AMERICAN PSYCHOLOGICAL ASSOCIATION

Emotion

© 2025 American Psychological Association
ISSN: 1528-3542

2026, Vol. 26, No. 3, 757–770
https://doi.org/10.1037/emo0001622

Aggregating Emotional Sequences Amplifies the Perception of Women as More Emotional Than Men

Megan Gorges[1], Roni Porat[2, 3], Jonas P. Schöne[4, 5, 6], and Amit Goldenberg[1, 7, 8]

[1] Harvard Business School, Harvard University
[2] Department of Political Science, Hebrew University of Jerusalem
[3] Department of International Relations, Hebrew University of Jerusalem
[4] Institute for Human-Centered AI, Stanford University
[5] Department of Psychology, Stanford University
[6] Department of Sociology, Stanford University
[7] Department of Psychology, Harvard University
[8] Digital, Data and Design Institute, Harvard University

The stereotype that women are more emotional than men is pervasive in Western culture, but little research has directly examined how this stereotype translates into judgments of emotionality. We propose that one way gender stereotypes shape judgments of emotionality is through the aggregation of emotional expressions, in which perceivers preferentially remember stereotype–congruent emotional stimuli and consequently overweight these stimuli when forming judgments. To test this, we conducted five studies ($N = 772$) during 2021–2025 among men participants. In Study 1, we validated the persistence of gender-emotion stereotypes. For Studies 2–5, we selected emotional expression stimuli that elicited no gender difference in ratings of emotionality at the single face level. Men participants saw sequences of male and female faces displaying emotional expressions ranging from neutral-to-angry (Study 2), neutral-to-happy (Study 3), and neutral-to-sad (Study 4) and were asked to indicate whether they considered the person in the sequence to be emotional or not. When men perceivers aggregated these stimuli (which exhibited no gender difference at the single face level), they were more likely to rate sequences of female faces as emotional. Furthermore, using a memory test we show that participants better remembered angry female faces within a sequence compared with angry male faces (Study 5), supporting the idea that aggregation of emotional information enables stereotypes to influence judgments via memory. This study reveals an important mechanism by which stereotypes are translated into emotionality judgments. We used only White stimuli faces and recruited only men participants, limiting generalizability.

Keywords: gender, stereotypes, emotion, social judgments, amplification

Supplemental materials: https://doi.org/10.1037/emo0001622.supp

The stereotype that women are more emotional than men—that they experience and express more emotion—is so strong and pervasive that it has been termed a "master stereotype" (Fabes & Martin, 1991; Prentice & Carranza, 2002; Shields, 2002, 2005). In a nationally representative survey conducted in 2018, 69% of Americans ascribed the trait "emotional" more to women than to men (Briton & Hall, 1995; Eagly et al., 2020; Fischer, 1993). Endorsement of such gender-emotion stereotypes is stronger in men than in women; given that men hold more positions of power in society, this stereotypical view likely has far-reaching consequences (Kray & Kennedy, 2017; Timmers et al., 2003). Surprisingly, despite the prevalence of the emotionality master stereotype, we know little about *how* this stereotype influences judgments of emotionality. In the current article, we argue that one way in which the emotionality master stereotype is translated into judgments is via the aggregation of emotional information. In natural exposure to emotional expressions, we often see sequences of emotional expressions over time that we need to summarize to form a judgment of emotionality (Goldenberg et al., 2022;

This article was published Online First November 13, 2025.

Maria Gendron served as action editor.

Megan Gorges https://orcid.org/0009-0000-6410-1773

Materials, data, and code can be found on the Open Science Framework at https://osf.io/4hf8x. A version of this article was released as a preprint on the Open Science Framework in February 2024. Some of the data and interpretations included in this article were presented at the annual meeting of the Society for Affective Science in 2023.

The authors are grateful to Jon Jachimowicz for reviewing an earlier draft of this article, and to members of the Organizational Behavior Lab at Harvard Business School and the Interdisciplinary Affective Science Lab at Northeastern University for their feedback.

Megan Gorges played a lead role in conceptualization, formal analysis, investigation, methodology, project administration, visualization, writing–original draft, and writing–review and editing. Roni Porat played a lead role in conceptualization and a supporting role in visualization, investigation, formal analysis, methodology, supervision, writing–original draft, and writing–review and editing. Jonas P. Schöne played a supporting role in conceptualization, formal analysis, and methodology. Amit Goldenberg played a lead role in conceptualization, investigation, methodology, project administration, and supervision and a supporting role in formal analysis, methodology, visualization, writing–original draft, and writing–review and editing.

Correspondence concerning this article should be addressed to Megan Gorges, Harvard Business School, Harvard University, Wyss Hall, 710 Soldiers Field Road, Boston, MA 02163, United States. Email: mgorges@hbs.edu

Krumhuber et al., 2013). We argue that when people aggregate emotional expressions, they are more likely to remember expressions that align with the emotionality master stereotype (Fiske & Neuberg, 1990; Higgins & Bargh, 1987), leading to biased judgments of emotionality. Shedding light on the connection between gender stereotypes and the judgment process is crucial, as judgments of emotionality can negatively impact how people perceive women's competence, legitimacy, and leadership potential (Brescoll, 2016; Brescoll & Uhlmann, 2008; Frasca et al., 2022).

Gender-Emotion Stereotypes and Judgments of Emotionality

Prior research has little to say about the way in which gender-emotion stereotypes impact how people summarize streams of emotional expressions into judgments of emotionality. Thus, we draw insights from the robust stream of research on how stereotypes are translated into judgments more broadly. Perceivers form impressions of targets based on a continuous negotiation between top-down influences from social categorization and stereotypes and the bottom-up aggregation of facial cues (Freeman & Ambady, 2011; Freeman et al., 2020). Perceivers are more likely to pay attention to and remember cues that align with stereotypes (Fiske, 1993; Macrae & Bodenhausen, 2000). We extend this theory to the case of gender and emotionality, arguing that the master emotionality stereotype will lead perceivers to preferentially remember women's intense emotional expressions, leading to biased judgments of emotionality.

One classic study by Barrett et al. (1998) provides an initial indication that stereotypes may induce gender bias in emotionality judgments when aggregating multiple emotional experiences. In that study, women and men were asked to provide both momentary ratings and global retrospective judgments of their emotions. There were no gender differences in the momentary ratings, but when making *global* judgments, women rated themselves as more emotional compared with men. This suggests the emotionality master stereotype may exert a stronger influence on judgments of emotionality when people aggregate emotional information over time. However, evidence for this idea comes from self-judgments in a single study, pointing to the need to examine the role of aggregation more closely.

Deciphering how stereotypes shape judgments of emotionality is further complicated by the question of *which* gender-emotion stereotypes exert the most influence on this process. In addition to the idea that women are judged as more emotional regardless of the emotion they are expressing, a separate stream of research has found that people perceive emotional expressions as more intense when they are consistent with gender stereotypes about discrete emotions. According to this competing theory, men are stereotyped to express anger and pride more often than women, and women are believed to express happiness, sadness, fear, love, and sympathy more often than men (Fabes & Martin, 1991; Plant et al., 2000). Accordingly, research in this stream has found that perceivers rate equivalent expressions of happiness as more intense on female faces, and expressions of anger as more intense on male faces (Adams et al., 2012; Algoe et al., 2000; Hess et al., 1997). These results seem to conflict with research on the emotionality master stereotype which finds that women are more likely to be judged as emotional when they express *any* emotion, including anger (Barrett & Bliss-Moreau, 2009). In the current article, we examine judgments of emotionality in reaction to multiple discrete emotions to provide further evidence in light of these incongruent findings.

Forming Judgments From Sequences of Emotional Expressions

In a real-world interaction, such as a conversation between colleagues or friends, a perceiver observing their conversation partner will often encounter a complicated stream of facial expressions representing the target's emotion. The perceiver may have the goal of assessing their partner's level of emotionality to help them decipher the emotional stakes and potential outcome of the interaction. To form this global judgment of their target's emotionality, the perceiver must aggregate across these multiple expressions of the target's emotion. Research on ensemble coding found that people are more likely to remember strong emotional expressions (Jackson et al., 2014; Lee & Cho, 2019), leading them to evaluate sequences of emotional expressions as more intense than their statistical average (Fredrickson, 2000; Goldenberg et al., 2022; Kahneman et al., 1993; Redelmeier & Kahneman, 1996).

When aggregating sequences of emotions, emotional expressions can be memorable not only because of their intensity (Goldenberg et al., 2022) and recency (Murdock, 1962) but also because of their alignment with certain stereotypes. Most research on implicit social cognition has found evidence for a strong congruent relationship between stereotypes and memory; stereotypes help people make sense of the dizzying array of information in their environment (Amodio & Devine, 2006), guiding them to notice and remember stereotype-consistent information (Dovidio et al., 2010; Frankenstein et al., 2020; Fyock & Stangor, 1994; Udeogu et al., 2022). We note that another stream of research has found the opposite—that people better remember stereotype-inconsistent information because it violates their expectations, prompting more extensive cognitive processing (e.g., Hastie & Kumar, 1979; Kroneisen & Bell, 2013). However, this argument primarily applies to situations where perceivers have weaker expectations about the target (e.g., due to experimentally induced schemas) and have time to process information extensively before forming a judgment (Stangor & McMillan, 1992). In the case of making real-time judgments based on a complex stream of emotional expressions against a backdrop of strong expectations due to gender-emotion stereotypes, perceivers' memories (and resulting judgments) are likely to be stereotype-*consistent* (Higgins & Bargh, 1987; Rojahn & Pettigrew, 1992). We argue that due to the master emotionality stereotype, perceivers will better remember women's intense emotional expressions compared with men's. This tendency will impact the aggregation process, leading to biased judgments of emotionality.

