The Effects of Face Masks on Emotion Interpretation in Socially Anxious Individuals

Lexie M. Mathis
Towson University, Department of Psychology

Previous research has found emotion interpretation biases in individuals with social anxiety (SA) when emotions are ambiguous. Additionally, research has shown that face masks impair emotion recognition. The present within-subjects, quasi-experimental study examined the impact of face masks on emotion identification in individuals with SA. After pre-screening using a brief SA scale (the SIAS-6, a shortened version of the Social Interaction Anxiety Scale), 92 undergraduate students with qualifying SA scores completed an emotion identification task. The task included images of masked and unmasked individuals expressing four emotions. Results indicated impaired accuracy for all types of masked trials. When examining incorrect choice selections, it was found that neutral masked faces were misinterpreted as sad or fearful significantly more than their unmasked counterparts, suggesting a mask-specific interpretation bias. In relation to previous research on SA and hypervigilance for threat faces, when considering all masked trials, accuracy was highest for masked angry trials. While this study was limited in that it did not include a non-anxious group, the results have interesting implications. Importantly, the present findings suggest that some previously identified SA-related emotion interpretation characteristics persist when viewing masked faces (e.g., hypervigilance for threats), while others do not (e.g., default bias for interpreting neutral as hostile). Future research should focus on elucidating the causes of the present results, which are important to this population and clinicians, as socially anxious individuals attribute greater social cost to negative emotions; thus, misinterpretation of emotions as negative may be detrimental to their mental health.

Keywords: social anxiety, face masks, emotion recognition, face perception, emotions

The COVID-19 pandemic (henceforth referred to as the pandemic), prompted widespread use of face masks in public settings. With this necessary protective measure, there were unintended social and mental health consequences. Masks impair emotion recognition (Carbon, 2020; Gori et al., 2021; Grundmann et al., 2021), and negatively impact perceptions of others, including perceived closeness (Grundmann et al., 2021), empathy (Wong et al., 2013), and friendliness (Wiesmann et al., 2021). Importantly, social anxiety (SA) affects the interpretation of emotions, including hypervigilance for threat faces (Klumpp & Amir, 2009; Mogg et al., 2004) and biases in emotion interpretation (Button et al., 2013; Gutiérrez-García & Calvo, 2017; Schofield et al., 2007; Yoon & Zinbarg, 2008). Considering these impacts, socially anxious individuals (SA individuals) may demonstrate unique patterns of emotion inference when viewing masked faces.

Social Anxiety and Emotion Interpretation

More specifically, regarding emotion interpretation biases, Gutiérrez-García and Calvo (2017) studied interpretations of ambiguous emotional faces and found that when compared to non-anxious controls, SA individuals were more likely to correctly detect low intensity displays of disgust and anger, more likely to incorrectly identify neutral faces as angry, and less likely to interpret neutral faces as happy or sad. These findings may reflect a greater ability or motivation of SA individuals to identify social threat emotions, possibly due to their hypervigilance for threat faces. Further, using an incidental learning task, Yoon and Zinbarg (2008) found that SA individuals demonstrate a default bias toward interpreting neutral faces as threatening. This bias may impact the mental health of SA individuals, as they also attribute excessive social cost to interacting with others displaying negative emotions of various intensities (i.e., they report that it would be excessively bad for them to interact with individuals displaying negative emotions, as compared to control participants' reports; Button et al., 2013; Schofield et al., 2007).

Impacts of Face Masks

Moreover, researchers have found evidence that emotion identification accuracy is impaired when masks are worn. Several experimental studies have examined this topic by having participants view photos of individuals and identify the facial expressions displayed, with and without the presence of masks in the photos (Carbon, 2020; Gori et al., 2021; Grundmann et al., 2021). All three of these recent studies found evidence that masks presented a challenge to emotion identification, as accuracy was lower for masked faces. This effect exists regardless of participant age, as it has been shown in children as young as age three (Gori et al., 2021) and in young, middle-aged, and older adults (Carbon, 2020; Gori et al., 2021; Grundmann et al., 2021).

Aside from impacting emotion identification, masks also have effects on social judgments and percep tions of others. In their experiment examining masks and emotion identification, Grundmann et al. (2021) also examined participants' social judgments of the image models. They found that participants' ratings of perceived closeness to the individual pictured were lower when the individual was wearing a mask. Further, Wiesmann et al. (2021) investigated an intervention to mitigate the negative effects of masks on the doctor-patient relationship. Their intervention, which consisted of smiling portrait photos placed on the chests of hospital staff members, was associated with significantly higher ratings of staff friendliness compared to when the staff wore masks alone. This suggests that patients perceive staff members as less friendly when masks are worn. Similar effects were found prior to the pandemic; in a randomized-controlled trial, Wong et al. (2013) examined the effect of medical face masks worn by physicians on the doctor-patient relationship. They found that patients perceived their doctors as significantly less empathetic when the doctors were masked versus not. Taken together, the results of these three studies indicate that masks negatively impact perceptions of others, including perceived closeness (Grundmann et al., 2021), empathy (Wong et al., 2013), and friendliness (Wiesmann et al., 2021). Considering the effects of SA on emotion interpretation, these impacts on perceptions of others may be exacerbated by the interpretation biases already present in SA.

