

Recognition of Masked Faces in the Era of the Pandemic: No Improvement Despite Extensive Natural Exposure



Erez Freud¹, Daniela Di Giammarino¹, Andreja Stajduhar¹,
R. Shayna Rosenbaum^{1,2}, Galia Avidan^{3,4}, and Tzvi Ganel³

¹Department of Psychology and the Centre for Vision Research, York University; ²Rotman Research Institute, Baycrest Health Sciences, Toronto, Canada; ³Department of Psychology, Ben-Gurion University of the Negev; and ⁴Department of Cognitive and Brain Sciences, Ben-Gurion University of the Negev

Psychological Science
2022, Vol. 33(10) 1635–1650
© The Author(s) 2022



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09567976221105459
www.psychologicalscience.org/PS


Abstract

Face masks, which became prevalent across the globe during the COVID-19 pandemic, have had a negative impact on face recognition despite the availability of critical information from uncovered face parts, especially the eyes. An outstanding question is whether face-mask effects would be attenuated following extended natural exposure. This question also pertains, more generally, to face-recognition training protocols. We used the Cambridge Face Memory Test in a cross-sectional study ($N = 1,732$ adults) at six different time points over a 20-month period, alongside a 12-month longitudinal study ($N = 208$). The results of the experiments revealed persistent deficits in recognition of masked faces and no sign of improvement across time points. Additional experiments verified that the amount of individual experience with masked faces was not correlated with the mask effect. These findings provide compelling evidence that the face-processing system does not easily adapt to visual changes in face stimuli, even following prolonged real-life exposure.

Keywords

face perception, COVID-19, inversion effect, learning and training, holistic processing, open data

Received 7/4/21; Revision accepted 4/15/22

Faces are among the most informative visual stimuli in human perception because they provide vital information for social interactions. Humans can extract significant information from brief exposure to a person's face, including their identity, gender, emotion, age, and race. Although it is clear that face-perception abilities rapidly develop from early childhood to adolescence (Pascalis et al., 2011), it is uncertain whether these abilities can improve later in life, particularly under conditions of suboptimal visual input (Pascalis et al., 2020; White, Kemp, Jenkins, Matheson, & Burton, 2014; Yovel et al., 2012). Here, in a large-scale naturalistic study, we focused our attention on the adult face-recognition system and investigated whether it could adapt and effectively handle the recognition of masked faces over a period of 20 months during the COVID-19 pandemic.

The COVID-19 pandemic introduced an unprecedented reality in which mask wearing became prevalent, and often mandatory, around the globe (Fig. 1). Previous research demonstrated that the occlusion of the lower part of the face hinders different aspects of face processing, including recognition of familiar faces (Carragher & Hancock, 2020), unfamiliar faces (Dhamecha et al., 2014; Freud et al., 2020; Gosselin & Schyns, 2001; Kret & de Gelder, 2012; Stajduhar et al., 2022), and emotional expressions (Grundmann et al., 2021). Notably, previous studies were conducted before, or soon after, face masks

Corresponding Author:

Erez Freud, York University, Department of Psychology and the Centre for Vision Research
Email: efreud@yorku.ca

became widely prevalent (i.e., first and second quarters of 2020), when participants did not have substantial experience with masked faces. Hence, an outstanding question is whether such extensive and prolonged experience with masked faces would change the way masked faces are perceived. In addition to its relevance and timeliness, this research also pertains to the plasticity of the matured human face-processing system.