The Present Research

The goal of this research was to further investigate how the master emotionality stereotype influences emotional judgments based on exposure to multiple emotional expressions. We conducted a survey study and four experiments (data gathered 2021–2025). Study 1, a descriptive survey, aimed to validate the persistence of the master

emotionality stereotype. In Studies 2 (focused on anger), 3 (focused on happiness), and 4 (focused on sadness), participants assessed emotionality in sequences of male and female angry, happy, and sad expressions.[1] Study 5 replicated Study 2 and used a memory task to investigate a potential mechanism for the gender effect on emotionality judgments.

We focus on two preregistered hypotheses. Our first hypothesis was that men perceivers would be more likely to judge a sequence of emotions expressed by a target female face as emotional compared to a target male face (even when no gender effect emerges in response to the individual emotional expressions composing the sequences). Our second hypothesis was that men perceivers would be more likely to remember emotionally intense female faces as having appeared in the sequence compared with emotionally intense male faces. In the Supplemental Material, we report findings for three additional hypotheses that replicate prior evidence of a sequential amplification effect (Goldenberg et al., 2022).

We limited participants to a single gender to simplify interpretation and chose to focus on testing men's perceptions because men have been shown to endorse gender-emotion stereotypes more strongly than women (Timmers et al., 2003). Given our goal of understanding how stereotypes impact emotionality judgments, we sought a participant population that would give us the best chance of seeing this effect in action. Furthermore, research on emotion perception has historically overrepresented women participants, often under assumptions (guided by the master emotionality stereotype) that women are more emotionally attuned or expressive than men; we aimed to help address this gap in the literature (Brody & Hall, 2008; Fischer & Manstead, 2000).

Study 1: Confirming the Persistence of the Stereotype That Women Are More Emotional

The goal of Study 1 was to ascertain if the stereotype that women are more emotional than men persists and if gender-emotion stereotypes vary for different discrete emotions.

Method

Participants

The study (and all others included in this article) was approved by a research ethics committee before data collection (Protocol No. IRB21-0015). We recruited 200 participants on Prolific, enabling us to detect a medium size effect which we anticipated based on prior studies with a similar design (Prentice & Carranza, 2002). We recruited only participants who self-identified as male on their Prolific profiles.[2] One participant reported identifying as nonbinary when they completed the survey and was excluded from the analysis. Only one participant failed an attention check, and results are reported including this participant. Our final sample consisted of 199 men participants (72% Caucasian; 82% had at least some college or vocational school; age: $M = 37.59$ years, $SD = 12.53$).

Procedure

Participants provided consent and then filled out a short survey measuring their opinions and cultural beliefs about emotions as well as their personality and demographic information. Participants were paid $1.00 for participating in the 6-min study.

Measures

We describe the measures relevant to the current article. Information about all other measures can be found in the Supplemental Material.

Emotionality Master Stereotype. Our measure of gender stereotypes about emotionality was adapted from Prentice and Carranza (2002). Participants responded to the question "Please indicate how typical you think each of the following characteristics is in adult American women (men)" using a 7-point Likert scale ranging from *not at all typical* to *very typical*, providing their ratings for "emotional" and "expresses emotion" for both women and men. Our focus was to measure emotionality in general (and thus report results here for the "emotional" item), but we found similar results for emotional expression reported in the Supplemental Material.

Discrete Emotion Gender Stereotypes. Our measure of discrete emotion gender stereotypes was adapted from Plant et al. (2000). Participants responded to the question "How often do you believe women (men) experience (express) the following emotions?" using a 7-point Likert scale ranging from *never* to *very frequently*, providing their ratings for anger, happiness, and sadness for both women and men and for both experience and expression of emotion. We selected these three emotions because they are the most studied in research on gender and emotion (e.g., Becker et al., 2007; Brescoll & Uhlmann, 2008; Hess et al., 2005) and include both positive and negative valence.

Demographics. Participants completed several demographic measures, including their gender, education, age, and race.

Results

To understand if the master emotionality stereotype persists, we conducted a paired t test to compare how typical participants thought being "emotional" was for women and men. Results indicated that the mean typicality for women being "emotional" ($M = 5.27$) was 1.71 points higher (on a scale from 1 to 7) than the mean typicality for men ($M = 3.56$), 95% CI [1.50, 1.93], $t(198) = 15.81$, $p < .001$, Cohen's $d = 1.12$ (see Figure 1).

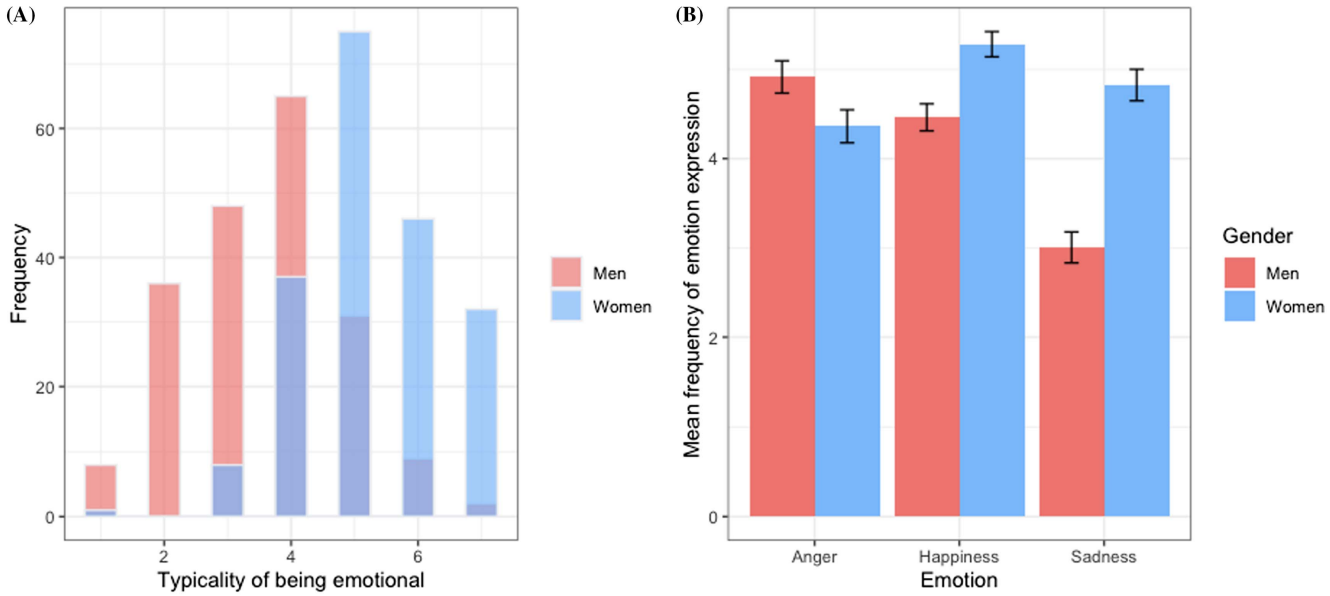
To ascertain if gender-emotion stereotypes vary by emotion, we ran a series of paired t tests comparing how frequently participants thought women and men experience and express anger, happiness, and sadness. Participants believed that men ($M = 5.09$) *experience* anger more frequently than women ($M = 4.68$), 95% CI [0.26, 0.56], $t(198) = 5.42$, $p < .001$, Cohen's $d = 0.38$. Similarly, participants believed that men ($M = 4.92$) *express*

[1] We ran an earlier iteration of Study 4 using different sadness stimuli (from the Radboud face set; Langner et al., 2010) and found no effect of face gender on emotionality judgments. However, we ran a separate study to gather participant impressions of all three sets of stimuli, and the Radboud sadness stimuli were rated as significantly less authentic than the anger and happiness stimuli. We later ran the iteration of Study 4 reported in this article using new sadness stimuli (from the IASLab face set). We report both earlier studies in the Supplemental Material.

[2] We sought to recruit participants who identified as men (gender identity) rather than focusing on biological sex (male). At the time we collected the data for Studies 1–3 and 5, Prolific only had a demographic category for sex. We worked around this by recruiting participants who identified as "male" on their profiles and then confirming their gender identity in the posttask survey.