Implications for the Present Study

While it is known that masks impair emotion recognition (Carbon, 2020; Gori et al., 2021; Grundmann et al., 2021) and induce more negative perceptions of others (Grundmann et al., 2021; Wiesmann et al., 2021; Wong et al., 2013), researchers have not yet identified specific populations which may be affected to a higher degree. As such, SA individuals may be at a higher risk of misinterpreting facial expressions when masks must be worn. Specifically, SA individuals interpret ambiguous faces as hostile (Gutiérrez-García & Calvo, 2017; Yoon & Zinbarg, 2008) - this may also be true when masks are worn, as they make expressions more uncertain. Those with SA attribute excessive social cost to interacting with individuals displaying negative emotions (Button et al., 2013; Schofield et al., 2007), and thus, misinterpretation of emotions may be detrimental to their mental health. This would be important for mental health professionals to be aware of when treating individuals with SA during the pandemic.

Based on evidence from prior research, the purpose of the present study was to examine emotion interpretation in individuals with SA when viewing masked and unmasked faces. A 2 (masking) X 4 (expression) within-subjects, quasi-experimental design was used to assess the following hypotheses:

H1: SA individuals will demonstrate lower accuracy for masked faces, with the lowest accuracy for neutral masked faces.

H2: Neutral faces will be most commonly confused with anger, with this being more likely to occur when faces are masked.

H3: SA scores will be negatively correlated with overall accuracy, and with accuracy for masked trials, while being positively correlated with the percentage of neutral faces confused as angry.

Method

Participants

Participants were 92 undergraduate psychology students at a large mid-Atlantic university with qualifying self-reported SA scores. All received course credit for participating. SA is especially prevalent in young adults and has increased during the pandemic (Hawes et al., 2021), making this a relevant population. Five participants were excluded, three for being outliers and two for other reasons (incomplete data and prior knowledge of the study), leaving 87 participants in analyses. A priori power analysis was used initially to inform the number of participants needed for a 2 x 4 within-subjects ANOVA, and post hoc power analysis indicated an achieved power of 88.08% with 87 participants for an effect size of $\eta 2p = 0.05$ (MorePower version 6.0.4; Campbell & Thompson, 2012).

A brief SA scale (SIAS-6; Peters et al., 2012) was used to pre-screen participants for social anxiety. Inclusion criteria consisted of a SIAS-6 score of at least 7 and a minimum age of 18 years old. The sample was diverse in race (42.5% White, 29.9% Black, 10.3% Hispanic/Latinx, 8.0% multiracial, 6.9% Asian, 2.3% other) and similar in age (M = 19.49, SD = 1.79). The majority of the sample identified as female (72.4%). This reflects the higher prevalence of SAD in females compared to males in the United States (National Institute of Mental Health, 2017), although other factors, such as major, likely contributed to this gender ratio as well. 4.5% identified as nonbinary or another gender identity.

Measures & Materials

Social Interaction Anxiety Scale

Mattick and Clarke's (1998) Social Interaction Anxiety Scale (SIAS) is a 19-item self-report measure of an individual's social interaction fear. A 6-item version (SIAS-6) was developed by Peters et al. (2012). The items are statements intended to measure anxiety levels when initiating and maintaining social interactions; for example, "I have difficulty making eye contact with others." Participants rate the degree to which each statement is characteristic of themselves on a 5-point Likert scale ranging from 0 (not at all) to 4 (extremely). The SIAS includes two reverse scored items. Total scores are calculated, with optimum cutoff scores being 39.5 for identifying SAD using the SIAS (Carleton et al., 2009) and 7 when using the SIAS-6 (Peters et al., 2012). The SIAS is both reliable and valid (Mattick & Clarke, 1998), and the SIAS-6 is comparable to the original scale (Peters et al., 2012).

Emotion Identification Task

For the emotion identification task, all participants viewed the same set of 96 images, which systematically differed in expressed emotion (happy, angry, fearful, neutral), mask-wearing (mask, no mask), and sex (male, female). Stimuli appeared individually in a Qualtrics survey (Qualtrics, Provo, UT), with the masked image block first (Gori et al., 2021) to avoid participant demand characteristics. Stimuli order was randomized within blocks. Because race can impact emotion identification (Kang & Lau, 2013; Tuminello & Davidson, 2011) through stereotyping and social categorization, only Caucasian faces were shown, as in previous work (e.g., Schofield et al., 2017). To mitigate these potential confounding effects, race was kept consistent throughout.

