. Participants could move the slider to gradually morph the emotion expressed by the face and select the expression they believed was initially displayed by the cue face. One slider anchor was the cue face itself, and all morphs on the slider were of the same individual. However, the emotion expressed by the other anchor depended on that trial's condition. For incongruent trials, this emotion was the emotion of the target stimulus, and for congruent trials, it was a randomly selected emotion other than that of the cue. For example, if an incongruent trial contained a facial expression of *anger* as a cue and either a facial expression of *sadness* or the word "sadness," then the subsequent slider would range from a face expressing *anger* to one expressing *sadness.* Conversely, if the cue was an expression of *anger* and the target was also an expression of *anger* or the word "anger," then the slider would range from a facial expression of anger to one expressing either *sadness* or *disgust*. The assignment of the cue face to the left or right anchor was pseudorandomized such that half of the trials had the cue face on the left, and the order of this assignment was randomized across trials. We created 9 intermediate morphs (at 10% increments) between each anchor using Abrosoft FantaMorph software, and the slider always began at the midpoint of the morphs (50% of each anchor emotion). We quantified *cue shift* as the distance on the scale between the face participants selected and the actual cue face computed as a percentage ranging from 0% (selecting the actual cue face) to 100% (selecting the opposite anchor). Participants could spend as long as they needed to complete this phase of the task.

Analyses of RT, sensitivity and alexithymia were identical to those of Study 1, with two modifications. First, visual similarity was no longer a factor in our design, so we did not include it in our RT or sensitivity analyses. Second, target type was analyzed as a between-subjects variable. We analyzed average cue shift and cue perception RTs for each condition using a 2 [target type] x 2 [congruence] repeated measures ANOVA.

Results and Discussion

Repetition-Priming Task. In general, Study 2 replicated patterns of Study 1 (see **Supplemental Figure S2**). Participants responded faster to congruent targets than incongruent targets [$F(1,69) = 26.72$, $p < .001$, $\eta_p^2 = .28$] and to face-word than face-face trials [$F(1,69) = 12.87$, $p = .001$, $d = 0.86$]. Again, an interaction indicated that priming occurred to a greater degree for face-word than face-face trials [$F(1,69) = 26.79$, $p < .001$, $\eta_p^2 = .28$]. However, with target type now split between-subjects, no priming occurred for face-face trials [$t(36) = 0.004$, $p = .997$, $d < 0.001$], though it did for face-word trials [$t(33) = -7.43$, $p < .001$, $d = 1.12$]. This absence of priming in Study 2's face-face condition may be due to the fact that participants in this condition never saw emotion words and hence did not explicitly learn that emotion types were constrained to *anger*, *disgust* and *sadness*. Without such restrictions on their conceptual space, participants may not have categorized faces as expressing these three emotions, a necessary step for conceptual priming. Hence, even the visual regularities of posed facial expressions are not sufficient for one exemplar to prime another if these expressions can't be translated into discrete types using emotion concepts (also see Gendron et al., 2012).

As in Study 1, sensitivity was lower for face-face trials (*M* = 0.72) than face-word trials (*M* = 1.11), [$t(69) = -3.48$, $p = .001$, $d = .84$] and higher alexithymia was associated with impaired performance for face-face trials [$r(35) = -.34$, $p = .04$] but not face-word trials [$r(32) = .009$, $p = .96$].

Cue Perception Task. There were no main effects or interactions for RTs on the cue perception task (*p*s > .79). However, two main effects emerged for our measure of cue shift (**Figure 4a).** First, greater cue shift occurred following incongruent targets (*M* = 26.08%) than congruent targets [*M* = 22.63%, $F(1,69) = 17.00$, $p < .001$, $\eta_p^2 = .20$], supporting our prediction that concepts activated by the target shift perceptual representations of the cue. Furthermore, the interaction between congruence and target type was not significant, indicating that the degree to which an incongruent target pushed participants' representation of the cue away from the actual cue face was approximately equal for both face-face and face-word trials [$F(1,69) = .01$, $p = .93$ $\eta_p^2 < .001$].

Second, cue shift was greater following face-face trials (*M* = 27.50%) than face-word trials [*M* = 21.04%, $F(1,69) = 12.13$, $p = .001$, $d = 0.84$]. In fact, average cue shift and sensitivity were negatively correlated across participants [$r(69) = -.47$, $p < .001$], suggesting that participants who experienced *less* cue shift were *more* sensitive in reporting whether cues and targets matched. Given these parallel results, we used a mediation approach to test whether "sharper" cue perceptions (less cue shift, meaning that representations were closer to the actual cue presented) in face-word trials contributed to the increased sensitivity exhibited in this condition. A mediation model with 1000 bootstrapped samples (Preacher & Hayes, 2004) yielded a significant indirect effect ($p = .02$; **Figure 4b**), suggesting that less cue shift in the face-word condition was partially, though not completely, responsible for higher sensitivity scores in this condition.

These results could reflect at least two mechanisms underlying the "label superiority effect" (Russell & Widen, 2002) we documented in Study 1. First, providing emotion concepts in face-word trials might "sharpen" representations of faces, decreasing cue shift and increasing sensitivity, while withholding emotion concepts in face-face trials might have the opposite effects: increasing cue shift and decreasing sensitivity. Alternatively, word targets may be less "noisy" distractors than face targets, allowing for more accurate representations of the cue, also decreasing cue shift and increasing *d'*.

Although our cue shift findings are consistent with both interpretations, the idea that words "sharpen" representations of target faces converges with data from Study 1 as well as extant evidence illustrating that words shape representations of faces. For instance, verbally labeling a morphed face shifts the perception of that face to appear more like a prototypical expression of the emotion label (Halberstadt & Niedenthal, 2001; Halberstadt, 2003; 2005) and providing emotion labels during emotion perception causes participants to misremember a morphed face as being more prototypical than its actual presentation (Roberson et al., 2007). These studies converge with our own RT, *d'* and cue shift data to suggest that conceptual processes can shape "sharper" representations of faces. Future work could further test this claim by assessing whether temporarily depriving participants of emotion concepts (e.g., through a cognitive manipulation called semantic satiation that temporarily inhibits access to the meaning of words (cf., Gendron et al., 2012) also "blurs" representations of emotion expressions.

