JSLHR

Research Article

Gender Differences in Identifying Facial, Prosodic, and Semantic Emotions Show Category- and Channel-Specific Effects Mediated by Encoder's Gender

Yi Lin,[a] Hongwei Ding,[a] and Yang Zhang[b]

Purpose: The nature of gender differences in emotion processing has remained unclear due to the discrepancies in existing literature. This study examined the modulatory effects of emotion categories and communication channels on gender differences in verbal and nonverbal emotion perception.

Method: Eighty-eight participants (43 females and 45 males) were asked to identify three basic emotions (i.e., happiness, sadness, and anger) and neutrality encoded by female or male actors from verbal (i.e., semantic) or nonverbal (i.e., facial and prosodic) channels.

Results: While women showed an overall advantage in performance, their superiority was dependent on specific types of emotion and channel. Specifically, women outperformed men in regard to two basic emotions (happiness and sadness) in the nonverbal channels and only the anger category with verbal content. Conversely, men did better for the anger category in the nonverbal channels and for the other two emotions (happiness and sadness) in verbal content. There was an emotion- and channel-specific interaction effect between the two types of gender differences, with male subjects showing higher sensitivity to sad faces and prosody portrayed by the female encoders.

Conclusion: These findings reveal explicit emotion processing as a highly dynamic complex process with significant gender differences tied to specific emotion categories and communication channels.

Supplemental Material: https://doi.org/10.23641/asha.15032583

The ability to interpret the emotional states of others and appropriately express one's own emotions forms an essential sociocognitive basis for the effective establishment, maintenance, and regulation of our interpersonal relationships (Fischer & Manstead, 2008; Lausen & Schacht, 2018). In verbal and nonverbal communication of emotions, gender has been repeatedly proposed as a key attribute that influences one's processing ability (J. A. Hall, 1978; Schirmer et al., 2002, 2005; Schmid et al., 2011; Thompson & Voyer, 2014). One area of theoretical and empirical contention centers around gender/sex differences in emotion cognition, their biological and social origin, and the implications in child development and clinical practice.

Theoretical explanations of gender development include evolutionary perspectives grounded in biological differences between men and women for survival and reproduction, and socialization theories that characterize gender roles based on societal division of labor and stereotypical expectations (see Chaplin, 2015, for a review). These theories often point to greater physical vulnerability of women and their primary role in child-rearing, which affords them an advantage in emotion processing (Babchuk et al., 1985; Brody & Hall, 2010; Eagly, 1987; Hampson et al., 2006; Hofmann et al., 2006). In line with this notion are a number of studies that demonstrated overall higher detectability of emotional information expressed by females (Belin et al., 2008; Collignon et al., 2010; Koeda et al., 2013; Lausen & Schacht, 2018; Pell, 2002; Scherer et al., 2001; Vasconcelos et al., 2017), and women's superiority in decoding emotions (Bonebright et al., 1996; Collignon et al., 2010; Donges et al., 2012; Forni-Santos & Osorio, 2015; Fujisawa & Shinohara, 2011; J. A. Hall,

[a]Speech-Language-Hearing Center, School of Foreign Languages, Shanghai Jiao Tong University, Shanghai, China
[b]Department of Speech-Language-Hearing Sciences & Center for Neurobehavioral Development, University of Minnesota Twin Cities, Minneapolis

Correspondence to Hongwei Ding: hwding@sjtu.edu.cn; and Yang Zhang: zhanglab@umn.edu

Editor-in-Chief: Bharath Chandrasekaran
Editor: Chao-Yang Lee

Received September 18, 2020
Revision received March 15, 2021
Accepted April 7, 2021
https://doi.org/10.1044/2021_JSLHR-20-00553

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

1978; J. A. Hall & Matsumoto, 2004; Hampson et al., 2006; Hoffmann et al., 2010; Kret & de Gelder, 2012; Lambrecht et al., 2014; Li et al., 2008; Menezes et al., 2017; Scherer & Scherer, 2011; Thayer & Johnsen, 2000; Vasconcelos et al., 2017).

However, female superiority in emotion cognition has not been conclusively established as it may depend on the emotional content of verbal or nonverbal cues in the context. According to the differential emotion theory (Ackerman et al., 1998), a set of emotions (e.g., joy, interest, sadness, anger, disgust, fear, and surprise) are discrete and distinguishable, which can orchestrate various changes in neurochemical processes, expressive behaviors, and subjective experiences. This has complicated the traditional evolutionary and social accounts for women's better emotion processing ability by giving rise to two predictions. Specifically, the attachment promotion hypothesis predicts a female advantage for all emotions, which enables their offspring to become securely attached, while the fitness threat hypothesis expects female superiority to be selective to negative emotions, which calls for action on women to eliminate the potential threat to infant survival (Hampson et al., 2006; Menezes et al., 2017; Thompson & Voyer, 2014). Empirical evidence suggests that the emotion-specific effects not only modulate the magnitude of women's superiority but can also change the direction of gender differences in emotion processing. Several meta-analyses revealed that emotional categories contribute to substantial variability of gender differences in emotion perception (J. A. Hall, 1978; McClure, 2000; Thompson & Voyer, 2014), with the observed effect sizes of female advantages in overall or single emotion perception ranging from $d = 1.56$ to $d = -0.87$ in the most recent meta-analytic study conducted by Thompson and Voyer (2014). Specifically, negative emotions (e.g., anger, sadness, fear, disgust, contempt) were demonstrated to produce marginally different effect sizes from positive emotions (e.g., happiness, interest), which were nonetheless significantly larger than "other emotions" (e.g., surprise, neutrality). Similarly, in emotion expression, happiness, sadness, and love tend to elicit better identification performances when encoded by women, whereas anger and pride are easier to identify when encoded by men (Becker et al., 2007; Bonebright et al., 1996; Calvo & Lundqvist, 2008; Hofmann et al., 2006; Johnson et al., 2011; Kelly & Hutson-Comeaux, 1999; Öhman et al., 2010; Plant et al., 2000).

Aside from the effects of distinct emotional categories, the communication channels through which emotions are presented have also produced mixed patterns of gender differences in emotion processing. In real-life interactions, emotions can be conveyed through various verbal (e.g., semantics) and nonverbal (e.g., facial expressions, body movements, prosody) channels. According to the sensory dominance theory, one communication channel can hold more perceptual advantages over others (Filippi et al., 2017; Kitayama & Ishii, 2002; Lin et al., 2020b; Pell et al., 2011; Spence et al., 2011), and this channel-specific effect tends to vary across different emotions (Hawk et al., 2009; Lambrecht et al., 2014; Williams et al., 2009), which can influence how men and women perform in emotion processing tasks. While relatively few studies reported a significant interaction effect on the detectability of emotions expressed by encoders of different genders, gender differences in decoding emotions are more susceptible to the interplay between emotion categories and communication channels. For example, women are more likely to display advantages when perceiving sadness, disgust, surprise in facial expressions (Forni-Santos & Osorio, 2015; Kret & de Gelder, 2012; Menezes et al., 2017), and understanding happiness, sadness, and fear in tone of voice (Bonebright et al., 1996; Lambrecht et al., 2014; Lausen & Schacht, 2018; Thompson & Voyer, 2014; Vasconcelos et al., 2017). Female superiority in decoding anger in nonverbal channels does not appear as consistently established as in verbal content (Kret & de Gelder, 2012; Lambrecht et al., 2014; Lausen & Schacht, 2018; Smith et al., 1989; Thompson & Voyer, 2014; Vasconcelos et al., 2017; Zupan et al., 2016).

Furthermore, the emotion- and channel-specific effects can interact to influence how well men and women can recognize emotions expressed by encoders of different genders. This sexually dimorphic effect has two competing variants. The own-gender hypothesis posits that it is easier to recognize facial or vocal expressions posed by the encoders of the same gender, which is more typically associated with aversive events and face recognition than other types of emotions and channels (Doi et al., 2010; Fulton et al., 2015; Herlitz & Lovén, 2013; Kret & de Gelder, 2012; Lewin & Herlitz, 2002). By contrast, the opposite-gender hypothesis advocates that individuals attend more strongly to emotional expressions by the opposite gender for better identification of potential mates, which is in accordance with the evolutionary argument for gender differences (Hofmann et al., 2006; Proverbio et al., 2010). This latter interpretation appears to be represented differently across more types of emotions and channels, such as happy, angry, fearful, and neutral facial expressions, as compared to alluring, erotic, and sad prosody (Chun et al., 2012; Cretser et al., 1982; Erwin et al., 1992; Ethofer et al., 2007; Hofmann et al., 2006; Lambrecht et al., 2014; Proverbio et al., 2010; Rahman et al., 2004). Interestingly, there are also some studies showing no significant interaction effect probably due to the influence of experimental (e.g., selected emotions and channels, stimulus presentation duration, measurement indices) and cultural (e.g., Western European vs. American vs. East Asian) settings, rendering it premature to reach a definitive conclusion without considering more methodological and sociocontextual factors (Calvo & Lundqvist, 2008; Kret & de Gelder, 2012; Palermo & Coltheart, 2004; Thayer & Johnsen, 2000; Thompson & Voyer, 2014).

A preliminary conclusion based on existing work on gender differences in encoding and decoding emotions is that the divergent results have left the mechanisms behind these differences unclear (Gohier et al., 2011; Lausen & Schacht, 2018). Despite a growing body of literature on the effects of emotion categories and communication channels on gender differences in emotion processing, the two domains of research are to some extent dissociated. On one hand, there is an unequal amount of attention on and a lack of comparisons among different communication channels in past literature on the emotion-specific effect. The effects of different emotion categories are much more often investigated in a single nonverbal channel (i.e., face or prosody) than the verbal one (Glenberg et al., 2009; Thompson & Voyer, 2014). Even within the context of nonverbal communication, visual facial displays are more often examined (Babchuk et al., 1985; Erwin et al., 1992; Forni-Santos & Osorio, 2015; Hampson et al., 2006; Hofmann et al., 2006; Menezes et al., 2017; Öhman et al., 2010; Rahman et al., 2004), while the processing of auditory prosodic signals is less understood (Chun et al., 2012; Ethofer et al., 2007; J. A. Hall, 1978; Hawk et al., 2009; Lausen & Schacht, 2018). On the other hand, there is no standardized method of developing emotional stimuli when the channel-specific effect is explored. Emotions included in existing studies are sometimes mixed with both basic (e.g., happy, sad, angry, fearful) and complex (e.g., alluring, erotic, embarrassed) emotional categories (Lambrecht et al., 2014; Lausen & Schacht, 2018). In addition, experimental control on the emotional stimuli in previous research tends to deviate from what we usually encounter in daily interactions. A common example is that emotional prosody is represented as nonlinguistic affective bursts (e.g., laughter, cries, sighs, groans, screams) to eliminate the potential confound of semantic content, which underrepresents the dynamic and complex nature of running speech (Belin et al., 2008; Collignon et al., 2008; Juslin & Laukka, 2001; Koeda et al., 2013; Vasconcelos et al., 2017).

