

Cognition and Emotion

ISSN: 0269-9931 (Print) 1464-0600 (Online) Journal homepage: www.tandfonline.com/journals/pcem20

Sex differences in emotion recognition: investigating the moderating effects of stimulus features

Yasaman Rafiee & Anne Schacht

To cite this article: Yasaman Rafiee & Anne Schacht (2023) Sex differences in emotion recognition: investigating the moderating effects of stimulus features, Cognition and Emotion, 37:5, 863-873, DOI: 10.1080/02699931.2023.2222579

To link to this article: https://doi.org/10.1080/02699931.2023.2222579

View supplementary material

Published online: 13 Jun 2023.

Submit your article to this journal

Article views: 799

View related articles

View Crossmark data

Citing articles: 14 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=pcem20

COGNITION AND EMOTION
2023, VOL. 37, NO. 5, 863–873
https://doi.org/10.1080/02699931.2023.2222579

Routledge
Taylor & Francis Group

Check for updates

Sex differences in emotion recognition: investigating the moderating effects of stimulus features*

Yasaman Rafiee [a,b] and Anne Schacht[a,b]

[a]Department for Cognition, Emotion and Behavior, Affective Neuroscience and Psychophysiology Laboratory, Institute of Psychology, University of Göttingen, Göttingen, Germany; [b]Leibniz ScienceCampus "Primate Cognition", Göttingen, Germany

ABSTRACT

Emotion recognition – a prerequisite for social interactions – varies among individuals. Sex differences have been proposed as a central source of individual differences, although the existing evidence is rather heterogeneous. In the current study ($N = 426$), we investigated the potential moderating effects of stimulus features, including modality, emotion specificity, and the sex of the encoder (referring to the sex of the actor) on the magnitude of sex differences in emotion recognition. Our findings replicated women's overall better emotion recognition, particularly evident for negative expressions (fear and anger) compared to men. This outperformance was observed across all modalities, with the largest differences for audiovisually expressed emotions, while the sex of the encoder had no impact. Given our findings, future studies should consider these and other potential moderator variables to better estimate sex differences.

ARTICLE HISTORY
Received 7 March 2023
Revised 30 May 2023
Accepted 1 June 2023

KEYWORDS
Sex differences; emotion recognition; facial expression; vocal expression; audiovisual

Emotion recognition plays a substantial role in shaping and maintaining interpersonal communication and social interactions. The underlying mechanisms of this ability are highly complex and yet not completely understood (Bänziger, 2016). In addition to contextual factors, several person-related factors have been proposed to determine emotion recognition, resulting in observable individual variations in this ability (see Lewis et al., 2016).

One of the most studied sources of individual differences in emotion recognition is the biological sex of the decoders (Connolly et al., 2019; Kret & De Gelder, 2012). The existing literature suggests that males and females differ in their skills in recognising emotional expressions (see Kret & De Gelder, 2012), with women usually outperforming men (for an overview see Hall, 1978; Hall, 1984; McClure, 2000; Thompson & Voyer, 2014). This finding seems to be consistent across several studies applying different experimental tasks (e.g. dynamic stimuli, morphs, high vs. low-intensity stimuli) (Hall, 1978, 1984; Thompson & Voyer, 2014) or even across different geographical regions (e.g. Hall et al., 2000). In this line, an extensive review by Hall (1978) and a meta-analysis by Thompson and Voyer (2014) revealed persistent findings concerning the role of sex as a source of individual differences in emotion recognition.

In contrast to these findings, some studies reported a lack of assumed sex differences in emotion recognition (e.g. Buck et al., 1974; Grimshaw et al., 2004; Hoffmann et al., 2010; Lambrecht et al., 2014; Lyusin & Ovsyannikova, 2016; Rahman et al., 2004; Wagner et al., 1986). This discrepancy may be due to methodological limitations, such as the lack of statistical power which is highly dependent on the sample size and the number of trials. Previous studies reported the

CONTACT Yasaman Rafiee yasaman.rfe@psych.uni-goettingen.de Institute of Psychology, University of Göttingen, Goßlerstraße 14, 37073 Göttingen

Supplemental data for this article can be accessed online at https://doi.org/10.1080/02699931.2023.2222579.

*Some of the data from the present study were presented at conferences and scientific meetings, and a preprint of the study was made available on PsyArXiv at the following link: https://psyarxiv.com/fuqjv.

© 2023 Informa UK Limited, trading as Taylor & Francis Group

effect size for observing sex differences in global emotion recognition (as overall performance across all modalities) as small to moderate (see Hall, 1978, 1984; Thompson & Voyer, 2014); therefore, high statistical power is required to detect the hypothesised sex differences. Moreover, some previous studies only concentrated on sex differences in global emotion recognition. Hence, there are yet unanswered questions concerning the potential moderating effects of different stimulus features that might also have led to some mixed evidence or inaccurate estimation of sex differences in emotion recognition. Among the most considered features were stimulus modality, emotion specificity, and sex of the encoder (referring to the sex of the actor) (see Dores et al., 2020; Kret & De Gelder, 2012; Thompson & Voyer, 2014). Ignoring these moderating effects risks underestimating presumably existing sex differences in emotion recognition (Lausen & Schacht, 2018), and in fact, the magnitude of sex differences in emotion recognition has been demonstrated to depend on such variables (Hall, 1978, 1984; Thompson & Voyer, 2014).