General Discussion

Across two studies, we observed consistent and novel evidence that conceptual processes contribute to emotion perception. Perceiving an emotion on a face primed the concept for that emotion, speeding responses to stimuli congruent with that concept. Additionally, we found that providing emotion concepts (through discrete emotion labels like "sadness") increases sensitivity for recognizing emotions, potentially by "sharpening" participants' representations of emotional faces. Finally, individuals with alexithymia—who have difficulty labeling emotions—exhibited emotion recognition difficulty only when trials did not include emotion labels. This suggests that impairments in alexithymia may stem from an inability to spontaneously apply emotion concepts. At the broadest level, these findings translate "New Look" ideas to emotion perception by demonstrating that the concepts a perceiver brings to bear on a situation may influence how they recognize emotions in others. Such an insight holds exciting implications for research in affective, social and clinical psychology.

With respect to affective science, these results add to growing evidence supporting constructivist theories of emotion by illustrating that concepts shape emotion not only in the intrapersonal domain (e.g., Lindquist & Barrett, 2008), but also the interpersonal domain. Indeed, our findings converge with mounting evidence that language plays a role in emotion perception. Extant work has shown that temporarily impairing access to emotion concepts in healthy individuals (Gendron et al., 2012; Lindquist et al., 2006), having permanently impaired access to conceptual knowledge through neurodegenerative disease (Lindquist et al., 2014) and cultural differences in concept knowledge (Gendron et al., 2014) all alter emotion perception. Similarly, studies on cross-modal priming indicate that simply viewing rapidly presented emotion expressions—even when they are supposed to be ignored—can influence how quickly people process emotion words (Carroll & Young, 2005; Baggott, Palermo, & Fox, 2011). The present studies add to this growing literature by demonstrating that healthy adults activate and apply emotion concepts *in the moment* during emotion perception and that these concepts shape representations of faces. As such, these findings support claims concerning the role of conceptual processes in emotion perception (Barrett et al. 2007; Lindquist & Gendron, 2013; Lindquist, Satpute & Gendron, in press; Adolphs, 2002) and stimulate new questions for research.

For example, our results may help to explain the effect of context on emotion perception. A growing body of research indicates that visual, social, and cultural contexts influence the emotions an individual perceives in a face (Aviezer et al., 2008; Aviezer, Trope, & Todorov, 2012; Hess, Adams, & Kleck, 2004; Jack et al., 2012; Gendron et al., 2014). The current framework suggests that these context-mediated emotion perception effects might involve concept use; that is, contextual factors might alter the concepts people access and apply to facial stimuli. Future research could test this hypothesis by assessing whether certain contexts automatically activate associated emotion concepts. Our findings also converge with an emerging model of social inference as a process of Bayesian cue integration (Zaki, 2013; Baker, Saxe & Tenenbaum, 2009; Ong, Zaki & Goodman, under review). In this view, perceivers compute the probability of a target experiencing an emotion by integrating information from multiple cues (e.g., the target's facial expression, vocalizations and situation).

In the clinical domain, our findings suggest that increasing accessibility and use of emotion concepts might be helpful for people high in alexithymia. In fact, the methods of Study 1 could be thought of as a form of "intervention" for alexithymia, as simply providing emotion labels on some trials circumvented emotion perception deficits for individuals high in alexithymia. Although others have already reported that alexithymia impairs emotion perception

(Cook, Brewer, Shah, & Bird, 2013; Lane, Sechrest, Riedel, Shapiro, & Kaszniak, 2000; Grynberg et al., 2012; Parker, Prkachin, & Prkachin, 2005; Prkachin, Casey, & Prkachin, 2009), here we extend this finding by demonstrating that providing emotion labels during an emotion perception task eliminates these difficulties. Given that alexithymia is comorbid with a host of psychological disorders (Taylor, 2000) and is specifically implicated in emotion perception difficulties in autism (Cook et al., 2013), training individuals with these disorders to recognize and apply emotion concepts might improve their ability to recognize others' emotions. Such training is effective in autism (Golan & Baron-Cohen, 2006; Silver & Oakes, 2001) and could be useful in other disorders.

Although findings presented here suggest that conceptual processes contribute to emotion perception, future research could enrich this model through three main contributions. First, research should focus on describing these processes at a mechanistic level, potentially using neuroimaging methods to explore how emotion concepts shape the neural representation of face percepts. Second, future studies should delineate the boundary conditions of conceptual processes in emotion perception by determining the maximal extent to which conceptual processes can bias and shape emotion perceptions (cf., Lindquist, Satpute & Gendron, under review). Finally, research should investigate the downstream social and psychological consequences of conceptual processes in emotion perception. For example, further research could explore whether individual differences in how people conceptualize emotions impacts real-world social outcomes such as interpersonal relationship quality.

However, one thing is abundantly clear: we can only hope to understand emotion perception when we take a "New Look" at the conceptual processes it involves. If activating emotion concepts shifts the speed, sensitivity and content of emotion perceptions, there are countless questions to answer concerning the breadth and depth of these processes. Furthermore, this approach to emotion perception can potentially provide insights that explain related findings in affective and social psychology and be useful in improving treatments for individuals with psychological illness.

Acknowledgements

The authors thank Stephen Frey and Christina Chwyl for help with data collection and stimuli preparation.

Author Contributions

ECN, KAL and JZ designed the experiments. ECN collected and analyzed data. ECN, KAL and JZ wrote the manuscript.

REFERENCES

Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, *1*(1), 21–62.

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*(7021), 68–72.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., … Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science*, *19*(7), 724–32.

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science (New York, N.Y.)*, *338*(6111), 1225–9.

Bagby, 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.

Baggott, S., Palermo, R., & Fox, A. M. (2011). Processing emotional category congruency between emotional facial expressions and emotional words. *Cognition & Emotion*, *25*(2), 369–79.

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*(3), 329-349.

