Facial and bodily actions speak louder than context: Inferring facial expressions of emotions in naturalistic images
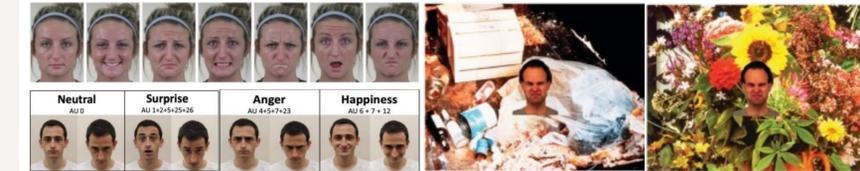
Ruoying Zheng[1] (r7zheng@ucsd.edu), Kevin Liu[1], Jiabin Zou[1], Chujun Lin[2,*]

[1]Department of Psychology, University of California San Diego; [2]Department of Psychology, Columbia University

UC San Diego

Introduction

Inferring others' emotions is fundamental for navigating the social world.
A combination of *cues from different channels* shapes emotion inferences.

Prior works: *isolated cue or artificial combined* cues (Du et al., 2014; Righart & de Gelder, 2008; Witkower & Tracy, 2022)

How do people utilize rich, multimodal cues in naturalistic contexts to interpret facial expressions of emotions?

Methods

Quantifying Naturalistic Cues from Large Datasets

Computational models + 1,093 U.S. representative participants quantified
4 channels · 16 fine-grained cues · 1125 naturalistic target persons
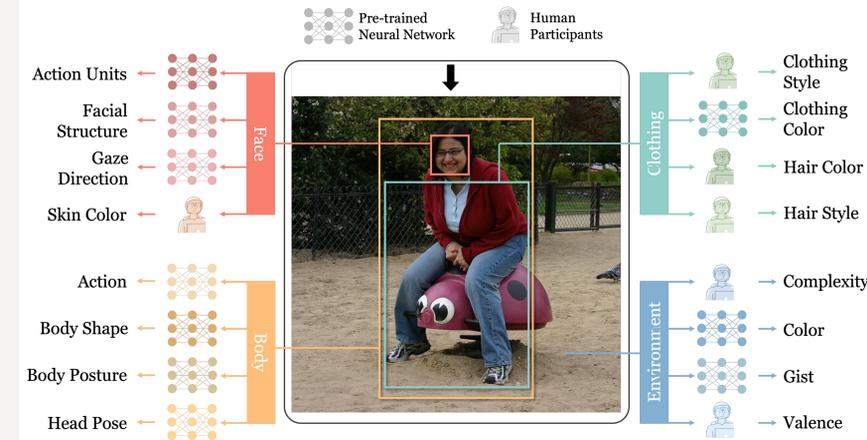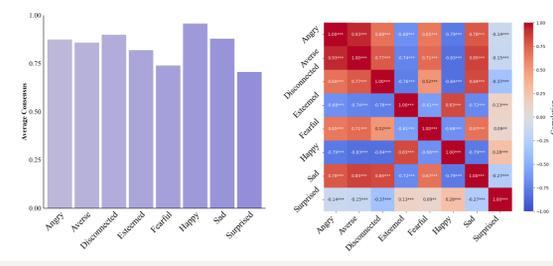