Figure 1

Gender Stereotypes of Emotionality and Discrete Emotions (Anger, Happiness, and Sadness)



Note. (A) Histogram based on Study 1 data showing that participants rated "emotional" as a characteristic more typical of women than men. Red shaded areas represent typicality for men, blue shaded areas represent typicality for women, and purple shaded areas represent the overlap of the two. (B) Bar chart of participant ratings of how often they believed men versus women express anger, happiness, and sadness. Error bars indicate 95% confidence intervals (±1.96 times the standard error). See the online article for the color version of this figure.

anger more frequently than women (M = 4.36), 95% CI [0.34, 0.77], $t(198)$ = 5.02, p < .001, Cohen's d = 0.36. Participants thought women (M = 5.20) *experience* happiness more frequently than men (M = 4.77), 95% CI [0.28, 0.57], $t(198)$ = 5.90, p < .001, Cohen's d = 0.42. Similarly, participants thought women (M = 5.28) *express* happiness more frequently than men (M = 4.46), 95% CI [0.65, 0.99], $t(198)$ = 9.62, p < .001, Cohen's d = 0.68. Participants perceived that women (M = 4.93) *experience* sadness more frequently than men (M = 4.50), 95% CI [0.26, 0.59], $t(198)$ = 5.04, p < .001, Cohen's d = 0.36. Similarly, participants thought women (M = 4.82) *express* sadness more frequently than men (M = 3.01), 95% CI [1.60, 2.04], $t(198)$ = 16.49, p < .001, Cohen's d = 1.17. This aligns with prior work theorizing that anger is associated with men, and happiness and sadness with women (Plant et al., 2000). These results validate the existence of gender-emotion stereotypes about both general emotionality and distinct emotions.

Study 2: Testing Gender Differences in Judgments of Emotionality With Sequences of Angry Expressions

The goal of Study 2 was to test whether sequences of female faces expressing anger are more likely to be judged as emotional compared with sequences of male faces.

Method

Participants

Sample size was determined based on data from a pilot study that suggested that 100 participants completing 50 trials each would provide almost 100% power (see the Supplemental Material for details). We recruited 100 participants on Prolific who self-identified as male on their Prolific profiles, but two participants reported identifying as female when they completed the posttask survey and were excluded from the analysis. All participants passed the attention check. Our final sample consisted of 98 men participants (62% Caucasian; 87% had at least some college or vocational school; age: M = 32.65 years, SD = 10.59).

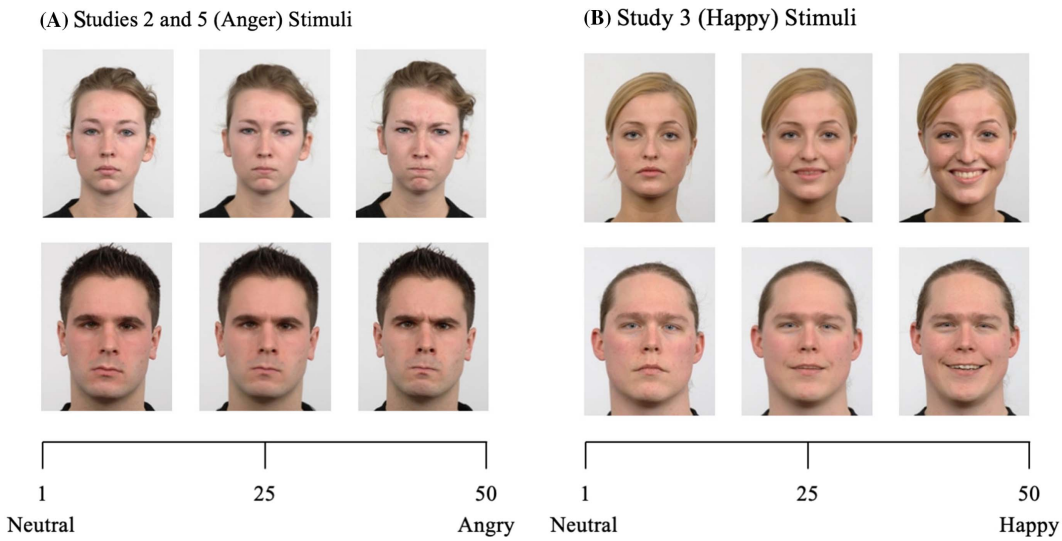
Materials

We used face morphs expressing different degrees of anger that were created and tested for a study on the evaluation of sequences of emotional expressions (Goldenberg et al., 2022). These stimuli were composed of six exemplar faces (three women and three men) from the Radboud face sample (Langner et al., 2010) expressing either neutral expressions or angry expressions. Goldenberg et al. (2022) used the software Fantamorph to create 48 intermediate emotional expressions between each identity's neutral and intensely angry expressions, resulting in morphed stimuli ranging from 1 (*neutral*) to 50 (*intensely angry*; see Figure 2 for example stimuli).

Stimuli Selection Pilot. Because we intended to test whether the master emotionality stereotype introduces bias into the process of aggregating sequences of emotional expressions, one of the most important preconditions of this experiment was that there would be no perceptual differences at the single face level between the male and female faces. If we had used emotional expression stimuli for which the individual female faces were more likely to be rated as emotional compared with the male faces, this would have prevented us from discerning whether any gender differences in emotionality judgments were due to the aggregation process or merely an artifact

Figure 2

Sample Stimuli Faces for Anger and Happiness



Note. A value of 25 corresponds to 50% intensity, and a value of 50%–100% intensity. The faces pictured come from the Radboud Faces Database and are free to be published in academic research articles. Study 4 used sadness stimuli from the Interdisciplinary Affective Science Laboratory Face Set that are not permitted to be published. See the online article for the color version of this figure.

of the stimuli. Therefore, to make sure that there were no gender-related differences at baseline of single face judgments of emotionality, we conducted a pilot experiment with the goal of selecting face identities that would generate equal emotionality ratings for women and men at the single face level (see Supplemental Material for full description and analysis). In brief, participants saw a single male or female face whose emotion intensity was randomly drawn from the face morphs described above, ranging from neutral to angry. After viewing each face, participants indicated if they judged the face they just saw as emotional or not emotional using a binary scale. This allowed us to assess each the probability that each face identity would be judged as emotional.

Our pilot analysis indicated that different identities were associated with different probabilities of being judged as emotional, but that these differences were unrelated to gender. We retained all six face identities because the probability of being judged as emotional was equal between the male and female faces. (When repeating this stimuli validation process for the happiness and sadness stimuli, we initially started with eight face identities and selected a subset of six identities for which the probability of being judged as emotional was equal for male and female faces.) This process ensured that any differences we observed at the sequence level would be a result of aggregation rather than characteristics of individual faces.

Procedure

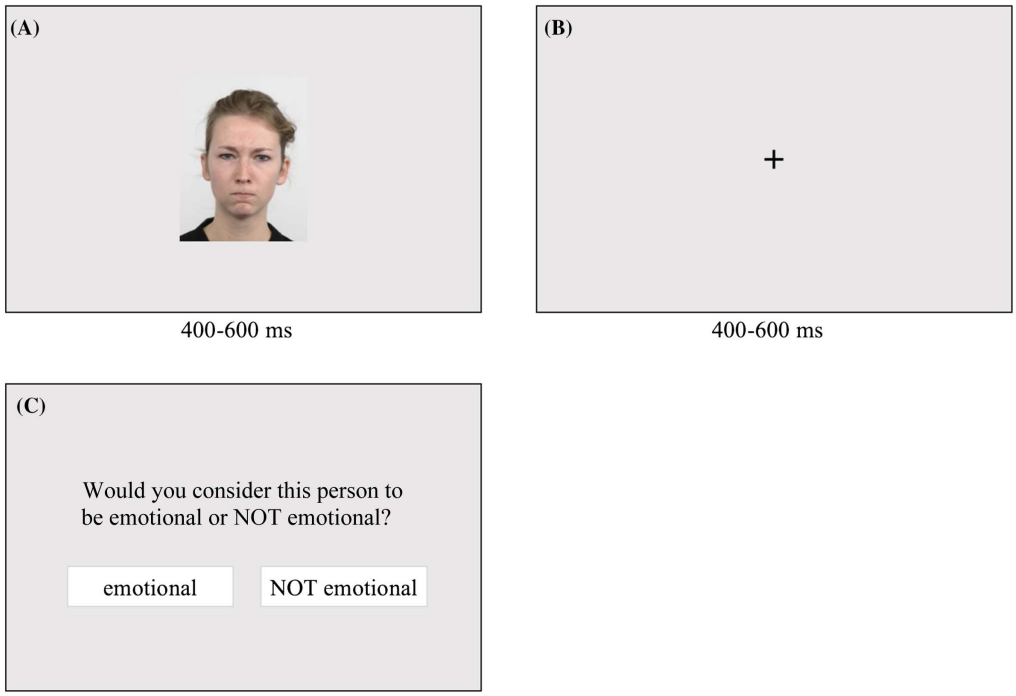
Participants completed a task that was adapted from a previous study that examined how people evaluate sequences of emotions (Goldenberg et al., 2022) but was modified to fit with the goals of the present study. Before starting the task, participants were told they would view a sequence of faces showing different degrees of emotions, judge the emotionality of those faces, and then complete a short survey. The structure of the task included three practice trials followed by 50 real trials.