Bar, M., Kassam, K. S.,* Ghuman, A. S.,* Boshyan, J., Schmid, A. M., Dale, A. M., Hämäläinen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., & Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences*, *103*(2), 449–454. [* authors contributed equally]

Barrett, L. F. (2011). Constructing Emotion, *20*, 359–380.

Barrett, L. F. & Lindquist, K. A. (2008). The embodiment of emotion. In G. R. Semin & E. R. Smith (Eds.), *Embodied grounding: social, cognitive, affective, and neuroscientific approaches.* New York: Cambridge University Press.

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.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.

Brosch, T., Pourtois, G., & Sander, D. (2010). The perception and categorisation of emotional stimuli: A review. *Cognition and Emotion*, *24*(3), 377-400.

Bruner, J. (1957). On Perceptual Readiness. *Psychological Review*, *64*(2), 123–152.

Bruner, J. S., Postman, L., & Rodrigues, J. (1951). Expectation and the Perception of Color. *The American Journal of Psychology*, *64*(2), 216–227.

Carmichael, L., Hogan, H. P., & Walter, A. A. (1932). An Experimental Study of the Effect of Language on the Reproduction of Visually Perceived Form. *Journal of Experimental Psychology*, *15*(1), 73–86.

Carroll, N. C., & Young, A. W. (2005). Priming of emotion recognition. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *58*(7), 1173–97.

Collins, A. M., & Loftus, E. F. (1975). A Spreading-Activation Theory of Semantic Processing. *Psychological Review*, *82*(6), 407–428.

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.

Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: results from the United Kingdom Health and Lifestyle Survey. *Psychology and aging*, *21*(1), 62–73.

Ekman, P. (1973). Cross-cultural studies of facial expression. In P. Ekman (Ed.), *Darwin and facial expression: A century of research in review* (pp. 169–222). New York: Academic.

Ekman, P., & Cordaro, D. (2011). What is Meant by Calling Emotions Basic. *Emotion Review*, *3*(4), 364–370.

Ekman, P., Friesen, W. V., & Hager, J. C. (Eds.). (2002) *Facial Action Coding System* [E-book]. Salt Lake City. UT: Research Nexus.

Gendron, M., Lindquist, K. A., & Barrett, L. F. (unpublished data). Ratings of IASLab facial expression stimuli. http://www.affective-science.org/

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture. *Emotion (Washington, D.C.)*, *14*(2), 251–62. doi:10.1037/a0036052

Gendron, M., Lindquist, K. a, Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion (Washington, D.C.)*, *12*(2), 314–25.

Gilbert, C. D., & Sigman, M. (2007). Brain states: top-down influences in sensory processing. *Neuron*, *54*(5), 677–96.

Golan, O., & Baron-Cohen, S. (2006). Systemizing empathy: teaching adults with Asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. *Development and psychopathology*, *18*(2), 591–617.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1). New York: Wiley.

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), e42429.

Halberstadt1, J. (2003). The paradox of emotion attribution: explanation biases perceptual memory for emotional expressions. *Current Directions in Psychological Science*, *12*(6), 197–201.

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

Hassin, R. R., Aviezer, H., & Bentin, S. (2013). Inherently Ambiguous: Facial Expressions of Emotions, in Context. *Emotion Review*, *5*(1), 60–65.

Hess, U., Adams, R. B., & Kleck, R. E. (2004). Facial appearance, gender, and emotion expression. *Emotion (Washington, D.C.)*, *4*(4), 378–88.

Jack, R. E., Garrod, O. G. B., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology : CB*, *24*(2), 187–92.

Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(19), 7241–4.

Lane, R. D., Sechrest, L., Riedel, R., Shapiro, D. E., & Kaszniak, a W. (2000). Pervasive emotion recognition deficit common to alexithymia and the repressive coping style. *Psychosomatic Medicine*, *62*(4), 492–501.

Levenson, R. W. (2011). Basic Emotion Questions. *Emotion Review*, *3*(4), 379–386.

Lindquist, K. A, Barrett, L. F., Bliss-Moreau, E., & Russell, J. a. (2006). Language and the perception of emotion. *Emotion (Washington, D.C.)*, *6*(1), 125–38.

Lindquist, K. A., & Barrett, L. F. (2008). Emotional complexity. In Lewis, M. D., Haviland-Jones, J. M., & Barrett, L. F. (Eds.), *Handbook of emotions* (pp. 513-530). New York, NY: Guildford.

Lindquist, K. A., & Gendron, M. (2013). What's in a Word? Language Constructs Emotion Perception. *Emotion Review*, *5*(1), 66–71.

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*. [*,† authors contributed equally].

Lindquist, K.A., Satpute, A. B., & Gendron, M. (in press). Does language do more than communicate emotion? *Current Directions in Psychological Science.*

Lupyan, G., Thompson-Schill, S. L., & Swingley, D. (2010). Conceptual penetration of visual processing. *Psychological Science*, *21*(5), 682–91.

Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences*, *110*(35), 14196–14201.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. *Basic processes in reading: Visual word recognition, 11,* 264-336.

Neisser, Ulric. (1967). *Cognitive psychology*. East Norwalk, CT: Appletoon-Century-Crofts.

Niedenthal, P. M., Winkielman, P., Mondillon, L., & Vermeulen, N. (2009). Embodiment of emotion concepts. *Journal of Personality and Social Psychology*, *96*(6), 1120–36. doi:10.1037/a0015574

Ong, D., Zaki, J., & Goodman, N. (under review). Affective cognition: Exploring lay theories of emotion.

Parker, P. D., Prkachin, K. M., & Prkachin, G. C. (2005). Processing of facial expressions of negative emotion in alexithymia: the influence of temporal constraint. *Journal of Personality*, *73*(4), 1087–107.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior research methods, instruments, & computers: a journal of the Psychonomic Society, Inc*, *36*(4), 717–31.

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.

Roberson, D., Damjanovic, L., & Pilling, M. (2007). Categorical perception of facial expressions: evidence for a “category adjustment” model. *Memory & Cognition*, *35*(7), 1814–29.

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: implications for neuropsychological models of aging. *Neuroscience and biobehavioral reviews*, *32*(4), 863–81.

Russell, J. A. (1991). Culture and the Categorization of Emotions, *110*(3), 426–450.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, *115*(1), 102–41.