Paradigm

475 U.S. representative participants rated 8 facial expressions of emotions
Angry · Averse · Disconnected · Esteemed · Fearful · Happy · Sad · Surprised
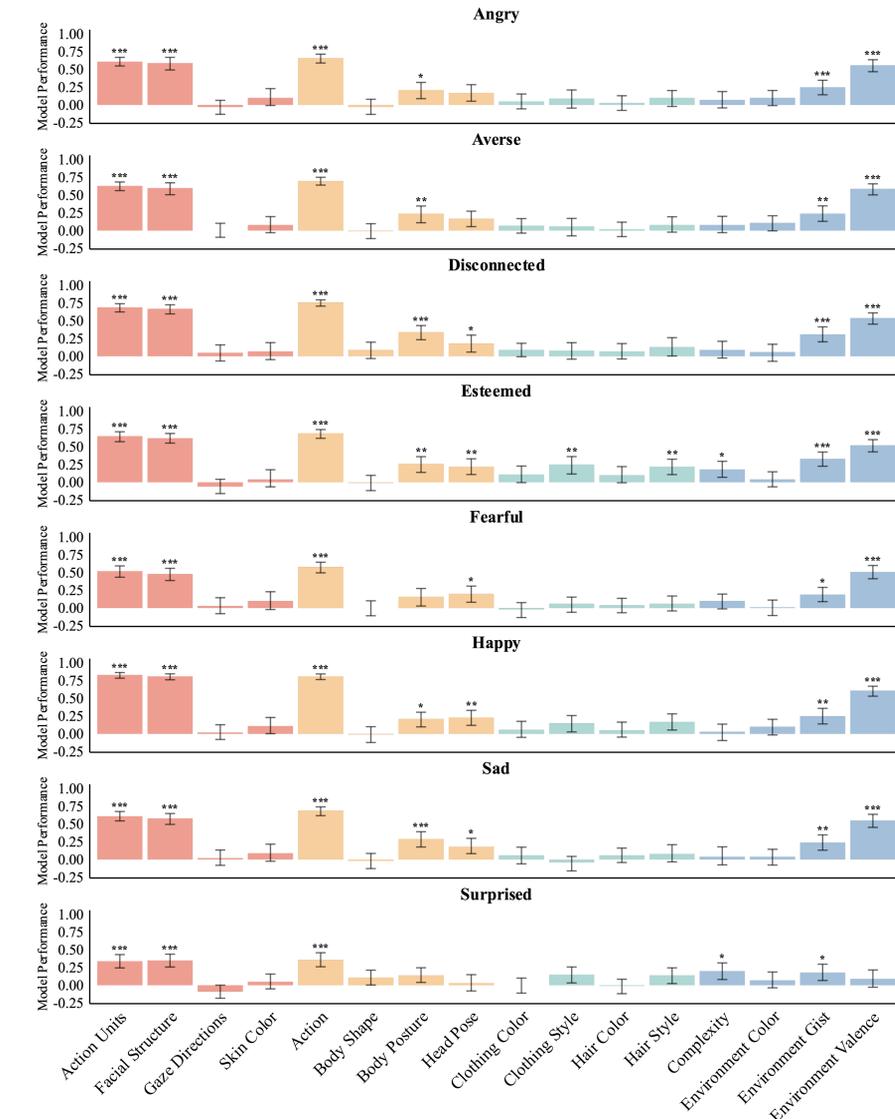

Results

- Participants showed *consistent* inferences in naturalistic contexts
- Emotion ratings revealed a clear valence-driven *correlational structure*

Facial, bodily, and contextual cues predicted naturalistic emotion inferences

Ridge Regression: $\text{Emotion} \sim \text{Feature}_1^{(cue_i)} + \text{Feature}_2^{(cue_i)} + \cdots + \text{Feature}_j^{(cue_i)}$


