Quantifying the influence of face masks on the interpretation of non-verbal emotional communication

Joost de Wildt
7185928
j.dewildt@students.uu.nl

Master's thesis
Credits: 27.5 ECTS

Master of Science Applied Cognitive Psychology (66604)

Utrecht University
Faculty of Social and Behavioral Sciences
Heidelberglaan 1
3584 CS Utrecht

Thesis supervisor
dhr. dr. S. (Surya) Gayet
Assistant Professor
s.gayet@uu.nl

Auditor
dhr. dr. L. (Leendert) van Maanen
Associate Professor
l.vanmaanen@uu.nl

28-06-2022

Universiteit Utrecht

Abstract

In the healthcare sector, good non-verbal communication of doctors towards their patients is proven to be of major importance. The broad use of face masks in the healthcare sector, however, complicates the transmission of non-verbal cues. Following the COVID-19 pandemic, several studies showed early evidence that interpreting emotional expressions was done less accurate when the expressor wore a mask. However, compared to the vast amount of research considering face masks' effectivity in reducing virus transmission, the number of studies regarding these communicative consequences pales into insignificance. More research is thus found to be necessary to be able to make a well-informed decision on whether wearing a mask is valuable. This study builds on the existing studies on the latter topic and takes a step towards quantifying the effect of face masks on the interpretation of non-verbal emotional communication. It does so by controlling for mask appearance, testing the influences of race and comparing effects of masks on distinguishing emotion valences with the same effects of showing only the upper face half. Results confirmed the negative effect of face masks on emotion recognition as seen in earlier studies. This effect was found to be similar to the effect of half faces, stressing its magnitude and implying that mask-wearing faces are encoded purely through the facial features in the upper face half. For Asian faces in particular, the effect of face masks was strongest. As people of Asian ethnicity are suggested to show lower intensity facial expressions, it is implied that high intensity emotion display in the facial features in the upper face half is important for nuancing masks' negative effect on emotional communication. In conclusion, limitations of this study and its implications for the healthcare sector are discussed.

Keywords: face masks, emotion recognition, facial expressions, race similarity, doctor-patient communication

Table of contents

Introduction

In the healthcare sector, good communication of doctors towards their patients is proven to be of major importance. On one hand, good communication of the doctor is claimed to be a large factor in patient satisfaction and increase the likelihood of a positive clinical outcome (Buckman, 2002; Marcinowicz et al., 2010; Wong et al., 2013). In the perspective of the doctor, on the other hand, failing to effectively communicate with the patient is linked to increased chances of a burnout and medicolegal issues. Non-verbal communication is particularly important in this context. In general, it has been found that roughly 80% of the concepts discussed in essential interpersonal communication are expressed through non-verbal cues (Bhat & Kingsley, 2020). Not only can non-verbal communication substitute verbal communication, it also complements and reinforces the verbal message. That is, the essence of spoken words is partially dependent on the supplemented non-verbal communication and can provoke a different response when accompanied by different non-verbal cues (Bhat & Kingsley, 2020). In a doctor-patient context, non-verbal cues are even believed to outweigh the verbal message in the case of contradicting verbal and non-verbal communication (Silverman & Kinnersley, 2010). As non-verbal communication 'spills' information that is not necessarily supposed to be shared, patients focus on the non-verbal cues that their doctor expresses and concentrate less on the accompanied verbal message (Ong et al., 1995). To illustrate, when a doctor has to convey disturbing news to a patient, showing reassuring non-verbal cues will likely reassure the patient, regardless of the severity of the situation (Bhat & Kingsley, 2020).

The prominence of non-verbal cues in this context is mainly due to their effectiveness in conveying (supportive) emotions. Non-verbal communication is shown to be the primary way to convey emotions and empathy (Roter et al., 2006; Vogel et al., 2018). Overall, merely 7% of emotional communication is done verbally, while 55% is performed through visual cues (Ong et al., 1995). Similarly, sympathetic traits such as warmth, likability and enthusiasm are believed to be more easily conveyed through non-verbal communication than through spoken words (Roter et al., 2006). As doctors are often expected to provide emotional support in their communication towards their patients (Ong et al., 1995), it is important to convey said emotions and traits clearly to facilitate a strong doctor-patient relationship and increase patient satisfaction. Overall, effective non-verbal communication of doctors towards their patients is thus of great importance in improving the provided healthcare. The effect of face masks usage on such communication is therefore a relevant topic to investigate for the healthcare sector, considering that doctors often wear a mask to prevent disease transmission. This thesis will look deeper into this topic to investigate the role that face masks have in healthcare quality.

Besides cues like body posture and tone of voice, facial expressions are important non-verbal prompts by which emotions are communicated (Botvinick et al., 2005). These facial expressions are typically formed through a combination of miniscule facial movements – or features – across different areas of the face, such as slightly raising an eyebrow (upper area) or clenching the jaw (lower area). The use of multiple facial areas for showing an expression implies that covering parts of the face may complicate the interpretation of (some of) these expressions, as it hides some facial features needed for showing that (Cohn et al., 2007). Having said that, it is not uncommon for doctors to wear a face mask while communicating with their patients. Especially during outbreaks of viruses like influenza, face masks are used even more in the healthcare sector (Chellamani et al., 2013; Lim et al., 2006). The reason for wearing a mask is that it should reduce the chance of the doctor transmitting a virus or other bacteria to the, often vulnerable, patient – or vice versa. A great deal of existing literature on face masks is therefore also dedicated to their ability to reduce such disease transmission (e.g. Ammon et al., 2000; Cowling et al., 2010; MacIntyre & Chughtai, 2015). However, to make a well-informed decision as to whether a face mask is worth wearing, its effectiveness should not be the only determinant – potential negative effects of wearing them must also be identified. Especially considering the earlier mentioned value of non-verbal emotional communication in a healthcare context, doctors should know the effect of wearing a mask on how their facial expressions are perceived by a potential patient.

Preceding research and current experimental setup

Extensive research into potential negative side effects of masks was largely missing until the outbreak of the COVID-19 pandemic. Similarly, few pre-pandemic studies exist that investigate the obstruction of facial expressions and its influence on emotional communication (e.g. Wong et al., 2013). As the COVID-19 pandemic led to widespread face mask usage over the entire world, the associated implications instantaneously became applicable (and relevant) to many more contexts than just the healthcare sector – leading to several new publications on this topic (e.g. Bani et al., 2021; Carbon, 2020; Ziccardi et al., 2022).

Findings of preceding research are mostly consistent in showing a decrease in emotion recognition accuracy when the face expressing that emotion is partially covered by a face mask. Having said that, this early evidence merely shows a fundamental hindering effect of face masks on emotion recognition, but lacks to investigate this effect more in-depth. As aforementioned, the number of studies regarding the communicative consequences pales into insignificance compared to the vast amount of research discussing masks' protective value. The studies that do exist show rather similar approaches in their experimental setup, which questions the generalizability of the conclusions drawn in this research. To be able to weigh the pros and cons of wearing a mask, establishing a more in-depth and generalizable quantification of face masks' influence on non-verbal communication thus seems a logical next step. This thesis will set a first step towards filling this research gap by building on the preceding studies and expanding their experimental setup through the following adaptations.

The experiments in previous studies (e.g. Bani et al., 2021; Carbon, 2020; Grundmann et al., 2021) seem to agree on using a similar adaptation of the Vienna Emotion Recognition Task (VERT-K; Derntl et al., 2008). Participants are typically exposed to facial stimuli showing a facial expression representing one of the six 'basic' emotions seen in facial expressions (sadness, anger, disgust, fear, happiness, surprise; Ekman & Friesen, 1971). Stimuli either show an unconcealed face or one containing a mask. Face masks' influence is subsequently determined by analyzing the variance in participants' performance in recognizing the specific emotions. Especially in a healthcare context, however, recognition of emotion valence (positive, negative or neutral) seems more relevant than identifying specific emotions. That is, the doctor's non-verbal communication typically has a reassuring or worrying role for the patient (Ong et al., 1995). Recognizing the presence of an emotion within the doctor's facial expression and whether this is concerning (negative) or assuring (positive) thus seems most important for patients. Investigating the ability to identify the presence of an emotion and distinguish emotions of positive and negative valence from each other in a face mask context will therefore be of interest in this study.

