

Inter-Rater Reliability of Emotional Expressions in Racial Minority Models

Natalie Laguer Torres¹, De’Jiah Edwards¹, Ishraj Bhandari¹, Jared Cortez¹, Sera Gonzalez¹, Kendra Seaman¹
¹ The University of Texas at Dallas



Introduction

- The **FACES database** (Ebner et al., 2010).
 - Images of six emotions: neutrality, sadness, disgust, fear, anger, and happiness
 - Younger, Middle-Aged, and Older face models
 - Caucasian only
- Present study is to replicate in two largest minority groups in the U.S.:
 - Black/African American face models
 - Hispanic/Latine face models
- Here we investigate the **inter-rater reliability** of emotional expressions portrayed by our racial minority models ($N = 7$).

Methods

Photo session protocol

Goal: Capture authentic emotional expressions

1. **Emotion Induction Phase:** Participants are guided to experience specific emotions.
2. **Personal Experience Phase:** Participants are encouraged to recall personal emotional experiences associated with the targeted emotions.
3. **Controlled Expression Phase:** Participants were instructed to display the targeted emotions based on standardized guidelines.

Each photo session yielded approximately 200 photographs per model.

Photo selection procedure

Goal: Identify the two images for each model that best represented each of the six facial expressions

Two independent raters answered the following questions:

1. “Is the emotion present?” (Yes = 1, No = 0).
2. “How pronounced is the emotion?” (not very pronounced = 1, moderately pronounced = 2, very pronounced = 3)
3. “Is the emotion blended?” (Blended = 1, Pure = 0).

Inter-rater reliability

Goal: Quantify the degree of agreement between two raters.

- **% Agreement** – percent of time raters agree, several limitations
- **Cohen's kappa** – measures *absolute* agreement
- **Kendall's W** – measures *extent* of agreement

Calculation/Statistic	Level of Agreement
≤ 0	None
0.40-0.59	Weak
0.60-0.79	Moderate
0.80-0.99	Strong
1	Perfect

Results

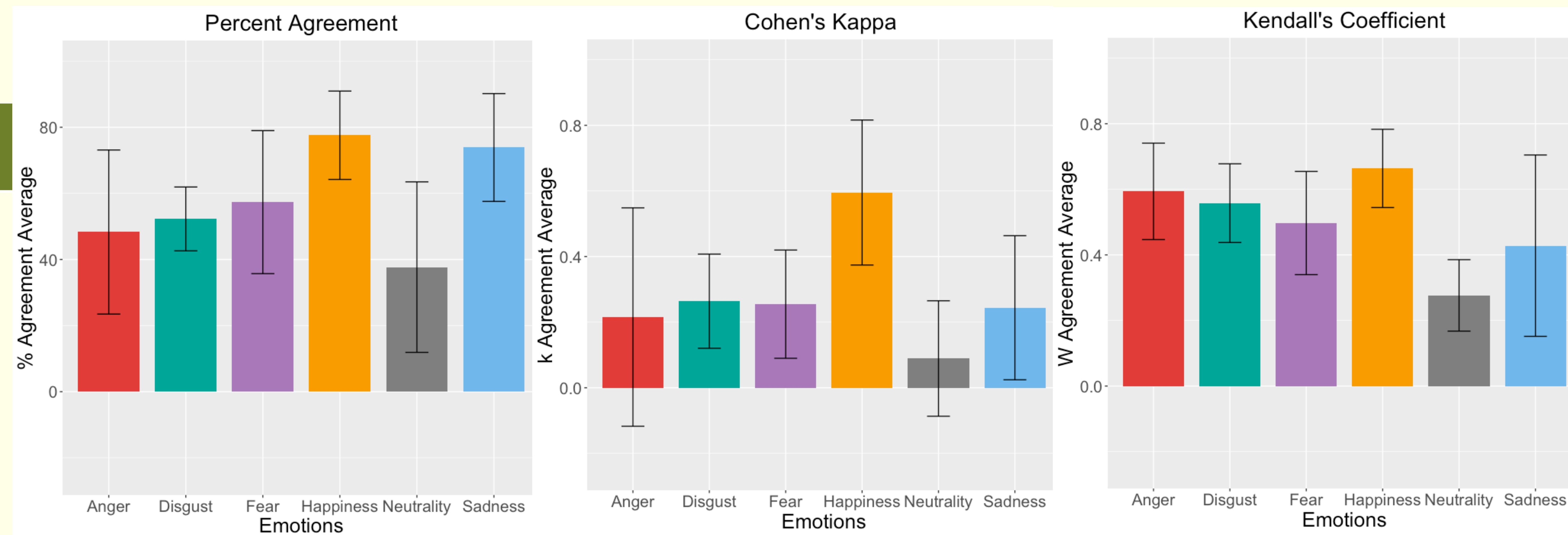
Question 1 – Presence of Target Emotion:

- Percent Agreement – average % is 78.30% ($SD = 5.71\%$)
- Cohen's kappa – average k value is $k = 0.40$ ($SD = 0.12$)

Moderate Agreement

Weak Agreement

Question 2 – Intensity of Emotion:



Most Agreement

Happiness

Moderate Agreement

Least Agreement

Neutrality

No Agreement

Question 3 – Blending of Emotions:

- Percent Agreement – average % is 78.08% ($SD = 5.61\%$)
- Cohen's kappa – average k value is $k = 0.49$ ($SD = 0.15$)

Moderate Agreement

Weak Agreement

Discussion

- $N =$ Only 7 of 36 facial models included in this sample.
- Preliminary findings show **Happiness** is most agreed on, but **Neutrality** is least agreed on.
- The best images for each face model/emotional expression will be validated by a larger, online sample.

- Validated photos will be shared on the Open Science Framework (OSF) for other researchers to use.
 - This will help diversify scientific research by providing stimuli that represent a more diverse population.
 - This project is an important first step in creating more inclusive psychological and neuroscience studies.