In each trial, participants were shown a sequence of facial expressions of the same person. The identity of the person was randomly selected from a pool of three male and three female faces that were selected in the piloting stage and that showed no gender differences in the probability of being judged as emotional at the single face level. The number of expressions displayed in each trial ranged between 1 and 12. The emotional expressions within a given trial ranged from 1 = *neutral* to 50 = *intensely angry*; the emotion intensity was normally distributed around a mean of 25.5. Each face was displayed in the center of the screen for 400, 500, or 600 ms (randomly selected). At the beginning of the sequence and before each face, a fixation cross appeared in the center of the screen for 400, 500, or 600 ms (randomly selected).

After viewing each sequence, participants were asked "Would you consider this person to be emotional or NOT emotional?" Participants indicated their response on a binary scale (emotional or not emotional; see Figure 3). We chose to measure emotionality as a binary measure for two reasons. First, we were interested in participants' global judgment of a certain sequence as emotional rather than aspects such as the average expressed emotion or the modal emotional response. Measuring emotionality as a binary scale allowed participants to use their own criteria to judge emotionality. Second, when a person's emotionality is discussed, it often seems to be framed around categorizing a person as emotional or not rather than considering emotionality on a continuous scale (in line with the broader phenomenon called the binary bias) and we wanted to reflect this tendency in our measure (Fisher & Keil, 2018).

After completing the task, participants filled out a survey including demographic information (age, gender, race, political orientation, education, and income) and additional measures gathered

Figure 3
Emotionality Judgment Task Stimuli



Note. (A) Participants viewed a sequence of between one and 12 emotional expressions (sequence length randomly varied between trials); each emotional expression appeared in the center of the screen for 400–600 ms (randomly varying between trials). (B) Between each emotional expression, participants saw a fixation cross in the center of the screen for 400–600 ms. (C) After viewing all emotional expressions in a given trial, participants were asked to judge whether the person in the sequence is emotional or not emotional. See https://osf.io/uqrxt for a video of an example trial. The face pictured comes from the Radboud Faces Database and is free to be published in academic research articles. See the online article for the color version of this figure.

for exploratory purposes (details available in the Supplemental Material). Participants were paid $2.80 for participating in the 20-min study.

Results

Our first hypothesis was that sequences of women's faces would be more likely to be judged as emotional compared with sequences of men's faces (Hypothesis 1 [H1]). We tested this hypothesis using a mixed generalized linear model, including random intercepts for participant identification number and trial number. We used the gender of the face shown in each sequence as the independent variable and participants' binary emotionality judgment (emotional or not emotional) as the dependent variable. The probability that a sequence was judged as emotional was significantly higher when the sequence showed a female face compared with a male face, in support of H1 ($b = 0.45$, 95% CI [0.33, 0.57], $SE = 0.06$, $z = 7.22$, $p < .001$; see Figure 4). In other words, sequences with female faces had a 61% probability of being judged as emotional, whereas sequences with male faces had a 39% probability of being judged as emotional.

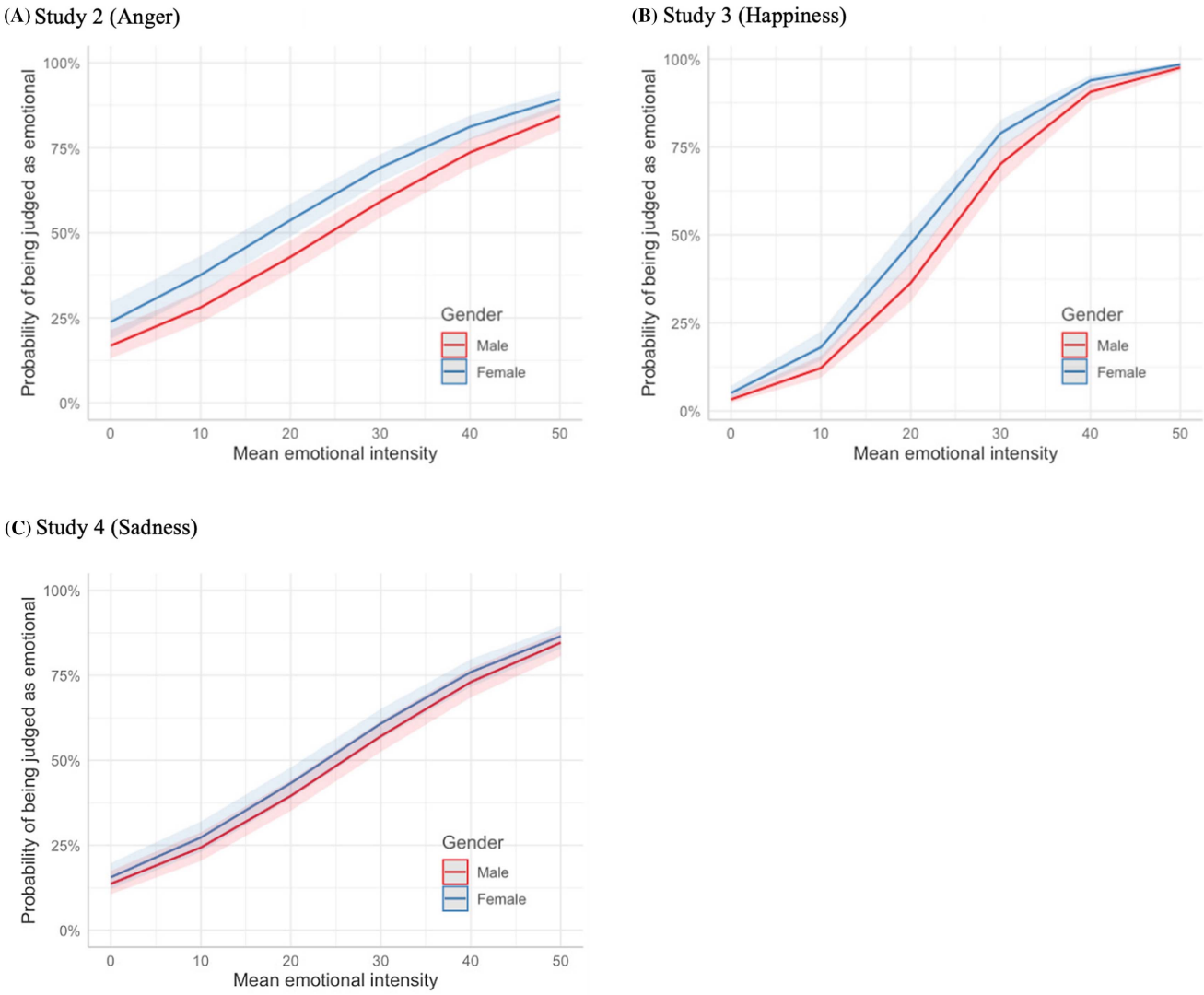
To summarize, sequences of angry female faces were more likely to be judged as emotional compared with male faces. Importantly, this tendency to judge sequences of women's emotional expressions as emotional (more often than men's) emerged even when there was no gender bias in judgments of emotionality in response to individual faces. These findings provide initial support for the idea that people make emotionality judgments in alignment with the emotionality master stereotype that women are more emotional than men. This result is the opposite of what we would expect to see if people made emotionality judgments guided by specific gender-emotion stereotypes (in which case women would be *less* likely to be judged as emotional when expressing anger which is stereotyped as masculine). However, it is not clear if this gender effect on emotionality judgments exists only for anger or if it extends to other emotions. It is possible that people are only more likely to judge women as emotional when they express negative emotions, or when they express emotions associated with men. We tested this possibility in Study 3 using expressions of happiness, a positive emotion associated with women (Plant et al., 2000).

Study 3: Testing Gender Differences in Judgments of Emotionality With Sequences of Happy Expressions

The goal of Study 3 was to replicate the effect of target gender on judgments of emotionality from Study 2 using expressions of happiness. Gender-emotion stereotypes about happiness differ from

Figure 4

Effect of Face Gender on Judgments of Emotionality



Note. These graphs were produced using mixed generalized linear models, including random intercepts for participant identification number and trial number. We used the gender of the face and the mean emotional intensity of the emotional expressions shown in the sequence as the independent variables and participants' binary emotionality judgment (emotional or not emotional) as the dependent variable. (A) Shows results from Study 2 (using angry expressions), (B) shows results from Study 3 (using happy expressions), and (C) shows results from Study 4 (using sad expressions). Shaded areas represent 95% confidence intervals (±1.96 times the standard error). See the online article for the color version of this figure.

those about anger (the focus of Study 2); women are expected to express happiness, whereas men are not.

Method

Participants

Consistent with Study 2, we estimated that 100 participants completing 50 trials each would provide sufficient power. We recruited 110 participants on Prolific (including 10 extra participants in case we needed to exclude any). We recruited only participants who self-identified as male on their profiles, but two participants reported identifying as nonbinary when they completed the posttask survey and were excluded from the analysis. We included the same attention check as in Study 2, and all participants passed. Our final sample included 108 men participants (73% Caucasian; 89% had at least some college or vocational school; age: $M = 40.24$ years, $SD = 12.28$).