Russell, J. A. (2003). Core Affect and the Psychological Construction of Emotion. *Psychological Review*, *110*(1), 145–172.

Russell, J. A., Bachorowski, J. A., & Fernandez-Dols, J. M. (2003). Facial and vocal expressions of emotion. *Annual review of psychology*, *54*, 329–49.

Russell, J. A., & Widen, S. C. (2002). A Label Superiority Effect in Children’s Categorization of Facial Expressions. *Social Development*, *11*(1), 30–52.

Silver, M., & Oakes, P. (2001). Evaluation of a New Computer Intervention to Teach People with Autism or Asperger Syndrome to Recognize and Predict Emotions in Others. *Autism*, *5*(3), 299–316.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc*, *31*(1), 137–49.

Steiger, J. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*(2), 245–251.

Somerville, L. H., & Whalen, P. J. (2006). Prior experience as a stimulus category confound: an example using facial expressions of emotion. *Social Cognitive and Affective Neuroscience*, *1*(3), 271–4.

Tracy, J. L., & Robins, R. W. (2008). The automaticity of emotion recognition. *Emotion*, *8*(1), 81–95.

Taylor, G. J. (2000). Recent developments in alexithymia theory and research. *The Canadian Journal of Psychiatry / La Revue canadienne de psychiatrie*, *45*(2), 134-142.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A, Marcus, J., Westerlund, A., & Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry research*, *168*(3), 242–9.

Whalen, P. J., Kagan, J., Cook, R. G., Davis, F. C., Kim, H., Polis, S., McLaren, D. G., Somerville, L. H., McLean, A. A., Maxwell, J. S., & Johnstone, T. (2004). Human Amygdala Responsivity to Masked Fearful Eye Whites. *Science*, *306*(December), 2061.

Wheatley, T., Weisberg, J., Beauchamp, M. S., & Martin, A. (2005). Automatic priming of semantically related words reduces activity in the fusiform gyrus. *Journal of Cognitive Neuroscience*, *17*(12),

Wise, T. N., Simpson, N., & Sheridan, M. J. (2000). Comparison of 26-item and 20-item versions of the Toronto Alexithymia Scale for psychiatric outpatients. *Psychological reports*, *87*(1), 127–32.

Zaki, J. (2013). Cue Integration: A Common Framework for Social Cognition and Physical Perception. *Perspectives on Psychological Science*, *8*(3), 296–312.

FIGURES

Figure 1

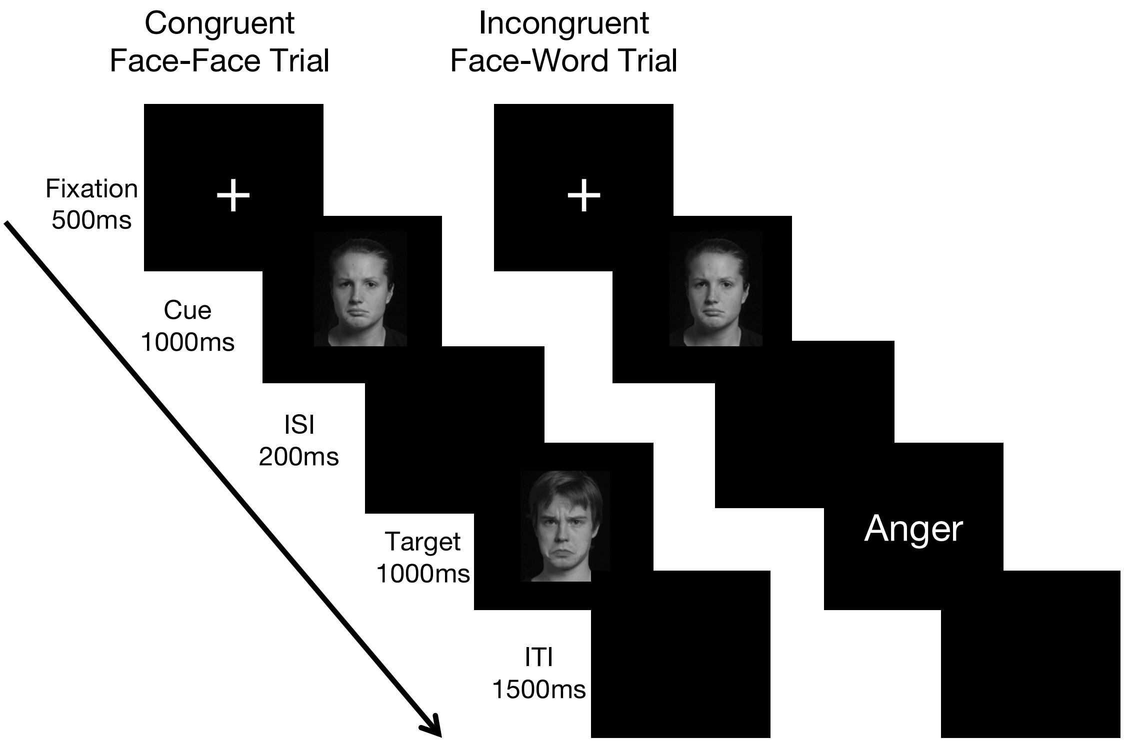


Fig. 1. Study 1 task schematic, displaying two of four possible conditions. Participants saw sequentially presented "cue" and "target" stimuli. While the target was onscreen, participants indicated whether or not it matched the emotion of the cue using a button press. Cue stimuli were always facial expressions of emotion. Targets could either be facial expressions of emotion or emotion category words (*target type*), and the emotions of the stimuli pair could either be identical or different (*congruence*).