The images were randomly selected from the Chicago Face Database (CFD; version 3.0; Ma et al., 2015), with each model appearing only once to avoid habituation. The CFD provides two types of happy images, however, only open-mouth happy images were used. CFD images are standardized in many ways, including characteristics of the models, environment, and photography (Ma et al., 2015). Models had multiple photos taken for each expression and independent raters identified the best photo of each. Neutral images were highly reliable based on norming data collected from a large and racially diverse sample; however, such data was not obtained for emotional images.

For the present study, Adobe Photoshop was used to digitally edit a randomly selected subset of images, adding a mask (906 x 644 pixels) that was tailored to the face of each individual, and adding faint shadows for realism. Masks were centered horizontally and placed vertically to cover the nasolabial-alar crease. An example of a CFD image before and after editing can be seen in Figure 1, and the CFD codes for all images used are available online at the link provided in the data availability statement. In the task, six multiple-choice options (happy, content, neutral, angry, fearful, and sad) were provided in a consistent order. Content and sad served as positive and negative distractor options to avoid ceiling effects and were intended to be somewhat distinguishable from the emotions shown. Two measures were obtained: accuracy and neutral trial incorrect choice selections (the percentage of responses for each incorrect option when the correct answer was neutral; this was used to determine if any interpretation biases were present).

Demographic Information

Within the Qualtrics survey were several demographic questions regarding age, gender, and race. All were open-ended questions, allowing participants to respond freely. Responses to the gender and race questions were later categorized. Categories were created based on the most common responses given. Gender categories were female, male, non-binary, and other, while race categories were White, Black, Hispanic/Latinx, multiracial, Asian, and other.

Procedure

After receiving IRB approval, SIAS-6 pre-screening occurred online. The researcher was blind to pre-screening scores during data collection. Each participant came to a university laboratory, provided written informed consent, and completed the Qualtrics survey on a university desktop computer. This included several open-ended demographic questions, the emotion identification task, and the SIAS. The researcher debriefed them, answered any questions, and provided mental health resources. Pre-screening and study data were combined, downloaded to Microsoft Excel, and analyzed using IBM SPSS Statistics 27. University COVID-19 precautions were complied with in the laboratory.

Results

In the present study, masking and expression of the image models were manipulated, and subsequently, SIAS total scores, emotion identification accuracy, and neutral trial incorrect choice selections were measured. Accuracy outcomes were checked for skewness and outliers before analysis. Three outliers were identified, having unusually low scores attributed to lack of attention. Unusually high scores were considered true data. While participants answered correctly in about three-quarters of all trials, accuracy for masked (M = 64.39%, SD = 8.30%) and unmasked (M = 86.71%, SD = 6.00%) trials notably differed. As for SA scores, the average SIAS score (M = 38.94, SD = 12.83) approached the cut-off score of 39.5, while the average SIAS-6 score (M = 11.31, SD = 3.95) exceeded the cut-off score of 7.

Prior to conducting the planned repeated measures ANOVAs, test assumptions were checked. Approximate normality was violated, as the dependent variables for each test were extremely skewed in some conditions. Transformations were attempted; however, some conditions became overcorrected. Thus, the aligned rank transform procedure (Salter & Fawcett, 1993), a nonparametric option which aligns and ranks data for each effect before conducting factorial ANOVAs, was utilized through the program ARTool (Elkin et al., 2021; Wobbrock et al., 2011). For each main effect and interaction effect, a separate ANOVA was conducted using the data aligned and ranked for that effect; only the effect for which the data was aligned was considered from each (Wobbrock et al., 2011). For instance, to examine the main effect of masking, the data was aligned and ranked for said effect using ARTool. Then, the appropriate ANOVA was conducted using this data, however, only the main effect of masking could be considered from these results. This process was repeated for each main effect and interaction. Mauchly's test was used to check the assumption of sphericity and Greenhouse-Geisser corrected values were reported if violated. For each simple effects analysis, an additional ANOVA was conducted using the data aligned and ranked for contrasts (Elkin et al., 2021).