Materials

We used emotional expression stimuli that were very similar to those used in Study 2, except with happy rather than angry faces. We

used morphed images drawn from eight face identities (four women and four men) created and tested by Goldenberg et al. (2022). We conducted a pilot (similar to the pilot in Study 2) to select six face identities (three women and three men) that showed no gender difference in judgments of emotionality when viewed as single images (see Supplemental Material for details).

Procedure

The procedure and measures were identical to those in Study 2.

Results

Our first hypothesis was that sequences featuring women's faces would be either more or less likely to be judged as emotional compared with sequences of men's faces (H1). For Study 3, our preregistered hypothesis about gender did not presume the direction of the effect; either male or female faces could be more likely to be judged as emotional when expressing happiness (see Supplemental Material for more details). We tested this hypothesis using the same model as in Study 2. The probability that a sequence was judged as emotional was significantly higher when the sequence showed a female face compared with a male face ($b = 0.37$, 95% CI [0.25, 0.49], $SE = 0.06$, $z = 6.15$, $p < .001$; see Figure 4). These results supported H1. In other words, female faces had a 59% probability of being judged as emotional, whereas male faces had a 41% probability of being judged as emotional.

Overall, our results with happiness expressions replicated those from Study 2 using anger expressions. Sequences of happy female faces were more likely to be judged as emotional compared with sequences of male faces. Because we see a gender effect for both angry and happy expressions, these results support the idea that people form judgments of emotionality in alignment with the stereotype that women are more emotional than men, rather than basing their judgment on emotion valence or discrete emotion gender stereotypes. In the next study, we tested if this effect extends to judgments of emotionality in reaction to expressions of sadness.

Study 4: Testing Gender Differences in Judgments of Emotionality With Sequences of Sad Expressions

The main goal of Study 4 was to replicate the effect of target gender on judgments of emotionality from Studies 2 and 3 using expressions of sadness, another emotion that is stereotyped as feminine (women are expected to express sadness, whereas men are not). A secondary goal of Study 4 was to test if effects vary by participant gender; previous studies used only men participants but here we include women participants as well.

Method

Participants

We estimated that 200 participants completing 50 trials each would provide sufficient power to detect both the effect of target gender on judgments of emotionality and any participant gender effects. In Studies 2 and 3, we were able to detect our effect of interest with approximately 100 men participants, so we doubled that number to have sufficient power to compare the effect between men and women participants. Prolific recruitment resulted in 212 participants (we included extra participants in case we needed to exclude any). We set quotas such that half the participants were men (self-identified as men on their profiles, including both cis and trans men) and half were women (self-identified as women, including both cis and trans women). Two participants were excluded for missing data (one lacked task data and one lacked posttask survey data). Three participants reported identifying as nonbinary or preferred not to indicate their gender when they completed the posttask survey and were excluded from the analysis. We included the same attention check as in Studies 2 and 3, and all participants passed. Our final sample included 207 participants (50% men, 61% Caucasian; 88% had at least some college or vocational school; age: $M = 39.52$ years, $SD = 12.33$).

Materials

We used sad emotional expression stimuli created in a similar manner to those used in Studies 2 and 3. We generated morphed images drawn from eight face identities (four women and four men) from the Interdisciplinary Affective Science Laboratory (IASLab) Face Set.[3] We used the software Fantamorph to create 48 intermediate emotional expressions between each identity's neutral and intensely sad expressions, resulting in morphed stimuli ranging from 1 (*neutral*) to 50 (*intensely sad*). We conducted a pilot (similar to the pilots in Studies 2 and 3) to select six face identities (three women and three men) that showed no gender difference in judgments of emotionality when viewed as single images (see Supplemental Material for details).

Procedure

The procedure and measures were identical to those in Studies 2 and 3, except participants were paid $3.40 for completing the 20-min task (compared with the $2.80 paid to the participants of Studies 2 and 3; pay was raised due to inflation in the intervening years).

Results

Our first hypothesis aimed to replicate the effects from Studies 2 and 3; we hypothesized that *participants who identify as men* would be more likely to judge sequences featuring women's faces as emotional compared with sequences of men's faces (H1). We tested this hypothesis using the same model as in Studies 2 and 3 and used a subset of the data that only included men participants ($n = 104$). The probability that a sequence was judged as emotional by men participants was significantly higher when the sequence showed a female face compared with a male face ($b = 0.12$, 95% CI [0.01, 0.24], $SE = 0.06$, $z = 2.10$, $p = .035$; see Figure 4). These results supported H1. In other words, female faces had a 53% probability of being judged as emotional by men participants, whereas male faces had a 47% probability of being judged as emotional.

Our second hypothesis was that *when including data from both women and men participants*, sequences featuring female faces would be more likely to be judged as emotional compared with

[3] Development of the IASLab Face Set was supported by the National Institutes of Health Director's Pioneer Award (No. DP1OD003312) to Lisa Feldman Barrett. More information is available online at https://www.affective-science.org.

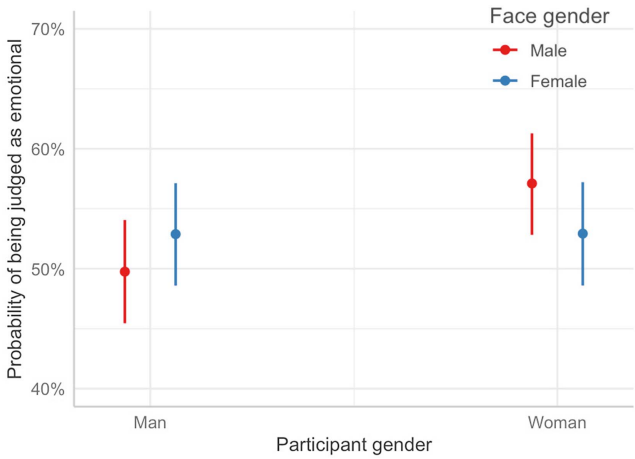
sequences of male faces (Hypothesis 2 [H2]). We tested this hypothesis with the same model we used to test H1 and used the full data set including both women and men participants (n = 207). When including all participants, there was no significant effect of face gender; the probability that a sequence was judged as emotional was not significantly different for sequences of female faces compared with sequences of male faces ($b = -0.02$, 95% CI [−0.10, 0.06], SE = 0.04, $z = -0.49$, p = .62). These results did not support H2.

Our third hypothesis was a preregistered exploratory analysis to test if the effect of face gender on the odds of a sequence being judged as emotional differed by participant gender (men vs. women); we did not hypothesize a direction for this effect (Hypothesis 3 [H3]). We tested this using a mixed generalized linear model with an interaction term, including random intercepts for participant identification number and trial number. The independent variable was an interaction term between the gender of the face shown in each sequence (face gender) and the gender of the perceiver (participant gender). The dependent variable was participants' binary emotionality judgment (emotional or not emotional). There was a significant interaction between face gender and participant gender ($b = -0.29$, 95% CI [−0.46, −0.13], $SE = 0.08$, $z = -3.48$, $p < .001$; see Figure 5).

To probe this interaction, we conducted simple effects analyses using estimated marginal means (*emmeans* package Version 1.10.7; Lenth, 2025). Women participants were significantly more likely to judge sequences featuring sad *male* faces as emotional (M = 0.57, 95% CI [0.53, 0.61]) compared with sequences of *female* faces (M = 0.53, 95% CI [0.49, 0.57]), $SE = 0.07$, $z = 2.81$, $p = .005$. In contrast (as found in the analysis for H1), men participants were significantly *less* likely to judge sequences featuring sad *male* faces as emotional (M = 0.50, 95% CI [0.45, 0.54]) compared with sequences of *female* faces (M = 0.53, 95% CI [0.49, 0.57]), SE = 0.05, $z = -2.11$, p = .017).

Figure 5

Interaction Effect of Face Gender and Participant Gender on Judgments of Emotionality



Note. Error bars represent 95% confidence intervals. See the online article for the color version of this figure.

To understand what might underlie the effect of participant gender on judgments of emotionality, we conducted exploratory analyses to test for participant gender differences in several self-reported measures of gender stereotypes included in the posttask survey (see Supplemental Material for complete analyses and results). These analyses revealed significant participant gender differences only for the Modern Sexism Scale (Swim et al., 1995). The results suggest that individual differences in modern sexism moderate the effect of face gender on judgments of emotionality. Lower levels of sexism were associated with a greater tendency to perceive sequences of male faces as emotional, whereas higher levels of sexism were associated with a greater tendency to perceive sequences of female faces as emotional. Given that men scored higher on modern sexism, this may underlie participant gender differences in judgments of emotionality. We explore this and other possible explanations in the discussion section.

When confining the sample to men participants, these results with sadness expressions replicated our findings from Studies 2 (anger) and 3 (happiness). Men participants were more likely to judge sequences of sad female faces as emotional compared with sequences of male faces. This finding supports the idea that regardless of the specific emotion expressed, male perceivers form judgments of emotionality in line with the master emotionality stereotype that women are more emotional than men.

