Language is a Unique Context for Emotion Perception

Cameron M. Doyle,[1] Maria Gendron,[2] & Kristen A. Lindquist[1]

1. Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill
2. Department of Psychology, Yale University

Address correspondence to:
Kristen A. Lindquist
Kristen.lindquist@unc.edu

Abstract

Access to words used to label emotion concepts (e.g., "disgust") facilitates perceptions of facial muscle movements as instances of specific emotions (see Lindquist & Gendron, 2013). However, it remains unclear whether the effect of language on emotion perception is unique or whether it is driven by language's tendency to evoke situational context. In two studies, we used a priming and perceptual matching task to test the hypothesis that the effect of language on emotion perception is unique to that of situational context. We found that participants were more accurate to perceptually match facial portrayals of emotion after being primed with emotion labels as compared to situational context or control stimuli. These findings add to growing evidence that language serves as context for emotion perception and demonstrate for the first time that the effect of language on emotion perception is not merely a consequence of evoked situational context.

Keywords: language, context, emotion perception, psychological construction

Imagine that you see an individual with raised eyebrows and a lowered jaw. How would you make meaning of those facial muscle movements as an instance of, say, surprise? Research suggests that situations can serve as helpful context for emotion perception (Aviezer, Dudarev, Bentin, & Hassin, 2011; Carroll & Russell, 1996). For example, situational context can influence the perceived meaning of posed facial expressions, such that participants are faster to categorize faces when they are accompanied by congruent emotional scenes (Righart & de Gelder, 2008). Developmentally, children have access to the visual scenes that accompany emotional facial portrayals from birth, yet there is little evidence that infants perceive facial portrayals as categorically distinct emotions (Ruba, Meltzoff, & Repacholi, 2019). Instead, children's ability to perceive facial portrayals as instances of discrete emotion categories is linked to their acquisition of emotion words such as "anger," "fear," "sadness," and "disgust" (Widen, 2013). Moreover, studies in adults show that language serves as a critical form of context that helps disambiguate the meaning of facial portrayals of emotion (Betz, Hoemann, & Barrett, 2019, see Lindquist & Gendron, 2013 for a review).

Despite substantial evidence for the role of language in emotion perception, it remains unclear whether the effect of language is unique or whether it is driven by language's tendency to evoke situational context. It could be that emotion words merely serve as cues for situations, and that situations themselves are more strongly associated with the ability to make meaning of facial portrayals as instances of emotion (Hess & Hareli, 2016). Alternatively, insofar as words cohere multiple representations of abstract concepts (Lindquist, MacCormack, & Shablack, 2015), it is possible that emotion words provide additional predictive information above and beyond the situational context. In the present studies, we used a within-subjects priming and

perceptual matching task to test the hypothesis that language is a unique context for emotion perception.

Method

Study 1

In Study 1, we used a novel task to test whether participants would be faster and/or more accurate in the perceptual matching of faces after being primed with words, scenes, or control stimuli. See Figure 1. On a given trial, online participants were primed with context in the form of an emotion label (e.g., "sadness" or "disgust"), an emotionally evocative scene, or a blank control stimulus for 250ms (*priming phase*). Following the context prime, participants viewed a target image of an individual depicting a facial portrayal of either sadness or disgust (*target phase*). Finally, participants viewed two images of that same individual depicting facial portrayals of sadness *and* disgust (*test phase*). Participants selected from the two options the target image that they had previously seen in the *target phase*. We took average accuracy scores and response latencies for each of the three within-subjects priming conditions and tested whether language serves as a more effective prime for perceptual matching of emotion on faces.

Participants

Participants were 174 online workers recruited via Amazon Mechanical Turk ($M_{age} = 39.33$, $SD_{age} = 13.36$, 59% female). Because emotion perception differs across languages and cultures (Gendron, Roberson, van der Vyver, & Barrett, 2014; Jack, Garrod, Yu, Caldara, & Schyns, 2012), eligibility was restricted to native English speakers who were born in the United States. Participants were compensated in a manner approved by the Institutional Review Board of the University of North Carolina at Chapel Hill.