Secondly, it is difficult to determine the actual strength of the noted effect on emotion recognition when only compared with the ideal (seeing an unconcealed face). To quantify the effect strength, both an ideal and a lower limit should be available to compare the emotion recognition in a mask context with. Only once has this been done before in this context (Grahlow et al., 2022). Therefore, besides measuring emotion recognition on faces with and without mask, stimuli showing only the upper face half will also be used. If emotion recognition is also measured with removed lower face halves, the influence of wearing a mask can be quantified along a scale with seeing the entire face as the ideal and the inability to see the lower half as the lower limit.

Third, existing studies develop face mask stimuli by photoshopping a standardized blue disposable face mask over half of the stimuli. However, it is therefore not sure whether the identified effects are due to the face masks' appearance rather than the presence of a face mask in general. In response, to solely measure the effect that wearing a face mask has on emotion recognition, it seems necessary to reduce chances of mask appearance interfering with the results by including different face masks in the stimuli (i.e. color, print, type of mask).

Lastly, stimuli used in the preceding studies were all retrieved from one of two databases of Caucasian models (FACES database; Ebner et al., 2010; or the Karolinska Directed Emotional Faces (KDEF) dataset; Lundqvist et al., 1998). Possible ethnical or racial differences in non-verbal emotional communication were therefore generally not considered. For a long time, the facial expressions associated with the six basic emotions tested in these studies have been suggested to be consistent across all cultures, making them universally interpretable (Chen & Jack, 2017; Wagner, 1990). However, there is an increasing body of literature claiming emotion recognition to be less accurate for faces from a different race than the observer (e.g. Elfenbein & Ambaday, 2002; Kito & Lee, 2004). Not to mention that the role of cultural or racial differences in a face mask context is yet to be discovered. Investigating a prospective interaction between the hindering effects of both face masks and race similarity on emotion recognition could give more insight into the strength of both effects. Moreover, with the growing emphasis on an inclusive, interracial society seen in recent years, exploring race's prospective role in the effect of face masks on emotion recognition seems relevant.

Hypotheses

Overall, wearing a mask seems to have a negative effect on emotion recognition in general (e.g. Bani et al., 2021; Grahlow et al., 2022; Grundmann et al., 2021). Adding the mentioned elements to the experiment, however, enables more thorough investigation of this effect. Based on existing literature, hypotheses are set up to facilitate this in-depth analysis.

Interpreting positive and negative emotions

Within the context of doctor-patient communication, presence and valence of the expressed emotion is considered most important – rather than identifying specific emotions. Overall, it seems as if classifying facial expressions as either showing a positive or negative emotion is easier than recognizing specific emotions. Whereas research agrees on multiple generic emotions of a negative valence (sadness, fear, disgust and anger), happiness is typically considered as being the only common positive emotion (Posamentier & Abdi, 2003). Accordingly, a positive facial expression is found to be easily recognized, while expressions of negative valence are more often wrongly interpreted for a different negative emotion (Dores et al., 2020; Wegrzyn et al., 2017). A similar pattern is seen in the preceding studies on face masks and emotion recognition. Whereas a hindering effect of face masks on emotion recognition accuracy is found, this effect is largely due to emotions of negative valence being confused with one another (Carbon, 2020; Grahlow et al., 2022). Although negative emotions may thus be confused with one another, positive and negative emotions seem to generally be correctly distinguished from each other – in both a face mask and an unconcealed face context. This implies that wearing a face mask does not influence the level to which negative and positive emotions can be distinguished from one another.

H1a: *A face mask does not increase difficulty in distinguishing positive and negative emotions from one another, compared to not wearing a mask.*

Having said that, the part of the observed hindering effect of mask wearing that is not caused by negative emotions being confused is largely due to participants being unable to identify any form of emotion. In other words, participants wrongly classified an emotional expressions as not containing any emotion. In the studies of Carbon (2020) and Grahlow et al. (2022), participants particularly struggled to discern emotion in stimuli of mask-wearing faces showing happy and sad emotional expressions. These results suggest that although positive and negative emotions can still be distinguished from one another in a mask context, the ability to identify the presence of emotion will be affected by a face mask.

H1b: *A face mask does increase difficulty in perceiving the presence of emotion, compared to not wearing a mask.*

Processing facial expressions

The added stimuli of the upper face half could not only give better insights into the strength of masks' effect on emotion recognition, it may also have several implications for understanding how facial expressions are processed and encoded. As referred to earlier, facial expressions are formed through a combination of attributes in different areas of the face (Cohn et al., 2007). Each of the six basic emotions is expressed through a specific set of these features (Wegrzyn et al., 2017). To illustrate, a smiling mouth and sparkling eyes are characteristics of happiness, whereas raised inner eyebrows and slanted mouth corners are typically associated with sadness. Early research found that humans instinctively encode facial expressions in a holistic manner, meaning that the integration of these features into the face is processed rather than the individual features (Calder et al., 2000; Calder & Jansen, 2005). Only when holistic processing is disrupted, the human mind is suggested to individually encode the available features required to identify the emotion associated with the facial expression (Tanaka et al., 2012). Individual feature processing does, however, cause less accurate emotion recognition than the holistic approach (Bombari et al., 2013).

Thus far, studies on facial processing have mainly found stimuli of half, scrambled, blurred or inverted faces to disrupt holistic evaluation (e.g. Bassili, 1979; Bombari et al., 2013; Collishaw & Hole, 2000). For example, Bassili (1979) observed how participants scored lower in an emotion recognition task when exposed to photos of only one half of the face than when the entire face was shown. Considering that a face mask conceals the lower half of a face (when worn correctly), it can be argued that wearing a mask causes similar effects as presenting only the upper half of the face. On the other hand, since the outline of the whole face remains intact, it is also possible that face masks actually cause less interference with emotion recognition than without the lower half of the face. In case of the latter, it can be concluded that faces containing a mask are still holistically evaluated, which is known to improve the extraction of a whole range of information from faces, such as identity, gender, attractiveness and trustworthiness (Abbas & Duchaine, 2008; Baudouin & Humphreys, 2006; Todorov et al., 2010; Young et al., 1987). Looking into the effect of face masks on holistic facial processing thus seems relevant for these reasons. By comparing the effect of face masks on emotion recognition with viewing only the upper half of a face, more insight can be gained into a face mask's impact on holistic face perception.

When comparing the individual studies done on face masks and half face stimuli, varying results regarding emotion recognition are noted. In various studies on individual feature processing, it appeared that the bottom area of the face was most important for accurate recognition of happiness (Bassili, 1979; Leppänen et al., 2007; Smith et al., 2005). Removing the lower half of a face caused participants to struggle with correctly identifying happiness, whereas excluding the upper half did not have this effect (Bassili, 1979). In contrary, the existing literature on face masks' effects showed that happiness was one of the emotions of which the recognition was least impeded in the mask condition (Carbon, 2020; Grahlow et al., 2022; McCrackin et al., 2022). These incongruencies first suggest that faces wearing a mask are not processed by their individual features and are still evaluated holistically. Secondly, it implies that removing the lower half of the face causes more confusion between negative and positive emotions.

H2: *A face mask does not decrease emotion recognition to the same extent that disregarding the bottom half of a face does.*

Race and the interpretation of facial expressions

Whereas a large body of research already confirmed the identification of same-race faces to be more easily recognized (e.g. McKone et al., 2019; Tanaka & Pierce, 2009), a growing number of studies now also claims facial expression encoding to be done more easily on same-race faces than on other-race faces. Yan et al. (2016), for example, found that Chinese participants made less errors with categorizing emotional expressions of faces of their own race than with those of Caucasian models – and vice versa. Similarly, Weathers et al. (2002) noted that African Americans were more successful at interpreting emotions when it considered faces of other African Americans, compared to those of Caucasians. Overall, humans are suggested to be more emotional attuned to faces of the same race as themselves than to other-race faces (Lee et al., 2008). This causes more cognitive functions to be necessary for encoding other-race faces.

Overall, processing other-race faces seems more cognitively intense, causing emotion recognition to be more difficult. In the same way, wearing a face mask is also suggested to be hindering for emotion recognition. It is therefore hypothesized that seeing a mask on an other-race face will cause some negative interaction between these two elements.