Clearly, observers extract information from both the mouth and the eye regions of faces to allow successful identification (Tardif et al., 2019). Yet previous research has demonstrated that not all face features are equally important for this purpose. Behavioral (Butler et al., 2010; Royer et al., 2018; Sekuler et al., 2004), neuropsychological (Caldara et al., 2005; DeGutis et al., 2012), and electrophysiological (Bentin et al., 2006) studies have emphasized the idea that face recognition relies more heavily on upper regions of the face, especially the eye region. Such reliance could not be attributed solely to the amount of distinguishable information conveyed by the eye region (Vinette et al., 2004). Greater reliance on the eyes has been associated with improved face recognition (Royer et al., 2018), whereas reduced, atypical processing of the eye region has been associated with developmental (DeGutis et al., 2012) and acquired (Caldara et al., 2005) prosopagnosia. Given that masks do not cover the critical parts of the face that include the eyes, one might predict that the effect of medical masks on face recognition could be attenuated over time and experience with masked faces, particularly following prolonged natural exposure.

Attenuation of the effect of masks nevertheless requires adaptation of the face-processing system to

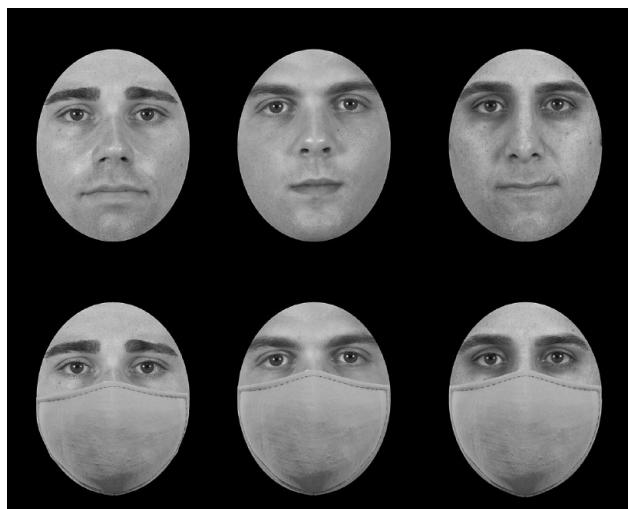


Fig. 1. Examples of faces with and without masks that are similar to the ones used in the experiment. Faces are reproduced with permission from the Chicago Face Database (Ma et al., 2015).

Statement of Relevance

Face recognition is one of the most skilled abilities in human perception. But what happens in the event of a sudden change, when faces appear dramatically different? This occurred during the COVID-19 pandemic: Faces were frequently occluded by masks, which led to deficits in our ability to recognize people. However, the ongoing nature of the pandemic provided an unprecedented opportunity to examine the malleability of the mature face-processing system; in other words, did recognition of masked faces improve? Here, we measured the effect of masks on face recognition in adults at six different time points over a 20-month period in a combined cross-sectional ($N = 1,732$) and longitudinal ($N = 208$) study. We found robust evidence for the absence of improvement in recognizing faces with masks, even when taking into account individuals' relative experience with masked faces. These findings provide compelling evidence that the face-processing system does not easily adapt to visual changes in face stimuli, even with prolonged, real-life exposure to altered faces.

visual changes (i.e., the occlusion of the mouth and nose areas). Previous research suggests that such adaptation could be challenging for the matured face-processing system (Pascalis et al., 2020). This view is supported by cross-sectional studies in groups of individuals who had unique experiences with faces and yet did not develop superior face-identification performance. For example, White and colleagues found that despite enduring, extensive experience with unfamiliar faces, border-control police officers showed poor ability to match such faces, and their performance was no different from that of naive controls (White, Kemp, Jenkins, et al., 2014). Similarly, obstetrician-gynecologist nurses performed poorly and comparably with controls on a task that required identification of newborn faces, despite their extensive experience with these faces (Yovel et al., 2012).

A more general, naturalistic example comes from studies of the other-race effect, which shares some similarities with the mask effect. In particular, humans show superior recognition performance for own-race faces compared with faces of a different race, and the processing of other-race faces is accomplished with reduced reliance on holistic processing compared with own-race faces (Michel et al., 2006; Tanaka et al., 2004; for a different view, see Wong et al., 2021). Notably,

extensive exposure to faces from other races in childhood (before age 12 years) can attenuate the other-race effect, but such attenuation is not typically observed in adults (McKone et al., 2019).