Emotional valence may shape cue and feature use in naturalistic emotion inferences
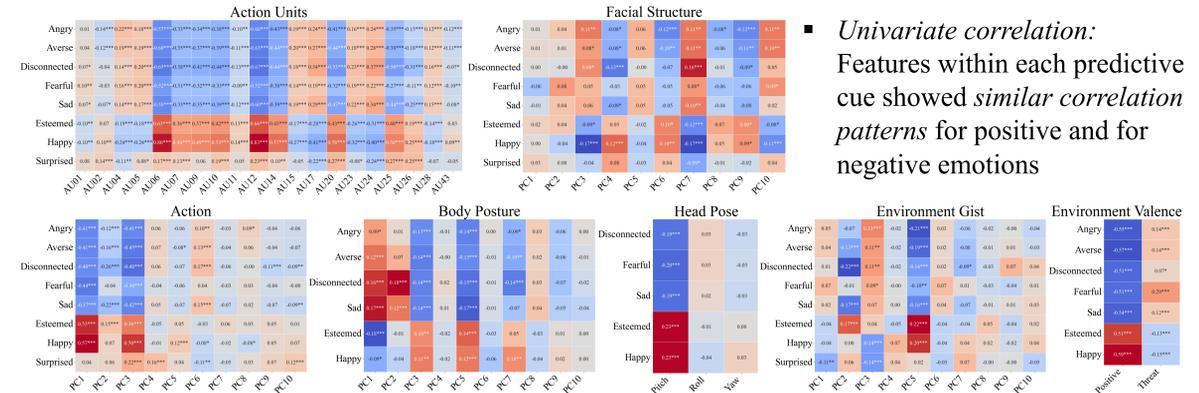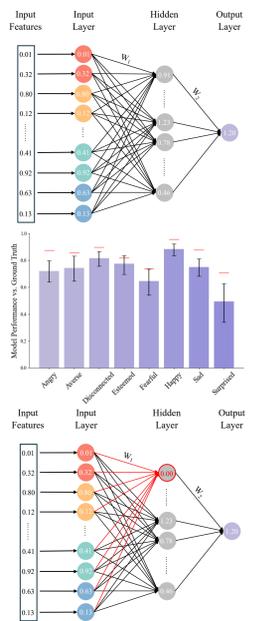

- *Univariate correlation:* Features within each predictive cue showed *similar correlation patterns* for positive and for negative emotions

Only facial and bodily actions carried unique information

Variation Partition

Most of the relevant cues in naturalistic contexts may play a supportive role
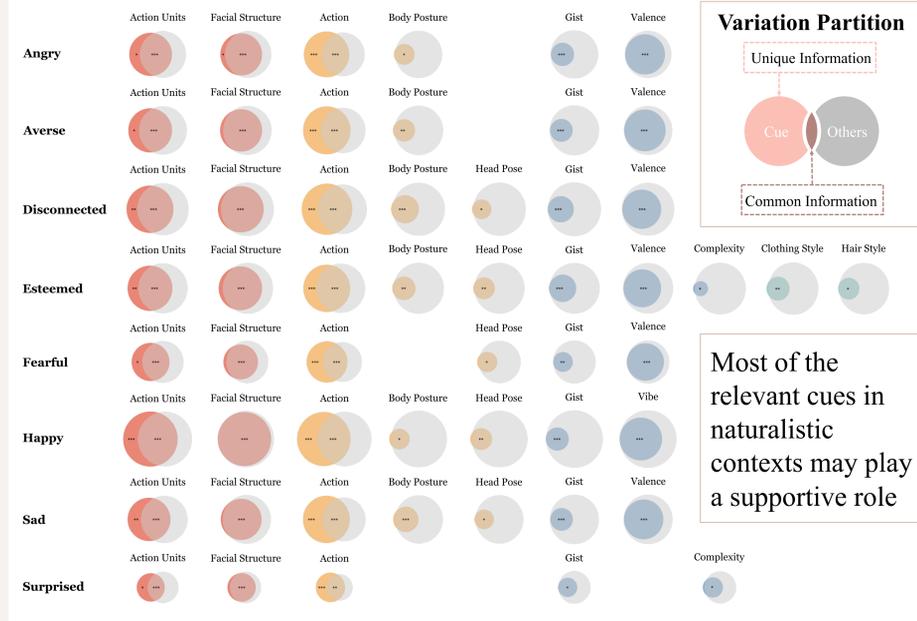
Ongoing Analysis

Conclusion

- Emotion inferences in naturalistic contexts rely on *multiple cues*, not just the face.
- Only facial muscle actions and bodily actions provide unique information, while other cues primarily reinforce these signals.
- Perceivers may integrate and reinterpret cues to construct a coherent understanding, organizing them along shared affective dimensions rather than discrete emotion categories.

References
Du, S., Tao, Y., & Martinez, A. M. (2014). Compound facial expressions of emotion. *Proceedings of the National Academy of Sciences, 111*(15).
Righart, R., & de Gelder, B. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective, & Behavioral Neuroscience, 8*(3), 264–272.
Witkower, Z., & Tracy, J. L. (2022). How and why head position changes the perception of facial expressions of emotion. *Emotion, 22*(5), 931–944.

More of my research