Given that online participants are more likely to yield problems due to inattention (Goodman, Cryder, & Cheema, 2013), we computed average response latencies as a metric of whether participants were attending to the task. Ten participants had extremely short (< 250ms) or extremely long (> 2000ms) average response latencies. We reasoned that participants with average response latencies of less than 250ms may have been making random selections in order to complete the task as quickly as possible, and participants with average response latencies of greater than 2000ms were likely not paying singular attention to the task. We removed these ten participants from analyses to ensure that our final sample only included attentive participants (final $N = 164$).

Stimuli

The present study employed three types of priming stimuli (language, situation, and control), as well as stimuli depicting facial portrayals of emotion. Language prime stimuli were the English-language emotion labels "Sadness" and "Disgust" presented in white font on a black background. Situation prime stimuli were emotionally evocative scenes drawn from the Nencki Affective Picture System (NAPS; Marchewka, Żurawski, Jednoróg, & Grabowska, 2014) presented on a black background. The NAPS images have been rated by a separate set of participants on the extent to which each image evokes a particular emotional response. We selected 30 images that were rated highly for their effectiveness in evoking feelings of sadness and disgust (15 images from each category). We chose to investigate the categories *Sadness* and *Disgust* because they tended to have less cross-classification of emotion ratings (i.e., there is more agreement among raters in terms of the ability of these images to evoke sadness and disgust). We also included a blank control stimulus, which was a white rectangle centered on a

black background. The purpose of including this control stimulus was so that it could be used as a baseline against which to compare the language and situation prime trials.

Facial portrayals of the English-language emotion categories *Sadness* and *Disgust* were drawn from the IASLab Face Set (https://www.affective-science.org/face-set.shtml). The IASLab Face Set contains images of different identities displaying non-caricatured facial muscle movements associated with English language emotion concepts. As with the NAPS images, the IASLab faces were rated by an independent sample of participants based on the extent to which they successfully portrayed particular emotions. We selected images of five identities who had the highest average ratings for their ability to portray facial muscle movements associated with sadness and disgust. Only female identities were used because they tended to have higher ratings than male identities, and we did not wish to introduce additional variance by including both female and male identities.

Procedure

Participants completed 120 trials of our novel priming and perceptual matching task. The task was programmed and presented using Inquisit 4.0 software (https://www.millisecond.com/), which has been shown to provide millisecond-level accuracy in response time data (De Clercq, Crombez, Buysse, & Roeyers, 2003). Each trial consisted of three phases: a within-subjects *priming phase,* a *target phase*, and a *test phase*. Each phase is described in detail below.

Priming Phase. Participants were primed with either a word (language prime trials) an emotionally evocative scene (situation prime trials) or a blank white rectangle (blank control trials) for 250ms. Language prime stimuli were the words "Sadness" or "Disgust," and situation prime stimuli were scenes meant to evoke feelings of sadness (e.g., a dying elderly person) or

disgust (e.g., an infected skin wound). Participants completed 40 language trials, 40 situation prime trials, and 40 blank control trials. Trials were randomly presented without replacement.

Target Phase. Following the *priming phase,* participants completed a *target phase* in which they viewed facial portrayals of *Sadness* or *Disgust* on a black background for 300ms. For 75% of the language and situation prime trials, facial portrayals were congruent with the word or scene primes that participants saw in the *priming phase*. The remaining 25%—as well as 25% of the blank control trials—were paired with neutral foils (i.e., an image from the IASLab set of one of the five identities displaying a neutral expression). Neutral foils were included in the *target phase* to ensure that participants engaged with the task. If all trials included congruent priming and target stimuli, participants might learn that they could simply select the facial portrayal that best matched the language or situation prime without paying attention to the target face.