Due to the multifaceted and highly relevant information that faces usually convey about individuals and their environment (Wilhelm et al., 2014), previous research mainly focused on sex differences in the recognition of *facial* expressions of emotion. However, the ecological validity of previous studies in this field is questionable, as emotions are usually expressed through more than one modality in natural interactions (de Gelder & Vroomen, 2000). Only a few studies attempted to investigate sex differences in recognising emotions expressed through other modalities, such as by means of auditory and audiovisual stimuli. Significantly better performance of women compared to men has been indicated for auditory stimuli such as emotional prosody (e.g. Lambrecht et al., 2014) and vocal expressions of emotion (e.g. Paulmann & Uskul, 2014). This pattern was replicated in a recent large-scale study ($N > 300$) by Lausen and Schacht (2018), although the magnitude of the sex differences was rather small. In another large-scale and cross-cultural study, Scherer et al. (2001) also found a small but significant main effect of sex, reflecting women's better performance in vocal emotion recognition compared to men. Similar to emotion recognition in the visual and auditory modalities, there is evidence for a female advantage in the recognition of audiovisual expressions of emotion (e.g. Hall, 1978; 1984).

Another conceivable moderator of sex differences in emotion recognition is emotion specificity (for an overview see Hall, 1978, 1984; Thompson & Voyer, 2014). Several studies reported an advantage for women over men in recognising negative emotions; however, this pattern was not observed in positive emotions (e.g. happy expressions) as men and women performed equally well in recognising them (Connolly et al., 2019; Hampson et al., 2006, 2021; Hoffmann et al., 2010; Montagne et al., 2005; Rotter & Rotter, 1988; for an overview see Thompson & Voyer, 2014). Negative emotions are in general more difficult to recognise, and therefore women's outperformance is more evident due to higher variability in recognising negative emotions (Thompson & Voyer, 2014). From an adaptive view, women's superiority in recognising negative emotions could play an important functional role in the survival of self and offspring. Babchuk et al. (1985) proposed the "primary caretaker" hypothesis, which claims that a women's advantage in emotion recognition arises from within the scope of their responsibility in childcare. In this line, Taylor et al. (2000) suggested that women, due to different parental investments, tend to have different stress responses to environmental threats, which are assumed to have been selectively evolved during evolution to increase the survival of self and offspring. Related to this is the "fitness threat hypothesis" which emphasises the relevance of recognition of negative emotions in reproductive success (Hampson et al., 2006). The "fitness threat hypothesis" gained preliminary support in a study by Hampson et al. (2006), showing women's quicker responses in recognising negative emotions compared to men. Interestingly, the previous childcare was not a predictive factor, which suggests that the women's advantage in recognising negative emotions is an evolved mechanism and is not learned through childcare experience. In addition, this finding was replicated in a recent study by Hampson et al. (2021), where they showed women's higher accuracy in recognising negative emotions in infants, toddlers, and adult faces compared to men.

Sex differences were studied not only in the ability to recognise but also to express emotions. Some studies suggested women's higher expressiveness and expression accuracy compared to men (e.g. Wagner et al., 1986) implying the potential moderating effect of the sex of the encoder on sex differences in emotion recognition (see Hall, 1984; Hall et al., 2000). Evidence is however mixed concerning the interaction of the sex of the decoder (referring to

the sex of the participant) and the sex of the encoder, with some studies showing better women's recognition of male expressions (Collignon et al., 2010; Davitz, 1964; Tagiuri, 1969; Thompson & Voyer, 2014) and others showing no moderating effects of the encoder's sex (Hall, 1978).

In sum, the partially inconsistent findings on sex differences in emotion recognition might not only be due to the disregard of the potential moderator variables described above. In addition, several earlier studies had potential methodological limitations. Remarkably, only a few studies investigated the hypothesised sex differences (a) within a large sample size and (b) by implementing enough experimental trials, both necessary for gaining appropriate statistical power to detect sex differences. Further, many previous studies were constrained to only one stimulus modality as well as to very few distinct emotions. Thus, their findings have restricted generalizability to fundamental variations in emotion recognition abilities (e.g. Collignon et al., 2010).

In light of these limitations, the present study aimed, first, to replicate indicated sex differences in emotion recognition, by studying a large-scale healthy population ($N = 426$) in a well-established study design, including a considerable number of trials (432/session). Second, we aimed to investigate to what extent the assumed sex differences might be moderated by different stimulus features, i.e. modality, emotion specificity, and sex of the encoder. We predicted:

(*H1*) a global advantage in emotion recognition for women over men;

(*H2*) an advantage in emotion recognition specifically for expressions of negative emotions (fear, anger, disgust, sadness) for women over men, indicated by an expected sex of decoder × emotion specificity interaction;

We attempted to explore the following hypotheses due to inconsistencies and lack of evidence:

(*H3*) the possible interaction effect between encoder's and decoder's sex in emotion (sex of encoder × sex of decoder);

(*H4*) if there are sex differences in recognizing emotions from visual, auditory, and audiovisual modalities (sex of decoder × stimulus modality).

Finally, we examined the possible confusion in the inferred and intended category of emotions, known as the confusion matrix, in men and women.

Methods

The data was obtained by merging two studies that were using the same set of stimuli and experimental procedure in two samples of healthy men (Lausen et al., 2020) and women (Rafiee et al., 2023). Both studies have been approved by the local ethics committee of the Institute of Psychology at Goettingen University. Each participant signed a written consent form according to the Declaration of Helsinki (DoH) ethical principles for human subjects before their participation. After completing the certain experiment, participants were reimbursed with a reasonable amount of money or course credits. The data and script can be found on Open Science Framework (OSF) under this link: https://osf.io/u2khx/.

Participants

Included participants were healthy, native German speakers with no history of psychiatric or neurological diseases, and had normal or corrected to normal vision and hearing, according to self-report. In total $N = 426$ ($M_{Age} = 24.2$, $SD_{Age} = 3.63$, Range = 18–36; 145 females) individuals were included in the analysis. Participants' characteristics are summarised in Table 1.