H3: *A face mask will have less effect on emotion recognition when it considers a same-race face than when worn by an other-race face.*

Method

Sample

Following Carbon (2020), the required sample size for this study was calculated through *a priori* power analysis performed in G*Power 3 – a stand-alone power analysis program commonly used in social and behavioral research (Faul et al., 2007). G*Power suggested a minimum sample size of *N* = 27 in order to identify a medium effect size ($d_z = 0.50$; Cohen, 1988) with a confidence interval of 95% ($\alpha = 0.05$) and power ($1-\beta$) of 0.80. Similar parameter values as Carbon (2020) were used. To minimize uncertainty about its outcome, it was decided not to consider all three measures (with mask, without mask and upper half of the face) and groups (negative, neutral and positive emotions) in this analysis. Instead, the *a priori* power analysis targeted a repeated measures *t*-test, only considering the two critical measures of this study (with and without mask).

Between May 10 and May 20 2022, a total of 25 participants (16 female) were gathered, with an age ranging from 19 to 79 (total: *M* = 29.4, *SD* = 17.0; female: *M* = 27.8, *SD* = 3.85; male: *M* = 32.2, *SD* = 6.77). A Shapiro Wilk test showed age to be non-normally distributed, W(25) = .55, $p < .001$, with a median age of 23 (female: *Mdn* = 23; male: *Mdn* = 23).

Considering the setup of the experiment, data was checked for outliers to ensure older participants did not pollute the data. No outliers were detected for mean emotion recognition accuracy in all three conditions, whereas two outliers were seen for mean response time in every stimuli condition. Given the within-subject design, however, data of the outliers was still considered in all the data analyses.

Twenty-four participants noted that they identified most as being white (Caucasian), whereas the remaining participant identified most as Asian. All participants reported normal or corrected-to-normal vision. Participants were recruited via personal invitations and through the SONA platform. On this platform, students of University Utrecht's Social and Behavioral Sciences faculty can receive credits for participating in experiments that are submitted on the platform. If applicable, participants thus received one credit per hour for participating in this study. As an incentive for the remaining participants, a €25 bol.com gift card was raffled.

Stimuli

The stimuli used in the experiment was retrieved from the racially diverse affective expression (RADIATE) face stimulus set (Conley et al., 2018). This open-access stimulus set includes photos of over hundred racially diverse models displaying sixteen different facial expressions. These include angry, disgusted, fearful, sad, happy, calm and neutral expressions with both an open and closed mouth and expressions of surprise and exuberant happiness to conclude the collection. All expressions are tested for their reliability, meaning that they are generally representative for the emotion that they are supposed to represent (Conley et al., 2018). For this study, expressions of anger and disgust were not used for the stimuli, as these were not considered as relevant emotions in a doctor-patient context. Photos of surprised models were also disregarded because of the difficulty to categorize this emotion as having either a positive or negative valence. Since this study investigates participants' ability to separate positive, negative and neutral emotions from one another, the used stimuli should be clearly separatable into these categories. Lastly, although the closed mouth expression of sadness shows a high reliability, the photos of an open mouthed sad expression appeared to be often confused as an expression of disgust (Conley et al. 2018). Removing the abovementioned expressions from consideration led to the use of fearful (open & closed mouth) and sad (closed mouth) expressions to represent negative emotions, the calm (open & closed) and neutral expressions (closed mouth) to represent neutral emotions and the expressions of happiness (open, closed & exuberant) to act as positive emotions. The open mouthed neutral expression was also disregarded to equalize the number of expressions used in every condition.

For the experiment, 216 photos of the RADIATE stimulus set were used. Of each race represented in the RADIATE set (black, Hispanic, Asian and white/Caucasian), six models (three male, three female models) were included in the stimuli. A total of nine photos were included for each of the 24 models, each portraying one of the facial expressions mentioned above. After removing the white background of the photos with Adobe Photoshop, a total of 216 extra stimuli were subsequently developed by removing the lower half of each face, followed by another 216 stimuli of the same faces covered with a face mask (see Figure 1). Nine images of different face masks were taken from the Google Image search engine, varying in color, print and type (i.e. 'standard' medical, fabric/reusable or FFP2). Using Adobe Photoshop, the masks were systematically photoshopped on the models to create diverse mask-expression combinations. To optimize the stimuli's realism, the masks were warped to the curve of the model's face and mask straps were drawn. Another 72 stimuli (24 per condition, one male and female model per race, one positive, negative and neutral expression per model) were conclusively developed in the same manner, which were used to let participants practice with the task before engaging in the actual experiment.

Procedure

As mentioned, the experiment was performed in a lab at University Utrecht campus in which most conditions could be controlled. Participants were sat behind a 25-inch iiyama monitor (1920 × 1080 resolution, 100 Hz refresh rate), with their head placed in a head movement stability tool positioned at 570 millimeter from the center of the screen. Each participant had the desk height adjusted in order to sit with a straightened back and their head comfortably resting in the stability tool. Before engaging in the actual experiment, participants first got shown a walkthrough of the task and performed a practice round with 12 random photos of the extra developed stimuli. From the practice round onward, the lighting of the lab was shut off to make sure the only light source in the room was the computer screen.

The actual experiment was set up as an adaptation of the Vienna Emotion Recognition Task (VERT-K; Derntl et al., 2008). Participants were exposed to a series of 432 stimuli (8.0° x 8.0°), presented with Matlab (The MathWorks, Natick, MA). The stimuli was divided into nine equal blocks of 48 photos, each block taking approximately two minutes. When having finished a block, participants were shown their overall accuracy score and response time for that 48 stimuli and were asked to take a self-paced break from the task to prevent them from tiring and losing focus. After completing the nineth (and last) block, participants were asked for any comments on the experiment, age, gender and the race they most identified with.

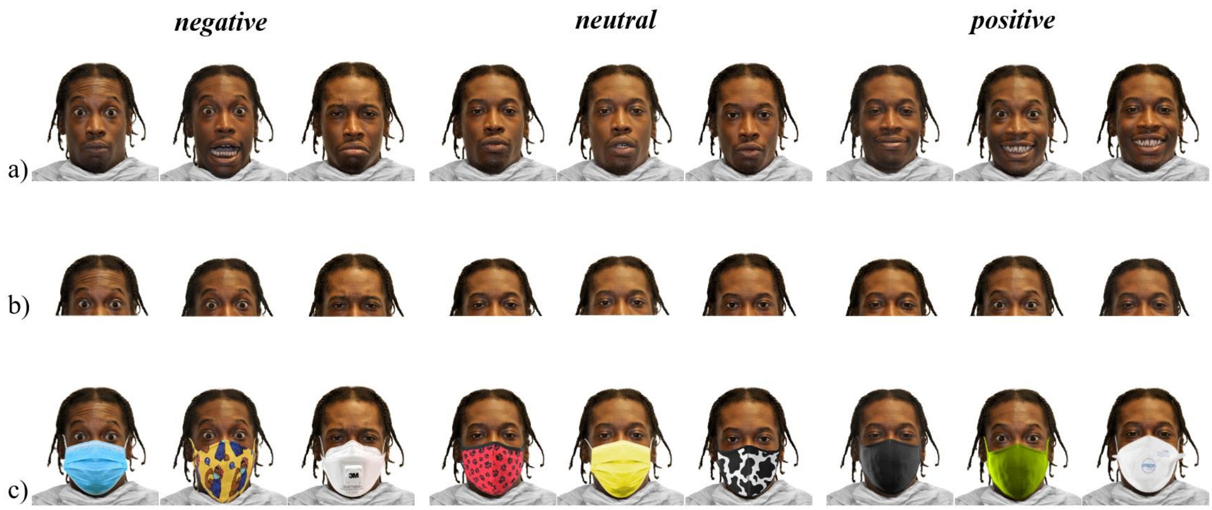

Figure 1. Example of the used stimuli; a black man showing 9 different facial expressions without face mask (a), with only the top half of his face (b) and while wearing a face mask (c). Each facial expression represents either a negative, neutral or positive emotion.