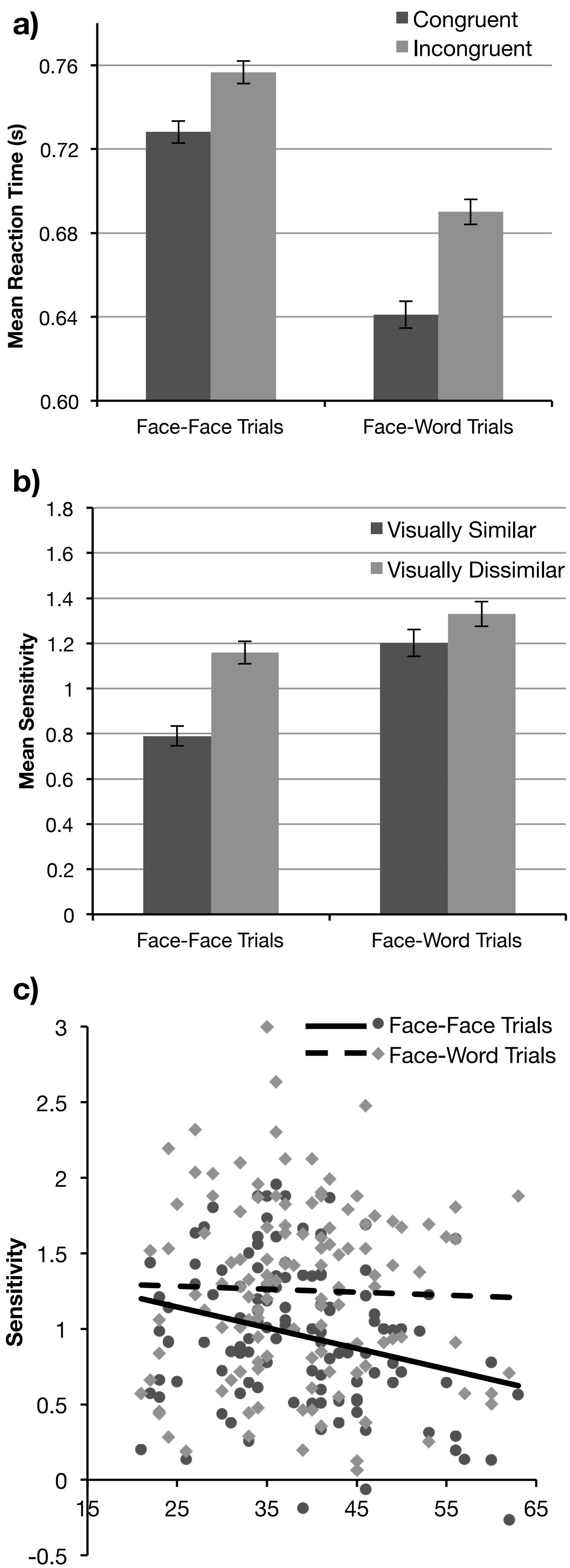


Figure 2

Fig. 2. a) Mean RTs for each condition; b) mean sensitivity scores for each target type condition, split by visual similarity sub-conditions; c) scatter plots of each participant's alexithymia scores and their average sensitivity for face-face and face-word pairs. Error bars represent standard error of the mean (SEM).

Figure 3

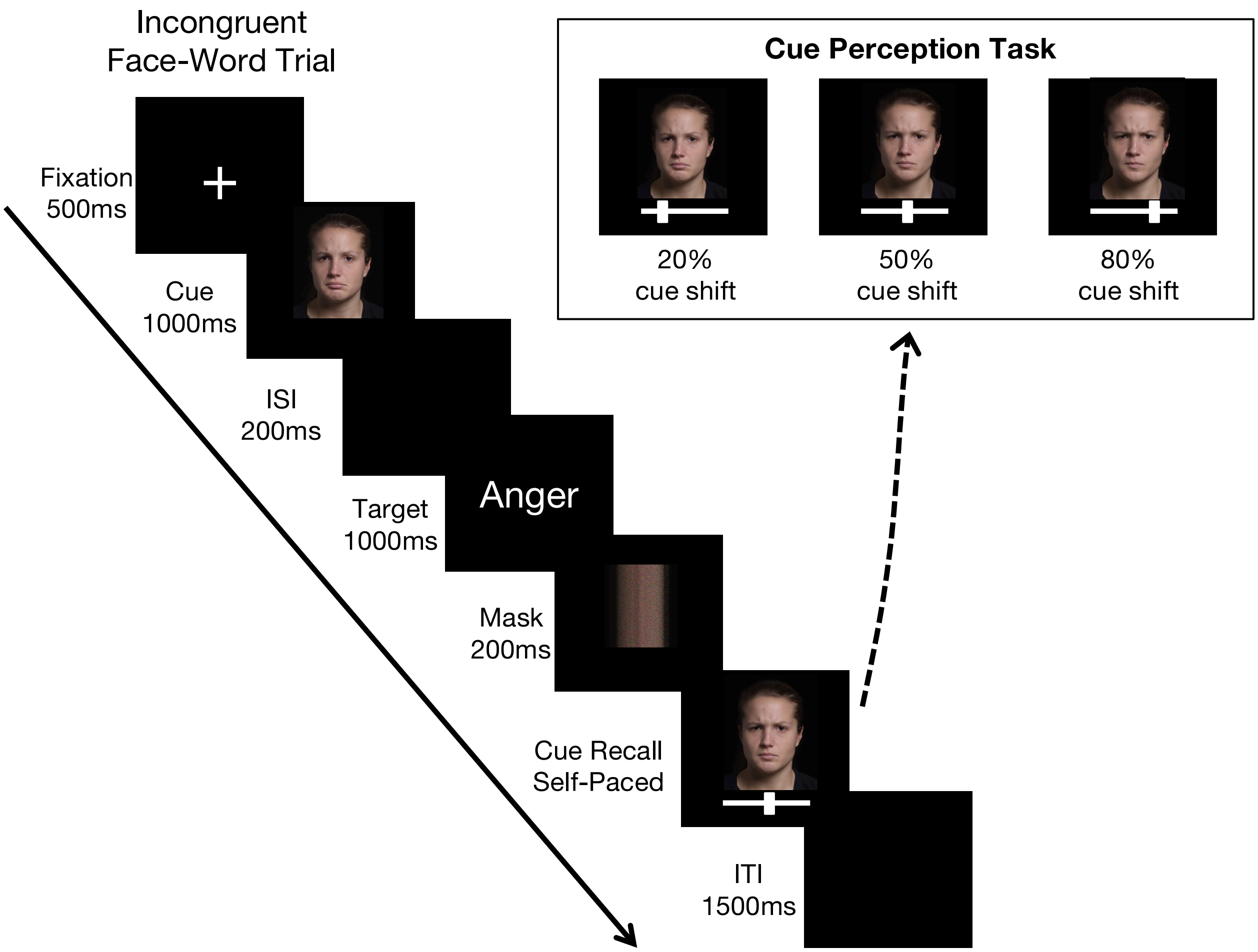


Fig. 3. Task schematic for Study 2, showing one of four possible conditions. Participants completed trials identical to those of Study 1, but following each trial, they completed the cue perception task. In this task, participants moved a slider to change the emotion expressed by the face until they found the image they believed had been shown as the cue face. Their selection was then translated into a percentage of cue shift. (ISI = Interstimulus interval; ITI = Intertrial interval.)