Procedure

The dataset was merged from two separate studies applying the same experimental design and task requests. The study collecting data from women was a longitudinal study (cf. Rafiee et al., 2023); however, we only included women's performance in their first experimental session to eliminate any effects of learning and practicing, respectively. Upon arrival, female participants filled out a computer-based screening

Table 1. Demographics of the participants.

	Female ($N = 145$)	Male ($N = 281$)
Age in years, mean (SD)	24.0 (3.39)	24.3 (3.75)
Relationship Status		
Single	47.6% (69)	54.1% (152)
In a relationship	46.2% (67)	40.2% (113)
Engaged	1.4% (2)	1.1% (3)
Married	2.1% (3)	1.4% (4)
Others	2.1% (3)	3.2% (9)
Hand		
Right	91% (132)	90% (252)
Left	9% (13)	10% (29)

Please note that the numbers in parentheses within the categories "Relationship Status" and "Hand Dominance" indicate the exact number of participants.

questionnaire related to their health and hormonal assessment (adapted from Jünger et al., 2018), followed by a German version of the Positive and Negative Affect Schedule (PANAS) questionnaire (Breyer & Bluemke, 2016). Next, saliva samples were collected prior the beginning of the emotion recognition task described below.

Men were recruited for one session only. Participants filled out a computer-based screening questionnaire related to their health and hormonal assessment (adapted from Schultheiss & Stanton, 2009), Multi-Motive Grid Questionnaire (MMG) (Sokolowski et al., 2000), and PANAS questionnaire (Breyer & Bluemke, 2016) at the beginning of the session. Next, their saliva samples were collected and then participants performed the emotion recognition task. Neither the obtained questionnaire data nor the hormonal measures extracted from saliva samples were in the scope of the current study.

Emotion recognition task

All participants performed a computer-based emotion recognition task, including visual, auditory, and audio-visual stimuli. The visual modality was extracted from Radboud face database (RaFD, Langner et al., 2010), and the auditory modality included affect burst, pseudo-words, and pseudo-sentences extracted from Montreal Affective Voices (Belin et al., 2008), Magdeburg Prosody Corpus (Wendt & Scheich, 2002), and Paulmann Prosodic Stimuli (Paulmann et al., 2008) respectively. In each experimental session, a total of 432 stimuli were presented in three different modalities (144 each) and containing of five emotional expressions (anger, fear, sadness, happiness, and disgust), and neutral expressions (24 per emotion × modality). The order of stimuli and each modality block was randomised across participants (see Lausen et al., 2020).

Trials started with a blank screen for 1000 ms followed by a fixation cross in the centre of the screen for 1000 ms. Following the presentation of the target stimulus (varying in duration between 319 and 4821 ms), a circular answer display appeared, containing all six (emotion) categories of interest, together with the selection cursor, in the centre of the display. The arrangement of the emotion labels was counterbalanced but remained constant throughout the task. Participants had to select an emotion category by using the mouse to move the cursor. They were required to select the corresponding emotion label as accurately and quickly as possible, although there was no time limitation for responding to stimuli. Emotion recognition accuracy and response time (relative to the onset of the answer display) were measured. Responses were considered hits when they matched the expected label with the name of the target expression. Before the main experiment, three practice trials were presented to familiarise the participants with the experimental procedure (Figure 1).

Statistical analysis

Using G*Power, the post-hoc power analysis revealed 99% statistical power for a small effect size ($f2 = 0.15$) in the population ($r = .15$; $\alpha = .05$; $1 - \beta = .99$) (Faul et al., 2007, 2009). We further conducted a comparison between our sample size and an estimated sample size obtained from a prior power analysis. The analysis indicated that a sample size of $N = 211$ participants was required for the same effect size ($f2 = 0.15$), with $\alpha = .05$ and $1 - \beta = .99$. Consequently, our sample size exceeded the required number of participants, providing evidence that our study possesses statistical power greater than $1 - \beta > .99$.

To analyse the data, R statistical software (version 4.0.5) and R studio (version 1.4.1106) were applied. A p-value of 0.05 was set as the cut-off criterion for two-tailed distributions. Data was preprocessed using the following packages, tidyverse (version 1.3.1), knitr (version 1.39), and dplyr (version 1.0.9). We used ggplot2 (version 3.3.6) and sjPlot (version 2.8.10) for data visualisation.

To test the first hypothesis – sex differences in global emotion recognition – we used the Mann-Whitney U test by comparing the relative frequency of hits between the two sexes. To examine the moderation effects of stimulus modality, sex of encoder, and emotion specificity with the sex of decoder (H2, exploratory analysis: H3, H4), we fitted one Generalised Linear Model with quasibinomial error structure and logit link function (McCullagh & Nelder, 1989), with the outcome variable as emotion recognition accuracy (correct vs. false), and stimulus features as predictors including stimulus modality, sex of encoder, and emotion specificity. Stimulus duration (log-transformed) was entered into the regression model as the control variable.