Trial outline

Participants were asked to focus on a fixation dot in the middle of the screen, which was replaced by a stimulus after 1000ms (see Figure 2). Different to the VERT-K procedure, each stimulus only appeared for 150 ms. As 120 and 160 ms is typically needed for holistic and individual feature processing of emotional expressions, respectively (Eimer & Holmes, 2002), an exposure time of 150 ms was chosen to draw out the most visceral answers. After showing the stimulus, participants were prohibited from answering for the first 300ms, after which they were given a maximum of two seconds to answer by using the keyboard. When response time exceeded the limit of 2000ms, participants were shown a statement saying they were too slow and the next trial was initiated. Following each answer, participants were shown whether their answer was correct. Although this arguably induced learnability of how different emotions were expressed by the models, the main aim was encouraging participants to attempt to provide accurate responses. The presentation of accuracy score and response time in between the blocks was done for the same purpose.

Experimental setup

As aforementioned, stimuli was developed into three conditions (i.e. face mask, half-face and whole-face stimuli). All participants viewed stimuli from all three of these conditions in a semi-randomized order to nuance a prospective learning effect. This entailed presenting no more than two photos of the same condition or three photos of the same facial expression successively

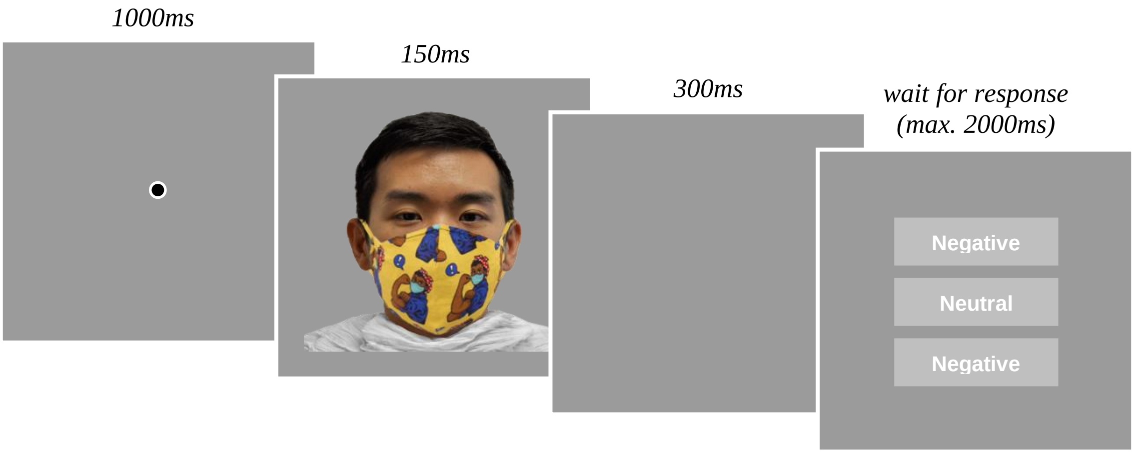

Figure 2. Visual representation of each trial. After focusing on the fixation dot (1000ms), a single stimulus was shown (150ms). Following an interval of 300ms, participants were subsequently given a maximum of 2000ms to answer which emotion valence they had identified in the stimulus.

(similarly to Grahlow et al., 2022). Also, to prevent recognition of certain models' expressions, each model-expression combination was shown only once; either as 'normal' photo, with face mask or as upper half of the face. The remaining stimulus set was balanced for these stimuli conditions and was doubled by showing each stimulus twice, in order to create a satisfactory stimulus set size. Responses were subsequently coded as Boolean integers (0 or 1) for incorrect and correct responses, respectively.

Analyses

Analyses were performed using either the integrated *t*-test function in Matlab or, when more in-depth testing was necessary, SPSS 27 (IBM SPSS Statistics). For all statistical analyses of emotion recognition accuracy, the NaN values of too slow responses were excluded from further analysis. Analyses regarding response time were only performed on reaction times of correct answers. Reaction times of under 100ms were also excluded from further analysis to account for conjecturing.

Hypothesis testing was mostly done through across-subject comparison of participants' average emotion recognition accuracy. Coded responses (0 or 1) were separated according the conditions within a relevant measure (e.g. face mask, half-face and whole-face stimuli for stimuli conditions or Asian, black, Hispanic and white for model races), after which the mean was taken for every condition within this measure to calculate within-subject accuracies. The accuracies of relevant condition pairs were subsequently compared at a group level through various Paired-Samples *t*-tests. Response time (in seconds) was similarly compared amongst each other through the use of Paired-Samples *t*-tests, if not specifically mentioned otherwise.

SPSS 27 was mostly used for performing a Repeated Measures ANOVA. To do so, data was categorized for the relevant measure in Matlab and exported to a compatible CSV file for SPSS. For all the performed ANOVAs where the assumption of sphericity was violated, Greenhouse-Geisser adjusted values were used.

Results

Face mask vs. no mask

First, the influence of face masks (compared to unconcealed faces) on the ability to distinguish negative and positive emotions from one another (H1a) was tested. A Paired-Samples *t*-test showed the share of negative expressions that were misinterpreted as being positive and vice versa was significantly larger with face mask stimuli (M = .061, SD = .030) than with stimuli showing the whole face (M = .018, SD = .018), $t(24) = 8.17$, $p < .001$. On average, positive and negative expressions were found to be confused almost three times more often with one another with face mask stimuli (M = 5.84, SD = 2.88) than with whole-face stimuli (M = 1.72, SD = 1.70). Participants also needed a longer response time with mask stimuli to correctly identify positive (mask: M = .41, SD = .11; whole face: M = .32, SD = .11), $t(24) = 6.05$, $p < .001$, and negative emotions (mask: M = .42, SD = .12; whole face: M = .35, SD = .11), $t(24) = 4.83$, $p < .001$. Overall, these analyses show that distinguishing positive and negative emotions from one another is more challenging when a mask is worn – therefore disproving H1a.

The ability to perceive the presence of an emotion in the facial expression (H1b) was investigated through a similar *t*-test. In this test, participants' error rate of perceiving positive or negative expression as being neutral – and vice versa – was compared between face mask stimuli and whole-face stimuli. It appeared that the average proportion of errors increased considerably when the stimuli contained a face mask (M = .17, SD = .055), compared to those showing the whole face (M = .11, SD = .051), $t(24) = 3.54$, $p = .001$. A subsequent *t*-test showed that the number of errors was nearly doubled with face mask stimuli (M = 26.2, SD = 8.02) compared to whole-face stimuli (M = 13.2, SD = 5.7). On top of that, participants required an overall longer response time to accurately identify any emotion with face mask stimuli (mask: M = .41, SD = .11; whole face: M = .34, SD = .10), $t(24) = 9.64$, $p < .001$. These results highlight that a face mask obstructs the ease of identifying the presence of emotion in a facial expression, regardless of it having a positive or negative valence – therefore supporting H1b.

All together, these results show a large increase in emotion recognition difficulty caused by a face mask. That is, the presence of a face mask in the stimuli doubled or tripled the number of errors made with identifying the presence or classifying the valence of an emotion in a facial expression, respectively.

Face mask vs. half face

Furthermore, participants' emotion recognition accuracy did not significantly differ between half-face stimuli (*M* = .76, *SD* = .050) and face mask stimuli, *t*(24) = .98, *p* = .34. Additional analyses did not find positive and negative emotions to be confused more often with one another in either of these conditions (mask: *M* = 5.8, *SD* = 2.9; half face: *M* = 5.9, *SD* = 3.1; *t*(24) = -.074, *p* = .94), nor did either condition appear to differ in participants' accuracy in recognizing a lack of emotion (mask: *M* =.86, *SD* = .074; half face: *M* = .84, *SD* = .087; *t*(24) = 1.20, *p* = .24). Response time was also found to not significantly differ between face mask (*M* = .41, *SD* = .11) and half-face stimuli (*M* = .43, *SD* = .12), *t*(24) = -1.87, *p* = .074. A final Repeated Measures ANOVA showed no main effect of stimuli type on how reaction time differed between the different emotion valences, *F*(1, 24) = 3.21, *p* = .086. Overall, in no way did dismissing the lower half of a face appear to disrupt emotion recognition more than adding a face mask did – therefore disproving H2.