Another interesting theoretical question is whether gender differences in emotion processing became evolutionarily adaptive at some point and thereafter modified social contexts, or emerged within a specific culture (Kret & de Gelder, 2012). Since most of the existing studies were conducted in Western European or American countries with a low-context culture, in which communication relies heavily on the literal meaning of words rather than social contexts, it remains to be tested how the observed effects are represented in a different contextual setting (E. T. Hall, 1989). To this end, we based our study in a Mandarin Chinese high-context culture, where contexts and interpersonal relationships are more important than actual words during communication, and investigated how emotion categories and communication channels shape gender differences in emotion expression and perception. We focused on three basic emotions (i.e., happiness, sadness, and anger), since the recognition of these emotions has remained more stable and accurate from an early age as compared to other basic emotions (e.g., fear, disgust, surprise) defined by Ekman (1992) and played a primary role in interpersonal interactions during developmental stages (Flom et al., 2008; Garcia & Tully, 2020; Lawrence et al., 2015). Participants were asked to identify stimuli of the three emotions and neutrality presented from three communication channels, namely, facial expressions, speech-embedded prosody, and words spoken with neutral tone of voice. Their performances were parsed by encoders' gender, decoders' gender, and their reciprocal interactions across different emotion categories and sensory channels.

Specifically, we made the following three hypotheses based on the evolutionary and social accounts of gender differences and existing empirical evidence related to the differential emotion theory and the sensory dominance effect. First, women would display an overall advantage in encoding emotions (Belin et al., 2008; Collignon et al., 2010; Koeda et al., 2013; Lausen & Schacht, 2018; Pell, 2002; Scherer et al., 2001; Vasconcelos et al., 2017), but this superiority effect would be emotion specific across three different channels: Happiness and sadness would be easier to identify when encoded by women, whereas anger would enjoy a perceptual advantage when encoded by men (Becker et al., 2007; Bonebright et al., 1996; Calvo & Lundqvist, 2008; Hofmann et al., 2006; Johnson et al., 2011; Kelly & Hutson-Comeaux, 1999; Öhman et al., 2010; Plant et al., 2000). Second, women would show an overall advantage in decoding emotions (Bonebright et al., 1996; Collignon et al., 2010; Donges et al., 2012; Forni-Santos & Osorio, 2015; Fujisawa & Shinohara, 2011; J. A. Hall, 1978; J. A. Hall & Matsumoto, 2004; Hampson et al., 2006; Hoffmann et al., 2010; Kret & de Gelder, 2012; Lambrecht et al., 2014; Li et al., 2008; Menezes et al., 2017; Scherer & Scherer, 2011; Thayer & Johnsen, 2000; Vasconcelos et al., 2017), especially for happiness and sadness in the nonverbal channels (Bonebright et al., 1996; Forni-Santos & Osorio, 2015; Kret & de Gelder, 2012; Lambrecht et al., 2014; Lausen & Schacht, 2018; Thompson & Voyer, 2014; Vasconcelos et al., 2017), and anger in verbal content (Smith et al., 1989). Finally, encoders and decoders' gender would interact with each other, showing an opposite gender bias that would be susceptible to influences from both emotions and channels (Chun et al., 2012; Cretser et al., 1982; Erwin et al., 1992; Ethofer et al., 2007; Hofmann et al., 2006; Lambrecht et al., 2014; Proverbio et al., 2010; Rahman et al., 2004). The findings from this hypothesis-driven study will contribute new data to our understanding of theoretical explanations on how the emotion- and channel-specific effects coincide with traditional evolutionary and social perspectives in shaping gender differences in emotion expression and perception. The behavioral results from normal adult participants will also have implications for further studies on the underlying mechanisms of emotion cognition as well as the pathogenesis, symptomatology, and intervention for a wide range of gender-related psychiatric affective disorders.

Method

Participants

The study was approved by the institutional review board at Shanghai Jiao Tong University in accordance with the Declaration of Helsinki for research involving human subjects. Written informed consent was obtained from all participants, who were paid for their time and involvement. Eighty-eight volunteers (43 females and 45 males) were recruited to take part in the study through an online advertisement. Female participants had a mean age of 23.8 years (SD = 2.4), and male participants were, on average, 23.7 years old (*SD* = 3.0). All participants were native speakers of Mandarin Chinese, who were undergraduate or graduate students. All had normal or corrected-to-normal vision, and they also had normal hearing as verified by standard audiological assessment for pure tones from 0.25 to 8 kHz (Koerner & Zhang, 2018). None had any history of speech, language, and hearing disorders or psychiatric illnesses.

To estimate the appropriate sample size, we conducted an a priori power analysis based on an earlier study examining the impact of sensory modality and emotional category on gender differences in emotion recognition (Lambrecht et al., 2014). We expected to see a medium effect size ($R^2 = .25$, Cohen's $f^2 = .33$), with $\alpha = .05$ and power $(1 - \beta) = .90$ (Ferguson, 2009). Power analysis in R (R Core Team, 2018) using the pwr package (Champely, 2020) indicated a total sample size of 67 for a significant four-way interaction among channels, emotions, and decoders and encoders' genders according to the third hypothesis, which suggested that the sample size in the current research had sufficient statistical power.

Stimuli

Each experimental stimulus conveyed one of the three basic emotions (i.e., happiness, sadness, and anger; Ekman, 1992) or neutrality. These emotions were expressed in three different communication channels, namely, facial expressions, tone of voice (prosody), and semantic content.

The face stimuli were taken from the Chinese Affective Picture System (Bai et al., 2005), a well-validated database with a standardized set of black-and-white photographs of Chinese actors portraying emotional or neutral facial expressions. The included photos were generated by 16 actors (eight women and eight men), who had received professional training in acting or directing. Their expressions were chosen based on the identification accuracy of emotional category and ratings of emotional intensity in an earlier norming study. A brief description of the norming study is presented at the end of this section. For more detailed explanations of the stimulus validation procedures and selection criteria, please refer to our recent study (Lin et al., 2020b).

The prosodic and semantic stimuli consisted of a series of disyllabic words, which were produced by four professionally trained speakers (two females and two males) in a quiet laboratory setting, and digitized at a sampling rate of 44.100 kHz with a 16-bit resolution. These words were common to native speakers of Mandarin Chinese, as verified by a word familiarity test in the norming procedure (see Supplemental Material S1 for the selected words for the prosodic and semantic stimuli and the word familiarity level). No significant difference was found in the familiarity rating of words in different emotional categories (all $p > .05$). Each word was pronounced 3 times by the four speakers, and the best ones were chosen for presentation based on the results of the norming study. For the prosodic stimuli, the speakers uttered semantically neutral disyllabic concrete Chinese words, the English translation of which are specified in the square brackets in the following examples (e.g., 肥皂 [soap]), 报纸 [newspaper]) in happy, sad, angry, and neutral prosody. For the semantic stimuli, the speakers enunciated the disyllabic words denoting happy (e.g., 欢乐 [cheerful]), sad (e.g., 抑郁 [depressed]), angry (e.g., 恼火 [annoyed]) emotions, and neutrality (e.g., 适中 [moderate]) in neutral tone of voice. The duration measures of the prosodic and semantic stimuli produced by female and male encoders are summarized in Tables 1 and 2, respectively. The mean f0 measures of the prosodic and semantic stimuli produced by female and male encoders are summarized in Tables 3 and 4, respectively.

There were 64 stimuli for each of the facial, prosodic, and semantic channels. The number of stimuli in each of the three channels was balanced between four emotion categories and between actors and actresses, yielding eight emotional stimuli per category for each encoder's gender in each channel (i.e., 64 stimuli = 8 faces/words × 4 emotions × 2 encoders). These stimuli went through a perceptual validation procedure in a norming study by 23 native speakers of Mandarin Chinese (11 women and 12 men, mean age ± *SD* = 23.7 ± 2.6), who did not participate in this study. All included stimuli received over 90% accuracy for emotion category recognition. All emotional (i.e., happy, sad, and angry) stimuli received an average rating above 3 for emotion intensity on a 7-point Likert scale (0 = *not intense*, 6 = *very intense*), and neutral stimuli had an intensity score around the midpoint on the scale. Only words with an average rating above 3 for familiarity on a 7-point Likert scale (0 = *not familiar*, 6 = *very familiar*) could be employed as the prosodic and semantic stimuli in the auditory modality. Table 5 shows the identification accuracy of emotional category and rating of emotional intensity for all included stimuli by female and male raters in the validation study. No significant difference was found in identification accuracy or emotional intensity for stimuli produced by female and male encoders in the same emotional category in each channel (all $p > .05$).

Procedure

The experiment was administered in a sound booth with the participant seated in a comfortable chair at around 70 cm from an LCD monitor. E-Prime (Version 2.0.8.22; Psychology Software Tools, 2012) was used for stimulus presentation. Emotional facial expressions were displayed in the center of an LCD screen over a constant white background, and the auditory stimuli of emotional prosody and

Table 1. Duration (milliseconds) of the prosodic stimuli.

Encoders' gender	Female		Male			
Emotion	*M*	*SD*	*M*	*SD*	*M/SD*	
Angry	1361.34	167.68	1233.78	226.59	1297.56	209.28
Happy	1066.14	93.88	1148.88	131.96	1107.51	121.76
Neutral	952.00	103.00	1066.67	142.72	1009.33	137.03
Sad	1912.63	164.70	1769.63	211.46	1841.13	202.56
M/SD	1323.02	396.05	1304.74	330.14	1313.88	364.70

Table 2. Duration (milliseconds) of the semantic stimuli.

Encoders' gender	Female		Male			
Emotion	*M*	*SD*	*M*	*SD*	*M/SD*	
Angry	941.00	116.65	1063.38	173.31	1002.19	159.90
Happy	1007.00	77.67	1075.99	94.57	1041.50	93.16
Neutral	987.00	98.43	1142.00	193.49	1064.50	171.96
Sad	1001.00	87.64	1091.08	107.36	1046.04	107.85
M/SD	984.00	99.61	1093.12	151.26	1038.56	139.20

Table 3. Mean f0 (Hz) of the prosodic stimuli.