To examine the absence of collinearity among predictors, Variance Inflation Factors (VIF; Field et al., 2012) was applied, using the function vif of the

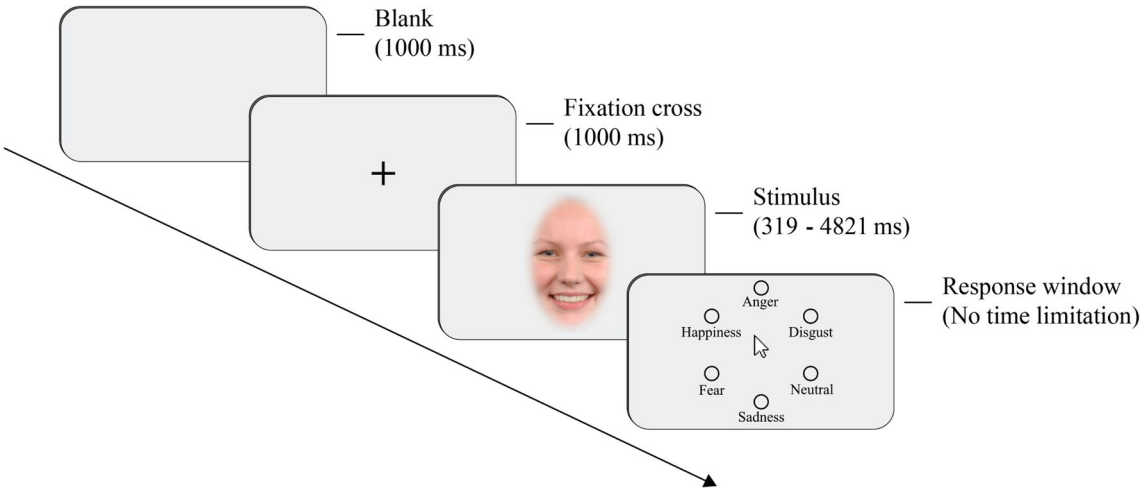

Figure 1. Schematic overview of an exemplary trial displaying happiness in visual modality.

package car (version 3.0.10; Fox & Weisberg, 2019). This showed no collinearity issue among predictors (maximum VIF: 1.059). In addition, model stability was assessed to find influential observations, which were not present in our model. The goodness of fit of the fitted model was assessed by likelihood ratio test by comparing the deviances of the fitted model with the minimal model lacking the predictor of interest.

Results

The non-parametric Mann-Whitney U indicated that the global emotion recognition accuracy was higher for women than for men ($W = 16{,}116$, $P < 0.001$). To further investigate the potential moderating effect of stimulus modality, sex of encoder, and emotion specificity on the magnitude of sex differences in emotion recognition, we fitted a generalised linear model (GLM), with the mentioned predictors and accuracy as the outcome variable. In addition, the relationship between emotion specificity and sex of decoder (emotion specificity × sex of decoder), stimulus modality and sex of decoder (stimulus modality × sex of decoder), and sex of encoder and sex of decoder (sex of encoder × sex of decoder) were entered as interaction terms. To control for the effect of stimulus duration on responses, we included duration (log-transformed – base e) as a control variable. The reference categories were visual for stimulus modality, neutral expression for emotion specificity, and male for both sex of the encoder and sex of the decoder.

The GLM revealed a significant moderating effect of emotion specificity on sex differences in emotion recognition. Women showed a better performance than men in recognising anger and fear but not sadness and disgust, partially supporting H2 (*female decoder* × *anger*: β = 0.302, SE = 0.09, 95% CI = [1.19–1.54], $t = 4.675$, OR = 1.35, $p < 0.001$; *female decoder* × *fear*: $\beta = 0.193$, SE =, 95% CI = [1.08–1.37], $t = 3.170$, OR = 1.21, $p = 0.002$; Figure 2, left panel). The model outcome did not show a significant interaction of either encoder's sex on sex differences in recognising emotions (Table 2). We also investigated the moderating effect of stimulus modality and sex differences on emotion recognition. The model showed a significant moderating effect of stimulus modality × sex of decoder in a way that women outperformed men in recognising emotions expressed in all modalities (*female decoder* × *auditory modality:* $\beta = -0.173$, SE = 0.03, 95% CI = [0.78–0.90], $t = -4.839$, OR = 0.84, $p < 0.001$; *female decoder* × *audiovisual modality:* $\beta = 0.123$, SE = 0.06, 95% CI = [1.03–1.24], $t = 2.519$, OR = 1.13, $p = 0.012$; Table 2 contains the GLM results). The most pronounced differences between the sexes were found in the ability to recognise emotions in the audiovisual modality, whereas the least pronounced differences were found in the recognition of emotions expressed in the auditory modality (note that the intercept corresponds to the performance of males in recognising neutral expressions

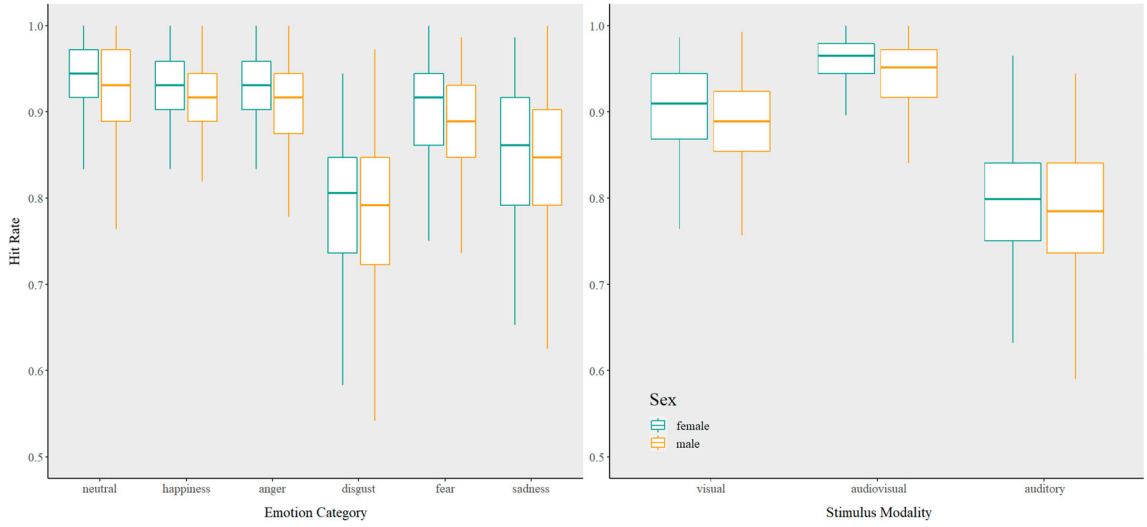

Figure 2. The boxplots depict the distribution of the mean hit rate (accuracy) with medians per emotion category (left panel) and modality (right panel), for female and male participants.

conveyed by male encoders in the visual modality. These findings are based on dummy coding and compared to the intercept).