Notably, across three studies, we find that men perceivers are more likely to judge *sequences* featuring female faces as emotional, even when no gender effect exists at the single face level. This hints that memory may play a role in judgments of emotionality, as suggested by a recent study (Goldenberg et al., 2022). When people view longer sequences of emotional expressions, they may be unable to hold all the expressions in mind when forming their judgment, so they selectively remember salient emotional expressions—those that are more intense and consistent with gender stereotypes. We examined this potential mechanism in the next study.

Study 5: Testing Memory as a Potential Mechanism by Which Sequences of Women's Angry Expressions Are More Likely to Be Judged as Emotional

The goal of Study 5 was to replicate the findings of Study 2 and explore a potential mechanism by which target gender affects judgments of emotionality. The study included two blocks; the first block was a shorter version of the emotionality judgment task used in Study 2, and the second was a memory task to test if men perceivers preferentially remember women's (compared with men's) emotionally intense anger expressions. We chose to use anger expressions (an emotion stereotyped as masculine) so we could discern whether men show enhanced memory for emotional expressions consistent with the master emotionality stereotype (in which case women's intense emotional expressions would be more memorable) or discrete emotion stereotypes (in which case men's intense emotional expressions would be more memorable).

Method

Participants

We chose our sample size based on a power analysis using data from Study 2; the results suggested that 100 participants would

provide almost 100% power (see Supplemental Material for details). We increased the target sample size beyond 100 to account for potential attrition and the addition of a memory task (and corresponding decrease in the number of trials in the emotionality judgment task from 50 to 25) and aimed for 165 participants; Prolific recruiting yielded 164 participants. We recruited only participants who self-identified as male on their Prolific profiles, but three participants reported identifying as female or nonbinary when they completed the posttask survey and were excluded from the analysis. One participant was missing posttask survey data and was excluded from the analysis. We included the same attention check as in Studies 2–4 and all participants passed. Our final sample included 160 men participants (74% Caucasian; 87% had at least some college or vocational school; age: $M = 38.95$ years, $SD = 13.90$).

Materials

Emotional expression stimuli were identical to those used in Study 2.

Procedure

Participants were told they would complete two blocks of tasks followed by a brief survey. Participants completed the first block, which was identical to the emotionality judgment task in Study 2 except with 25 trials instead of 50 (preceded by three practice trials).

In the second block, participants were told that the goal was to examine whether they could remember the emotional expressions of the faces they saw in each sequence. Participants were shown sequences of eight emotional expressions in each trial; the number of expressions was held constant because variation in sequence length would complicate interpretation of the results of the memory test. After viewing the sequence of eight expressions, participants were asked to recall which of two faces appeared in the sequence, choosing between a true target expression that was shown in the sequence and a false target expression that was not shown. In each trial, the emotional intensity of the false target expression was determined by taking the midpoint of the two expressions in the sequence with the largest intensity difference. For example, if the sequence in a trial showed an emotional expression with intensity Level 17 and another expression with intensity Level 45 and that was the biggest difference in that sequence, the false target expression would be set at intensity Level 31. The true and false target facial expressions appeared on the screen immediately after the final fixation cross of the sequence, and participants had as much time as they needed to indicate which face was the correct one that appeared in the sequence. Participants completed 25 trials in this memory block (preceded by three practice trials).

After completing the memory block, participants filled out a survey including demographic information (age, gender, race, political orientation, education, and income) and additional measures gathered for exploratory purposes (details available in the Supplemental Material). Participants were paid $4.20 for participating in the 30-min study.

Results

We first examined the data from the emotionality judgment block, and successfully replicated the results from Study 2. Our first hypothesis was that sequences of women's faces would be more likely to be judged as emotional compared with sequences of men's faces (H1). We tested this hypothesis using the same model as in Study 2 and found that the probability that a sequence was judged as emotional was significantly higher when the sequence showed a female face compared with a male face, in support of H1 ($b = 0.20$, 95% CI [0.06, 0.33], $SE = 0.07$, $z = 2.89$, $p = .004$; see Figure 6A). In other words, female faces had a 55% probability of being judged as emotional, whereas male faces had a 45% probability of being judged as emotional.

Next, we turned our attention to the memory block. Our second hypothesis was that men participants would better remember women's emotionally intense faces compared with men's (H2). To test this hypothesis, we used a mixed generalized linear model with random intercepts for participant identification number and trial number. The independent variable was the interaction between the gender of the target face (male or female) and the intensity of the true target expression, and the dependent variable was whether or not the participant correctly identified the true target expression. Results supported H2; the interaction between face gender and expression intensity was significant such that men participants were more likely to correctly remember intense female faces compared with intense male faces ($b = 0.01$, 95% CI [0.002, 0.02], $SE = 0.005$, $z = 2.50$, $p = .013$; see Figure 6B).

In sum, these results show that men perceivers preferentially remember emotionally intense faces, and they are even more likely to remember emotionally intense faces when they are female rather than male. This suggests that memory may be one potential mechanism by which people tend to judge female faces as emotional; given that they preferentially remember emotionally intense female facial expressions, they may be more likely to draw on these expressions when forming their emotionality judgments.

Transparency and Openness

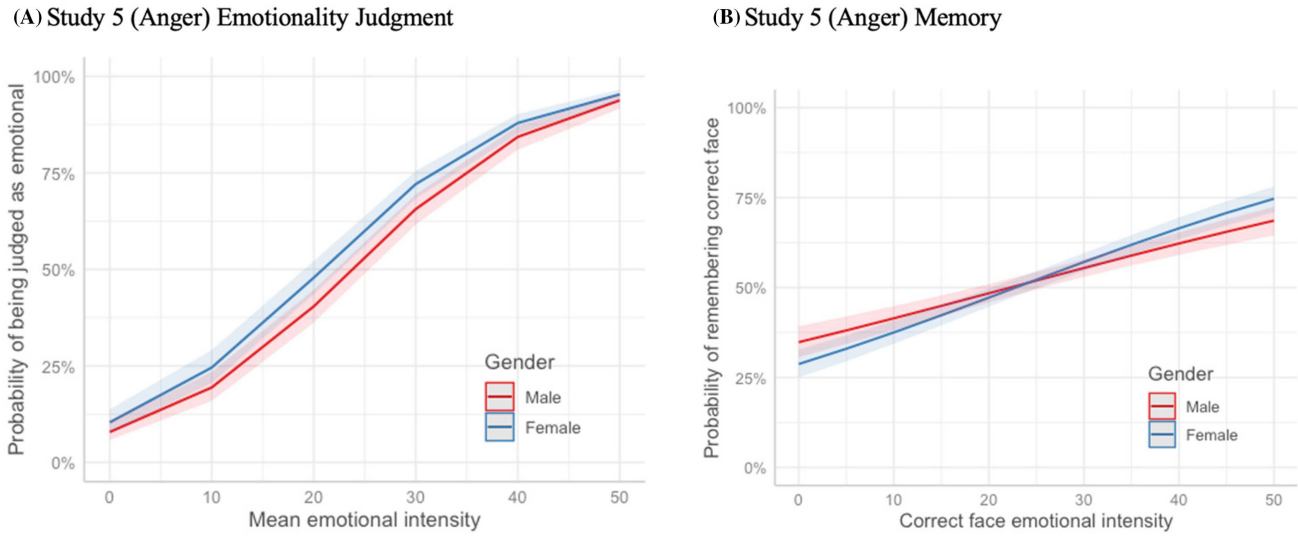
We report how we determined our sample sizes, data exclusions, and all measures in the studies (reported either in the article or in the Supplemental Material). Data, analysis code, and survey materials for all studies are available at https://osf.io/4hf8x. There is not a preregistration for Study 1 because it was exploratory; all other studies were preregistered at the following links in Study 2 at https://osf.io/aq8hj, Study 3 at https://osf.io/nc5kt, Study 4 at https://osf.io/fkx3g, and Study 5 at https://osf.io/kg97j. Data were analyzed using R, Version 4.4.3 (R Core Team, 2025).

General Discussion

There is ample evidence that women are stereotypically believed to be more emotional than men, but much less is known about the process by which a perceiver judges a target as emotional. Across four studies, we showed that the master emotionality stereotype persists (Study 1) and that one way this stereotype influences judgments of emotionality is through the process of aggregating multiple emotional expressions (Studies 2–4). We found that men perceivers were more likely to judge women as emotional compared with men based on sequences of angry, happy, and sad expressions, even though there was no gender difference in judgments of emotionality based on the individual emotional expressions composing these sequences. In our final study, we found that men

Figure 6

Interaction Effect of Face Gender and Emotional Intensity on Memory



Note. Panel A displays results from the Study 5 emotionality judgment block, showing that sequences of angry emotional expressions on female faces were more likely to be judged as emotional compared with male faces across the range of emotional intensity levels. Panel A was produced using a mixed generalized linear model, including random intercepts for participant identification number and trial number. We used the gender of the face and the mean emotional intensity of the emotional expressions shown in the sequence as the independent variables and participants' binary emotionality judgment (emotional or not emotional) as the dependent variable. Panel B displays results from the Study 5 memory block, showing the significant interaction between the emotional intensity of the face shown in the sequence and the gender of that face; participants were more likely to remember and correctly identify emotionally intense female faces compared with emotionally intense male faces. Shaded areas represent 95% confidence intervals (±1.96 times the standard error). See the online article for the color version of this figure.

perceivers were more likely to remember women's intense emotional expressions compared with men's.