By contrast, in recent years, different groups have demonstrated some level of plasticity in the matured face-processing system. Training programs led to modest improvements in face-perception abilities, particularly when observers were asked to individuate faces and received feedback on their decisions (White, Kemp, Jenkins, & Burton, 2014; Yovel et al., 2012). Additionally, adults with congenital prosopagnosia show improved face-perception abilities following a 13-week perceptual training program. The improvement was found to persist for at least 3 months, and there is some evidence for generalization to new faces (Corrow et al., 2019; for a review, see DeGutis et al., 2014). Finally, focused training programs have been shown to be successful in reducing the other-race effect in adults (Tanaka & Pierce, 2009), but this effect is specific to perceptual processing and does not appear to extend to memory (McGugin et al., 2011). Nevertheless, many of these studies are limited in terms of the number of participants, the nonecological nature of the training regime, and, perhaps most importantly, the duration and quality of exposure.

Thus, an outstanding question is whether consistent, naturalistic exposure to masked faces is sufficient to elicit adaptation of the face-processing system in adulthood. Here, we addressed this question by evaluating the ability to perceive masked faces in the era of the COVID-19 pandemic. We used the well-established Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006; Russell et al., 2009). Since it was introduced, the CFMT has been used extensively to evaluate face-processing capabilities (e.g., Bobak et al., 2016; DeGutis et al., 2013; Wilmer et al., 2010), to identify individuals with superior face-processing abilities (Russell et al., 2009), and to identify individuals who suffer from prosopagnosia (Avidan et al., 2011; Duchaine & Nakayama, 2006). Other studies have demonstrated that the CFMT has high reliability (Bowles et al., 2009) and is correlated with other face-perception tests (DeGutis et al., 2013; Stacchi et al., 2020), with self-rated face-recognition ability (Bowles et al., 2009), and, importantly, with naturalistic assessments of face-perception abilities (Balas & Saville, 2017). Thus, the CFMT is a well-suited tool for estimating face-processing abilities in the era of the pandemic.

We first characterized face-recognition abilities for masked and nonmasked faces at a critical time point, just when mask wearing became highly prevalent around the globe (Freud et al., 2020; May 2020), and continued to track performance in a combined cross-sectional and longitudinal design over a period of 20 months. Each

participant completed the CFMT with upright and inverted faces to measure possible qualitative changes in the processing of masked faces.

Method

Participants

A total of 1,768 participants were tested at six different time points: May 2020 (accuracy data were previously reported by Freud et al., 2020), September 2020, January 2021, May 2021, September 2021, and January 2022. All participants were recruited online (<https://www.prolific.co/>) and were compensated for their time (~\$6 CAD for 30 min). Twenty-eight participants (1.5%) with an average reaction time (RT) greater than 10 s were removed from the analysis. Because of a technical error, participants in September 2021 were allowed to complete several sessions of the experiment. We identified all participants who completed the experiment more than once and discarded their data from the repeated sessions. The large sample size at each time point was used to ensure sufficient statistical power to identify any changes in face-perception abilities across time points. Table 1 includes a summary of the participants' information at each time point and across the control experiments. Figure 2 shows the age distribution across time points as well as the number of participants per country of residence. Notably, demographic variables were largely similar across time points. Note that for one time point (September 2021), there was a large group of participants from South Africa (relative to other time points). To account for this discrepancy, we have repeated the analyses reported below, excluding these participants. Similar results were obtained.

We recruited two additional groups of participants ($n = 300$, September 2021; $n = 283$, October 2021), who completed control experiments that were designed to explore the relationship between reported experience with masks and ability to recognize masked faces. Finally, in January 2022, we reached out to participants who had completed the experiment in January 2021 and asked them to retake the experiment (longitudinal testing). A total of 209 participants completed the retest session.