Test Phase. The final phase of each trial was a *test phase* in which participants viewed two images of the same identity that they had seen in the *target phase* and engaged in a two-alternative forced choice task where they selected the facial portrayal they had seen in the *target phase.* One of the images was the exact face they had seen the individual making in the *target phase,* and the other image was a facial portrayal of either sadness or disgust. For example, if the target face had been a facial portrayal of sadness, the two choices in the *test phase* would be that same image of sadness, and an image of the same identity depicting a facial portrayal of disgust. On trials in which the target face had been a neutral foil, the neutral foil they had seen in the *target phase* was paired with an image of that same identity portraying either sadness or disgust. Images were presented side-by-side on a black background, and participants selected which face they had seen in the *target phase* in a self-paced manner. Participants indicated their choice using a button press on their keyboard (i.e., "1" key for the face on the left and "9" key for the face on

the right). Image placement for the target face (left v. right side of the screen) was randomized across trials. Following their selection, participants were presented with a white fixation cross on a black background for 500ms which signaled that the next trial was about to begin.

Figure 1

Schematic of trial procedure used in Studies 1 and 2
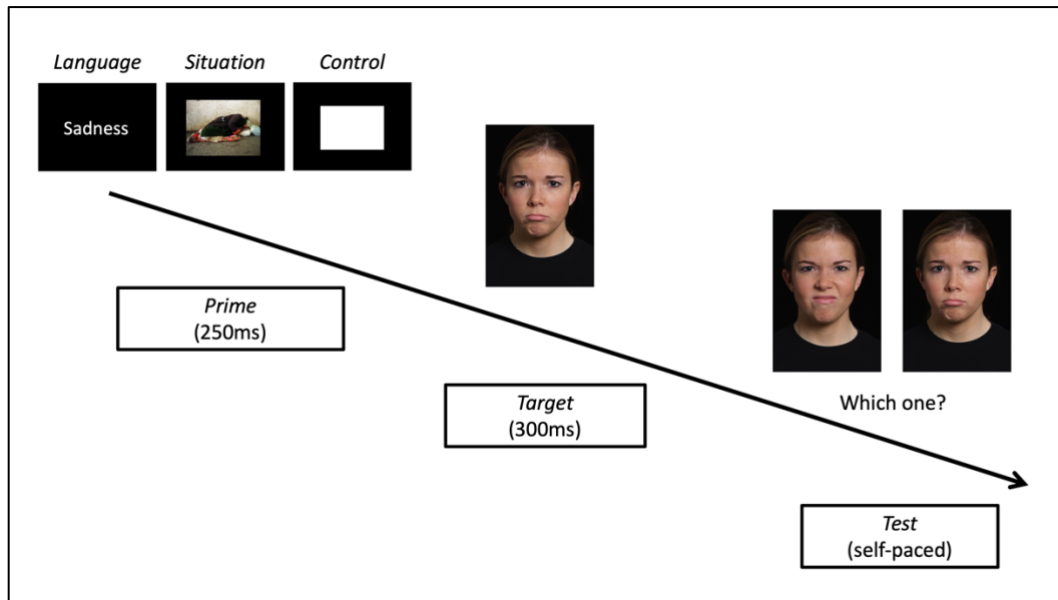
Data Analysis

Prior to data analysis, we excluded all trials in which the target image was a neutral foil because we had no substantive interest in these trials. As noted above, neutral foils were included simply to ensure that participants were engaging with the task (i.e., paying attention to the faces presented in the *target phase* rather than simply basing their choices in the *test phase* on the language or situation primes presented in the *priming* phase). After removing the neutral foil trials, we were left with 90 trials from each participant (30 language prime trials, 30 situation prime trials, and 30 blank control trials).