Face mask vs. no mask: Influence of race similarity

For the investigation of the role of race similarity in face masks' relation with emotion interpretation (H3), participants' emotion recognition accuracies in the different conditions were separated into that of same-race and other-race models. Regarding whole-face stimuli, no effect of race similarity was found. First, emotion recognition accuracy with same-race faces (*M* = .90, *SD* = .076) did appear to not significantly differ with participants' accuracy with other-race faces (*M* = .89, *SD* = .046), *t*(24) = -.59, *p* = .56. Also, no difference in response time was found between same-race faces (*M* = .37, *SD* = .11) and other-race faces (*M* = .35, *SD* = .10), *t*(24) = -1.98, *p* = .059. For face mask stimuli, nonetheless, accuracy appeared to be significantly higher with same-race faces (*M* = .80, *SD* = .096) than with other-race faces (*M* = .77, *SD* = .059), *t*(24) = -2.08, *p* = .048. Response time, on the other hand, did also not appear to differ significantly between same-race and other-race faces with face mask stimuli (same race: *M* = .42, *SD* = .12; other race: *M* = .41, *SD* = .11), *t*(24) = .91, *p* = .37.

A more in-depth Paired-Samples *t*-test was then performed, comparing the average accuracy decrease of same-race face mask stimuli (as apposed to whole-face stimuli) with the same decrease of other-race models wearing a mask. Despite the contradicting results in the earlier analyses, this test showed the difference in emotion recognition accuracy of mask and whole-face stimuli to not significantly differ between other-race faces (*M* = -.12, *SD* = .074) and same-race faces (*M* = -.10, *SD* = .12), *t*(24) = 1.17, *p* = .25. A subsequent Repeated-Measures ANOVA confirmed this finding by showing an insignificant interaction effect of condition (face mask or whole-face stimuli) and race similarity (same-race or other-race faces) on emotion recognition accuracy, *F*(1, 24) = .0040, *p* = .25, η_p^2 = .054. That is to say, the influence of face masks on accuracy was not strengthened or weakened by whether it concerned same-race or other-race faces (with relation to the observer) – therefore disproving H3.

Half face vs. no mask: Influence of race similarity

Besides the analysis of the hypotheses, similar analyses as above were performed for the investigation of race similarity effects with the half-face stimuli. As other-race faces are more often processed by individual facial features, disrupting holistic processing is proposed to have less effect on these faces than on same-race faces (Cassidy et al., 2011; Michel et al., 2008). The lack of difference in face masks' effect between same-race and other-race faces therefore suggests that holistic processing is not disrupted by the presence of a mask. However, the analysis into the effect of race similarity with half-face stimuli also showed no differences in half-face stimuli's effect on emotion recognition accuracy between same-race faces and other-race faces – even though holistic processing is deliberately disrupted with these stimuli. As opposed to whole-face stimuli, the emotion recognition accuracy decrease caused by half-face stimuli did not appear to differ between same-race faces (*M* = -.11, *SD* = .11) and other-race faces (*M* = -.14, *SD* = .063), *t*(24) = 1.72, *p* = .099. Similarly, no interaction effect was found of condition (half-face or whole-face stimuli) and race similarity on emotion recognition accuracy, *F*(1, 24) = 2.95, *p* = .099, η_p^2 = .11.

To confirm, the role of race similarity in the accuracy difference between whole-face and face mask stimuli was compared with that same role in the accuracy difference between whole-face and half-face stimuli. A Paired-Samples *t*-test compared the difference between accuracy decrease with other-race faces and accuracy decrease with same-race faces for face mask stimuli on one hand (*M* = -.025, *SD* = .11) and that same difference in accuracy decrease for half-face stimuli on the other (*M* = -.032, *SD* = .094). This test showed these race similarity effects not to be statistically different from one another, *t*(24) = .39, *p* = .70. A Repeated Measures ANOVA again showed no interaction effect of condition (face mask or half-face stimuli) and race similarity (same-race or other-race faces) on emotion recognition accuracy, F(1, 24) = .0080, *p* = .93, $\eta_p^2 < .001$.

In other words, the influence of disrupted holistic processing (by means of half-face stimuli) on accuracy was also not strengthened or weakened by race similarity – as was suggested in literature. Any conclusions concerning facial processing of mask-wearing faces can therefore not be drawn based on the results belonging to the investigation of H3.

Model race: Individual investigation of different races

Another additional analysis was done to look further into the effect of model race on participants' ability to interpret emotional facial expressions. Since race similarity did only appear to influence emotion recognition with face mask stimuli, more tests were performed to analyze this effect more in-depth. In particular, which race couples (observer race and model race) are more compatible than others. However, as the gathered data only contained one participant identifying with another race than white (Asian), merely data of participants identifying as being white was considered. Therefore, rather than investigating effects of race similarity, the face masks' effect on emotion recognition are measured across the different races to determine prospective deviations. A Repeated Measures ANOVA revealed a significant interaction effect of race and stimulus condition (whole-face, half-face and face mask stimuli) on emotion recognition accuracy, $F(6, 138) = 2.22$, $p = .045$. Considering the lack of interaction between overall race similarity and stimuli condition, the latter interaction effect implies that obstructing the face (showing only the upper half or partially concealing the face with a face mask) has more effect on white observers' emotion recognition accuracy for some races than for others – which is confirmed by Figure 3.

A series of simple effects tests were subsequently performed in the form of Paired-Samples *t*-tests comparing same-race (white) face accuracy with participants' accuracy with each of the other races individually (Asian, black and Hispanic). The findings of these tests supported the earlier found interaction effect and revealed how that effect was mainly caused by Asian face stimuli. Multiple Paired-Samples *t*-tests were performed to compare the accuracy decreases caused by face mask stimuli (compared to whole-face stimuli) for the different races. Emotion recognition appeared to decrease significantly more for Asian faces ($M = -.16$, $SD = .11$) than for white faces ($M = -.10$, $SD = .11$), $t(23) = 2.89$, $p = .008$, whereas no difference in accuracy decrease was found between white and black faces ($M = -.11$, $SD = .11$), $t(23) = .37$, $p = .71$, or white and Hispanic faces ($M = -.12$, $SD = .97$), $t(23) = .68$, $p = .50$. Confirmatory, when the trials of Asian faces were excluded from the *t*-test, participants' emotion recognition with face mask stimuli did not appear to significantly differ between same-race faces ($M = .81$, $SD = .097$) and other-race faces ($M = .78$, $SD = .068$), $t(23) = 1.560$, $p = .13$. Overall, these results show that the earlier found differences in emotion recognition accuracy between same-race and other-race faces is not due to race (dis)similarity, but is explained by the strong effect that a face mask has on white participants' accuracy when it is worn by an Asian face.

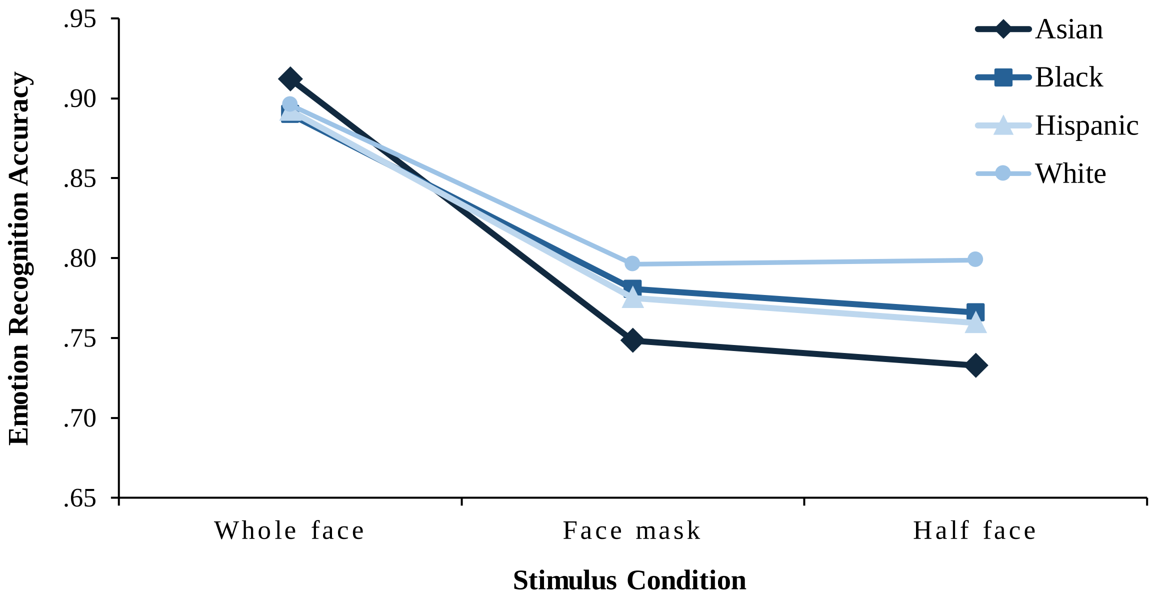

Figure 3. White participants' average emotion recognition accuracy with each stimuli type (face mask, half-face or whole-face stimuli), shown for all four races represented in the stimulus set. The intersection of lines indicates the accuracy decrease due to face obstruction is larger with Asian faces than with Black, Hispanic or White faces – proving an interaction effect of stimulus condition and race.