Subsequently, we compared the main model with a model lacking sex of decoder and its interactions with each of the stimulus features. The likelihood ratio test revealed that the main model, including sex and its interaction with stimulus features, predicts the outcome variable significantly better (*Deviance* = 236.91, df = *9*, $p < 0.001$). Furthermore, the main model was compared to a null model omitting only the interaction terms. Again, the main model explains the outcome variable significantly better than the null model (*Deviance* = 109.736, df = *8*, $p < 0.001$). Overall, these comparisons indicate that not only sex of decoder but also features of the stimulus are significant predictors of emotion recognition.

The results of the confusion matrix indicated that the least recognised expression of emotion was disgust in the auditory modality and the most recognised emotion expression was happiness in the visual modality in men and women (see Tables3a and 3b in the supplementary document).

Table 2. Results of Generalised Linear Model investigating the association of stimulus features and sex differences in emotion recognition, *N* = 426 (females = 145). Note, fem = female.

	Estimates	*SE*	*t*	*p*	*OR*	*95% CI*
Decoder's sex [fem]	0.158	0.06	2.915	**0.004**	1.17	1.05–1.30
Emotion [Anger]	−0.283	0.03	−8.035	**<0.001**	0.75	0.70–0.81
Emotion [Disgust]	−1.264	0.01	−39.986	**<0.001**	0.28	0.27–0.30
Emotion [Fear]	−0.563	0.02	−16.651	**<0.001**	0.57	0.53–0.61
Emotion [Happiness]	−0.086	0.03	−2.347	**0.019**	0.92	0.85–0.99
Emotion [Sadness]	−0.895	0.01	−27.215	**<0.001**	0.41	0.38–0.44
Encoder's sex [fem]	0.22	0.02	12.557	**<0.001**	1.25	1.20–1.29
Modality [Audiovisual]	0.722	0.05	27.892	**<0.001**	2.06	1.96–2.17
Modality [Auditory]	−0.736	0.01	−37.108	**<0.001**	0.48	0.46–0.50
Stimulus duration (log- transformed)	−0.002	0.01	−0.175	0.861	1.00	0.98–1.02
Decoder's sex [fem] × Emotion [Anger]	0.302	0.09	4.675	**<0.001**	1.35	1.19–1.54
Decoder's sex [fem] × Emotion [Disgust]	−0.066	0.05	−1.195	0.232	0.94	0.84–1.04
Decoder's sex [fem] × Emotion [Fear]	0.193	0.07	3.170	**0.002**	1.21	1.08–1.37
Decoder's sex [fem] × Emotion [Happiness]	0.084	0.07	1.296	0.195	1.09	0.96–1.24
Decoder's sex [fem] × Emotion [Sadness]	0.033	0.06	0.573	0.567	1.03	0.92–1.16
Decoder's sex [fem] × Encoder's sex [fem]	0.044	0.03	1.382	0.167	1.04	0.98–1.11
Decoder's sex [fem] × Modality [Audiovisual]	0.123	0.06	2.519	**0.012**	1.13	1.03–1.24
Decoder's sex [fem] × Modality [Auditory]	−0.173	0.03	−4.839	**<0.001**	0.84	0.78–0.90

Discussion

Emotion recognition, as an essential ability for social beings like humans, varies among individuals. One of the most frequently discussed sources of individual differences in emotion recognition has been the sex of the decoder, with a trend for women performing better than men. In contrast, there were contradictory findings indicating no sex differences (e.g. Buck et al., 1974; Hoffmann et al., 2010; Lambrecht et al., 2014; Wagner et al., 1986), likely due to low statistical power such as insufficient sample size and trial number. The effect size reported in previous studies tended to be small to moderate, indicating that high statistical power is required to detect the effect of interest. In addition, other potential moderating effects of stimulus features on the magnitude of sex differences in emotion recognition are understudied, which potentially lead to inaccurate estimation of sex differences. The present study aimed to fill this gap by investigating the moderating effects of stimulus modality, emotion specificity, and sex of the encoder on sex differences in emotion recognition, in a large-scale healthy population ($N = 426$). Going beyond previous studies focusing on emotion recognition in faces, this study also used vocal and bimodal expressions with the aim of increasing ecological validity.

Sex differences in emotion recognition

In line with previous studies, our findings indicate that women are more accurate in recognising emotions. In the current study, the obtained effect size concerning sex differences in global emotion recognition was rather small ($d = .17$), which also corresponds with the reported effect size ($d = .19$) in an extensive meta-analysis including 215 studies (see Thompson & Voyer, 2014). Our findings also indicated that stimulus features including emotion specificity and stimulus modality significantly contribute to variance in effect size and thus supports the proposed necessity for considering these factors as predictors of sex differences in emotion recognition.