Accuracy

To test H1, a 2 (masking: mask, no mask) x 4 (expression: happy, neutral, angry, fearful) repeated measures ANOVA was performed on accuracy (Figure 2). Significant main effects of masking, $F(1, 86) = 644.70$, $p < .001$, $\eta 2p = .88$, and expression, $F(2.52, 216.57) = 35.59$, $p < .001$, $\eta 2p = .29$, $\varepsilon = .84$, were found. These effects were qualified by a significant interaction, $F(3, 258) = 90.56$, $p < .001$, $\eta 2p = .51$. A simple effects analysis revealed that masks significantly impaired accuracy for happy, $F(1, 86) = 741.70$, $p < .001$, angry, *F(1, 86) = 138.74, p < .001, and fearful trials, F(1, 86) = 495.51,* $p < .001$, but not neutral trials. Frequent confusion with one distractor option may have caused the lack of an effect on accuracy for neutral trials. Here, the first part of H1, which stated that accuracy would be lower for masked faces, was supported, however, the second part of H1, which predicted that accuracy would be lowest for neutral masked trials, was not supported. Interestingly, when comparing accuracy among all masked trials, accuracy was highest for masked angry trials (M = 79.50, SD = 13.13), suggesting that hypervigilance was present for ambiguous threat faces.

Neutral Trial Incorrect Choice Selections

To test H2, a 2 (masking: mask, no mask) X 5 (expression choice: happy, content, angry, fearful, sad) repeated measures ANOVA was performed on the percentages of incorrect responses for each option (Figure 3). Notably, seven participants were accurate for all neutral masked trials, unmasked trials, or both, meaning that $n = 80$, as participants with missing data were excluded here. Significant main effects of masking, $F(1, 79) = 69.75$, $p < .001$, $\eta 2p = .47$, and expression choice, $F(2.43, 192.24) = 145.94$, $p < .001$, $\eta 2p = .65$, $\varepsilon = .61$ were found. These effects were qualified by a significant interaction effect, $F(1.91, 151.07) = 27.06$, $p < .001$, $\eta 2p = .26$, $\varepsilon = .48$. A simple effects analysis revealed a significant effect of masking on selecting content, $F(1, 79) = 15.96$, $p < .001$, fearful, $F(1, 79) = 17.42$, p < .001, or sad, $F(1, 79) = 18.86$, $p < .001$, rather than neutral. Specifically, when incorrect, participants tended to choose content instead of neutral for unmasked trials, while they tended to choose fearful and/or sad instead of neutral for masked trials. There was no effect of masking on selecting happy or angry rather than neutral. Contrary to H2, there was a nonsignificant trend toward selecting angry for neutral unmasked faces.

Relationships Among Variables

Lastly, to test H3, Spearman's correlations among SA scores, accuracy outcomes, and neutral trial incorrect choice selections were calculated. "SA scores" refers to both SIAS and SIAS-6 scores. "Accuracy outcomes" refers to accuracy for each masking condition. No correlations were significant among SA scores and accuracy outcomes or among SA scores and neutral trial incorrect choice selections. As such, H3 was not supported. SIAS and SIAS-6 scores had a moderate positive correlation, $r = .47$, $p < .001$.

Discussion

This study aimed to determine whether SA individuals demonstrate emotion interpretation biases when viewing masked faces, as previous research had not examined the effects of masks in populations that may be differentially affected by the associated emotion recognition challenges. The present study's SA participants demonstrated impaired accuracy for identifying the emotions of masked individuals, while having high accuracy for masked angry trials and misinterpreting neutral faces as sad and/or fearful significantly more when masked than not. Response tendencies were identified by examining confusions of neutral faces, but it is unknown if tendencies when viewing masked faces are specific to SA or if they are more generally applicable.

Accuracy

Consistent with previous research regarding the impact of masks on emotion recognition (Carbon, 2020; Gori et al., 2021; Grundmann et al., 2021), accuracy was impaired for masked trials. Accuracy was not significantly different for neutral trials based on masking, however, accuracy for neutral unmasked trials was lower than accuracy for other unmasked trials. This may have been due to one of the distractor emotions used, namely "content," being similar to "neutral." As seen in Figure 3, more than half of incorrect responses for neutral unmasked trials were content ($M = 54.78\%$).

In addition, among all types of masked trials, the present study's SA participants displayed the highest accuracy for masked angry trials. This is consistent with previous research which found that SA participants correctly identified low intensity displays of disgust and anger more often than non-anxious controls (Gutiérrez-García & Calvo, 2017), as participants in both studies identified ambiguous expressions of anger correctly. If this is specific to SA individuals, it may imply greater motivation or ability to identify ambiguous expressions of anger. Further, this relates to prior findings of hypervigilance for threat faces by SA individuals (Klumpp & Amir, 2009; Mogg et al., 2004), as anger is a social threat emotion.