Discussion

The purpose of this study was to examine and quantify the influence of face masks on the ability to interpret non-verbal emotional communication, specifically in a doctor-patient context. This influence was investigated through a within-subject adaptation of the VERT-K (Derntl et al., 2008), during which participants' ability to distinguish different emotion valences from each other was tested with three types of stimuli. Recognition accuracy and response time was measured for unobstructed faces, faces covered by a mask and faces of which only the upper face half was shown. The effect of the presence of a mask was subsequently compared along a scale with unobstructed faces as the ideal and only the upper face half being available as the lower limit. In contrary to earlier studies on this topic, the stimuli used in this study contained face masks varying in type (i.e. 'standard' medical, fabric/reusable or FFP2) and appearance and included faces of various races, to draw more generalizable conclusions. Stimulus material was retrieved from the RADIATE database, which contains faces of Asian, black, Hispanic and white models.

Despite the alterations of this experiment compared to previous research, results showed a similar negative influence of face masks on emotion recognition as was found in these earlier studies (e.g. Bani et al., 2021; Carbon, 2020; Ziccardi et al., 2022). That is to say, recognition of the emotion valence within a facial expression was done worse and slower when the model displaying the emotion wore a mask instead of their entire face being visible. To gain a better understanding of the abovementioned effect, the effect of face masks was investigated for each emotion valence individually. The negative effect of face masks on emotion recognition noted in previous research was often due to negative emotions being confused with other negative emotions, rather than with positive emotions (Carbon, 2020; Grahlow et al., 2022). It was therefore reasoned that face masks would only cause difficulty in identifying the presence of emotion and not in distinguishing different emotion valences. It appeared, however, that wearing a mask both obstructs the ability to distinguish emotion valences and recognize the presence of emotion. For both parts of emotion recognition, face masks respectively induced three and two times the number of misinterpretations as against faces that were unconcealed. This negative influence turned out to be so strong, that emotion recognition was equally as hindered as when the lower half of the face is completely removed from consideration.

This strong negative influence of face masks could mean two things. First, there is a possibility that this intensity is a result of the already hindered facial processing attributable to partially covering the face, and the distraction caused by the presence of a mask on top of that. Several participants noted such a distracting effect with the face mask stimuli. If attention is first drawn towards the mask before perceiving the facial expression, this could complicate intuitive encoding of the emotion in the already short exposure time. Especially in this experiment, where not only plain blue masks were used like in earlier studies (e.g. Bani et al., 2021; Carbon et al., 2020), alternating different-looking masks may have caused such undesirable distraction. On the other hand, these findings could suggest that wearing a face mask is simply processed the same as having half a face. That is to say, the bottom half of the face is completely discarded during facial expression processing of a face wearing a mask. Holistic processing would therefore be disrupted, causing the expression to be perceived through individual feature processing (Tanaka et al., 2012) – as would occur with half faces. Follow-up research is needed, however, to draw conclusions more concretely about these two possibilities.

Further analysis into the effects of face masks considered the role of race in this effect and, more specifically, how emotion recognition with face masks varies between faces of the same race as the observer and differently raced faces. It appeared that the negative effect of face masks on emotion recognition was not affected when the faces were of a different race than the observers. As previous research claims disrupted holistic processing has less effect on emotion recognition with other-race faces than with same-race faces (Cassidy et al., 2011; Rodger et al., 2010), these findings would suggest that holistic processing is not disrupted by a mask. That being said, race dissimilarity was neither found to nuance the effect of deliberately disrupting holistic processing by not showing the lower face half. As emotion recognition was thus comparable for same- and other-race faces in both an unaffected holistic processing and a disrupted holistic processing setting, the abovementioned results do not lend themselves to drawing conclusions on the way faces wearing a mask are processed.

All in all, no effect of overall race similarity was found to give clarification about the way facial expressions are processed when a mask is worn. When looking into the differences between same-race and each other-race individually, however, face masks appeared to be particularly harmful for white observers' emotion recognition when they were worn by Asian faces. When worn by Asian faces, the negative effect of masks on emotion recognition was found to be stronger than when worn by white (same-race) faces. Such a difference in effect was not found for either black or Hispanic faces, when compared to white faces. Similarly, not seeing the lower face half had a stronger effect on white observers' emotion recognition when it considered an Asian upper face half (compared to seeing the whole face) than when it was a white upper face half. Similar differences were again not found with black or Hispanic faces.

Possibly this is due to the premise that people of an Asian ethnicity typically communicate emotional expressions at a lower intensity than black, Hispanic or Caucasian (white) people do (Cai et al., 2022; Matsumoto, 1993). Montirosso et al. (2010) already suggested that facial expression features that distinctively indicate a certain emotion become less explicit when the expression is conveyed at a low intensity, which causes poorer recognition accuracy. In other words, lower intensity causes less information to be visible in each facial feature for observers to base their interpretation of the expression on. Especially when a part of these facial features is concealed and emotion recognition is done through individual feature processing, this could cause a large decrease in emotion recognition performance. Given that face masks conceal the lower face half, it could therefore be reasoned that the low amount of available information in upper face halves of Asian people causes masks' effect on emotion recognition to be particularly strong for Asian faces. This implies that mask-wearing faces are processed through their individual facial features, rather than in a holistic manner. Not only does this show the strength of face masks' negative effect on emotion recognition, it also suggests that training to express emotions through the individual facial features in the upper face half could help in reducing (or even overcoming) the abovementioned effect.

Implications

Overall, the findings of this study show that face masks are extremely harmful for non-verbal emotional communication. As illustrated at the beginning of this study, effective non-verbal communication of doctors towards their patient is of great importance for patient satisfaction and a good clinical outcome (Marcinowicz et al., 2010). Not only does this play a part in building a strong doctor-patient relationship, it also helps improving the provided healthcare. Having said that, the findings of this study indicate that this can be considerably obstructed when doctors choose (or are obliged) to wear a mask while communicating with their patient. Wearing a mask was found to increase the number of misinterpretations of facial expressions with two to three times. Moreover, the time needed for accurate interpretation increased by more than 20% when a mask was worn. Given the potential negative impact on the provided healthcare this could have, doctors should be careful in their decision to wear a face mask.

The results of this study also strongly suggest that the facial area covered by a mask is completely disregarded when encoding the face, disrupting holistic processing. Emotion recognition was equally hindered when a face mask was worn as when the observer only got to see the upper face half. This implies that observers base their interpretation only on the individual facial features seen in the upper area of the face. This knowledge is valuable for the understanding of facial processing, as it takes a first step in understanding what is necessary for holistic processing to be disrupted. However, most importantly, these findings imply that the hindered communication due to face masks may be nuanced by training doctors on increasing the non-verbal cues they express in the upper half of the face. As implied by the analysis of Asian faces, face masks' effect on emotion recognition is subject to the intensity with which the emotion is expressed and the information that is available in the upper face half. Given that all six basic emotions (sadness, anger, disgust, fear, happiness, surprise) have distinct features in the upper half area of the face (Wegrzyn et al., 2017), exaggerating these individual features is therefore suggested to improve emotion recognition when a mask is worn.

Limitations and recommendations for future research

This study did not go without limitations. First of all, as was expected, all participants noted experiencing a learning effect during the experiment. Considering the relatively long duration of the experiment and the large stimulus set, it was considered valuable to provide participants with feedback on their accuracy to encourage answering as accurately as possible. As a result, however, participants mentioned gradually becoming more aware of how certain emotions were expressed by the models. A Repeated Measures ANOVA showed emotion recognition accuracy to increase over the course of the experiment, $F(8, 192) = 7.21$, $p < .001$, proving this learning effect. Presumably, this entailed participants increasingly basing their answers on recognition (which facial features correspond to which emotion valence) instead of their genuine interpretation of the facial expression. Be that as it may, since the sequence in which stimuli were presented was (semi-)randomized for each participant and each model-expression combination was only presented twice throughout the entire experiment, any learning effect should not be able to have a noticeable influence on possible patterns observed in the data. It is, nonetheless, suggested for future research to further investigate how the negative effect of a face mask on emotion recognition alters over time – due to learnability of facial expressions. Getting a better understanding of how (quickly) emotion recognition with mask-wearing faces improves with the correct feedback can have important implications not only for the healthcare sector, but for any sector that makes use of face masks frequently.