We used R Statistical Software to analyze participants' accuracy across all 90 trials, as well as their response latencies for trials in which they chose the correct face in the *test phase.*

We first computed an average of participant's accuracy scores (i.e., the percentage of trials in which they chose the correct face) for each of the three priming conditions (i.e., language, situation, and control). We then conduced a two-way analysis of variance to investigate differences in the extent to which the language, situation, and control primes facilitated accurate perceptual matching of the faces viewed in the *target phase*, and whether accuracy differed by emotion category. Following our analyses of participants' accuracy, we computed average response latencies for accurate trials and conducted another two-way analysis of variance assessing whether response latencies differed by priming condition or emotion category. Data and code used for analysis are publicly available at https://osf.io/v83xr/.

Results and Discussion

An analysis of accuracy scores revealed a main effect of priming condition, $F(2, 163) =$ 6.48, $p = .002$, $\eta_p^2 = .02$. Post hoc analyses using Tukey's Honestly Significant Difference (HSD) test revealed that accuracy was greater in the language priming condition ($M = 93.5\%$) relative to situation (89.2%) and control (90.5%), ps $< .001$. There was a marginally significant difference between situation and control, $p = .096$. There was also a main effect of emotion category, $F(1, 163) = 7.11$, $p = .008$, $\eta_p^2 = .01$ such that accuracy was greater for disgust trials ($M = 92.5\%$) relative to sadness trials ($M = 89.7\%$). There was no interaction between priming condition and emotion category, $p = .135$ (see Figure 2).

An analysis of response latencies for correct trials revealed an unpredicted pattern of effects. There was no main effect of priming condition ($p = .317$), and no main effect of emotion category ($p = .125$). There was, however, a significant interaction between priming condition and emotion category, $F(2, 163) = 3.40$, $p = .034$, $\eta_p^2 = .01$. Post hoc analyses using Tukey's HSD

test revealed that when primed with situational context, participants were faster to respond accurately on disgust trials ($M = 851$ms) as compared to sadness trials ($M = 921$ms), $p < .001$.

Study 1 showed that when facial portrayals of emotion are primed with linguistic context, participants are more accurate to perceptually match those faces in a two-alternative forced choice task. However, response latencies for images primed by situational context differ by emotion category. Visual comparison of the situation prime images from each category suggested that the sadness images were more visually complex (i.e., there were more visually salient points of interest present in the scene), whereas disgust images tended to have only one central image. For example, a sad image of an elderly person in a hospital room has more salient points of interest than a disgusting image of an infected skin wound. This difference in complexity may have contributed to increased response latencies for sadness as compared to disgust images on situation prime trials. In addition, the finding that participants were more accurate for disgust relative to sadness trials may also be related to this difference in visual complexity. Although there was not a significant interaction between priming condition and emotion category for accuracy, the pattern of means was consistent with the interpretation that sad images might have been more visually complex; the difference in accuracy between disgust and sadness trials was greatest for the situation prime condition.

A second caveat of Study 1 is that we used 15 situation primes per emotion category, but only one language prime (i.e., the words "Sadness" and "Disgust" themselves). It is possible that the main effect of priming condition on accuracy is confounded by this difference in the number of stimulus items across priming conditions. Participants may have quickly learned that there were only two different words used, but there were substantially more images that they would have had to attend to. In this way, accuracy may have been greater in the language priming

condition simply because the task was less attentionally burdensome to participants. We attempted to control for both of these issues in Study 2.

Study 2

In Study 2, we replicated and extended Study 1 using images from the recently published Complex Affective Scene Set (COMPASS; Weierich, Kleshchova, Rieder, & Reilly, 2019), which controls for the degree of visual complexity present in the images. We additionally included synonyms for emotion words so that we had the same number of primes for each category (i.e., five language primes and five situation primes per category). Synonyms for emotion words were selected from a list of thesaurus entries for "sadness" and "disgust." With the exception of these changes to situation and language prime stimuli, all other aspects of the Study 2 procedure were identical to Study 1.