Incorrect Response Tendencies for Neutral Trials

As aforementioned, neutral faces were confused as fearful and/or sad more frequently when faces were masked than not. These response tendencies are particularly interesting and may reflect participants' personal feelings during the pandemic. Notably, the mean percentage of neutral masked faces confused as sad ($M = 57.32\%$) was larger than the percentage confused as fearful ($M = 5.62\%$). Perhaps the tendency to misinterpret neutral masked faces as sad could be explained by the projection of depressive symptoms. There is reason to speculate such, as depression increased in prevalence since the pandemic began. Early in the pandemic, the prevalence of moderate to severe depression in U.S. college students was approximately 36.2% (Lee et al., 2021), while prior to the pandemic, 25.67% of U.S. college students had scores indicative of possible major depression (Acharya et al., 2018). Moreover, depression is the most common comorbid condition for individuals suffering from SAD, with 35-70% having this comorbidity based on several clinical studies (Koyuncu et al., 2019). Given these statistics, the prevalence of depression in this sample was potentially high. If so, the sad response tendency may reflect participants projecting their own negative emotions onto the neutral faces viewed, particularly when ambiguous due to masking.

Importantly, projection of one's own affective state when interpreting the emotions of others is supported by research. For example, Trilla et al. (2021) observed the phenomenon of emotional egocentricity when participants' affective states were manipulated before an emotion perception task. Participants were more likely to judge faces as sad when they were experiencing sadness as opposed to happiness, indicating that projection affects the perception of facial expressions. Furthermore, in a study comparing the responses of depressed, remitted, and healthy control individuals, Leppänen et al. (2004) found an impairment in the recognition of neutral faces. Specifically, depressed and remitted patients tended to mistake neutral faces for emotional faces (e.g., sad or happy), although false happy responses occurred more in the group of remitted patients. Given this strong evidence for perception of neutral faces as sad in those with depressive symptoms, along with the phenomenon of emotional egocentricity shown by Trilla et al. (2021) and the prevalence rates of depression mentioned previously, projection of depressive emotions may have impacted interpretations of neutral masked faces in the present study.

In contrast, several prior studies have shown a bias for interpreting neutral faces as threatening in SA in dividuals (Gutiérrez-García & Calvo, 2017; Yoon & Zinbarg, 2008), which was not replicated here. The present results illustrate that SA participants most frequently confused neutral faces as sad when masked, while Gutiérrez-García and Calvo (2017) found that SA individuals were less likely than non-anxious controls to interpret neutral faces as sad. This inconsistency suggests a mask-specific effect, however, it is unknown whether this is generalized or SA-based. Although the neutral threat bias, such as that shown by Yoon and Zinbarg (2008), was not present for neutral unmasked trials either, the presence of prior masked trials may have impacted this, and thus, the previous findings are not refuted. Further, the methodology may not have allowed for enough SA variation to produce similar results. The present study recruited participants with a minimum score, whereas previous studies selected extreme scores from a participant pool (Gutiérrez-García & Calvo, 2017), used median split grouping (Klumpp & Amir, 2009), or clinician diagnosis grouping (Mogg et al., 2004).

Limitations

The present study had several important limitations. First, it did not include a low- or non-anxious group due to the limited scope of this study, being that it served as a first-year project leading up to a master's thesis. Inclusion of a control group would aid in elucidating whether the identified mask-related response tendencies are specific to SA. Second, the researcher lacked access to a clinical sample, relying on a brief SA scale for pre-screening, which was not as highly correlated with the full scale as expected (Peters et al., 2012), suggesting measurement problems. Third, ecological validity should be considered, as static images are unlike real life, where movement and contextual information facilitate emotion interpretation.

Implications and Future Research

Despite the limitations, the present findings contribute information to the research gap on populations that may be uniquely affected by the challenges associated with mask-wearing. As predicted, emotion identification accuracy was impaired for masked trials, however, neutral masked trials were misinterpreted differently than expected (as sad rather than angry). These findings can inform clinicians about how those with SA interpret the emotions of masked individuals, contributing to an improved understanding of pandemic-related effects when providing treatment. The mental health effects of the incorrect response tendencies identified here should be considered, as individuals with SA attribute excessive social cost to negative emotions, including fear and sadness (Button et al., 2013). Importantly, neutral masked faces were most often confused as sad, with this occurring significantly more than when unmasked. If SA individuals often misinterpret neutral masked faces as sad, this may cause greater anxiety for these individuals when in social situations due to their attribution of excessive social cost to interacting with individuals displaying sadness. Furthermore, these findings may bring increased awareness to possible differential effects of masks on other neurodivergent populations. Future research could examine the effects of masks on emotion interpretation in other populations typically affected by interpretation biases to determine whether the known effects are altered by the presence of masks.