All experiments were performed according to relevant guidelines and regulations of the ethics review board at York University. All participants provided informed consent. Data and analysis code for all experiments are available on OSF (<https://osf.io/tq92h/>) under CC-By Attribution 4.0 International license.

Materials

The extended version of the CFMT (Duchaine & Nakayama, 2006; Russell et al., 2009) was used to assess

Table 1. Demographic Details Across Time Points and Conditions

Time point and condition	Men (<i>n</i>)	Women (<i>n</i>)	Age (in years)
May 2020			
Masked	70	76	25.7 (7.7)
Nonmasked	72	75	25.4 (7.2)
September 2020			
Masked	67	77	26.4 (8.8)
January 2021			
Masked	120	128	25.4 (7.6)
Nonmasked	123	124	27.1 (9.7)
May 2021			
Masked	74	78	24.8 (6.8)
Nonmasked	73	76	26.0 (8.4)
September 2021			
Masked	62	65	27.4 (10.0)
Nonmasked	38	71	26.7 (7.4)
September 2021 (within-subjects CFMT)	150	150	26.3 (7.8)
October 2021 (GFMT: masked version)			
Masked	141	142	26.3 (7.8)
January 2022 (longitudinal testing)			
Masked	49	53	27.5 (8.5)
Nonmasked	55	52	29.3 (10.1)
January 2022			
Masked	70	78	27 (8.8)
Nonmasked	71	76	26.9 (8.5)

Note: In the age column, values shown are means (standard deviations are given in parentheses). CFMT = Cambridge Face Memory Test; GFMT = Glasgow Face Matching Test.

face-perception abilities. The standard CFMT includes three phases (total of 72 trials) with increasing levels of difficulty. The first phase (easy) involves learning to recognize six unfamiliar male faces from three different viewpoints and then testing recognition of these faces in a three-alternative forced-choice task. The second phase involves a refresher in which the six faces are presented simultaneously from one (frontal) viewpoint followed by testing from novel viewpoints and lighting conditions. The third (difficult) phase is similar to the second phase but includes test images with added visual noise. All faces were of Caucasian men with neutral facial expression. Images were cropped to include only the internal features of the face. The long form of the CFMT includes an additional 30 trials with an even higher level of difficulty in which novel images varying in pose, emotional expression, and amount of information available are presented. This latter part is typically used to identify super-recognizers (individuals with extraordinarily high face-recognition memory; Russell et al., 2009), whereas the first three phases are more

sensitive to detecting basic performance and potential deficits in face-perception abilities (Murray & Bate, 2020). For the current study, we limited the analyses to the standard form of the CFMT (Levels 1–3).

For the last two time points (September 2021, January 2022) and for the control experiments, we included an experience with masked faces (EMF) questionnaire (Table 2). The questionnaire consists of four scales: (a) level of experience that each participant had with masked faces (Experience scale; Items 1, 2, 7, and 8), (b) extent to which regulations about mask wearing were enforced in the participant's country of residence (Regulation scale; Items 3 and 4), (c) subjective difficulty in recognizing masked faces (Subjective scale; Items 5 and 6), and (d) continuous evaluation of experience with masked faces (Item 9).

To obtain a more objective measurement of presumed experience with masked faces, we also asked participants to indicate whether they work in person, work from home (remotely), or are unemployed. Participants who were tested in January 2022 (Table 1)