Encoders' gender	Female		Male			
Emotion	*M*	*SD*	*M*	*SD*	*M/SD*	
Angry	198.10	37.11	139.45	22.69	168.78	42.49
Happy	272.92	30.42	219.44	59.09	246.18	54.07
Neutral	196.58	21.03	119.61	9.51	158.09	41.81
Sad	178.99	21.27	124.48	7.17	151.74	31.54
M/SD	211.65	45.90	150.75	51.61	181.20	57.56

semantics were presented binaurally over Sennheiser HD280 PRO headphones at 70 dB SPL.

The experiment consisted of three subtasks, namely, an emotional face identification task, an emotional prosody identification task, and an emotional semantics identification task. There were a total of 192 trials separated into three blocks, each of which contained 64 trials of emotional stimuli presented in one of the facial, prosodic, or semantic channels. Each trial began with a fixation cross in the center of the screen for 1,100 ms, followed by a target expression of emotion. Participants were asked to identify the emotion portrayed by the actor or actress by pressing one of the four emotion-coded keys on a keyboard ("v" for happy, "b" for sad, "n" for angry, and "m" for neutral) as quickly as possible. Both accuracy and response time were collected. Reaction time was measured from the onset of stimulus presentation until the key press. After the response was made, the blank screen was displayed for 1,000 ms before the next trial began. The presentation order of the blocks was fully randomized across participants, and the trial order in each block was pseudorandomized to ensure that none of the two subsequent trials presented emotion of the same category. Participants started the experiment after familiarizing themselves with the experimental procedure by completing eight practice trials in each block with 100% accuracy. They were allowed to have a short rest for up to 30 s after running every 32 trials.

Statistical Analyses[1]

A series of linear mixed-effects models in R (Version 3.6.1) with the lme4 package were employed for data analysis (Bates, Mächler, et al., 2015). Accuracy data were transformed into rationalized arcsine unit (rau) to mitigate the ceiling effects (Studebaker, 1985), and reaction time data were log-transformed to undermine the effect of positive skewness (Baayen & Milin, 2010). The transformed accuracy and reaction time data were respectively entered as dependent variables in the mixed-effects models. The fixed categorical factors consisted of communication channel (three levels: facial, prosodic, and semantic), emotion category (four levels: angry, happy, neutral, and sad), encoders' gender (two levels: female and male), and decoders' gender (two levels: female and male). The facial channel, angry stimuli, and female encoders/decoders were set as the baseline level for communication channel, emotion category, and encoders' and decoders' gender, respectively. Where necessary, pairwise planned comparisons were conducted to verify differences between genders in encoding as well as in decoding, between the prosodic and semantic channels, and between emotional categories. Specifically, we applied Tukey's post hoc tests with $\alpha = .05$ in the emmeans package (Lenth, 2020) for all pairwise contrasts in case of a significant main effect or interaction, and selectively reported the contrasts of significant main effects and highest level interactions involving differences of encoders' or decoders' gender in text. Individual decoders and test items were entered as random factors for intercepts. The models with intercepts, coefficients, and error terms for accuracy and reaction time analyses are specified as follows[2]:

$$\begin{aligned}&\text{Accuracy(rau)}_{ij}/\log(\text{Reaction time})_{ij} = \beta_0 + \\&(\beta_1 \times \text{channel}) + (\beta_2 \times \text{emotion}) + (\beta_3 \times \text{encoder}) + \\&(\beta_4 \times \text{decoder}) + (\beta_5 \times \text{channel} \times \text{emotion}) + \\&(\beta_6 \times \text{channel} \times \text{encoder}) + (\beta_7 \times \text{emotion} \times \text{encoder}) + \\&(\beta_8 \times \text{channel} \times \text{decoder}) + (\beta_9 \times \text{emotion} \times \text{decoder}) + \\&(\beta_{10} \times \text{encoder} \times \text{decoder}) + \beta_{11} \times \text{channel} \times \text{decoder} \times \\&\text{encoder})+(\beta_{12} \times \text{emotion} \times \text{decoder} \times \text{encoder}) + \\&(\beta_{13} \times \text{channel} \times \text{emotion} \times \text{decoder}) + (\beta_{14} \times \text{channel} \times \\&\text{emotion} \times \text{encoder}) + (\beta_{15} \times \text{channel} \times \text{emotion} \times \\&\text{encoder} \times \text{decoder}) + b_{0i} + b_{1j} + \varepsilon_{ij}\end{aligned} \quad (1)$$

In these models, β_0 represented the intercept, which was the predicted outcome when all other predictors were equal to 0.

[1]No speed accuracy trade-off was found when examining the relationship between accuracy and reaction time by a Pearson correlational analysis ($r = -.053$, $p = .312$). Thus, we separately analyzed accuracy and RT data using linear mixed-effects models.

[2]Given that our research hypotheses anticipated differences in the decoding/encoding of emotions in different communication channels as a function of the encoders'/decoders' gender, we started with the maximal model with the most random-effects components in the linear mixed-effects analyses, including emotion by encoders' gender (Hypothesis 1), emotion and channel interaction by decoders' gender (Hypothesis 2), and Emotion × Channel interaction by decoders and encoders' gender (Hypothesis 3). These random-effects structures did not improve the model fit but produced singular fits and nonconverging models, suggesting that the models were over parameterized (Bates, Kliegl et al., 2015). Thus, we reduced the random-effects structures to eliminate the problem of singularity and nonconvergence (Barr et al., 2013). We selected the highest nonsingular converging models as the final models for data analyses.

Table 4. Mean f0 (Hz) of the semantic stimuli.

Encoders' gender	Female		Male			
Emotion	*M*	*SD*	*M*	*SD*	*M/SD*	
Angry	185.25	19.17	124.68	12.71	154.96	34.37
Happy	229.04	26.81	121.15	4.39	175.10	57.26
Neutral	192.72	20.40	122.35	9.54	157.53	38.62
Sad	225.21	30.34	121.16	5.20	173.18	56.40
M/SD	208.05	31.28	122.34	8.76	167.75	51.28

β_1, β_2… β_{15} represented the coefficients for the four fixed factors and their interactions, respectively. These coefficients reflected how much the outcome variable changed relative to a unit of change in the corresponding predictors. The random intercepts were represented as b_{0i} and b_{1j}, where *i* varied according to decoder participants and *j* varied according to test items. An error term (ε) was also included to account for the distance between the predicted value and the actual data point (i.e., residual).

Results

Accuracy

Overall, participants completed the experiment with considerably high accuracy (M = 95.7%, SD = 9.0%). Table 6 summarizes the mixed-effects models for accuracy data.

Linear mixed-effects analyses on accuracy in raus showed no main effect of encoders' gender ($p > .05$), but significant main effects of sensory channel, $\chi^2(2) = 20.83$, $p < .001$; emotion category, $\chi^2(3) = 102.70$, $p < .001$; and decoders' gender, $\chi^2(1) = 6.45$, $p = .01$. Notably, female decoders were slightly more accurate in emotion recognition than their male counterparts ($\hat{\beta}_4 = 2.75$, *standard error* [*SE*] = 1.08, $t = 2.56$, $p = .01$, Cohen's $d = 0.16$). There were also significant interactions between channel and emotion, $\chi^2(6) = 96.82$, $p < .001$; between channel and encoders' gender, $\chi^2(2) = 44.77$, $p < .001$; and emotion and encoders' gender, $\chi^2(2) = 10.70$, $p = .01$. More importantly, we found significant three-way interactions among channel, emotion, and decoders' gender, $\chi^2(6) = 14.25$, $p = .03$, Cohen's $f^2 = .21$, and among channel, emotion, and encoders' gender, $\chi^2(6) = 100.25$, $p < .001$, $f^2 = .27$, which are illustrated in Figures 1 and 2, respectively.

The pairwise comparison of female and male decoders' performances under each emotion category in the three tasks (see Supplemental Material S2) revealed that female decoders achieved significantly higher accuracy than their male counterparts only when identifying words with angry semantics ($\hat{\beta}_{13} = 6.65$, $SE = 2.58$, $t = 2.58$, $p = .01$, $d = 0.41$), happy ($\hat{\beta}_{13} = 8.03$, $SE = 2.58$, $t = 3.11$, $p = .002$, $d = 0.49$) and sad ($\hat{\beta}_{13} = 6.01$, $SE = 2.58$, $t = 2.33$, $p = .02$, $d = 0.37$) tone of voice, and sad ($\hat{\beta}_{13} = 6.13$, $SE = 2.58$, $t = 2.38$, $p = .02$, $d = 0.38$) faces.

Table 5. Identification accuracy of emotional category and rating of emotional intensity for the stimuli adopted in the experiment by female and male raters in the norming study.

Stimulus type	Emotion category	Rater's gender	Identification accuracy of emotional category		Rating of emotional intensity	
			M	*SD*	*M*	*SD*
Face	Angry	Female	92.59%	1.73%	4.41	0.41
		Male	93.98%	2.07%	4.59	0.72
	Happy	Female	93.98%	2.07%	4.64	0.47
		Male	92.59%	1.73%	4.40	0.72
	Sad	Female	92.13%	1.31%	4.14	0.60
		Male	93.06%	1.96%	3.93	0.33
	Neutral	Female	92.59%	1.73%	3.48	0.23
		Male	91.78%	0.31%	3.34	0.30
Prosody	Angry	Female	95.65%	3.55%	4.65	0.29
		Male	96.14%	3.80%	4.89	0.31
	Happy	Female	92.27%	3.42%	4.44	0.33
		Male	92.75%	3.55%	4.02	0.29
	Sad	Female	93.24%	4.63%	4.23	0.18
		Male	94.20%	2.90%	4.57	0.14
	Neutral	Female	93.24%	2.98%	3.27	0.22
		Male	92.27%	1.81%	3.28	0.22
Semantics	Angry	Female	96.74%	2.88%	5.05	0.28
		Male	96.74%	3.61%	5.04	0.24
	Happy	Female	97.83%	3.07%	4.77	0.35
		Male	96.86%	2.85%	4.68	0.34
	Sad	Female	96.74%	4.21%	4.34	0.51
		Male	97.28%	3.03%	4.51	0.42
	Neutral	Female	95.65%	3.77%	3.08	0.17
		Male	96.20%	2.61%	3.02	0.17

Note. Participants identified the emotional category of the stimuli and rated the intensity on a 7-point scale (0 = *not intense*, 6 = *very intense*).