In addition, the higher accuracy shown for masked angry trials may reflect greater motivation or ability of SA individuals to identify these social threat emotions, which can be explained by their hypervigilance for threat faces (Klumpp & Amir, 2009; Mogg et al., 2004). These results reinforce prior findings of hypervigilance for ambiguous threat faces, demonstrating that this effect is also applicable to situations when SA individuals view masked angry faces. However, the previously identified neutral threat bias, which is characteristic of SA, was not reproduced here, suggesting that this effect may not be present when viewing faces during the pandemic. Instead, a different emotion interpretation bias in which neutral masked faces were most often misinterpreted as sad, may have taken precedence. This bias may result from the projection of depressive emotions, as depression is the most common comorbidity of SA and its prevalence increased during the pandemic. Future research should examine the role of both SA and depression in interpreting the emotions of masked faces and should prioritize the inclusion of a low- or non-anxious control group.

Acknowledgements

The author wishes to thank Dr. Kim Shifren for her continued mentorship and valuable insight, Dr. Justin Buckingham for consultation on study design, Dr. Jeff Kukucka for statistical suggestions, and Ryan Rosenthal for his critical reviewal of the manuscript and support throughout.

Declaration of Interest Statement

The author reports no conflict of interest.

Data Availability Statement

The data that support the present findings are openly available in Open Science Framework at http://doi.org/10.17605/OSF.IO/VM4DC.

References

Acharya, L., Jin, L., & Collins, W. (2018). College life is stressful today – Emerging stressors and depressive symptoms in college students. *Journal of American College Health, 66(7)*, 655–664. https://doi.org/10.1080/07448481.2018.1451869

Button, K., Lewis, G., Penton-Voak, I., & Munafò, M. (2013). Social anxiety is associated with general but not specific biases in emotion recognition. *Psychiatry Research, 210(1),* 199–207. https://doi.org/10.1016/j.psychres.2013.06.005

Campbell, J. I. D., & Thompson, V. A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. B*ehavior Research Methods, 44(4),* 1255–1265. https://doi.org/10.3758/s13428-012-0186-0

Carbon, C. C. (2020). Wearing Face Masks Strongly Confuses Counterparts in Reading Emotions. *Frontiers in Psychology, 11,* 1–8. https://doi.org/10.3389/fpsyg.2020.566886

Carleton, R. N., Collimore, K. C., Asmundson, G. J. G., McCabe, R. E., Rowa, K., & Antony, M. M. (2009). Refining and validating the social interaction anxiety scale and the social phobia scale. *Depression and Anxiety, 26(2),* E71–E81. https://doi.org/10.1002/da.20480

Elkin, L. A., Kay, M., Higgins, J. J., & Wobbrock, J. O. (2021). An Aligned Rank Transform Procedure for Multifactor Contrast Tests. *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST '21)*, 754–768.

Gori, M., Schiatti, L., & Amadeo, M. B. (2021). Masking Emotions: Face Masks Impair How We Read Emotions. *Frontiers in Psychology, 12,* 1–9. https://doi.org/10.3389/fpsyg.2021.669432

Grundmann, F., Epstude, K., & Scheibe, S. (2021). Face masks reduce emotion-recognition accuracy and perceived closeness. *PLoS ONE, 16(4),* 1–18. https://doi.org/10.1371/journal.pone.0249792

Gutiérrez-García, A., & Calvo, M. G. (2017). Social anxiety and threat-related interpretation of dynamic facial expressions: Sensitivity and response bias. *Personality and Individual Differences, 107*, 10–16. https://doi.org/10.1016/j.paid.2016.11.025

Hawes, M. T., Szenczy, A. K., Klein, D. N., Hajcak, G., & Nelson, B. D. (2021). Increases in Depression and Anxiety Symptoms in Adolescents and Young Adults during the COVID-19 Pandemic. *Psychological Medicine,* 1–9. https://doi.org/10.1017/S0033291720005358

Kang, S. M., & Lau, A. S. (2013). Revisiting the outgroup advantage in emotion recognition in a multicultural society: Further evidence for the in-group advantage. *Emotion, 13(2),* 203–215. https://doi.org/10.1037/a0030013

Klumpp, H., & Amir, N. (2009). Examination of vigilance and disengagement of threat in social anxiety with a probe detection task. Anxiety, *Stress and Coping, 22(3),* 283–296. https://doi.org/10.1080/10615800802449602

Koyuncu, A., İnce, E., Ertekin, E., & Tükel, R. (2019). Comorbidity in social anxiety disorder: Diagnostic and therapeutic challenges. In *Drugs in Context* (Vol. 8, pp. 1–13). Bioexcel Publishing LTD. https://doi.org/10.7573/dic.212573

Lee, J., Jeong, H. J., & Kim, S. (2021). Stress, Anxiety, and Depression Among Undergraduate Students during the COVID-19 Pandemic and their Use of Mental Health Services. *Innovative Higher Education, 46(5),* 519–538. https://doi.org/10.1007/s10755-021-09552-y

Leppänen, J. M., Milders, M., Bell, J. S., Terriere, E., & Hietanen, J. K. (2004). Depression biases the recognition of emotionally neutral faces. *Psychiatry Research, 128(2)*, 123–133. https://doi.org/10.1016/j.psychres.2004.05.020

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47(4),* 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy, 36(4)*, 455–470.