Figure 4

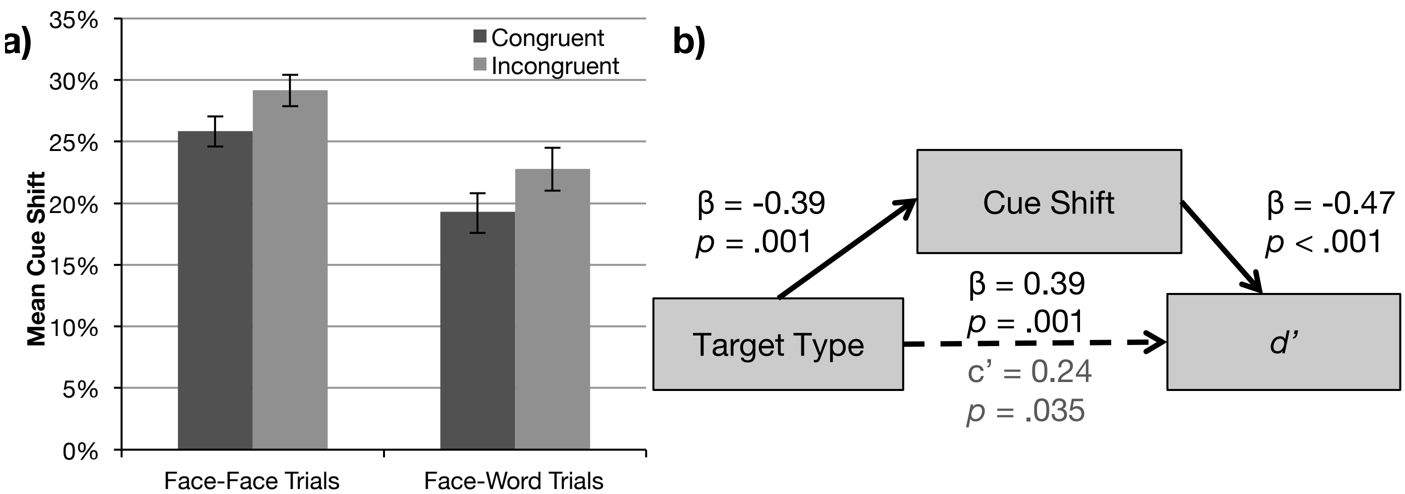


Fig. 4. a) Mean cue shift for each condition, error bars represent SEM; b) mediation model testing whether target type influences sensitivity because it reduces cue shift, with standardized betas.

SUPPLEMENTAL MATERIALS

Study 1 Method Supplement

Participants were recruited from Stanford University and the surrounding area for an "Emotion Perception Study." The emotion perception task (coded in MATLAB) included a total of 160 trials, evenly split into two blocks, and participants completed 8 practice trials before each block. Each trial began with a 500ms fixation cross, followed by the presentation of a "cue" face for 1000ms. Next, a blank screen was displayed for 200ms before the "target" appeared for another 1000ms. While the target was on the screen, participants indicated whether or not the cue and target matched using the keys "a" or "l." To control for handedness, the designation of "a" and "l" to responses of "match" and "no-match" were reversed between blocks, and initial mappings were counterbalanced across participants. Following the judgment stimulus, a blank screen was presented for 1500ms between trials. If participants did not respond while the judgment stimulus was displayed, the worlds "No Response" appeared onscreen for the first 1000ms of the intertrial interval. We selected only negative faces from the emotions proposed to be biologically "basic" (see Ekman & Cordaro, 2011). We did not include positive emotions since "happiness" is often considered the only "basic" positive emotion and including a single positive emotion amongst several negative emotions can inflate participants' performance for those trials relative to others (e.g., Lindquist et al. 2006). Using normed ratings of intensity (Lindquist, Gendron, & Barrett, unpublished data), we ensured that all stimuli were rated as at least moderately intense (≥ 2.5 on a 5-point scale). We used all four expressions (anger, disgust, fear and sadness) made by 59 models. The expressions from 2 other models were combined to complete our stimuli set.