Table 6. Summary of mixed-effects models for accuracy data.

Model	Target effect	χ^2	*df*	R^2_{acc}	Cohen's f^2_{acc}
1	Chan	20.83***	2	.06	.06
2	Emo	102.70***	3	.10	.11
3	Gen (En)	2.20	1	.10	.12
4	Gen (De)	6.45*	1	.10	.12
5	Chan × Emo	96.82***	6	.15	.17
6	Chan × Gen (En)	44.77***	2	.16	.20
7	Emo × Gen (En)	10.70*	3	.17	.20
8	Chan× Gen (De)	.70	2	.17	.20
9	Emo × Gen (De)	6.99	3	.17	.21
10	Gen (En) × Gen (De)	.00	1	.17	.21
11	Chan × Gen (En) × Gen (De)	1.21	2	.17	.21
12	Emo × Gen (En) × Gen (De)	.95	3	.17	.21
13	Chan × Emo × Gen (De)	14.25*	6	.18	.21
14	Chan × Emo × Gen (En)	100.25***	6	.21	.27
15	Chan × Emo × Gen (De) × Gen (En)	2.80	6	.22	.27

Note. *df* = degree of freedom; acc = accuracy; Chan = Channel; Emo = Emotion; Gen (En) = Encoders' gender; Gen (De) = Decoders' gender.

*$p < .05$. ***$p < .001$.

We also compared emotions encoded by male and female actors in the three channels (see Supplemental Material S2). For the angry emotion, no significant difference was found between female and male actors in all three channels ($p > .05$). For the happy emotion, stimuli portrayed by female actors were more accurately identified than those produced by males in the facial ($\hat{\beta}_{14} = 6.67$, $SE = 2.39$, $t = 2.79$, $p = .005$, $d = 0.42$) and prosodic ($\hat{\beta}_{14} = 12.66$, $SE = 2.39$, $t = 5.29$, $p < .001$, $d = 0.80$) channels, whereas men-produced stimuli yielded higher accuracy in the semantic channel ($\hat{\beta}_{14} = -22.19$, $SE = 2.39$, $t = -9.28$, $p < .001$, $d = -1.40$). Similarly, women-encoded sad stimuli elicited better performances in the facial ($\hat{\beta}_{14} = 6.70$, $SE = 2.39$, $t = 2.80$, $p = .005$, $d = 0.42$) and prosodic ($\hat{\beta}_{14} = 9.92$, $SE = 2.39$,

Figure 1. Identification accuracy of four types of emotional stimuli presented in three different communication channels by female and male decoders. Mean accuracy is displayed in the bar charts, with error bars showing 95% confidence intervals. Asterisks mark the significance level: *$p < .05$; **$p < .01$.

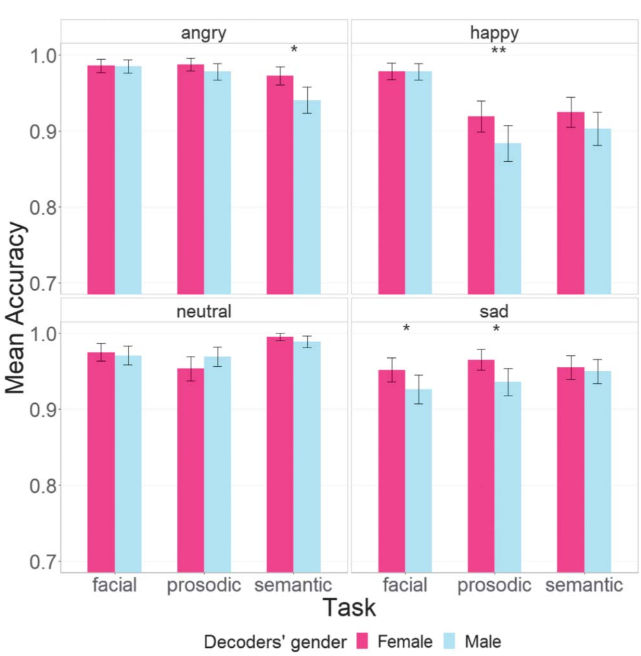
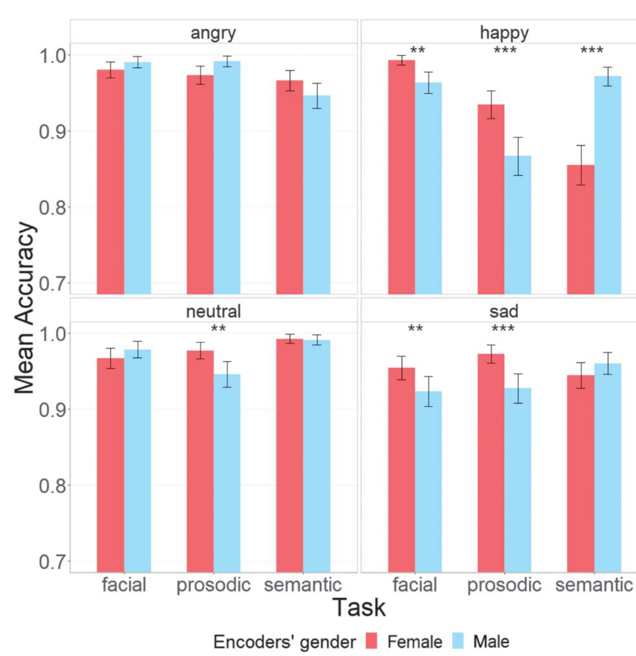


Figure 2. Identification accuracy for stimuli produced by female and male encoders in three different communication channels. Mean accuracy is displayed in the bar charts, with error bars showing 95% confidence intervals. Asterisks mark the significance level: **$p < .01$; ***$p < .001$.



$t = 4.15$, $p < .001$, $d = 0.62$) channels irrespective of decoders' gender, and no significant difference was found between encoders of the two genders in the semantic channel ($p > .05$). For neutrality, stimuli produced by females triggered significantly more accurate responses only in the prosodic channel ($\hat{\beta}_{14} = 6.23$, $SE = 2.39$, $t = 2.60$, $p = .009$, $d = 0.39$), and no significant difference was found in facial and semantic channels.

Reaction Time

When analyzing reaction time data, we excluded responses over 2 *SD*s from the mean time (3.9%) and incorrect responses (3.6%; Baayen & Milin, 2010; Chien et al., 2017). Table 7 summarizes the mixed-effects models for reaction time data.

Linear mixed-effects analyses on log-transformed reaction time showed no main effect of decoders' gender ($p > .05$), but main effects of channel, $\chi^2(2) = 2126.28$, $p < .001$; emotion, $\chi^2(3) = 806.82$, $p < .001$; and encoders' gender, $\chi^2(3) = 6.24$, $p = .01$. Notably, emotions encoded by females triggered slightly more rapid responses in emotion recognition than those by males ($\hat{\beta}_3 = -.01$, $SE = .01$, $t = -2.59$, $p = .01$, $d = -0.04$). There were also significant interactions between channel and emotion, $\chi^2(6) = 239.73$, $p < .001$; channel and encoders' gender, $\chi^2(2) = 39.22$, $p < .001$; emotion and encoders' gender, $\chi^2(3) = 44.32$, $p < .001$; encoders' gender and decoders' gender, $\chi^2(1) = 5.30$, $p = .02$; and among channel, emotion, and encoders' gender, $\chi^2(6) = 218.99$, $p < .001$. More importantly, we found a significant interaction among channel, emotion, and encoders and decoders' gender, $\chi^2(6) = 14.52$, $p = .02$, $f^2 = .42$, which is illustrated in Figure 3.

A pairwise comparison was conducted to parse out the four-way interaction effect (see Supplemental Material S3). Anger elicited significantly faster responses when encoded by the actors than actresses in the facial channel for both female ($\hat{\beta}_{15} = .09$, $SE = .03$, $t = 3.29$, $p = .001$, $d = 0.25$) and male ($\hat{\beta}_{15} = .18$, $SE = .03$, $t = 6.72$, $p < .001$, $d = 0.52$) observers. Actor-encoded angry stimuli also exhibited a processing advantage in the prosodic channel regardless of decoders' gender (female decoders: $\hat{\beta}_{15} = .06$, $SE = .03$, $t = 2.42$, $p = .02$, $d = 0.19$; male decoders: $\hat{\beta}_{15} = .08$, $SE = .03$, $t = 3.28$, $p = .001$, $d = 0.25$). However, words with angry semantics were more rapidly identified when spoken by the actresses for both female ($\hat{\beta}_{15} = -.11$, $SE = .03$, $t = -4.16$, $p < .001$, $d = -0.33$) and male ($\hat{\beta}_{15} = -.11$, $SE = .03$, $t = -4.27$, $p < .001$, $d = -0.34$) decoders. In terms of happiness, actress-encoded stimuli triggered faster responses compared with those portrayed by the actors in the facial (female decoders: $\hat{\beta}_{15} = -.15$, $SE = .03$, $t = -5.86$, $p < .001$, $d = -0.46$; male decoders: $\hat{\beta}_{15} = -.09$, $SE = .03$, $t = -3.43$, $p < .001$, $d = -0.26$) and prosodic channels (female decoders: $\hat{\beta}_{15} = -.19$, $SE = .03$, $t = -6.92$, $p < .001$, $d = -0.56$; male decoders: $\hat{\beta}_{15} = -.11$, $SE = .03$, $t = -4.05$, $p < .001$, $d = -0.33$). By contrast, words with happy semantics were recognized faster when spoken by the actors for both female ($\hat{\beta}_{15} = .10$, $SE = .03$, $t = 3.49$, $p < .001$, $d = 0.28$) and male ($\hat{\beta}_{15} = -.09$, $SE = .03$, $t = 3.36$, $p < .001$, $d = 0.27$) decoders. For the sad emotion, both female ($\hat{\beta}_{15} = .07$, $SE = .03$, $t = 2.74$, $p = .006$, $d = 0.22$) and male ($\hat{\beta}_{15} = .14$, $SE = .03$, $t = 5.22$, $p < .001$ $d = 0.41$) decoders made faster responses when encountering emotional semantics expressed by the actors than actresses. There was also an interaction between encoders' and decoders' gender for the perception of sadness in the facial and prosodic channels: While male decoders made faster responses when identifying sad faces ($\hat{\beta}_{15} = -.06$, $SE = .03$, $t = -2.35$, $p = .02$, $d = -0.19$) and prosody ($\hat{\beta}_{15} = -.09$, $SE = .03$, $t = -1.97$, $p < .05$, $d = -0.15$) portrayed by the actresses than those by the actors, there was no such significant difference in reaction time among female decoders ($p > .05$). Neutral stimuli were identified