Participants

Participants were 174 online workers recruited via Amazon Mechanical Turk ($M_{age} =$ 38.63, $SD_{age} =$ 12.95, 50% female). As in Study 1, eligible participants were native English speakers who were born in the United States. Participants were again compensated in a manner approved by the Institutional Review Board of the University of North Carolina at Chapel Hill. Three participants had extreme average response latencies (i.e., less than 250ms or greater than 2000ms). These participants were removed from analyses to ensure that our final sample only included attentive participants (final $N =$ 171).

Stimuli

As in Study 1, we used three types of priming stimuli (language, situation, and control), as well as target stimuli depicting facial portrayals of emotion. Study 2 stimuli included the same blank control stimulus and the same target images of five identities drawn from the IASLab Face

Set. However, the language and situation prime stimuli used in Study 2 differed from that of

Study 1. Rather than simply using the English-language emotion labels "Sadness" and "Disgust"

as language primes, we added four synonyms for each emotion category. This change was made

in an effort to match the number of stimuli used in each of the experimental conditions (i.e., the

language prime condition and the situation prime condition). Words used as language primes

during sadness trials in Study 2 were "Sadness," "Misery," "Despair," "Sorrow," and "Anguish."

Words used as language primes during disgust trials in Study 2 were "Disgust," "Repulsion,"

"Nausea," Sickening," and "Aversion." As in Study 1, the language primes were presented in

white font on a black background.

In addition to the above changes to language prime stimuli, we also altered situation

prime stimuli used in Study 2. We were concerned that our findings in Study 1 may have been

affected by potential differences across emotion categories in the visual complexity of the

images used. We thus used images from the COMPASS set (Weierich et al., 2019), which

controls for the degree of visual complexity in images. In contrast to the NAPS images used in

Study 1, the COMPASS images were not previously rated by a separate sample of participants

based on the extent to which they evoked discrete feelings such as sadness and disgust. We thus

normed the stimuli using a separate sample of online workers recruited from Amazon

Mechanical Turk. Participants ($N = 100$) were native English speakers who were born in the

United States ($M_{age} = 38.41$, $SD_{age} = 12.45$, 45% female). Participants rated the COMPASS

images by choosing which of several emotion categories (i.e., Anger, Disgust, Fear, Happiness,

Sadness, and Surprise) each image evoked. For a given image, we computed the frequency with

which each emotion was selected. If the majority of participants rated a given image as evoking

an instance of sadness, it was classified as a "sadness image," whereas if the majority of

participants rated a given image as evoking an instance of disgust, it was classified as a "disgust image," and so on. We selected the five images that were rated most frequently as evoking feelings of sadness, and the five images that were rated most frequently as evoking feelings of disgust. Thus, we had five situation prime stimulus items for each emotion category that were matched in terms of visual complexity.

Data Analysis

Prior to data analysis, we again removed the trials in which the target face had been a neutral foil, leaving us with 90 trials per participant (30 language prime trials, 30 situation prime trials, and 30 blank control trials). As in Study 1, we analyzed participants' accuracy across all 90 trials, as well as their response latencies for trials in which they chose the correct face in the *test phase.* Study 2 also afforded us an additional opportunity to assess differences in accuracy for the language prime condition based on whether the word was a "basic-level" emotion label (i.e., "Sadness" or "Disgust") or a "subordinate-level" emotion label (i.e., one of the four synonyms for each emotion). "Basic-level" categories are those that are learned first during language acquisition, used most frequently in discourse, and are more generalizable than subordinate-level categories (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). In emotion, basic-level categories may name the most frequently occurring or prototypical representation of a facial portrayal of emotion. We thus predicted that basic-level emotion labels would serve as superior primes for emotion categories as compared to subordinate-level categories.