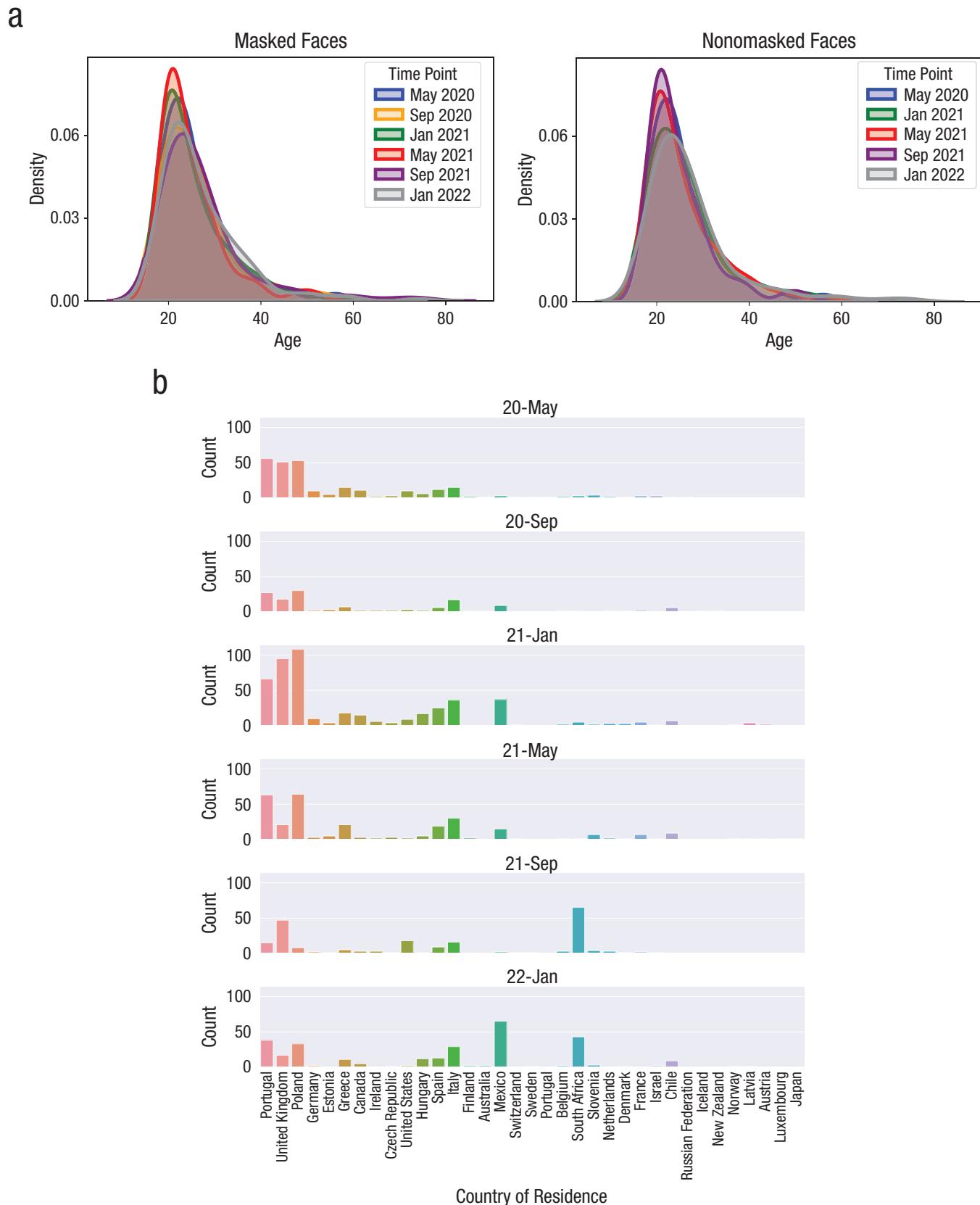


Fig. 2. Demographic properties of participants. The demographic properties were similar across the four testing time points. (a) Density histogram of participants' age. (b) Distribution of participants' country of residence.