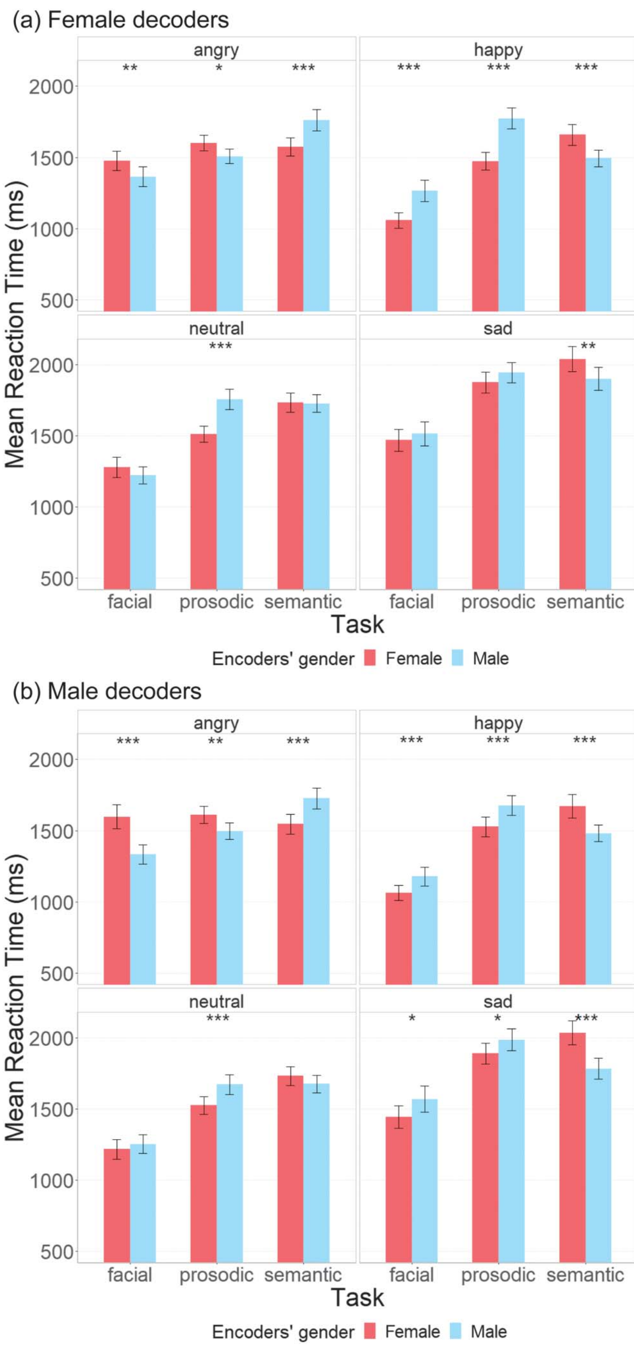
Table 7. Summary of mixed-effects models for reaction time data.

Model	Target effect	χ^2	*df*	R^2_{rt}	Cohen's f^2_{rt}
1	Chan	2126.28***	2	.23	.13
2	Emo	806.82***	3	.27	.30
3	Gen (En)	6.24*	1	.27	.37
4	Gen (De)	.18	1	.27	.37
5	Chan × Emo	239.73***	6	.28	.37
6	Chan × Gen (En)	39.22***	2	.29	.40
7	Emo × Gen (En)	44.32***	3	.29	.40
8	Chan× Gen (De)	.91	2	.29	.40
9	Emo × Gen (De)	.80	3	.29	.40
10	Gen (En) × Gen (De)	5.30*	1	.29	.40
11	Chan × Gen (En) × Gen (De)	.86	2	.29	.40
12	Emo × Gen (En) × Gen (De)	2.30	3	.29	.40
13	Chan × Emo × Gen (De)	5.32	6	.29	.40
14	Chan × Emo × Gen (En)	218.99***	6	.30	.40
15	Chan × Emo × Gen (De) × Gen (En)	14.52*	6	.30	.42

Note. *df* = degree of freedom; rt = reaction time; Chan = Channel; Emo = Emotion; Gen (En) = Encoders' gender; Gen (De) = Decoders' gender.

*$p < .05$. ***$p < .001$.

Figure 3. Reaction time for stimuli produced by female and male encoders in three different communication channels by (a) female and (b) male decoders. Mean reaction time is displayed in the bar charts, with error bars showing 95% confidence intervals. Asterisks mark the significance level: *$p < .05$; **$p < .01$; ***$p < .001$.



faster when produced by female than male actors in the prosodic channel irrespective of decoders' gender (female decoders: $\hat{\beta}_{15} = -.14$, $SE = .03$, $t = -5.37$, $p < .001$, $d = -0.42$; male decoders: $\hat{\beta}_{15} = -.09$, $SE = .03$, $t = -3.49$, $p < .001$, $d = -0.27$). No significant difference was found between male and female decoders in the facial and semantic channels ($p > .05$).

Discussion

This study investigated how emotion categories and communication channels influenced gender differences in emotion expression and perception. Specifically, we compared how emotional (i.e., angry, happy, sad) signals that were encoded by female and male actors/speakers in three sensory channels (i.e., face, prosody, and semantics) were decoded by the observers/listeners of the two genders in a high-context Chinese culture. Overall, our results revealed that emotion identification performances are modulated by encoders' gender, decoders' gender, and their reciprocal interactions. In addition, these striking differences in men and women's abilities to express and identify emotions significantly vary across emotion categories and communication channels, suggesting that both variables are essential when understanding the nature of gender differences in emotion expression and perception. These major findings will be discussed in detail in the following three subsections.

Identification Performance by Encoders' Gender

As predicted in the first hypothesis, there was a main effect of encoders' gender in reaction time, showing an overall female advantage in emotion expression. However, beyond our expectation, such a gender difference in encoding emotions varies across not only emotion categories but also sensory channels. The results showed that nonverbal expressions (i.e., face and prosody) of both happiness and sadness were more accurately and rapidly identified when encoded by females, but verbal content (i.e., semantics) of the corresponding emotion category was more perceptually salient when encoded by males. Conversely, angry faces and prosody triggered faster responses when encoded by males, but angry semantics yielded more rapid identification when encoded by females. These results indicate that identifications of different categories of emotions encoded by female and male actors vary depending on verbal and nonverbal communication channels.

Our findings provide robust evidence for the emotion-specific effects on gender differences in emotional expressiveness, which is consistent with traditional gender-stereotypical norms of emotional expression and recent empirical investigations on facial affect and emotional prosody. According to the literature, women express prosocial emotions (e.g., happiness) and internalizing negative emotions (e.g., sadness) more frequently and intensely than men, whereas men are typically more associated with emotions that imply power and dominance (e.g., anger; Becker et al., 2007; Bonebright et al., 1996; Calvo & Lundqvist, 2008; Chaplin, 2015; Collignon et al., 2010; Fischer, 2000; Fischer & LaFrance, 2014; Glenberg et al., 2009; Hess et al., 2007; Hofmann et al., 2006; Johnson et al., 2011; Kelly & Hutson-Comeaux, 1999; Öhman et al., 2010; Plant et al., 2000). One relevant account for why men and women differ in how well they can express particular types of emotion is the distinctive social roles and accompanying positions of the two genders (see Chaplin [2015] and Fischer & LaFrance [2014] for reviews).

Interestingly, such emotion-specific abilities of expressing emotions among men and women are represented differently between verbal and nonverbal channels in our study. This novel finding suggests that gender differences in encoding emotions are also channel specific. It is possible that female and male expressivity in verbal and nonverbal channels could be mediated by the emotional arousal/intensity of the stimuli. The facial expressions and prosody of happiness and sadness could be more easily identified when portrayed by women despite lower emotional arousal/intensity (compared with those expressions of anger), since female advantages in emotional expression have been widely acknowledged across studies on nonverbal interactions (Belin et al., 2008; Chaplin, 2015; Collignon et al., 2010; Cortes et al., 2017; Fischer et al., 2018; Fischer & LaFrance, 2014; J. A. Hall, 1984; Hoffmann et al., 2010; Koeda et al., 2013; Lausen & Schacht, 2018; Thompson & Voyer, 2014; Vasconcelos et al., 2017). For angry faces and prosody, however, female superiority in emotional expression could be overpowered by highly arousing negative stimuli, which are more likely to be associated with men when presented in nonverbal channels (Vasconcelos et al., 2017). For the encoding of nonverbal emotional information, men displayed greater advantages in the expression of word meaning, especially words denoting happiness and sadness. It is likely that, unlike nonverbal communication that favors women as they are more competent in demonstrating subtle changes in stimulus properties (e.g., facial features and acoustic variations), verbal communication relies on objective and global semantic processing (Fischer et al., 2018; Hoffmann et al., 2010; Schmid et al., 2011). In this case, happy and sad words that are less emotionally arousing could enjoy greater perceptual advantages even when uttered by men, whereas women could manifest their expressivity only when conveying emotional semantics with higher arousal/intensity (i.e., anger).

Identification Performance by Decoders' Gender

By examining the effect of decoders' gender on emotion identification, we confirmed our second hypothesis that women were generally more accurate in recognizing emotional signals than men. This female superiority effect was especially prominent for the processing of sad faces, happy and sad prosody, and angry semantics, though each experimental condition displayed a trend for women to outperform men in identifying emotional (i.e., happy, sad and angry) signals. These results are in agreement with several reviews and meta-analyses on gender differences, in which most of the included studies have shown an overall female advantage in decoding emotions albeit small to moderate differences (J. A. Hall, 1978, 1984; Kret & de Gelder, 2012; Thompson & Voyer, 2014). From an evolutionary perspective, the fact that women are superior in recognizing a wide range of emotions displayed in different channels could be explained by different biological competence and social roles of the two genders: Women are "wired" from birth to be more perceptually sensitive than men to emotional signals, which enables them to better detect emotions in regard to caretaking, romantic relations, and socialization (Brody & Hall, 2010; Collignon et al., 2010; Fischer et al., 2018; J. A. Hall, 1978).