Results and Discussion

An analysis of accuracy scores in Study 2 revealed a main effect of priming condition, $F(2, 170) = 4.39$, $p = .013$, $\eta_p^2 = .01$. Post hoc analyses using Tukey's HSD test revealed that, as predicted, accuracy was greater in the language prime condition ($M = 92.6\%$) relative to situation

($M = 90.0\%$) and control ($M = 90.9\%$), ps $< .001$ and $.023$, respectively. There was no difference in accuracy between situation and control, $p = .313$. In contrast to Study 1, there was no main effect of emotion category, $p = .465$. Finally, there was no interaction between priming condition and emotion category, $p = .706$ (see Figure 2).

Figure 2

Means and 95% confidence intervals for participants' accuracy in Studies 1 and 2
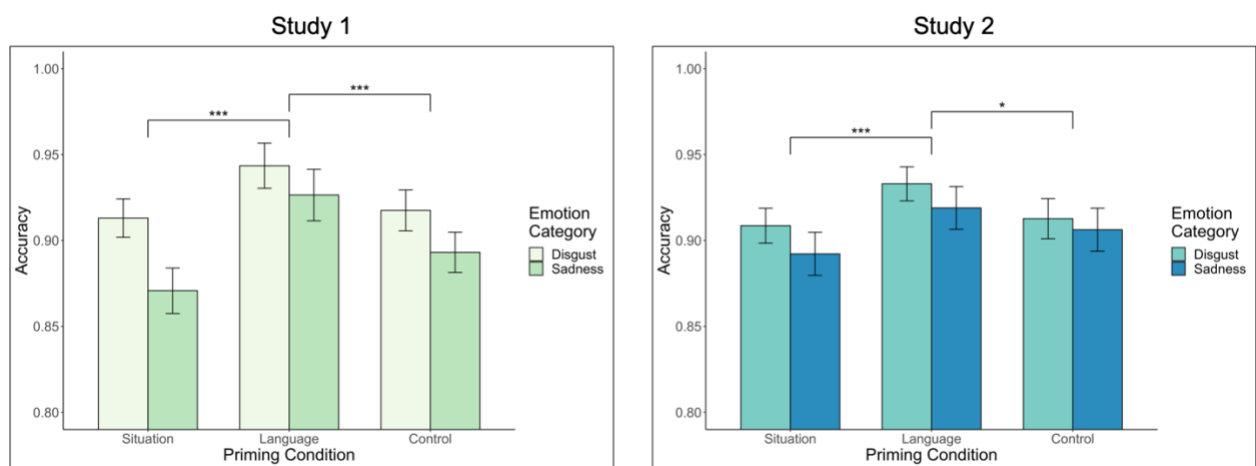
Note: *** indicates p < .001 and * indicates p < .05

An analysis of response latencies for correct trials in Study 2 showed that there was no main effect of priming condition ($p = .221$) and no main effect of emotion category ($p = .330$). Critically, in contrast to Study 1, we found that when using stimuli that are matched for visual complexity, we no longer had a significant interaction between priming condition and emotion category, $p = .294$.

The design of Study 2 also afforded an exploratory analysis of the differential impact of basic-level v. subordinate-level emotion labels on accurate perceptual matching of facial portrayals of emotion. We conducted a one-way analysis of variance comparing accuracy across the situation, control, basic-level word, and subordinate-level word priming conditions. As

predicted, we found a significant effect of priming condition, $F(3, 170) = 10.62$, $p < .001$, $\eta_p^2 =$ .06. Post hoc analyses using Tukey's HSD test revealed that accuracy was greater in the basic-level word prime condition ($M = 94.2\%$) relative to the situation prime ($M = 90.0\%$) and control conditions ($M = 90.9\%$), ps $< .001$. Accuracy was marginally greater in the basic-level word condition as compared to the subordinate-level word condition, $p = .051$. Accuracy was greater in the subordinate-level word condition as compared to the situation condition, $p = .031$. Finally, there was no difference in accuracy between the situation and control conditions, nor between the subordinate-level word and control conditions, ps $= .660$ and $.378$, respectively.