Analysis of Study 1 Stimuli

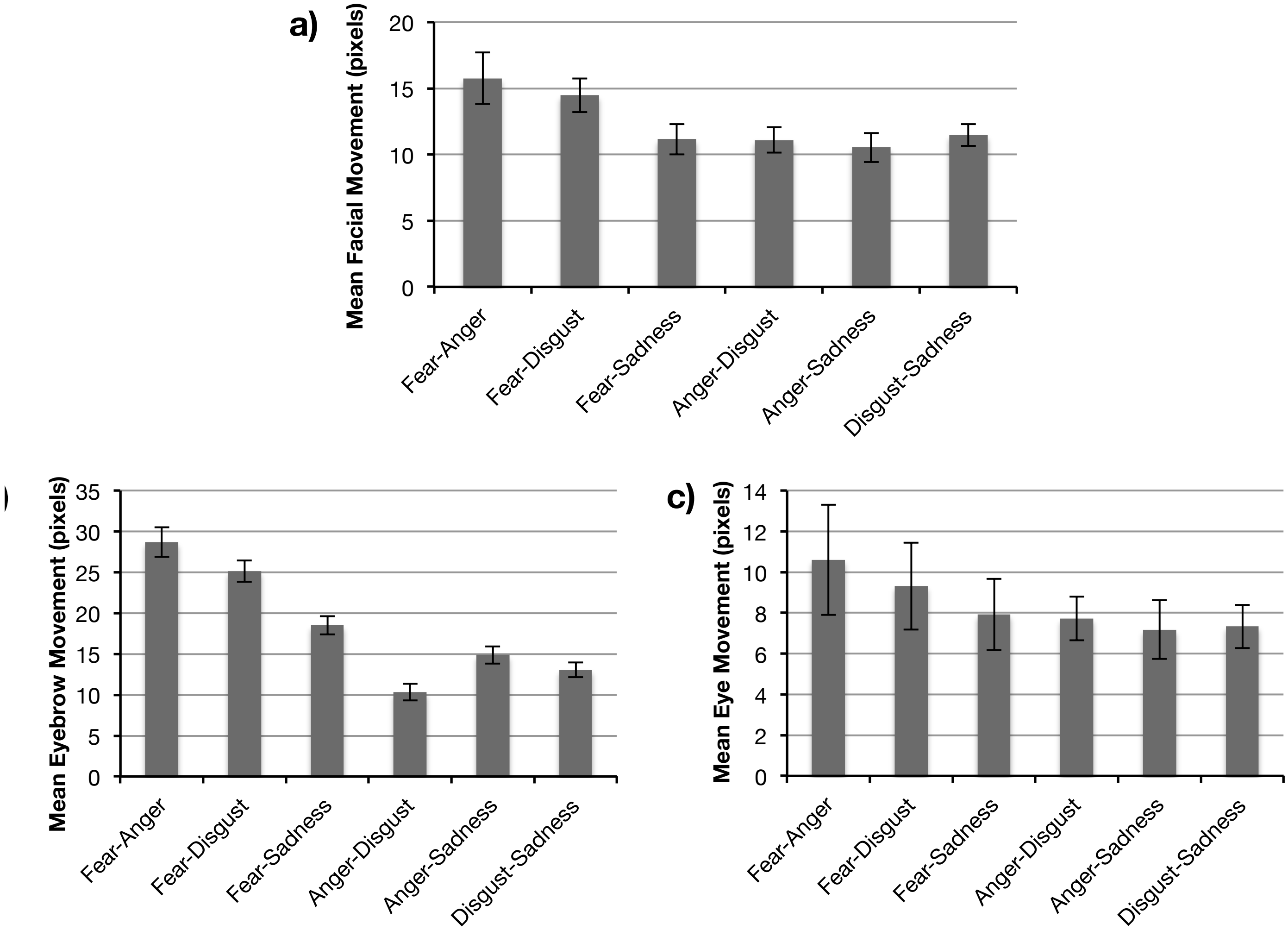
As cited in the article's main text, several studies demonstrate that posed expressions of fear differ substantially from other faces, particularly because they involve widened eyes and visible sclera (Ekman, Friesen, & Hager, 2002). For example, Jack, Garrod, & Schyns (2014) demonstrate that fear and surprise faces are perceived as distinct from other expressions because they involve raised upper eyelids early in the expression; Adolphs, et al. (2005) show that attention to the eyes is critical for identifying fear expressions; Somerville & Whalen (2006) discovered that undergraduate participants report hardly ever seeing fear faces in their everyday lives; and neuroimaging studies demonstrate that visual attention (modulated via amygdala projections to visual cortex) is disproportionately directed to fearful faces (Whalen, et al., 2004), especially when the gaze is direct (Adams et al., 2003). Hence, fear faces seem to be visually and psychologically dissimilar from those of other emotions.

To further corroborate this position, we performed a metric analysis of the facial stimuli used in Study 1. We used Abrosoft FantaMorph to define the location of critical facial structures on stimuli from the 59 models for whom we used all four emotion expressions. Using a total of 74 points, we identified the location of the eyebrows (16 points), eyes (22 points), nose (16 points) and mouth (20 points) in an x/y coordinate system. We then used coordinate geometry to quantify the distance each model's facial features had to move for them to transition between each emotion expression. This resulted in six *facial movement* measures for each model (i.e., fear-anger, fear-disgust, fear-sadness, anger-disgust, anger-sadness and disgust-sadness).

One-way Repeated Measures ANOVAs indicated that *facial movement* scores were not equal across emotion pairs [$F(5,290) = 8.35$, $p < .001$, $\eta_p^2 = .13$, **Figure S1a**]. Visual inspection of these data suggested that *facial movement* may be greater when transitioning between fear and a non-fear emotion than when transitioning between two non-fear emotions. We thus computed two *facial movement* composite scores for each model, one which averaged *facial movement* for emotion pairs involving fear (i.e., fear-anger, fear-disgust or fear-sadness) and one for pairs not involving fear (i.e., anger-disgust, anger-sadness or disgust-sadness). A paired-samples t-test indicated that overall *facial movement* is indeed higher when models must transition from fear to a non-fear emotion than when they transition between two non-fear emotions [$t(58) = 3.39$, $p = .001$, $d = .32$].