Our article makes two important contributions to the domain of affect and gender. First, our consistent results across anger, happiness, and sadness support the idea that judgments of emotionality are guided by the master emotionality stereotype, rather than gender stereotypes about specific emotions. A stream of research has focused on how gender stereotypes about discrete emotions (e.g., the belief that men express more anger and women express more happiness and sadness) translate into stereotype-consistent emotional expressions being perceived as more intense (Adams et al., 2012; Algoe et al., 2000; Hess et al., 1997). If these discrete gender-emotion stereotypes (e.g., the stereotype that anger is a masculine emotion) guided the aggregation of emotional expressions into judgments of emotionality, we would have expected perceivers to be more likely to judge sequences of men's angry faces as emotional. By demonstrating that men perceivers are more likely to judge women's faces as emotional not only when expressing happiness and sadness (both stereotypically feminine emotions) but also when expressing anger (a stereotypically masculine emotion), we provide evidence for the argument that the master emotionality stereotype exerts a distinct and more powerful influence than discrete gender-emotion stereotypes when men perceivers are judging a target's emotionality. Indeed, the effect of target gender was stronger for anger than for happiness and sadness. Together, our findings hint that while discrete gender-emotion stereotypes may shape perceptions of *emotion intensity*, they may not influence *judgments of emotionality* in predictable ways.

Second, our findings indicate that over and above any tendency for specific stimuli characteristics to elicit judgments of emotionality (e.g., the masculinity or femininity of the face, as explored in Barrett & Bliss-Moreau, 2009), the very process of aggregating across a sequence of emotional expressions enables stereotypes to influence judgments of emotionality via a memory mechanism. Much of the research on stereotypes and perceptions of women's emotions has examined judgments based on single emotional expressions, but more often in daily interactions people are exposed to multiple emotional expressions. This research suggests one way to close this gap, highlighting how the common and automatic process of drawing on memory to summarize emotional information provides an opportunity for stereotypes to impact judgment. Our theorizing integrates insights from theories of ensemble coding (Goldenberg et al., 2022; Jackson et al., 2014) and stereotype-consistent memory (Dovidio et al., 2010; Fyock & Stangor, 1994), pointing to the possibility that aggregation could play a key role in how other stereotypes (such as racial or ethnic stereotypes) impact judgments.

Constraints on Generality and Future Directions

The generalizability of this research is subject to three key limitations that invite further investigation. First, we used only men participants for Studies 1–3 and 5, and in Study 4 our effect of interest holds only for men participants. Women participants were more likely to judge sequences of sad male faces as emotional, whereas men participants were more likely to judge sequences of sad female faces as emotional. We argue this may be because men

harbor stronger gender-emotion stereotypes than women (Timmers et al., 2003), and this translates into more stereotyped judgments of emotionality, According to social role theory (Eagly & Wood, 2012) and hegemonic masculinity frameworks (Connell & Messerschmidt, 2005), men are often socialized to prioritize emotional restraint and to distance themselves from sensitivity to others' emotional states—particularly those perceived as inconsistent with masculine norms. As a result, men may be more likely to show gender-congruent biases in emotion recognition (e.g., greater attunement to women's emotions or discomfort with men's emotional expression). This explanation is tentatively supported by exploratory analyses in Study 4 (full details reported in the Supplemental Material) showing that men participants scored higher on the Modern Sexism Scale (Swim et al., 1995) and that stronger endorsement of modern sexism was related to a higher likelihood of judging sequences of female faces as emotional. It is also possible that this perceiver gender effect is due to in-group bias; perceivers may be more accurate and less influenced by stereotypes when judging emotions expressed by members of their own gender group (Weisbuch & Ambady, 2008). Additional research is needed to test these competing theories.

A second limitation is that our stimuli were created from posed emotional expressions and presented devoid of context. The stimuli showed sequences of faces with dynamically changing muscle movements meant to represent the sequential expression of a single emotion. Future research should use stimuli with naturally flowing emotional expressions that may span multiple emotions (e.g., in videos) and situate stimuli in a variety of social contexts, which we know influence how people perceive emotional expressions on men's and women's faces (Fischer et al., 2013; Kelly & Hutson-Comeaux, 1999).

Third, we used only White faces in our stimuli materials to simplify interpretation. Gender-emotion stereotypes and reactions to facial expressions vary based on the social group membership of the target face; by focusing on only White faces, we could more cleanly infer the relationship between stereotypes and judgments (Durik et al., 2006; Motro et al., 2022; Xie et al., 2021). If the studies were replicated using stimuli from other social groups, we would expect the specific results to vary but the higher level theoretical argument to remain consistent—that observers preferentially remember stereotype-consistent emotional expressions and thus are more likely to judge sequences containing such expressions as emotional. We encourage future research to expand beyond White faces and use racially diverse stimuli to understand how the intersection of racial, ethnic, and gender stereotypes shape judgments of emotionality.

We also call upon future research to further investigate limitations to internal validity. In the current research we categorized target gender in a binary way (male or female) based on the gender of the person from whom the emotional expression stimuli were created. While this approach is well-founded in the sense that people do spontaneously infer gender as categorical (Campanella et al., 2001), we encourage future research to manipulate gender cues in target faces (e.g., their relative masculinity and femininity; Sutherland et al., 2015) to assess their impact on judgments of emotionality.

Conclusion

Being judged as emotional carries negative consequences for women, diminishing their perceived status, legitimacy, and leadership potential (Carnevale et al., 2019; Frasca et al., 2022). In this article, we highlight how the aggregation of emotional expressions over time enables stereotypes to shape biased judgments of emotionality. Understanding this additional source of bias can inform efforts to prevent stereotype-based judgments of emotionality and reduce their harmful impact.

References

Adams, R. B., Jr., Nelson, A. J., Soto, J. A., Hess, U., & Kleck, R. E. (2012). Emotion in the neutral face: A mechanism for impression formation? *Cognition and Emotion*, *26*(3), 431–441. https://doi.org/10.1080/02699931.2012.666502

Algoe, S. B., Buswell, B. N., & DeLamater, J. D. (2000). Gender and job status as contextual cues for the interpretation of facial expression of emotion. *Sex Roles*, *42*(3–4), 183–208. https://doi.org/10.1023/A:1007087106159

Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology*, *91*(4), 652–661. https://doi.org/10.1037/0022-3514.91.4.652

Barrett, L. F., & Bliss-Moreau, E. (2009). She's emotional. He's having a bad day: Attributional explanations for emotion stereotypes. *Emotion*, *9*(5), 649–658. https://doi.org/10.1037/a0016821

Barrett, L. F., Robin, L., Pietromonaco, P. R., & Eyssell, K. M. (1998). Are women the "more emotional" sex? Evidence from emotional experiences in social context. *Cognition and Emotion*, *12*(4), 555–578. https://doi.org/10.1080/026999398379565

Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology*, *92*(2), 179–190. https://doi.org/10.1037/0022-3514.92.2.179

Brescoll, V. L. (2016). Leading with their hearts? How gender stereotypes of emotion lead to biased evaluations of female leaders. *The Leadership Quarterly*, *27*(3), 415–428. https://doi.org/10.1016/j.leaqua.2016.02.005

Brescoll, V. L., & Uhlmann, E. L. (2008). Can an angry woman get ahead? Status conferral, gender, and expression of emotion in the workplace. *Psychological Science*, *19*(3), 268–275. https://doi.org/10.1111/j.1467-9280.2008.02079.x

Briton, N. J., & Hall, J. A. (1995). Beliefs about female and male nonverbal communication. *Sex Roles*, *32*(1–2), 79–90. https://doi.org/10.1007/BF01544758

Brody, L. R., & Hall, J. A. (2008). Gender and emotion in context. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 395–408). Guilford Press.