In terms of the emotion-specific effects, we replicated previous observations of women's merits in identifying happiness in the auditory prosodic channel (Bonebright et al., 1996; Diaconescu et al., 2013; Fujisawa & Shinohara, 2011; Lambrecht et al., 2014). We also revealed female advantages in the recognition of negative emotions (i.e., sadness and anger), which is consistent with existing research using various types of facial, prosodic, and semantic stimuli (Deng et al., 2016; Forni-Santos & Osorio, 2015; Gohier et al., 2011; Li et al., 2014; Menezes et al., 2017; Smith et al., 1989; Vasconcelos et al., 2017; Zupan et al., 2016). Since the fitness threat hypothesis advocates an exclusive negative bias of women in emotion processing, the attachment promotion hypothesis appears to better accommodate our data (Hampson et al., 2006; Menezes et al., 2017; Thompson & Voyer, 2014). Interestingly, though women demonstrated a more generalized advantage in emotion perception across both positive and negative emotion categories, this female superiority effect was only observed in particular combinations of emotion and channel in the accuracy measure. Thus, it warrants further investigation under what circumstances, in what measures, and to what extent these two predictions of the differential emotion theory are established (Ackerman et al., 1998).

Another intriguing finding in line with our prediction is that the emotion-specific effects on female advantages in emotion perception are represented in verbal and nonverbal channels in a differentiated manner. Happiness and sadness bear greater processing advantages for women when presented in nonverbal channels, whereas anger is more perceptually salient only when denoted in verbal content. One conceivable explanation for the absence of gender-specific advantages in decoding anger in nonverbal channels is that humans are innately sensitive to socially salient and attention-catching stimuli (e.g., screams, alarm calls) in the environment irrespective of gender (Fox et al., 2000; Lausen & Schacht, 2018). It is also plausible that interlocutors tend to show their dissatisfaction implicitly and avoid direct confrontations as suggested by anger during nonverbal communication in a high-context Chinese culture (E. T. Hall, 1989). Of particular note is that this channel-specific variation in female decoders' perception of different emotion categories coincides with our observations of encoders' gender, suggesting a correspondence between these two types of gender differences. Up till now, less clear has been the extent to which and the reason why emotion expression and perception associate with each other, which warrants examination in future studies.

Interplay Between Encoders' and Decoders' Gender

Our final hypothesis on the interaction between encoders' and decoders' gender was partially supported by evidence of sadness processing in the two nonverbal channels. While male decoders made significantly faster responses when identifying sad faces and prosody portrayed by the actresses than the actors, female decoders showed a similar perceptual pattern but their performance did not reach statistical significance and demonstrated smaller effect sizes. Our results provided limited support to the opposite-gender hypothesis in facial (Erwin et al., 1992; Hofmann et al., 2006; Kim et al., 2019; Rahman et al., 2004) and vocal (Cretser et al., 1982; Erwin et al., 1992; Ethofer et al., 2007; Lambrecht et al., 2014) emotion processing among male decoders. Such an opposite-gender bias is selective to sad faces and prosody, which has again concurred with the fact that gender differences are modulated by the interaction between emotion categories and communication channels (Hawk et al., 2009; Lambrecht et al., 2014; Thompson & Voyer, 2014; Williams et al., 2009).

Notably, in our results, the opposite-gender effect was only established for male decoders, while females exhibited a tendency to respond rapidly to the sad faces and prosody portrayed by the encoders of the same gender. The performance of women decoders did not reach significance possibly due to their overall superiority in emotion perception, which may undermine the effect triggered by encoders' gender. Another possible interpretation bears on gender norms of the expression of sadness, which is a sign of vulnerability such that it is more female stereotyped than male stereotyped. While females are inclined to endorse a unisex standard of sadness, males tend to hold double criteria with higher acceptability of women than men to express this type of emotion (Cretser et al., 1982; Fischer et al., 2013; Fischer & LaFrance, 2014; Glenberg et al., 2009; Kelly & Hutson-Comeaux, 1999; Plant et al., 2000; Warner & Shields, 2007). Furthermore, contextual and methodological differences might also account for why we did not observe a bidirectional gender bias for both female and male decoders as indicated in some earlier studies (Chun et al., 2012; Ethofer et al., 2007; Hoffmann et al., 2010). For instance, our study involved participants with a high-context Mandarin Chinese cultural background and more types of communication channels and different categories of basic emotions compared with previous ones, which may affect statistical sensitivity. It is possible that cultural influences may socialize men and women to act based on certain social expectations and thereby modify the meaning of emotional expressions as a function of both encoders' and decoders' genders (Kret & de Gelder, 2012). Thus, to what extent the interplay between encoders' and decoders' gender is emotion and channel specific requires further scrutiny to verify generalizability of the findings.

Implications, Limitations, and Future Studies

This study demonstrates how the emotion- and channel-specific effects converge to uncover the nature of gender differences in emotion expression and perception in a high-context Chinese culture. Our findings suggest a complex interplay of endogenous and exogenous variables in emotion processing, which sheds new light on the theoretical debates and underlying mechanisms of emotion cognition. Future studies can investigate how other types of individual differences in demographic (e.g., age) and psychosocial (i.e., attention, memory, language, personality traits) characteristics are associated with emotion processing performances under various interactional circumstances (Kidd et al., 2018; Lausen & Schacht, 2018). Our study also provides significant insights into exploring the pathogenesis and symptomatology of a wide range of mental disorders. Emotion processing deficits have been repeatedly recognized as a core symptom in several psychiatric illnesses that appear to be gender related, such as higher incidences of schizophrenia and autism among males, and mood disorders among females (Cahill, 2006; Deng et al., 2016; Feldman et al., 2018; Lin et al., 2018, 2020a). As the current research has evidenced robust gender differences in emotion processing across sensory channels and emotion categories among healthy subjects, it would be worthwhile to examine under what conditions and to what extent patients with these gender-related disorders show abnormality, which has the potential to facilitate disease prognosis and intervention in clinical practice. For example, it remains to investigate whether higher autistic prevalence among males, which often results in their social isolation, is related to their inferior abilities in expressing and perceiving prosocial and internalizing emotions in nonverbal communication as observed for happiness and sadness in this study.

One novel important finding of our study was that gender differences were observed in the accuracy measure in some combinations of emotion and channel, whereas in other experimental conditions, gender differences were only found in the reaction time measure. For instance, the Channel × Emotion × Decoders' Gender interaction was not significant in the measurement of response time. Instead, the four-way interaction among channel, emotion, and decoders and encoders' gender was significant in the measurement of reaction time but not accuracy. For the encoding of emotions, stimuli produced by men and women displayed a similar processing pattern but differed in significance levels and effect sizes between accuracy and reaction time data. It is likely that some potentially significant differences that might have had a larger magnitude were underestimated, despite efforts made to mitigate the influence of performance ceiling in statistical analyses. Consistent with previous cross-channel/modal investigations (Filippi et al., 2017; Lin et al., 2020b), these results suggest that the accuracy and response time measure may not exhibit the same profile in emotion processing, which underlies the necessity to integrate them for a more comprehensive interpretation of the results. Future efforts could employ new methods by combining behavioral experiments and neurophysiological techniques with higher temporal and spatial resolution (e.g., event-related potentials and functional magnetic resonance imaging), which can bring us closer to a detailed mechanistic understanding of the existing theoretical accounts on gender differences and emotion processing—in what measures, conditions, or contexts the theories are supported and in what measures, conditions, or contexts they diverge and may require alternative explanations.

There are some limitations in our study. First, our stimulus design involved a small number of encoders, especially for the auditory stimuli of emotional prosody and semantics. There could also be distinct perceptual and acoustic differences between acted and authentic emotional vocal expressions across cultures (Anikin & Lima, 2018; Cowen et al., 2019; Kitayama & Ishii, 2002). In addition, though we based our study in a high-context Mandarin Chinese culture, our decoder samples were about the same age and attended the same highly selective university in Shanghai. It is highly possible that students living in such an international city may have extensive accesses to western cultures. Thus, it remains to be examined to what extent the current findings and discussions can be extended to other cultural contexts. It would be more appropriate for future studies to involve a larger size and wider range of populations (Amorim et al., 2019; Lima et al., 2014; Lima & Castro, 2011), and test both Eastern and Western participants in a comparative approach, which can contribute to our understanding of gender differences in emotion processing in cross-linguistic and cross-cultural settings. Moreover, emotional information was conveyed through face, prosody, and semantics as three independent channels in our study. Though one may encounter real-life situations when they decode information through a single channel (e.g., talking on the phone, listening to news broadcast, or looking at one's face when he/she is not speaking; Hawk et al., 2009), it is more often the case that emotional signals are simultaneously transmitted through multiple channels. Thus, it is necessary to explore whether the gender differences observed in our unisensory emotion identification tasks can be extended to a multisensory communication setting where participants need to disentangle the interactions among different sensory channels (Filippi et al., 2017; Lin et al., 2020b; Thompson & Voyer, 2014).

Conclusions

The current research examined how well decoders of different genders can identify different categories of emotions encoded by female and male actors through verbal and nonverbal channels. The results indicated that women are generally more superior in emotion processing, but this gender superiority effect is constrained by channels and emotions: Happiness and sadness are better decoded and encoded by females in the nonverbal (i.e., facial and prosodic) channels, whereas anger is more perceptually salient in verbal (i.e., semantic) content; for male encoders, however, anger enjoys greater perceptual advantages when expressed in the nonverbal channels, whereas happiness and sadness are more prominent when denoted in verbal content. In addition, there is an emotion- and channel-specific interaction between the two types of gender. Male decoders tend to demonstrate more pronounced differences than female decoders when responding to sad faces and prosody portrayed by female encoders compared with the ones encoded by males. These findings clearly demonstrate that gender differences in emotion processing are modulated by emotion categories and communication channels, which has important implications for theoretical accounts on gender differences in encoding and decoding emotions, and practical applications in developing gender-appropriate diagnosis and intervention.

Acknowledgments

H.W.D. and Y.Z. were supported by the major program of National Fund of Philosophy and Social Science of China (18ZDA293). Y.Z. additionally received support from the University of Minnesota's Brain Imaging Grant and Grand Challenges Exploratory Research Grant for international collaboration. We also thank the editor for providing insightful comments and suggestions.