Secondly, the ecological validity of photoshopping the face masks onto the stimuli is debatable. Due to the global COVID-19 pandemic, people were forced to get accustomed to communicating while wearing face masks. Results of this study show that this has not resolved the hindering effect of face masks on the interpretation of the observer, but it has possibly changed how people wearing the mask express their emotions. That is to say, they learned to dramatize their facial expression to overcome the restrains of a mask and effectively convey that expression. Photoshopping a mask onto an originally unconcealed face does not, however, account for this possibility. It is therefore necessary for a follow-up study on this topic to compose a new stimulus set by making photos of people already wearing a face mask. Using a plain mask is recommended (as this is better editable) to provide the possibility to edit and control for mask appearance.

In addition, the sample used in this experiment contained a non-normally distributed age, with a median age of 23. In an earlier study on this topic, Grundmann et al. (2021) found emotion recognition of old observers to suffer significantly more from face masks than that of younger observers. To make the results of this study even more generalizable, it is therefore recommended to do follow-up research with a similar approach, but which also controls for any discrepancies between emotion recognition of participants of various ages.

Concerning the investigation of race similarity, only one participant identified with a different race than white (Caucasian). Although this did not have any consequence for the investigation of race similarity's general effect, deeper analysis into observer-expressor race couples was only done with data of these white participants. Conclusions following this analysis are thus merely based on the data of white subjects and no conclusions on general race similarity effects can be drawn. That is, the found results may be due to visual differences between the models instead of actual ingroup/outgroup effects. Whereas this does give useful insights into the effect of face masks on non-verbal communication and has important implications for doctors of various races, it is recommended for a following study to acquire a sample with a variation of ethnicities. By doing so, other observer-expressor race couples can be examined and more generalizable conclusions can be drawn on both face masks and their interaction with race similarity.

Despite these limitations, the results of this study stress the great extent to which wearing a face mask obstructs non-verbal emotional communication. Especially in the healthcare sector, where it is crucial for this form of communication to be performed effectively, this has severe implications. Whereas preceding literature on face masks was mostly dedicated to its protective value by means of reducing disease transmission, the findings of this study help in weighing up the benefits and drawbacks of wearing a mask. Future research on face masks should therefore also focus on expanding this balance of pros and cons for doctors to be able to make an even better informed decision on whether to wear a mask. For now, however, doctors should be careful with their decision to wear a mask during the consultations with their patients and avoid its use whenever possible.

Ethics approval

Before starting the experiment, ethical approval was obtained from the Ethical Review Board (FERB) of the Faculty of Social and Behavioral Sciences of Utrecht University (Ref. no.: 22-1536).

References

Abbas, Z. A., & Duchaine, B. (2008). The role of holistic processing in judgments of facial attractiveness. Perception, 37(8), 1187–1196. https://doi.org/10.1068/p5984

Ammon, A., Reichart, P. A., Pauli, G., & Petersen, L. R. (2000). Hepatitis B and C among Berlin dental personnel: Incidence, risk factors, and effectiveness of barrier prevention measures. Epidemiology and Infection, 125(2), 407–413. https://doi.org/10.1017/s0950268899004537

Bani, M., Russo, S., Ardenghi, S., Rampoldi, G., Wickline, V., Nowicki, S., & Strepparava, M. G. (2021). Behind the mask: Emotion recognition in healthcare students. Medical Science Educator, 31(4), 1273–1277. https://doi.org/10.1007/s40670-021-01317-8

Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. Journal of Personality and Social Psychology, 37(11), 2049–2058. https://doi.org/10.1037/0022-3514.37.11.2049

Baudouin, J. Y., & Humphreys, G. W. (2006). Configural information in gender categorisation. Perception, 35(4), 531–540. https://doi.org/10.1068/p3403

Bhat, B. V., & Kingsley, M. K. (2020). Effective non-verbal communication. In S. C. Parija & B. V. Adkoli (Eds.), Effective Medical Communication: The A, B,C, D, E of it (pp. 39–47). Springer. https://doi.org/10.1007/978-981-15-3409-6_4

Bombari, D., Schmid, P. C., Schmid Mast, M., Birri, S., Mast, F. W., & Lobmaier, J. S. (2013). Emotion recognition: The role of featural and configural face information. Quarterly Journal of Experimental Psychology, 66(12), 2426–2442. https://doi.org/10.1080/17470218.2013.789065

Botvinick, M., Jha, A. P., Bylsma, L. M., Fabian, S. A., Solomon, P. E., & Prkachin, K. M. (2005). Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. NeuroImage, 25(1), 312–319. https://doi.org/10.1016/j.neuroimage.2004.11.043

Buckman, R. (2002). Communications and emotions: Skills and effort are key. BMJ, 325(7366), 672. https://doi.org/10.1136/bmj.325.7366.672

Cai, J., Meng, Z., Khan, A. S., Li, Z., O’Reilly, J., & Tong, Y. (2022). Probabilistic attribute tree structured convolutional neural networks for facial expression recognition in the wild. IEEE Transactions on Affective Computing, 13, 1–14. https://doi.org/10.1109/taffc.2022.3156920

Calder, A. J., & Jansen, J. (2005). Configural coding of facial expressions: The impact of inversion and photographic negative. Visual Cognition, 12(3), 495–518. https://doi.org/10.1080/13506280444000418

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. Journal of Experimental Psychology: Human Perception and Performance, 26(2), 527–551. https://doi.org/10.1037/0096-1523.26.2.527

Carbon, C. C. (2020). Wearing face masks strongly confuses counterparts in reading emotions. Frontiers in Psychology, 11, 566886. https://doi.org/10.3389/fpsyg.2020.566886

Cassidy, K. D., Quinn, K. A., & Humphreys, G. W. (2011). The influence of ingroup/outgroup categorization on same- and other-race face processing: The moderating role of inter- versus intra-racial context. Journal of Experimental Social Psychology, 47(4), 811–817. https://doi.org/10.1016/j.jesp.2011.02.017

Chellamani, K. P., Veerasubramanian, D., & Vignesh Balaji, R. S. (2013). Surgical face masks: Manufacturing methods and classification. Journal of Academia and Industrial Research, 2(6), 320–324.

Chen, C., & Jack, R. E. (2017). Discovering cultural differences (and similarities) in facial expressions of emotion. Current Opinion in Psychology, 17, 61–66. https://doi.org/10.1016/j.copsyc.2017.06.010

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates.

Cohn, J. F., Ambadar, Z., & Ekman, P. (2007). Observer-based measurement of facial expression with the Facial Action Coding System. In J. A. Coan & J. J. B. Allen (Eds.), Handbook of Emotion Elicitation and Assessment (pp. 203–221). Oxford University Press.

Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. Perception, 29(8), 893–909. https://doi.org/10.1068/p2949

Conley, M. I., Dellarco, D. V., Rubien-Thomas, E., Cohen, A. O., Cervera, A., Tottenham, N., & Casey, B. (2018). The racially diverse affective expression (RADIATE) face stimulus set. Psychiatry Research, 270, 1059–1067. https://doi.org/10.1016/j.psychres.2018.04.066

Cowling, B. J., Zhou, Y., Ip, D. K. M., Leung, G. M., & Aiello, A. E. (2010). Face masks to prevent transmission of influenza virus: A systematic review. Epidemiology and Infection, 138(4), 449–456. https://doi.org/10.1017/s0950268809991658

Derntl, B., Kryspin-Exner, I., Fernbach, E., Moser, E., & Habel, U. (2008). Emotion recognition accuracy in healthy young females is associated with cycle phase. Hormones and Behavior, 53(1), 90–95. https://doi.org/10.1016/j.yhbeh.2007.09.006

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

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES — A database of facial expressions in young, middle-aged, and older women and men: Development and validation. Behavior Research Methods, 42(1), 351–362. https://doi.org/10.3758/brm.42.1.351

Eimer, M., & Holmes, A. (2002). An ERP study on the time course of emotional face processing. Neuroreport, 13(4), 427–431. https://doi.org/10.1097/00001756-200203250-00013

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. Journal of Personality and Social Psychology, 17(2), 124–129. https://doi.org/10.1037/h0030377

Elfenbein, H. A., & Ambady, N. (2002). Is there an in-group advantage in emotion recognition? Psychological Bulletin, 128(2), 243–249. https://doi.org/10.1037/0033-2909.128.2.243

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

Grahlow, M., Rupp, C. I., & Derntl, B. (2022). The impact of face masks on emotion recognition performance and perception of threat. PLOS ONE, 17(2), e0262840. https://doi.org/10.1371/journal.pone.0262840

Grundmann, F., Epstude, K., & Scheibe, S. (2021). Face masks reduce emotion-recognition accuracy and perceived closeness. PLoS ONE, 16(4), e0249792. https://doi.org/10.1371/journal.pone.0249792

Kito, T., & Lee, B. (2004). Interpersonal perception in Japanese and British observers. Perception, 33(8), 957–974. https://doi.org/10.1068/p3471

Lee, K. U., Khang, H. S., Kim, K. T., Kim, Y. J., Kweon, Y. S., Shin, Y. W., Kwon, J. S., Ho, S. H., Garfinkel, S. N., Chae, J. H., & Liberzon, I. (2008). Distinct processing of facial emotion of own-race versus other-race. NeuroReport, 19(10), 1021–1025. https://doi.org/10.1097/wnr.0b013e3283052df2

Leppänen, J. M., Kauppinen, P., Peltola, M. J., & Hietanen, J. K. (2007). Differential electrocortical responses to increasing intensities of fearful and happy emotional expressions. Brain Research, 1166, 103–109. https://doi.org/10.1016/j.brainres.2007.06.060

Lim, E. C. H., Seet, R. C. S., Lee, K. H., Wilder-Smith, E. P. V., Chuah, B. Y. S., & Ong, B. K. C. (2006). Headaches and the N95 face-mask amongst healthcare providers. Acta Neurologica Scandinavica, 113(3), 199–202. https://doi.org/10.1111/j.1600-0404.2005.00560.x

Lundqvist, D., Flykt, A., & Öhman, A. (1998). Karolinska Directed Emotional Faces (KDEF) [Dataset]. APA PsycTests. https://doi.org/10.1037/t27732-000

MacIntyre, C. R., & Chughtai, A. A. (2015). Facemasks for the prevention of infection in healthcare and community settings. BMJ, 350, h694. https://doi.org/10.1136/bmj.h694

Marcinowicz, L., Konstantynowicz, J., & Godlewski, C. (2010). Patients' perceptions of GP non-verbal communication: A qualitative study. British Journal of General Practice, 60(571), 83–87. https://doi.org/10.3399/bjgp10X483111

Matsumoto, D. (1993). Ethnic differences in affect intensity, emotion judgments, display rule attitudes, and self-reported emotional expression in an American sample. Motivation and Emotion, 17(2), 107–123. https://doi.org/10.1007/bf00995188

McCrackin, S. D., Capozzi, F., Mayrand, F., & Ristic, J. (2022). Face masks impair basic emotion recognition. Social Psychology. https://doi.org/10.1027/1864-9335/a000470

McKone, E., Wan, L., Pidcock, M., Crookes, K., Reynolds, K., Dawel, A., Kidd, E., & Fiorentini, C. (2019). A critical period for faces: Other-race face recognition is improved by childhood but not adult social contact. Scientific Reports, 9(1). https://doi.org/10.1038/s41598-019-49202-0

Michel, C., Corneille, O., & Rossion, B. (2010). Holistic face encoding is modulated by perceived face race: Evidence from perceptual adaptation. Visual Cognition, 18(3), 434–455. https://doi.org/10.1080/13506280902819697

Montirosso, R., Peverelli, M., Frigerio, E., Crespi, M., & Borgatti, R. (2010). The development of dynamic facial expression recognition at different intensities in 4- to 18-year-olds. Social Development, 19(1), 71–92. https://doi.org/10.1111/j.1467-9507.2008.00527.x

Ong, L., de Haes, J., Hoos, A., & Lammes, F. (1995). Doctor-patient communication: A review of the literature. Social Science & Medicine, 40(7), 903–918. https://doi.org/10.1016/0277-9536(94)00155-m

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. Neuropsychologia, 49(5), 1226–1235. https://doi.org/10.1016/j.neuropsychologia.2011.02.021

Posamentier, M. T., & Abdi, H. (2003). Processing faces and facial expressions. Neuropsychology Review, 13(3), 113–143. https://doi.org/10.1023/a:1025519712569

Rodger, H., Kelly, D. J., Blais, C., & Caldara, R. (2010). Inverting faces does not abolish cultural diversity in eye movements. Perception, 39(11), 1491–1503. https://doi.org/10.1068/p6750

Roter, D. L., Frankel, R. M., Hall, J. A., & Sluyter, D. (2006). The expression of emotion through nonverbal behavior in medical visits. Mechanisms and outcomes. Journal of General Internal Medicine, 21(S1), S28–S34. https://doi.org/10.1111/j.1525-1497.2006.00306.x

Silverman, J., & Kinnersley, P. (2010). Doctors' non-verbal behaviour in consultations: Look at the patient before you look at the computer. British Journal of General Practice, 60(571), 76–78. https://doi.org/10.3399/bjgp10x482293

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. Psychological Science, 16(3), 184–189. https://doi.org/10.1111/j.0956-7976.2005.00801.x

Tanaka, J. W., Kaiser, M. D., Butler, S., & le Grand, R. (2012). Mixed emotions: Holistic and analytic perception of facial expressions. Cognition & Emotion, 26(6), 961–977. https://doi.org/10.1080/02699931.2011.630933

Tanaka, J. W., & Pierce, L. J. (2009). The neural plasticity of other-race face recognition. Cognitive, Affective, & Behavioral Neuroscience, 9(1), 122–131. https://doi.org/10.3758/cabn.9.1.122

Todorov, A., Loehr, V., & Oosterhof, N. N. (2010). The obligatory nature of holistic processing of faces in social judgments. Perception, 39(4), 514–532. https://doi.org/10.1068/p6501

Vogel, D., Meyer, M., & Harendza, S. (2018). Verbal and non-verbal communication skills including empathy during history taking of undergraduate medical students. BMC Medical Education, 18, 157. https://doi.org/10.1186/s12909-018-1260-9

Wagner, H. L. (1990). The spontaneous facial expression of differential positive and negative emotions. Motivation and Emotion, 14(1), 27–43. https://doi.org/10.1007/bf00995547

Weathers, M. D., Frank, E. M., & Spell, L. A. (2002). Differences in the communication of affect: Members of the same race versus members of a different race. Journal of Black Psychology, 28(1), 66–77. https://doi.org/10.1177/0095798402028001005

Wegrzyn, M., Vogt, M., Kireclioglu, B., Schneider, J., & Kissler, J. (2017). Mapping the emotional face. How individual face parts contribute to successful emotion recognition. PLoS ONE, 12(5), e0177239. https://doi.org/10.1371/journal.pone.0177239

Wong, C. K. M., Yip, B. H. K., Mercer, S., Griffiths, S., Kung, K., Wong, M. C. S., Chor, J., & Wong, S. Y. S. (2013). Effect of facemasks on empathy and relational continuity: A randomised controlled trial in primary care. BMC Family Practice, 14(200), 1–7. https://doi.org/10.1186/1471-2296-14-200

Yan, X., Andrews, T. J., Jenkins, R., & Young, A. W. (2016). Cross-cultural differences and similarities underlying other-race effects for facial identity and expression. Quarterly Journal of Experimental Psychology, 69(7), 1247–1254. https://doi.org/10.1080/17470218.2016.1146312

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. Perception, 16(6), 747–759. https://doi.org/10.1068/p160747

Ziccardi, S., Crescenzo, F., & Calabrese, M. (2022). “What is hidden behind the mask?” facial emotion recognition at the time of COVID-19 pandemic in cognitively normal multiple sclerosis patients. Diagnostics, 12(1), 47. https://doi.org/10.3390/diagnostics12010047