Mogg, K., Philippot, P., & Bradley, B. P. (2004). Selective Attention to Angry Faces in Clinical Social Phobia. *Journal of Abnormal Psychology, 113(1),* 160–165. https://doi.org/10.1037/0021-843X.113.1.160

National Institute of Mental Health. (2017). *Social Anxiety Disorder*. https://www.nimh.nih.gov/health/statistics/social-anxiety-disorder

Peters, L., Sunderland, M., Andrews, G., Rapee, R. M., & Mattick, R. P. (2012). Development of a short form Social Interaction Anxiety (SIAS) and Social Phobia Scale (SPS) using nonparametric item response theory: The SIAS-6 and the SPS-6. *Psychological Assessment, 24(1)*, 66–76. https://doi.org/10.1037/a0024544

Salter, K. C., & Fawcett, R. F. (1993). The ART test of interaction: A robust and powerful rank test of interaction in factorial models. *Communications in Statistics: Simulation and Computation, 22(1)*, 137–153. https://doi.org/https://doi.org/10.1080/03610919308813085

Schofield, C. A., Coles, M. E., & Gibb, B. E. (2007). Social anxiety and interpretation biases for facial displays of emotion: Emotion detection and ratings of social cost. *Behaviour Research and Therapy, 45(12)*, 2950–2963. https://doi.org/10.1016/j.brat.2007.08.006

Trilla, I., Weigand, A., & Dziobek, I. (2021). Affective states influence emotion perception: evidence for emotional egocentricity. *Psychological Research, 85(3),* 1005–1015. https://doi.org/10.1007/s00426-020-01314-3

Tuminello, E. R., & Davidson, D. (2011). What the face and body reveal: In-group emotion effects and stereotyping of emotion in African American and European American children. *Journal of Experimental Child Psychology, 110*(2), 258–274. https://doi.org/10.1016/j.jecp.2011.02.016

Wiesmann, M., Franz, C., Sichtermann, T., Minkenberg, J., Mathern, N., Stockero, A., Iordanishvili, E., Freiherr, J., Hodson, J., Habel, U., & Nikoubashman, O. (2021). Seeing faces, when faces can't be seen: Wearing portrait photos has a positive effect on how patients perceive medical staff when face masks have to be worn. *PLoS ONE, 16(5),* 1–9. https://doi.org/10.1371/journal.pone.0251445

Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. (2011). The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only ANOVA Procedures. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '11)*, 143–146.

Wong, C. K. M., Yip, B. H. K., Mercer, S., Griffiths, S., Kung, K., Wong, M. C., Chor, J., & Wong, S. Y. (2013). Effect of facemasks on empathy and relational continuity: a randomised controlled trial in primary care. *BMC Family Practice, 14*(**200**), 1–7. http://www.biomedcentral.com/1471-2296/14/200

Yoon, K. L., & Zinbarg, R. E. (2008). Interpreting Neutral Faces as Threatening Is a Default Mode for Socially Anxious Individuals. *Journal of Abnormal Psychology, 117*(3), 680–685. https://doi.org/10.1037/0021-843X.117.3.680