References

Ackerman, B. P., Abe, J. A. A., & Izard, C. E. (1998). Differential emotions theory and emotional development. In M. F. Mascolo & S. Griffin (Eds.), *What develops in emotional development?* (pp. 85–106). Springer. https://doi.org/10.1007/978-1-4899-1939-7_4

Amorim, M., Anikin, A., Mendes, A. J., Lima, C. F., Kotz, S. A., & Pinheiro, A. P. (2019). Changes in vocal emotion recognition across the life span. *Emotion, 21*(2), 315–325. https://doi.org/10.1037/emo0000692

Anikin, A., & Lima, C. F. (2018). Perceptual and acoustic differences between authentic and acted nonverbal emotional vocalizations. *Quarterly Journal of Experimental Psychology, 71*(3), 622–641. https://doi.org/10.1080/17470218.2016.1270976

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*(2), 12–28. https://doi.org/10.21500/20112084.807

Babchuk, W. A., Hames, R. B., & Thompson, R. A. (1985). Sex differences in the recognition of infant facial expressions of emotion: The primary caretaker hypothesis. *Ethology and Sociobiology, 6*(2), 89–101. https://doi.org/10.1016/0162-3095(85)90002-0

Bai, L., Ma, H., Huang, Y. X., & Luo, Y. J. (2005). The development of native Chinese affective picture system: A pretest in 46 college students. *Chinese Mental Health Journal, 19*(11), 719–722.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious mixed models*. arXiv.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

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

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal affective voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods, 40*(2), 531–539. https://doi.org/10.3758/BRM.40.2.531

Bonebright, T. L., Thompson, J. L., & Leger, D. W. (1996). Gender stereotypes in the expression and perception of vocal affect. *Sex Roles, 34*(5–6), 429–445. https://doi.org/10.1007/BF01547811

Brody, L. R., & Hall, J. A. (2010). *Gender, emotion, and socialization*. Springer. https://doi.org/10.1007/978-1-4419-1465-1_21

Cahill, L. (2006). Why sex matters for neuroscience. *Nature Reviews Neuroscience, 7*(6), 477–484. https://doi.org/10.1038/nrn1909

Calvo, M. G., & Lundqvist, D. (2008). Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behavior Research Methods, 40*(1), 109–115. https://doi.org/10.3758/BRM.40.1.109

Champely, S. (2020). *pwr: Basic functions for power analysis*. R package. https://CRAN.R-project.org/package=pwr

Chaplin, T. M. (2015). Gender and emotion expression: A developmental contextual perspective. *Emotion Review, 7*(1), 14–21. https://doi.org/10.1177/1754073914544408

Chien, Y.-F., Sereno, J. A., & Zhang, J. (2017). What's in a word: Observing the contribution of underlying and surface representations. *Language and Speech, 60*(4), 643–657. https://doi.org/10.1177/0023830917690419

Chun, J. W., Park, H. J., Park, I. H., & Kim, J. J. (2012). Common and differential brain responses in men and women to nonverbal emotional vocalizations by the same and opposite sex. *Neuroscience Letters, 515*(2), 157–161. https://doi.org/10.1016/j.neulet.2012.03.038

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research, 1242,* 126–135. https://doi.org/10.1016/j.brainres.2008.04.023

Collignon, O., Girard, S., Gosselin, F., Saint-Amour, D., Lepore, F., & Lassonde, M. (2010). Women process multisensory emotion expressions more efficiently than men. *Neuropsychologia, 48*(1), 220–225. https://doi.org/10.1016/j.neuropsychologia.2009.09.007

Cortes, D. S., Laukka, P., Lindahl, C., & Fischer, H. (2017). Memory for faces and voices varies as a function of sex and expressed emotion. *PLOS ONE, 12*(6), Article e0178423. https://doi.org/10.1371/journal.pone.0178423

Cowen, A. S., Laukka, P., Elfenbein, H. A., Liu, R., & Keltner, D. (2019). The primacy of categories in the recognition of 12 emotions in speech prosody across two cultures. *Nature Human Behaviour, 3*(4), 369–382. https://doi.org/10.1038/s41562-019-0533-6

Cretser, G. A., Lombardo, W. K., Lombardo, B., & Mathis, S. (1982). Reactions to men and women who cry: A study of sex differences in perceived societal attitudes versus personal attitudes. *Perceptual and Motor Skills, 55*(2), 479–486. https://doi.org/10.2466/pms.1982.55.2.479

Deng, Y., Chang, L., Yang, M., Huo, M., & Zhou, R. (2016). Gender differences in emotional response: Inconsistency between experience and expressivity. *PLOS ONE, 11*(6). Article e0158666. https://doi.org/10.1371/journal.pone.0158666

Diaconescu, A. O., Hasher, L., & McIntosh, A. R. (2013). Visual dominance and multisensory integration changes with age. *NeuroImage, 65,* 152–166. https://doi.org/10.1016/j.neuroimage.2012.09.057

Doi, H., Amamoto, T., Okishige, Y., Kato, M., & Shinohara, K. (2010). The own-sex effect in facial expression recognition. *NeuroReport, 21*(8), 564–568. https://doi.org/10.1097/WNR.0b013e328339b61a

Donges, U.-S., Kersting, A., & Suslow, T. (2012). Women's greater ability to perceive happy facial emotion automatically: Gender differences in affective priming. *PLOS ONE, 7*(7). Article e41745. https://doi.org/10.1371/journal.pone.0041745

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Psychology Press.

Ekman, P. (1992). Are there basic emotions? *Psychological Review, 99*(3), 550–553. https://doi.org/10.1037/0033-295X.99.3.550

Erwin, R. J., Gur, R. C., Gur, R. E., Skolnick, B., Mawhinney-Hee, M., & Smailis, J. (1992). Facial emotion discrimination: I. Task construction and behavioral findings in normal subjects. *Psychiatry Research, 42*(3), 231–240. https://doi.org/10.1016/0165-1781(92)90115-J

Ethofer, T., Wiethoff, S., Anders, S., Kreifelts, B., Grodd, W., & Wildgruber, D. (2007). The voices of seduction: Cross-gender effects in processing of erotic prosody. *Social Cognitive and Affective Neuroscience, 2*(4), 334–337. https://doi.org/10.1093/scan/nsm028

Feldman, J. I., Dunham, K., Cassidy, M., Wallace, M. T., Liu, Y., & Woynaroski, T. G. (2018). Audiovisual multisensory integration in individuals with autism spectrum disorder: A systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews, 95,* 220–234. https://doi.org/10.1016/j.neubiorev.2018.09.020

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40*(5), 532–538. https://doi.org/10.1037/a0015808

Filippi, P., Ocklenburg, S., Bowling, D. L., Heege, L., Güntürkün, O., Newen, A., & de Boer, B. (2017). More than words (and faces): Evidence for a Stroop effect of prosody in emotion word processing. *Cognition and Emotion, 31*(5), 879–891. https://doi.org/10.1080/02699931.2016.1177489

Fischer, A. H. (2000). *Gender and emotion: Social psychological perspectives*. Cambridge University Press. https://doi.org/10.1017/CBO9780511628191

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., Kret, M. E., & Broekens, J. (2018). Gender differences in emotion perception and self-reported emotional intelligence: A test of the emotion sensitivity hypothesis. *PLOS ONE, 13*(1). Article e0190712. https://doi.org/10.1371/journal.pone.0190712

Fischer, A. H., & LaFrance, M. (2014). What drives the smile and the tear: Why women are more emotionally expressive than men. *Emotion Review, 7*(1), 22–29. https://doi.org/10.1177/1754073914544406

Fischer, A. H., & Manstead, A. S. R. (2008). Social functions of emotion. In M. Lewis, J. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed.). Guilford.

Flom, R., Gentile, D. A., & Pick, A. D. (2008). Infants' discrimination of happy and sad music. *Infant Behavior & Development, 31*(4), 716–728. https://doi.org/10.1016/j.infbeh.2008.04.004

Forni-Santos, L., & Osorio, F. L. (2015). Influence of gender in the recognition of basic facial expressions: A critical literature review. *World Journal of Psychiatry, 5*(3), 342–351. https://doi.org/10.5498/wjp.v5.i3.342

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition & Emotion, 14*(1), 61–92. https://doi.org/10.1080/026999300378996

Fujisawa, T. X., & Shinohara, K. (2011). Sex differences in the recognition of emotional prosody in late childhood and adolescence. *Journal of Physiological Sciences, 61*(5), 429–435. https://doi.org/10.1007/s12576-011-0156-9

Fulton, E. K., Bulluck, M., & Hertzog, C. (2015). Orienting to face expression during encoding improves men's recognition of own