Study 2 replicated and extended the findings of Study 1 by better controlling the complexity of disgust- v. sadness-congruent scenes and using an equal number of stimuli across the experimental priming conditions (i.e., language and situation). Study 2 also demonstrated that the specific effect of basic-level emotion labels (i.e., "Sadness" and "Disgust") on accuracy for perceptual matching of facial portrayals of emotion is stronger than that of subordinate-level labels (e.g., "Sorrow" and "Repulsion").

General Discussion

In two studies, we demonstrate that language exerts a unique effect on emotion perception. These findings add to constructionist hypotheses that words for emotion categories help people acquire, cohere, and use stored predictions about emotion concepts to make meaning of facial portrayals of emotion (Doyle & Lindquist, 2017, 2018). Words are thought to serve as special cues to their referents because they do not vary across instances of the categories they label (Edmiston & Lupyan, 2015). Although many instances of the category anger can look quite different from one another, the word "anger" refers to each instance invariably. This is perhaps how language binds multiple representations of emotion perception together as members of the

same category (Lindquist et al., 2015). For example, the word "sadness" is associated with representations of the feelings, situations, behaviors and cognitions experienced during instances of sadness. A word may thus prime a broad set of predictions (including the visual situation, but also behaviors, vocalizations, sounds, smells, etc.) that can be used to make meaning of a percept. In contrast, a visual scene, although in the same modality as visual facial portrayals of emotion, may be inferior because it does not prime such a broad set of predictions. We also found that basic-level categories were better primes than subordinate-level categories. Basic-level categories are those thought to be most cognitively efficient for categorization because maximize the similarity of within category members relative to other category members (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Future work should examine how individual differences in emotion word use differentially predict emotion perception.

Limitations

These findings are the first to our knowledge to demonstrate that language serves as a unique context for emotion perception above and beyond merely evoking situational context. However, this work is not without its limitations. First, it is possible that language primes emotion perception because the word stimuli were easier to process than the scene stimuli. However, we did not observe a main effect of priming condition on response latencies for accurate trials, and prior research demonstrates that scene stimuli can be processed in less than 100ms (Lowe, Rajsic, Ferber, & Walther, 2018), calling this interpretation into question. Moreover, language is a more abstract stimulus than situational context and invokes multi-modal representations that are inconsistent with the modality of the target stimulus. It thus stands to reason that language could cause relatively *more* interference in perception of facial portrayals than a visual scene. Second, it is possible that language served as a better prime for emotion

perception because people implicitly used emotion category labels to respond to the task. We mitigated this possibility by using a task that does not explicitly require language (i.e., participants performed a perceptual matching task that did not use emotion category labels). This is consistent with past work demonstrating that emotion words even impact perceptual priming of facial portrayals of emotion, which occurs in visual cortex (Gendron, Lindquist, Barsalou, & Barrett, 2012).

Conclusion

These findings add to growing evidence that language serves as context for emotion perception and demonstrate for the first time that the effect of language on emotion perception is not merely a consequence of evoked situational context. Future research should investigate the extent to which language and situational context may interact during the perception of emotion on faces.

References

Aviezer, H., Dudarev, V., Bentin, S., & Hassin, R. R. (2011). The Automaticity of Emotional Face-Context Integration. *Emotion*, *11*(6), 1406–1414. https://doi.org/10.1037/a0023578.The

Betz, N., Hoemann, K., & Barrett, L. F. (2019). Words Are a Context for Mental Inference. *Emotion*, *19*(8), 1463–1477.

Carroll, J. M., & Russell, J. A. (1996). Do Facial Expressions Signal Specific Emotions? Judging Emotion From the Face in Context. *Journal of Personality and Social Psychology*, *70*(2), 205–218.