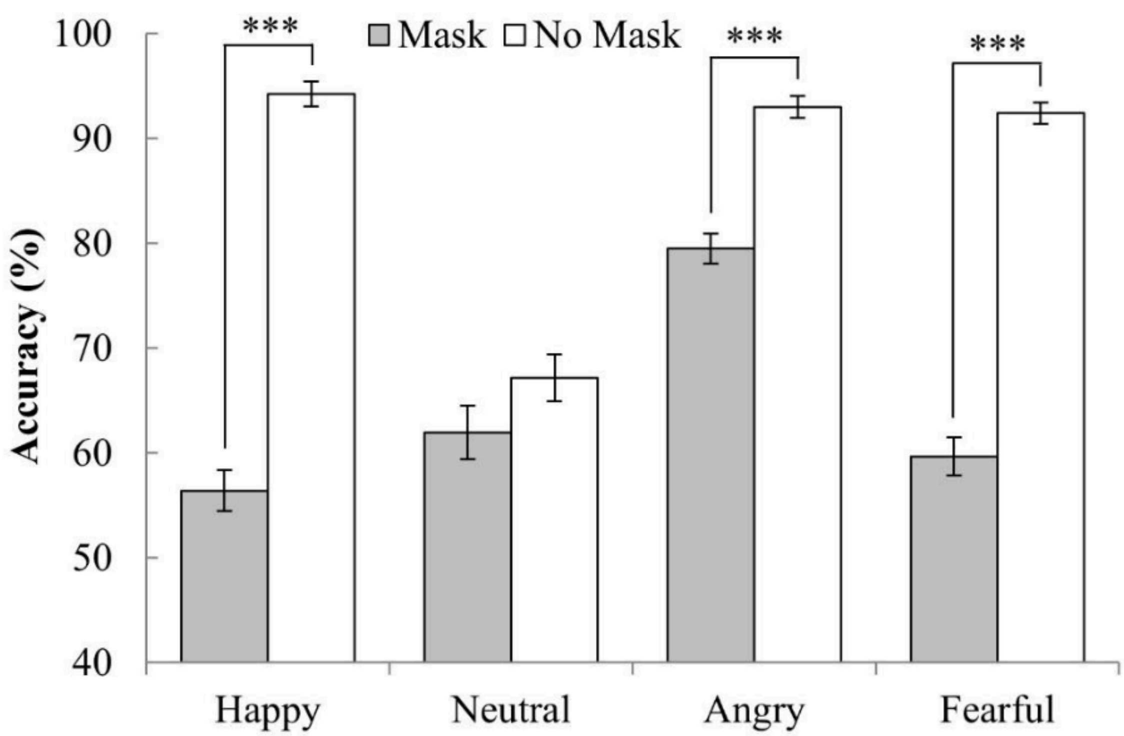
Figure 1

Emotion Identification Task Image Examples

Note. Although the same individual is shown here, this is for illustrative purposes only, as no individuals were included more than once in the task. For this individual, the expression shown is neutral and the mask-edited image was used in the task. CFD stimulus code: CFD-WF-211-001-N. (a) original CFD image. (b) mask-edited CFD image.

Figure 2

Emotion Identification Accuracy Among Masking and Expression Conditions

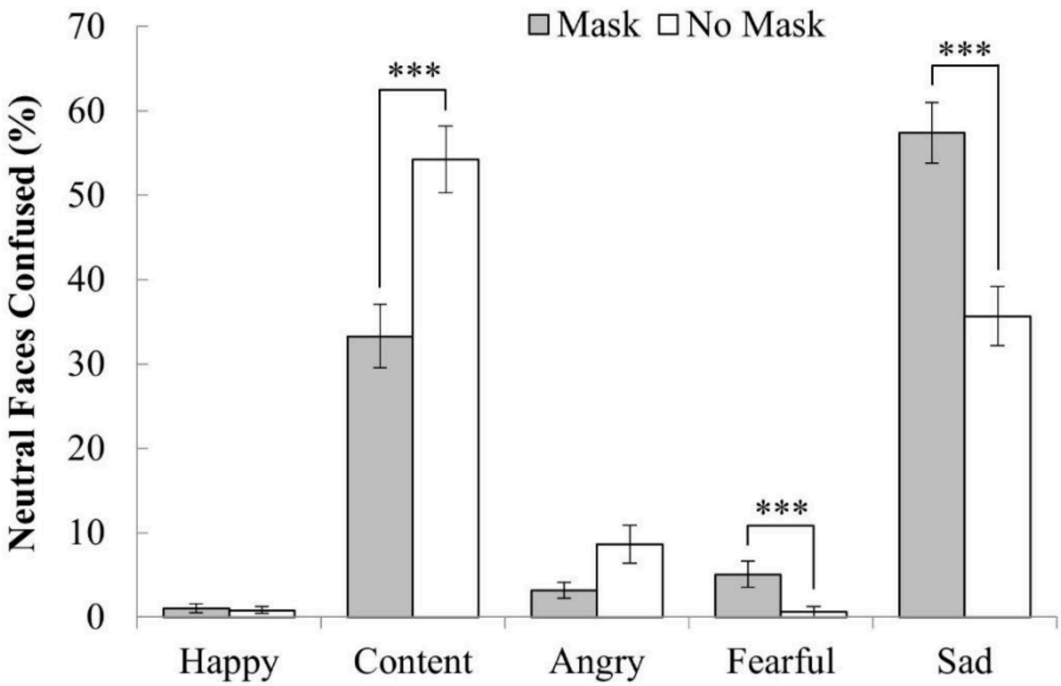


Note. $n = 87$.

$^{***}p < .001$.

Figure 3

Neutral Trial Incorrect Choice Selections



Note. $n = 80$.

***$p < .001$.