Campanella, S., Chrysochoos, A., & Bruyer, R. (2001). Categorical perception of facial gender information: Behavioural evidence and the face-space metaphor. *Visual Cognition*, *8*(2), 237–262. https://doi.org/10.1080/13506280042000072

Carnevale, A. P., Smith, N., & Peltier Campbell, K. (2019). *May the best woman win? Education and bias against women in American politics*. The Georgetown University, Center on Education and the Workforce. https://cew.georgetown.edu/wp-content/uploads/Women_in_Politics.pdf

Connell, R. W., & Messerschmidt, J. W. (2005). Hegemonic masculinity: Rethinking the concept. *Gender & Society*, *19*(6), 829–859. https://doi.org/10.1177/0891243205278639

Dovidio, J. F., Hewstone, M., Glick, P., & Esses, V. M. (2010). Prejudice, stereotyping and discrimination: Theoretical and empirical overview. In J. F. Dovidio, M. Hewstone, & P. Glick (Eds.), *The SAGE handbook of prejudice, stereotyping and discrimination* (pp. 3–28). Sage. https://doi.org/10.4135/9781446200919.n1

Durik, A. M., Hyde, J. S., Marks, A. C., Roy, A. L., Anaya, D., & Schultz, G. (2006). Ethnicity and gender stereotypes of emotion. *Sex Roles*, *54*(7–8), 429–445. https://doi.org/10.1007/s11199-006-9020-4

Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *American Psychologist*, *75*(3), 301–315. https://doi.org/10.1037/amp0000494

Eagly, A. H., & Wood, W. (2012). Social role theory. In P. van Lange, A. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories in social psychology* (pp. 458–476). Sage Publications. https://doi.org/10.4135/9781446249222.n49

Fabes, R. A., & Martin, C. L. (1991). Gender and age stereotypes of emotionality. *Personality and Social Psychology Bulletin*, *17*(5), 532–540. https://doi.org/10.1177/0146167291175008

Fischer, A. H. (1993). Sex differences in emotionality: Fact or stereotype? *Feminism & Psychology*, *3*(3), 303–318. https://doi.org/10.1177/0959353593033002

Fischer, A. H., Eagly, A. H., & Oosterwijk, S. (2013). The meaning of tears: Which sex seems emotional depends on the social context. *European Journal of Social Psychology*, *43*(6), 505–515. https://doi.org/10.1002/ejsp.1974

Fischer, A. H., & Manstead, A. S. (2000). The relation between gender and emotion in different cultures. In A. H. Fischer (Ed.), *Gender and emotion: Social psychological perspectives* (pp. 71–94). Cambridge University Press. https://doi.org/10.1017/CBO9780511628191.005

Fisher, M., & Keil, F. C. (2018). The binary bias: A systematic distortion in the integration of information. *Psychological Science*, *29*(11), 1846–1858. https://doi.org/10.1177/0956797618792256

Fiske, S. T. (1993). Social cognition and social perception. *Annual Review of Psychology*, *44*(1), 155–194. https://doi.org/10.1146/annurev.ps.44.020193.001103

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). Academic Press. https://doi.org/10.1016/S0065-2601(08)60317-2

Frankenstein, A. N., McCurdy, M. P., Sklenar, A. M., Pandya, R., Szpunar, K. K., & Leshikar, E. D. (2020). Future thinking about social targets: The influence of prediction outcome on memory. *Cognition*, *204*, Article 104390. https://doi.org/10.1016/j.cognition.2020.104390

Frasca, T. J., Leskinen, E. A., & Warner, L. R. (2022). Words like weapons: Labeling women as emotional during a disagreement negatively affects the perceived legitimacy of their arguments. *Psychology of Women Quarterly*, *46*(4), 420–437. https://doi.org/10.1177/03616843221123745

Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition and Emotion*, *14*(4), 577–606. https://doi.org/10.1080/026999300402808

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, *118*(2), 247–279. https://doi.org/10.1037/a0022327

Freeman, J. B., Stolier, R. M., & Brooks, J. A. (2020). Dynamic interactive theory as a domain-general account of social perception. In B. Gawronski (Ed.), *Advances in experimental social psychology* (Vol. 61, pp. 237–287). Academic Press. https://doi.org/10.1016/bs.aesp.2019.09.005

Fyock, J., & Stangor, C. (1994). The role of memory biases in stereotype maintenance. *British Journal of Social Psychology*, *33*(3), 331–343. https://doi.org/10.1111/j.2044-8309.1994.tb01029.x

Goldenberg, A., Schöne, J., Huang, Z., Sweeny, T. D., Ong, D. C., Brady, T. F., Robinson, M. M., Levari, D., Zaki, J., & Gross, J. J. (2022). Amplification in the evaluation of multiple emotional expressions over time. *Nature Human Behaviour*, *6*(10), 1408–1416. https://doi.org/10.1038/s41562-022-01390-y

Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology*, *37*(1), 25–38. https://doi.org/10.1037/0022-3514.37.1.25

Hess, U., Adams, R., Jr., & Kleck, R. (2005). Who may frown and who should smile? Dominance, affiliation, and the display of happiness and anger. *Cognition and Emotion*, *19*(4), 515–536. https://doi.org/10.1080/02699930441000364

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior*, *21*(4), 241–257. https://doi.org/10.1023/A:1024952730333

Higgins, E. T., & Bargh, J. A. (1987). Social cognition and social perception. *Annual Review of Psychology*, *38*(1), 369–425. https://doi.org/10.1146/annurev.ps.38.020187.002101

Jackson, M. C., Linden, D. E., & Raymond, J. E. (2014). Angry expressions strengthen the encoding and maintenance of face identity representations in visual working memory. *Cognition and Emotion*, *28*(2), 278–297. https://doi.org/10.1080/02699931.2013.816655

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science*, *4*(6), 401–405. https://doi.org/10.1111/j.1467-9280.1993.tb00589.x

Kelly, J. R., & Hutson-Comeaux, S. L. (1999). Gender-emotion stereotypes are context specific. *Sex Roles*, *40*(1–2), 107–120. https://doi.org/10.1023/A:1018834501996

Kray, L. J., & Kennedy, J. A. (2017). Changing the narrative: Women as negotiators—And leaders. *California Management Review*, *60*(1), 70–87. https://doi.org/10.1177/0008125617727744

Kroneisen, M., & Bell, R. (2013). Sex, cheating, and disgust: Enhanced source memory for trait information that violates gender stereotypes. *Memory*, *21*(2), 167–181. https://doi.org/10.1080/09658211.2012.713971

Krumhuber, E. G., Kappas, A., & Manstead, A. S. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review*, *5*(1), 41–46. https://doi.org/10.1177/1754073912451349

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Lee, H. J., & Cho, Y. S. (2019). Memory facilitation for emotional faces: Visual working memory trade-offs resulting from attentional preference for emotional facial expressions. *Memory & Cognition*, *47*(6), 1231–1243. https://doi.org/10.3758/s13421-019-00930-8

Lenth, R. (2025). *emmeans: Estimated marginal means, aka least-squares means* (R package Version 1.10.7) [Computer software]. https://CRAN.R-project.org/package=emmeans

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, *51*(1), 93–120. https://doi.org/10.1146/annurev.psych.51.1.93

Motro, D., Evans, J. B., Ellis, A. P. J., & Benson, L., III. (2022). Race and reactions to women's expressions of anger at work: Examining the effects of the "angry Black woman" stereotype. *Journal of Applied Psychology*, *107*(1), 142–152. https://doi.org/10.1037/apl0000884

Murdock, B. B., Jr. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*(5), 482–488. https://doi.org/10.1037/h0045106

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

Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly*, *26*(4), 269–281. https://doi.org/10.1111/1471-6402.t01-1-00066

R Core Team. (2025). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain*, *66*(1), 3–8. https://doi.org/10.1016/0304-3959(96)02994-6

Rojahn, K., & Pettigrew, T. F. (1992). Memory for schema-relevant information: A meta-analytic resolution. *British Journal of Social Psychology*, *31*(2), 81–109. https://doi.org/10.1111/j.2044-8309.1992.tb00958.x

Shields, S. A. (2002). *Speaking from the heart: Gender and the social meaning of emotion*. Cambridge University Press.

Shields, S. A. (2005). The politics of emotion in everyday life: "Appropriate" emotion and claims on identity. *Review of General Psychology*, *9*(1), 3–15. https://doi.org/10.1037/1089-2680.9.1.3

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin*, *111*(1), 42–61. https://doi.org/10.1037/0033-2909.111.1.42

Sutherland, C. A., Young, A. W., Mootz, C. A., & Oldmeadow, J. A. (2015). Face gender and stereotypicality influence facial trait evaluation: Counter-stereotypical female faces are negatively evaluated. *British Journal of Psychology*, *106*(2), 186–208. https://doi.org/10.1111/bjop.12085

Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. A. (1995). Sexism and racism: Old-fashioned and modern prejudices. *Journal of Personality and Social Psychology*, *68*(2), 199–214. https://doi.org/10.1037/0022-3514.68.2.199

Timmers, M., Fischer, A., & Manstead, A. (2003). Ability versus vulnerability: Beliefs about men's and women's emotional behaviour. *Cognition and Emotion*, *17*(1), 41–63. https://doi.org/10.1080/02699930302277

Udeogu, O. J., Frankenstein, A. N., Sklenar, A. M., Urban Levy, P., & Leshikar, E. D. (2022). Predicting and remembering the behaviors of social targets: How prediction accuracy affects episodic memory. *BMC Psychology*, *10*(1), Article 96. https://doi.org/10.1186/s40359-022-00801-z

Weisbuch, M., & Ambady, N. (2008). Affective divergence: Automatic responses to others' emotions depend on group membership. *Journal of Personality and Social Psychology*, *95*(5), 1063–1079. https://doi.org/10.1037/a0011993

Xie, S. Y., Flake, J. K., Stolier, R. M., Freeman, J. B., & Hehman, E. (2021). Facial impressions are predicted by the structure of group stereotypes. *Psychological Science*, *32*(12), 1979–1993. https://doi.org/10.1177/09567976211024259

Received October 11, 2024
Revision received April 28, 2025
Accepted October 11, 2025 ■