De Clercq, A., Crombez, G., Buysse, A., & Roeyers, H. (2003). A simple and sensitive method to measure timing accuracy. *Behavior Research Methods, Instruments, and Computers*, *35*(1), 109–115. https://doi.org/10.3758/BF03195502

Doyle, C. M., & Lindquist, K. A. (2017). Language and emotion: Hypotheses on the constructed nature of emotion perception. In J.-M. Fernandez-Dols & J. A. Russell (Eds.), *The Science of Facial Expression*. New York: Oxford University Press.

Doyle, C. M., & Lindquist, K. A. (2018). When a word is worth a thousand pictures: Language shapes perceptual memory for emotion. *Journal of Experimental Psychology: General*, *147*(1), 62–73. https://doi.org/10.1037/xge0000361

Edmiston, P., & Lupyan, G. (2015). What makes words special? Words as unmotivated cues. *Cognition*, *143*, 93–100. https://doi.org/10.1016/j.cognition.2015.06.008

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

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*, *14*(2), 251–262. https://doi.org/10.1037/a0036052

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data Collection in a Flat World: The

Strengths and Weaknesses of Mechanical Turk Samples. *Journal of Behavioral Decision

Making*, *26*(3), 213–224. https://doi.org/10.1002/bdm.1753

Hess, U., & Hareli, S. (2016). The Impact of Context on the Perception of Emotions. In C. Abell

& J. Smith (Eds.), *The Expression of Emotion: Philosophical, Psychological and Legal

Perspectives* (pp. 199–218). Cambridge: Cambridge University Press.

https://doi.org/10.1017/CBO9781316275672.010

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), 7241–7244.

https://doi.org/10.1073/pnas.1200155109

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., MacCormack, J. K., & Shablack, H. (2015). The role of language in emotion:

predictions from psychological constructionism. *Frontiers in Psychology*, *6*, 444.

https://doi.org/10.3389/fpsyg.2015.00444

Lowe, M. X., Rajsic, J., Ferber, S., & Walther, D. B. (2018). Discriminating scene categories

from brain activity within 100 milliseconds. *Cortex*, *106*, 275–287.

https://doi.org/10.1016/j.cortex.2018.06.006

Marchewka, A., Żurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective

Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality,

realistic picture database. *Behavior Research Methods*, *46*, 596–610.

https://doi.org/10.3758/s13428-013-0379-1

Righart, R., & de Gelder, B. (2008). Recognition of facial expressions is influenced by emotional

scene gist. *Cognitive Affective & Behavioral Neuroscience*, *8*(3), 264–272.

https://doi.org/10.3758/CABN.8.3.264

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects

in natural categories. *Cognitive Psychology*, *8*(3), 382–439. https://doi.org/10.1016/0010-

0285(76)90013-X

Ruba, A. L., Meltzoff, A. N., & Repacholi, B. M. (2019). How Do You Feel? Preverbal Infants

Match Negative Emotions to Events. *Developmental Psychobiology*, *55*(6), 1138–1149.

Weierich, M. R., Kleshchova, O., Rieder, J. K., & Reilly, D. M. (2019). The Complex Affective

Scene Set (COMPASS): Solving the Social Content Problem in Affective Visual Stimulus

Sets. *Collabra: Psychology*, *5*(1), 53. https://doi.org/10.1525/collabra.256

Widen, S. C. (2013). Children's Interpretation of Facial Expressions: The Long Path from

Valence-Based to Specific Discrete Categories. *Emotion Review*, *5*(1), 72–77.

https://doi.org/10.1177/1754073912451492

Declarations

Conflicts of interest/Competing interests: The authors declare no conflicts/competing interests.

Ethics approval: This research was conducted in a manner approved by the Institutional Review Board of the University of North Carolina at Chapel Hill.

Consent to participate: All participants provided informed consent.

Consent for publication: N/A

Availability of data and material: Data are publicly available at https://osf.io/v83xr/.

Code availability: Code used for data analysis is publicly available at https://osf.io/v83xr/.