In our second hypothesis, we predicted women's advantage specifically in recognising negative emotions (anger, disgust, fear, and sadness). As hypothesised, women were more accurate in recognising anger and fear than men; however, this pattern was not observed in the recognition of disgust and sadness. We suggest two potential explanations for not observing sex differences in the recognition of disgust and sadness expressions.

First, our results may have been influenced by the study design (e.g. exposure time, response time limitation) and the stimulus material used. The recognition rate of disgust and sadness expressions was found to be the lowest among the emotional expressions respectively (see Table 2), with poor performance of both men (hit rates: disgust = 0.78; sadness = 0.83) and women (hit rates: disgust = 0.78; sadness = 0.84). This likely led to no significant differences in recognition performance between the sexes. However, there were significant differences in the recognition rates for anger and fear expressions between female (hit rates: anger = 0.93; fear = 0.89) and male decoders (hit rates: anger = 0.89; fear = 0.87; see also supplementary material, Tables 3a and 3b). Even, using the same stimuli can also result in varying outcomes probably due to different study designs. For example, a study using the same visual stimuli (RaFD) in a large sample of 1249 participants (15.4% male) with a 10-second exposure time and 1-second response time indicated sex differences only in the recognition of anger (and contempt) expressions (Dores et al., 2020).

Second, the higher survival priority of fear and anger recognition relative to other emotions may be another explanation for the discrepancy in results. Recognising fear and anger conveys information about potential threats and dangers, making their recognition highly relevant to survival (e.g. Skuse, 2003). Although disgust expressions also carry (biologically) important information, its level of indicated danger might be perceived as less immediate and severe compared to fear and anger in the context of laboratory studies. Related to this point, many previous studies have shown sex differences mostly in the recognition of anger expressions, likely due to its high saliency, while sex differences in recognising other negative emotions have been inconsistently reported. A meta-analysis by Thompson and Voyer (2014) showed that only anger had a significant impact on the variability of sex differences in emotion recognition.

Like emotion specificity and in line with previous studies (Hall, 1978, 1984; Thompson & Voyer, 2014), stimulus modality had a significant moderating effect on the magnitude of sex differences in the current study. Women performed better across all modalities compared to men and, remarkably, the largest difference in hit rates was observed when

comparing women's performance in audiovisual modality to men. These findings support Collignon et al.'s (2010) study, where women were not only better at recognising unimodal emotional expressions (visual and auditory), but also at integrating visual and auditory information, leading to their advantage in recognising audiovisually expressed emotions.

Our study did not reveal that the interaction between sex of the encoder and sex of the decoder influences emotion recognition, in line with some previous reports (e.g. Hall, 1978; Harris et al., 2016; Lausen & Schacht, 2018; Thayer & Johnsen, 2000), but contrary to others (e.g. Collignon et al., 2010; Davitz, 1964; Tagiuri, 1969; Thompson & Voyer, 2014). The heterogenous findings might be due to utilising a different set of stimuli and or the authenticity of the presented stimuli. Although we used an established and validated stimulus set in our study, it could be argued that the moderating influence of encoder's sex on emotion recognition depends to some extent on other contextual factors. Emotions expressed by female encoders in our study were recognised with greater accuracy, suggesting a potential advantage in their expressive abilities compared to male encoders; however, null findings regarding the interaction effect of encoder and decoder's sex in emotion recognition call into question the own-sex bias in emotion recognition.

Overall, our study demonstrated that emotion specificity and stimulus modality, but not the sex of the encoder have a moderating effect on sex differences in emotion recognition. Therefore, we, recommend including both variables in future studies of emotion recognition in order to obtain a more precise assessment of sex differences in this domain.

Implication

Sex differences in emotion recognition have been consistently replicated in numerous studies, including the current study. The methodological strengths of this study support the reliability of the effect, although it is small in magnitude. The small effect may not have significant implications in everyday life. However, it may have a greater implication in clinical settings and in the diagnosis of gender-specific psychiatric disorders. For example, considering sex differences in emotion recognition is important when designing diagnostic tools for disorders with different prevalence in women and men. Evidence suggests that autism is often diagnosed at a later age in females than males autism due to their higher functioning in social interaction (Wood-Downie et al., 2021), highlighting the need to consider sex differences in the implementation of diagnostic instruments.

Limitations and outlook

One of the potential limitations of the current study may be the combination of two datasets collected in two large studies (see Lausen et al., 2020; Rafiee et al., 2023). Consequently, the proportion of men and women in the current study was imbalanced ($N = 426$, 145 women), due to the strict inclusion criteria for recruiting women in the study by Rafiee et al. (2023) that focused on ovulatory cycle shifts in emotion recognition. A gender-balanced sample should be preferred in future studies to control for the main and interaction effects of encoder and decoder sex in emotion recognition in future studies (Dores et al., 2020; Lausen & Schacht, 2018). Another limitation relates to the stimulus material: In the audiovisual condition, we presented static faces combined with congruent emotional prosody. A combination of dynamic facial expressions with emotional prosody could further increase ecological validity (Collignon et al., 2010). In addition, implementing a different range of intensities of the emotional expressions could have sex differences become more evident. For instance, Hoffmann et al. (2010) did not find sex differences in the recognition of facial expressions with 100% intensity, i.e. when the expressions could be easily recognised; however, they did find sex differences with women outperforming in the recognition of only subtle facial expressions.

Some studies have suggested that the multiple-choice format has low ecological validity as real-life situations often do not provide a list of potential responses (Georgopoulos et al., 2022). This might lead to draw inferences about the displayed emotion not elicited by the stimulus itself (e.g. Betz et al., 2019; Georgopoulos et al., 2022; Russell, 1994). However, the multiple-choice format has several advantages, including reduced missing data due to the requirement to choose a label, decreased response variability, and easier data coding (Dores et al., 2020).