gender faces. *Acta Psychologica, 161,* 18–24. https://doi.org/10.1016/j.actpsy.2015.08.005
Garcia, S. E., & Tully, E. C. (2020). Children's recognition of happy, sad, and angry facial expressions across emotive intensities. *Journal of Experimental Child Psychology, 197,* 104881. https://doi.org/10.1016/j.jecp.2020.104881
Glenberg, A. M., Webster, B. J., Mouilso, E., Havas, D., & Lindeman, L. M. (2009). Gender, emotion, and the embodiment of language comprehension. *Emotion Review, 1*(2), 151–161. https://doi.org/10.1177/1754073908100440
Gohier, B., Senior, C., Brittain, P. J., Lounes, N., El-Hage, W., Law, V., Phillips, M. L., & Surguladze, S. A. (2011). Gender differences in the sensitivity to negative stimuli: Cross-modal affective priming study. *European Psychiatry, 28*(2), 74–80. https://doi.org/10.1016/j.eurpsy.2011.06.007
Hall, E. T. (1989). *Beyond culture*. Anchor.
Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin, 85*(4), 845–857. https://doi.org/10.1037/0033-2909.85.4.845
Hall, J. A. (1984). *Nonverbal sex differences: Communication accuracy and expressive style*. Johns Hopkins University Press.
Hall, J. A., & Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion, 4*(2), 201–206. https://doi.org/10.1037/1528-3542.4.2.201
Hampson, E., van Anders, S. M., & Mullin, L. I. (2006). A female advantage in the recognition of emotional facial expressions: Test of an evolutionary hypothesis. *Evolution and Human Behavior, 27*(6), 401–416. https://doi.org/10.1016/j.evolhumbehav.2006.05.002
Hawk, S. T., van Kleef, G. A., Fischer, A. H., & van der Schalk, J. (2009). "Worth a thousand words": Absolute and relative decoding of nonlinguistic affect vocalizations. *Emotion, 9*(3), 293–305. https://doi.org/10.1037/a0015178
Herlitz, A., & Lovén, J. (2013). Sex differences and the own-gender bias in face recognition: A meta-analytic review. *Visual Cognition, 21*(9–10), 1306–1336. https://doi.org/10.1080/13506285.2013.823140
Hess, U., Adams, R. B., Jr., & Kleck, R. E. (2007). When two do the same, it might not mean the same: The perception of emotional expressions shown by men and women. In U. Hess & P. Philippot (Eds.), *Group dynamics and emotional expression* (pp. 33–50). Cambridge University Press. https://doi.org/10.1017/CBO9780511499838.003
Hoffmann, H., Kessler, H., Eppel, T., Rukavina, S., & Traue, H. C. (2010). Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men. *Acta Psychologica, 135*(3), 278–283. https://doi.org/10.1016/j.actpsy.2010.07.012
Hofmann, S. G., Suvak, M., & Litz, B. T. (2006). Sex differences in face recognition and influence of facial affect. *Personality and Individual Differences, 40*(8), 1683–1690. https://doi.org/10.1016/j.paid.2005.12.014
Johnson, K. L., McKay, L. S., & Pollick, F. E. (2011). He throws like a girl (but only when he's sad): Emotion affects sex-decoding of biological motion displays. *Cognition, 119*(2), 265–280. https://doi.org/10.1016/j.cognition.2011.01.016
Juslin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion, 1*(4), 381–412. https://doi.org/10.1037/1528-3542.1.4.381
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
Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences, 22*(2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006
Kim, H. E., Hong, Y. J., Kyeong, S., & Kim, J. J. (2019). Different neural substrates at first impressions of same-sex and opposite-sex faces in women. *Neuroscience Letters, 709,* 134389. https://doi.org/10.1016/j.neulet.2019.134389
Kitayama, S., & Ishii, K. (2002). Word and voice: Spontaneous attention to emotional utterances in two languages. *Cognition and Emotion, 16*(1), 29–59. https://doi.org/10.1080/0269993943000121
Koeda, M., Belin, P., Hama, T., Masuda, T., Matsuura, M., & Okubo, Y. (2013). Cross-cultural differences in the processing of non-verbal affective vocalizations by Japanese and Canadian listeners. *Frontiers in Psychology, 4,* 105. https://doi.org/10.3389/fpsyg.2013.00105
Koerner, T. K., & Zhang, Y. (2018). Differential effects of hearing impairment and age on electrophysiological and behavioral measures of speech in noise. *Hearing Research, 370,* 130–142. https://doi.org/10.1016/j.heares.2018.10.009
Kret, M. E., & de Gelder, B. (2012). A review on sex differences in processing emotional signals. *Neuropsychologia, 50*(7), 1211–1221. https://doi.org/10.1016/j.neuropsychologia.2011.12.022
Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2014). Gender differences in emotion recognition: Impact of sensory modality and emotional category. *Cognition & Emotion, 28*(3), 452–469. https://doi.org/10.1080/02699931.2013.837378
Lausen, A., & Schacht, A. (2018). Gender differences in the recognition of vocal emotions. *Frontiers in Psychology, 9,* 882. https://doi.org/10.3389/fpsyg.2018.00882
Lawrence, K., Campbell, R., & Skuse, D. (2015). Age, gender, and puberty influence the development of facial emotion recognition. *Frontiers in Psychology, 6,* 761. https://doi.org/10.3389/fpsyg.2015.00761
Lenth, R. (2020). *emmeans: Estimated marginal means, aka least-squares means.* R Package.
Lewin, C., & Herlitz, A. (2002). Sex differences in face recognition: Women's faces make the difference. *Brain and Cognition, 50*(1), 121–128. https://doi.org/10.1016/S0278-2626(02)00016-7
Li, H., Yuan, J., & Lin, C. (2008). The neural mechanism underlying the female advantage in identifying negative emotions: An event-related potential study. *NeuroImage, 40*(4), 1921–1929. https://doi.org/10.1016/j.neuroimage.2008.01.033
Li, Y., Gu, F., Zhang, X., Yang, L., Chen, L., Wei, Z., Zha, R., Wang, Y., Li, X., Zhou, Y., & Zhang, X. (2014). Cerebral activity to opposite-sex voices reflected by event-related potentials. *PLOS ONE, 9*(4), Article e94976. https://doi.org/10.1371/journal.pone.0094976
Lima, C. F., Alves, T., Scott, S. K., & Castro, S. L. (2014). In the ear of the beholder: How age shapes emotion processing in nonverbal vocalizations. *Emotion, 14*(1), 145–160. https://doi.org/10.1037/a0034287
Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion, 11*(5), 1021–1031. https://doi.org/10.1037/a0024521
Lin, Y., Ding, H., & Zhang, Y. (2018). Emotional prosody processing in schizophrenic patients: A selective review and meta-analysis. *Journal of Clinical Medicine, 7*(10), 363. https://doi.org/10.3390/jcm7100363
Lin, Y., Ding, H., & Zhang, Y. (2020a). Multisensory integration of emotion in schizophrenic patients. *Multisensory Research, 33*(8), 865–901. https://doi.org/10.1163/22134808-bja10016
Lin, Y., Ding, H., & Zhang, Y. (2020b). Prosody dominates over semantics in emotion word processing: Evidence from cross-channel and cross-modal Stroop effects. *Journal of Speech,*

Language, and Hearing Research, 63(3), 896–912. https://doi.org/10.1044/2020_JSLHR-19-00258

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin, 126*(3), 424–453. https://doi.org/10.1037/0033-2909.126.3.424

Menezes, C. B., Hertzberg, J. C., das Neves, F. E., Prates, P. F., Silveira, J. F., & Vasconcellos, S. J. L. (2017). Gender and the capacity to identify facial emotional expressions. *Estudos de Psicologia, 22*(1), 1–9. https://doi.org/10.22491/1678-4669.20170001

Öhman, A., Juth, P., & Lundqvist, D. (2010). Finding the face in a crowd: Relationships between distractor redundancy, target emotion, and target gender. *Cognition & Emotion, 24*(7), 1216–1228. https://doi.org/10.1080/02699930903166882

Palermo, R., & Coltheart, M. (2004). Photographs of facial expression: Accuracy, response times, and ratings of intensity. *Behavior Research Methods, Instruments, & Computers, 36*(4), 634–638. https://doi.org/10.3758/BF03206544

Pell, M. D. (2002). Evaluation of nonverbal emotion in face and voice: Some preliminary findings on a new battery of tests. *Brain and Cognition, 48*(2), 499–504. https://doi.org/10.1006/brcg.2001.1406

Pell, M. D., Jaywant, A., Monetta, L., & Kotz, S. A. (2011). Emotional speech processing: Disentangling the effects of prosody and semantic cues. *Cognition and Emotion, 25*(5), 834–853. https://doi.org/10.1080/02699931.2010.516915

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

Proverbio, A. M., Riva, F., Martin, E., & Zani, A. (2010). Neural markers of opposite-sex bias in face processing. *Frontiers in Psychology, 1,* 169. https://doi.org/10.3389/fpsyg.2010.00169

Psychology Software Tools. (2012). *E-Prime 2.0*. https://www.pstnet.com

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rahman, Q., Wilson, G. D., & Abrahams, S. (2004). Sex, sexual orientation, and identification of positive and negative facial affect. *Brain and Cognition, 54*(3), 179–185. https://doi.org/10.1016/j.bandc.2004.01.002

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32*(1), 76–92. https://doi.org/10.1177/0022022101032001009

Scherer, K. R., & Scherer, U. (2011). Assessing the ability to recognize facial and vocal expressions of emotion: Construction and validation of the emotion recognition index. *Journal of Nonverbal Behavior, 35*(4), 305–326. https://doi.org/10.1007/s10919-011-0115-4

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research, 14*(2), 228–233. https://doi.org/10.1016/S0926-6410(02)00108-8

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2005). On the role of attention for the processing of emotions in speech: Sex differences revisited. *Cognitive Brain Research, 24*(3), 442–452. https://doi.org/10.1016/j.cogbrainres.2005.02.022

Schmid, P. C., Mast, M. S., Bombari, D., & Mast, F. W. (2011). Gender effects in information processing on a nonverbal decoding task. *Sex Roles, 65*(1), 102–107. https://doi.org/10.1007/s11199-011-9979-3

Smith, K. C., Ulch, S. E., Cameron, J. E., Cumberland, J. A., Musgrave, M. A., & Tremblay, N. (1989). Gender-related effects in the perception of anger expression. *Sex Roles, 20*(9/10), 487–499. https://doi.org/10.1007/BF00288197

Spence, C., Parise, C., & Chen, Y. C. (2011). The Colavita visual dominance effect. In M. M. Murray & M. T. Wallace (Eds.), *Frontiers in the neural bases of multisensory processes* (pp. 523–550). CRC Press. https://doi.org/10.1201/9781439812174-34

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *Journal of Speech and Hearing Research, 28*(3), 455–462. https://doi.org/10.1044/jshr.2803.455

Thayer, J., & Johnsen, B. H. (2000). Sex differences in judgement of facial affect: A multivariate analysis of recognition errors. *Scandinavian Journal of Psychology, 41*(3), 243–246. https://doi.org/10.1111/1467-9450.00193

Thompson, A. E., & Voyer, D. (2014). Sex differences in the ability to recognise non-verbal displays of emotion: A meta-analysis. *Cognition & Emotion, 28*(7), 1164–1195. https://doi.org/10.1080/02699931.2013.875889

Vasconcelos, M., Dias, M., Soares, A. P., & Pinheiro, A. P. (2017). What is the melody of that voice? Probing unbiased recognition accuracy with the Montreal affective voices. *Journal of Nonverbal Behavior, 41*(3), 239–267. https://doi.org/10.1007/s10919-017-0253-4

Warner, L. R., & Shields, S. A. (2007). The perception of crying in women and man: Angry tears, sad tears, and the "right way" to cry. In P. Philippot & U. Hess (Eds.), *Group dynamics and emotional expression* (pp. 92–117). Cambridge University Press. https://doi.org/10.1017/CBO9780511499838.006

Williams, L. M., Mathersul, D., Palmer, D. M., Gur, R. C., Gur, R. E., & Gordon, E. (2009). Explicit identification and implicit recognition of facial emotions: I. Age effects in males and females across 10 decades. *Journal of Clinical and Experimental Neuropsychology, 31*(3), 257–277. https://doi.org/10.1080/13803390802255635

Zupan, B., Babbage, D., Neumann, D., & Willer, B. (2016). Sex differences in emotion recognition and emotional inferencing following severe traumatic brain injury. *Brain Impairment, 18*(1), 36–48. https://doi.org/10.1017/BrImp.2016.22