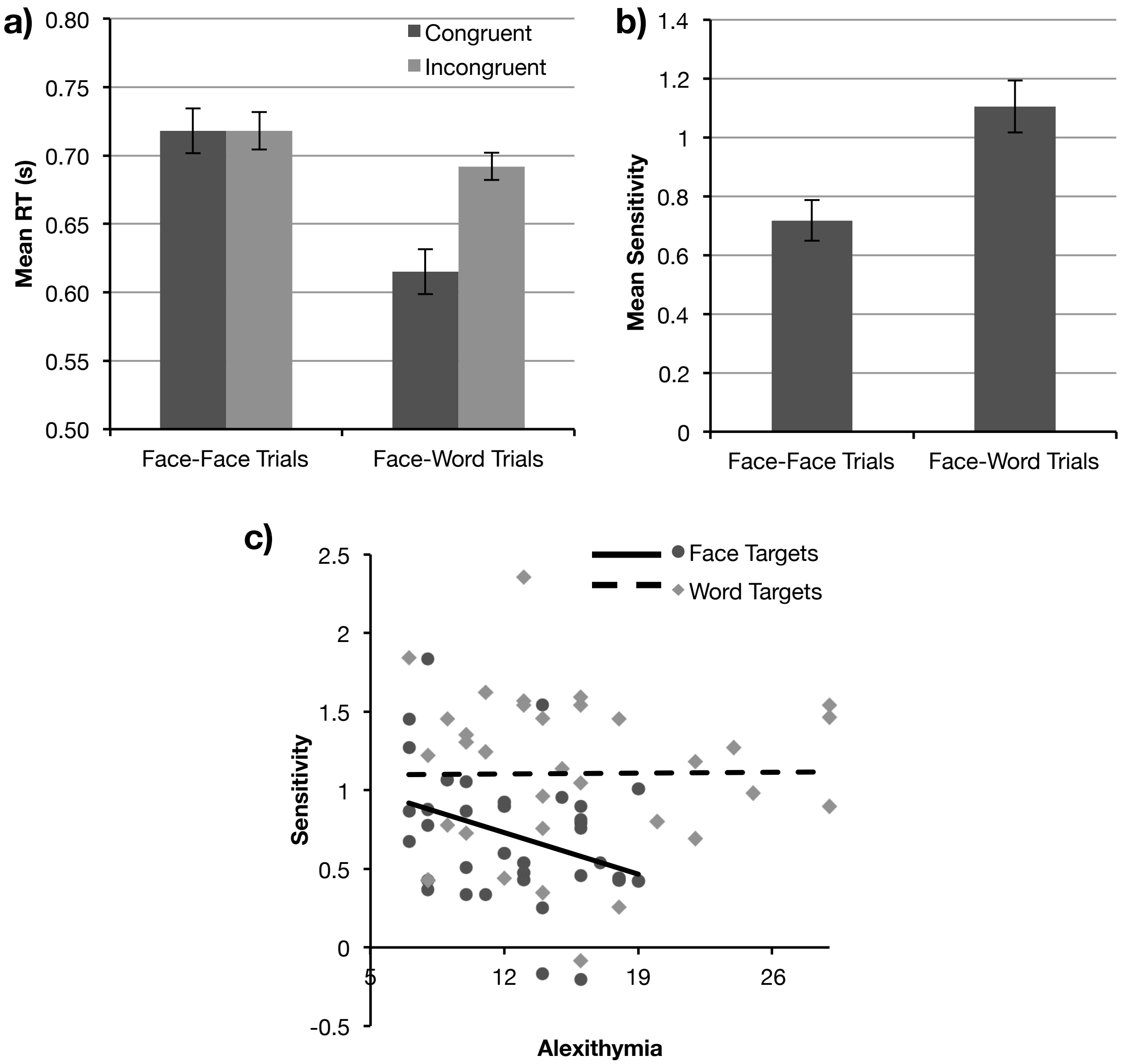
Follow-up analyses focused on testing whether the eyebrow and eye regions were driving these dissimilarities across expressions. Hence, we created *eyebrow movement* and *eye movement* scores by calculating the distance that points surrounding the eyebrows and eyes had to travel to transition between each emotion. As expected, one-way repeated measures ANOVAs demonstrated that average *eyebrow movement* [$F(5,290) = 37.24$, $p < .001$, $\eta_p^2 = .39$, **Figure S1b**] and *eye movement* [$F(5,290) = 3.80$, $p = .002$, $\eta_p^2 = .39$, **Figure S1c**] scores were significantly different across emotion pairings. We then averaged *eyebrow movement* and *eye movement* scores for pairs involving fear and those not involving fear. Paired-samples t-tests revealed that both the eyebrows [$t(58) = 7.25$, $p < .001$, $d = .91$] and the eyes [$t(58) = 2.53$, $p = .01$, $d = .22$] moved more when models transitioned from fear to a non-fear emotion than when they transitioned between two non-fear emotions. Hence, a metric analysis of our own stimuli suggests that fear faces are less *visually similar* to expressions of other emotions, particularly due to movement in the eye region. These results corroborate our claim that fear expressions are dissimilar from other expressions and should consequently be analyzed separately in Study 1 and excluded in Study 2.

Supplemental Figure 1 (Analysis of Study 1 Stimuli)



Supplemental Fig. 1. Average *facial movement* required for models to transition between each emotion used as stimuli in Study 1. Transitions involving fear are grouped on the left illustrate the increased dissimilarity between fear and other emotions. Results are split according to facial region: a) movement of all points on the face, b) movement of points around the eyebrows only and c) movement of points around the eyes only. Error bars represent SEM.

Supplemental Figure 2 (Analyses of Study 2 Repetition-Priming Task)



Supplemental Fig. 2. Analyses of dependent variables in Study 2's Repetition-Priming Task. a) Mean RTs for each condition; b) mean sensitivity scores for face-face and face-word pairs; c) scatter plot showing each participant's average sensitivity score and their alexithymia score, with separate trend lines for participants in the face-face and face-word conditions. Error bars represent SEM.

Supplemental Table 1: Means and standard deviations of signal detection measures

Condition	HR	FAR	*d'*	*c*
Study 1				
Face-Face Trials	0.68 (0.13)	0.33 (0.13)	0.96 (0.47)	-0.01 (0.31)
Visually Similar	-	0.39 (0.15)	0.79 (0.50)	-0.10 (0.33)
Visually Dissimilar	-	0.27 (0.15)	1.16 (0.56)	0.09 (0.33)
Face-Word Trials	0.64 (0.14)	0.21 (0.11)	1.26 (0.60)	0.24 (0.28)
Visually Similar	-	0.23 (0.12)	1.20 (0.66)	0.21 (0.28)
Visually Dissimilar	-	0.19 (0.12)	1.33 (0.61)	0.28 (0.31)
Study 2				
Face-Face Trials	0.57 (0.14)	0.31 (0.12)	0.72 (0.42)	0.16 (0.31)
Face-Word Trials	0.63 (0.11)	0.24 (0.11)	1.11 (0.52)	0.21 (0.23)

Notes: HR = Hit Rate, FAR = False Alarm Rate, *d'* = sensitivity, *c* = bias, numbers in parentheses are standard deviations.