To achieve a coherent understanding of the moderating effects of stimulus features in sex differences in emotion recognition, we encourage future studies to include more variation in stimulus intensities and to implement a wider range of emotional expressions

(e.g. complex emotions), more lifelike stimuli (e.g. authentic expressions), as well as different stimulus modalities (e.g. bodily expression). As a prerequisite, the generation and validation of new sets of stimuli would be recommended.

Conclusion

The present study replicated significant differences in emotion recognition between men and women, indicating that sex is a crucial factor in explaining individual variations in this ability. Remarkably, our study indicates that the characteristics of the stimuli, such as emotion specificity and stimulus modality, play a role in the magnitude of these sex differences. Exploring possible moderating factors that influence sex differences in emotion recognition could provide a deeper understanding of both variation and commonalities across individuals.

Acknowledgements

We express our gratitude to the student assistants who assisted with data collection, Adi Lausen for supplying data from male subjects, Edmund Henniges for programming the experiment, Francesco Grassi for helping with data preprocessing, and Roger Mundry for his statistical support. We further would like to thank all participants who took part in the experiments.

Author contributions

The study was conceptualised and designed by Y.R. and A.S. The project was supervised by A.S. Y.R, analysed the data and drafted the manuscript. Both authors contributed by commenting and revising the final manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by funding from the German Academic Exchange Service (DAAD), Deutsche Forschungsgemeinschaft; Grant/Award Number: Project number 254142454/GRK 2070, and the Leibniz-Gemeinschaft through support for the Leibniz ScienceCampus Primate Cognition (W45/2019).

Data availability statement

The data that support the findings of this study are openly available on Open Science Framework at https://osf.io/u2khx/.

ORCID

Yasaman Rafiee http://orcid.org/0000-0002-5958-7876

References

Babchuk, W. A., Hames, R., & 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

Bänziger, T. (2016). Accuracy of judging emotions. In J. A. Hall, M. S. Mast, & T. V. West (Eds.), *The social psychology of perceiving others accurately* (pp. 23–51). Cambridge University Press. https://doi.org/10.1017/CBO9781316181959.002

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

Betz, N., Hoemann, K., & Barrett, L. F. (2019). Words are a context for mental inference are a contextfor mental inference. *Emotion*, *19*(8), 1463–1477. https://doi.org/10.1037/emo0000510

Breyer, B., & Bluemke, M. (2016). *Deutsche version der positive and negative affect schedule PANAS (GESIS panel)*. Zusammenstellung Sozialwissenschaftlicher Items Und Skalen. https://doi.org/10.6102/zis242.

Buck, R. W., Miller, R. E., & Caul, W. F. (1974). Sex, personality, and physiological variables in the communication of emotion via facial expression. *Journal of Personality and Social Psychology*, *30*(4), 587–596. https://doi.org/10.1037/h0037041

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

Connolly, H. L., Lefevre, C. E., Young, A. W., & Lewis, G. J. (2019). Sex differences in emotion recognition: Evidence for a small overall female superiority on facial disgust differences in emotion recognition: Evidence for a small overall female superiority on facial disgust. *Emotion*, *19*(3), 455–464. https://doi.org/10.1037/emo0000446. Epub 2018 May 21. PMID: 29781645.

Davitz, J. R. (1964). A review of research concerned with facial and vocal expressions of emotion. In J. R. Davitz (Ed.), *The communication of emotional meaning* (pp. 13–30). McGraw-Hill.

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, *14*(3), 289–311. https://doi.org/10.1080/026999300378824

Dores, A. R., Barbosa, F., Queirós, C., Carvalho, I. P., & Griffiths, M. D. (2020). Recognizing emotions through facial expressions: A largescale experimental study emotions through facial expressions: A largescale experimental study. *International Journal of Environmental Research and Public Health Journal of Environmental Research and public Health*, 17(20), 7420. https://doi.org/10.3390/ijerph17207420

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and

regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. SAGE Publications.

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Georgopoulos, M. A., Brewer, N., Lucas, C. A., & Young, R. L. (2022). Speed and accuracy of emotion recognition in autistic adults: The role of stimulus type, response format, and emotion. *Autism Research*, *15*(9), 1686–1697. https://doi.org/10.1002/aur.2713

Grimshaw, G. M., Bulman-Fleming, M. B., & Ngo, C. (2004). A signal-detection analysis of sex differences in the perception of emotional faces. *Brain and Cognition*, *54*(3), 248–250. https://doi.org/10.1016/j.bandc.2004.02.029

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*. JohnsHopkins University Press.

Hall, J. A., Carter, J. D., & Horgan, T. G. (2000). Gender differences in nonverbal communication of emotion. In A. H. Fischer (Ed.), *Gender and emotion: Social psychological perspectives* (pp. 97–117). Cambridge University Press. https://doi.org/10.1017/CBO9780511628191.006

Hampson, E., 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

Hampson, E., Istasy, P., Owais, S., Chow, J., Howidi, B., & Ouellette, S. (2021). Sex differences in the recognition of children's emotional expressions: A test of the fitness threat hypothesis. *Evolutionary Psychological Science*, *7*(1), 45–60. https://doi.org/10.1007/s40806-020-00254-w

Harris, D. A., Hayes-Skelton, S. A., & Ciaramitaro, V. M. (2016). What's in a face? How face gender and current affect influence perceived emotion. *Frontiers in Psychology*, *7*, 1468. https://doi.org/10.3389/fpsyg.2016.01468

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. Epub 2010 Aug 21. PMID: 20728864. https://doi.org/10.1016/j.actpsy.2010.07.012

Jünger, J., Kordsmeyer, T., Gerlach, T. M., & Penke, L. (2018). Fertile women evaluate male bodies as more attractive, regardless of masculinity. *Evolution and Human Behavior*, *39*(4), 412–423. https://doi.org/10.1016/j.evolhumbehav.2018.03.007

Kret, M. E., & De Gelder, B. A. (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. Epub 2012 Jan 8. PMID: 22245006.

Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2014). Gender differences in emotion recognition: Impact of sensory modality and emotional category. *Cognition and Emotion*, *28*(3), 452–469. https://doi.org/10.1080/02699931.2013.837378. Epub 2013 Oct 24. PMID: 24151963.

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

Lausen, A., Broering, C., Penke, L., & Schacht, A. (2020). Hormonal and modality specific effects on males' emotion recognition ability. *Psychoneuroendocrinology*, *119*, https://doi.org/10.1016/j.psyneuen.2020.104719

Lausen, A., & Schacht, A. (2018). Gender differences in the recognition of vocal emotions. *Frontiers in Psychology*, *9*(9), 882. https://doi.org/10.3389/fpsyg.2018.00882. PMID: 29922202; PMCID: PMC5996252.

Lewis, G. J., Lefevre, C. E., & Young, A. W. (2016). Functional architecture of visual emotion recognition ability: A latent variable approach. *Journal of Experimental Psychology: General*, *145*(5), 589–602. https://doi.org/10.1037/xge0000160. Epub 2016 Mar 17. PMID: 26986040.

Lyusin, D., & Ovsyannikova, V. (2016). Measuring two aspects of emotion recognition ability: Accuracy vs. sensitivity. *Learning and Individual Differences*, *52*, 129–136. https://doi.org/10.1016/j.lindif.2015.04.010

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. PMID: 10825784.

McCullagh, P., & Nelder, J. (1989). *Generalized Linear Models* (2nd ed.). Chapman & Hall. ISBN: 9780412317606.

Montagne, B., Kessels, R., Frigerio, E., de Haan, E., & Perrett, D. (2005). Sex differences in the perception of affective facial expressions: Do men really lack emotional sensitivity? *Cognitive Processing*, *6*(2), 136–141. https://doi.org/10.1007/s10339-005-0050-6

Paulmann, S., Pell, M. D., & Kotz, S. A. (2008). How aging affects the recognition of emotional speech. *Brain and Language*, *104*(3), 262–269. https://doi.org/10.1016/j.bandl.2007.03.002

Paulmann, S., & Uskul, A. K. (2014). Cross-cultural emotional prosody recognition: Evidence from Chinese and British listeners. *Cognition and Emotion*, *28*(2), 230–244. https://doi.org/10.1080/02699931.2013.812033. Epub 2013 Jul 17. PMID: 23862740.

Rafiee, Y., Stern, J., Ostner, J., Penke, L., & Schacht, A. (2023). Does emotion recognition change across phases of the ovulatory cycle? *Psychoneuroendocrinology*, *148*, 105977, https://doi.org/10.1016/j.psyneuen.2022.105977

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

Rotter, N. G., & Rotter, G. S. (1988). Sex differences in the encoding and decoding of negative facial emotions. *Journal of Nonverbal Behavior*, *12*(2), 139–148. https://doi.org/10.1007/BF00986931

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, *115*(1), 102–141. https://doi.org/10.1037/0033-2909.115.1.102

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

Schultheiss, O. C. O., & Stanton, S. J. (2009). Assessment of salivary hormones. In E. Harmon-Jones & J. S. Beer (Eds.), *Methods in social neuroscience* (pp. 17–44). Guilford Press.

Skuse, D. (2003). Fear recognition and the neural basis of social cognition. *Child and Adolescent Mental Health*, *8*(2), 50–60. https://doi.org/10.1111/1475-3588.00047

Sokolowski, K., Schmalt, H.-D., Langens, T. A., & Puca, R. M. (2000). Assessing achievement, affiliation, and power motives all at once: The multi-motive grid (MMG). *Journal of Personality Assessment*, *74*(1), 126–145. https://doi.org/10.1207/S15327752JPA740109

Tagiuri, R. (1969). Person perception. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (2nd ed., Vol. 3, pp. 634–655). Addison-Wesley.

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychological Review*, *107*(3), 411–429. https://doi.org/10.1037/0033-295X.107.3.411. PMID: 10941275.

Thayer, J. F., & Johnsen, B. (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 and Emotion*, *28*(7), 1164–1195. https://doi.org/10.1080/02699931.2013.875889

Wagner, H. L., MacDonald, C. J., & Manstead, A. S. R. (1986). Communication of individual emotions by spontaneous facial expressions. *Journal of Personality and Social Psychology*, *50*(4), 737–743. https://doi.org/10.1037/0022-3514.50.4.737

Wendt, B., & Scheich, H. (2002, April). The "magdeburger prosodie-korpus". *Speech Prosody*, 11–13. https://www.isca-speech.org/archive/icslp_2002/index.html

Wilhelm, O., Hildebrandt, A., Manske, K., Schacht, A., & Sommer, W. (2014). Test battery for measuring the perception and recognition of facial expressions of emotion. *Frontiers in Psychology*, *5*, 404. https://doi.org/10.3389/fpsyg.2014.00404

Wood-Downie, H., Wong, B., Kovshoff, H., Mandy, W., Hull, L., & Hadwin, J. A. (2021). Sex/gender differences in camouflaging in children and adolescents with autismm. *Journal of Autism and Developmental Disorderss*, *51*(4), 1353–1364. https://doi.org/10.1007/s10803-020-04615-z