Table 2. Experience With Masked Faces (EMF) Questionnaire (List of Items)

Please rate the degree to which you agree or disagree with the following eight statements: 1 (*strongly disagree*) to 7 (*strongly agree*)

1. I have a lot of experience encountering people who wear masks.
2. I spend a lot of time in public places where mask wearing is common.
3. Mask wearing is strongly enforced in my country of residence.
4. Most people follow mask mandates in my country.
5. I have difficulties recognizing familiar people (i.e., family members, colleagues, friends) when they are wearing masks.
6. When people wear masks, I have difficulties deciding whether they are familiar to me or not.
7. Most people who I encounter everyday are wearing masks.
8. In my daily life, I need to recognize a lot of people who wear masks.
9. What is the proportion of people you know who you encounter wearing a mask (0%–100%)?

also completed the 20-item prosopagnosia index (PI20) to provide another validated measure of subjective face-perception abilities (Shah, Sowden, et al., 2015).

Lastly, to test for a possible effect of mask exposure on face perception, in October 2021, we also collected data using a different paradigm in which external face cues were available. In particular, we used the short version of the Glasgow Face Matching Test (GFMT; Burton et al., 2010). In this test, participants are shown 40 pairs of faces, photographed in full-face view but with different cameras, and are asked to make same/different judgments. Masks were graphically added to all faces.

Procedure

For the main experiment, each participant was randomly assigned to one of two groups. The first group completed the original CFMT (faces without masks), whereas the second group completed a modified version of the CFMT in which an identical face mask was added to all faces (Fig. 1). To explore participants' style of processing faces with and without masks, we had each participant complete the test twice, once with upright faces and once with inverted faces. Block order (upright or inverted) was counterbalanced between participants. For the September 2020 testing point, only one order was employed (upright faces followed by inverted faces) because of a technical error. Accuracy scores (0–72) for upright and inverted faces were computed and served as the main dependent variable. We

also analyzed RTs for correct trials. For the control experiments, participants completed the masked and nonmasked versions of the CFMT (only the upright version) or the masked version of the GFMT. Statistical analyses were conducted using JASP (JASP Team, 2020) and in-house codes written in Python.

Results

Cross-sectional analysis

We explored the extent to which persistent, natural exposure to masked faces facilitated recognition abilities of these faces. To this end, we asked participants to complete the CFMT (Duchaine & Nakayama, 2006) with upright and inverted faces (within-subjects). The faces were either masked or nonmasked (between-subjects) across six different time points (between-subjects) over a period of 20 months, during which face-mask wearing became prevalent on a daily basis in the era of the pandemic.

Figure 3 shows the group averages across conditions on the standard CFMT for each of the six time points. The results show a persistent deficit in the ability to recognize masked faces. In addition, for all time points, the face-inversion effect was reduced for masked faces, pointing to a qualitative difference in the processing of these faces, which again persisted over time.

To statistically validate these results, we first employed an analysis of variance (ANOVA) with time points, orientation, gender, and group (masked, nonmasked). Note that for this analysis, we included only five time points because we did not have data for the nonmasked-faces condition for September 2020. The ANOVA demonstrated main effects of group (nonmasked > masked), $F(1, 1570) = 214, p < .001, \eta_p^2 = .12$, orientation (upright > inverted), $F(1, 1570) = 2846, p < .001, \eta_p^2 = .64$, and gender (female > male), $F(1, 1570) = 66.4, p < .001, \eta_p^2 = .04$. An interaction was found between orientation and group, with a reduced inversion effect for masked faces, $F(1, 1570) = 232.2, p < .001, \eta_p^2 = .129$, that was consistent across time points (three-way interaction: $F < 1$). The reduced inversion effect could not be easily accounted for by a floor effect for the inverted masked faces because the average CFMT score was well above chance level (33%, 24 points) even for these faces. Importantly, although we found a small effect for time point, $F(4, 1570) = 3.065, p = .016, \eta_p^2 = .008$ (with a higher average for the May 2021 and January 2022 samples), there was no interaction between time point and group, $F(4, 1570) = 1.095, p = .357, \eta_p^2 = .003$. Similar results were observed when participants' age served as a covariate. These results provide evidence for a persistent decrement in face-perception abilities across time

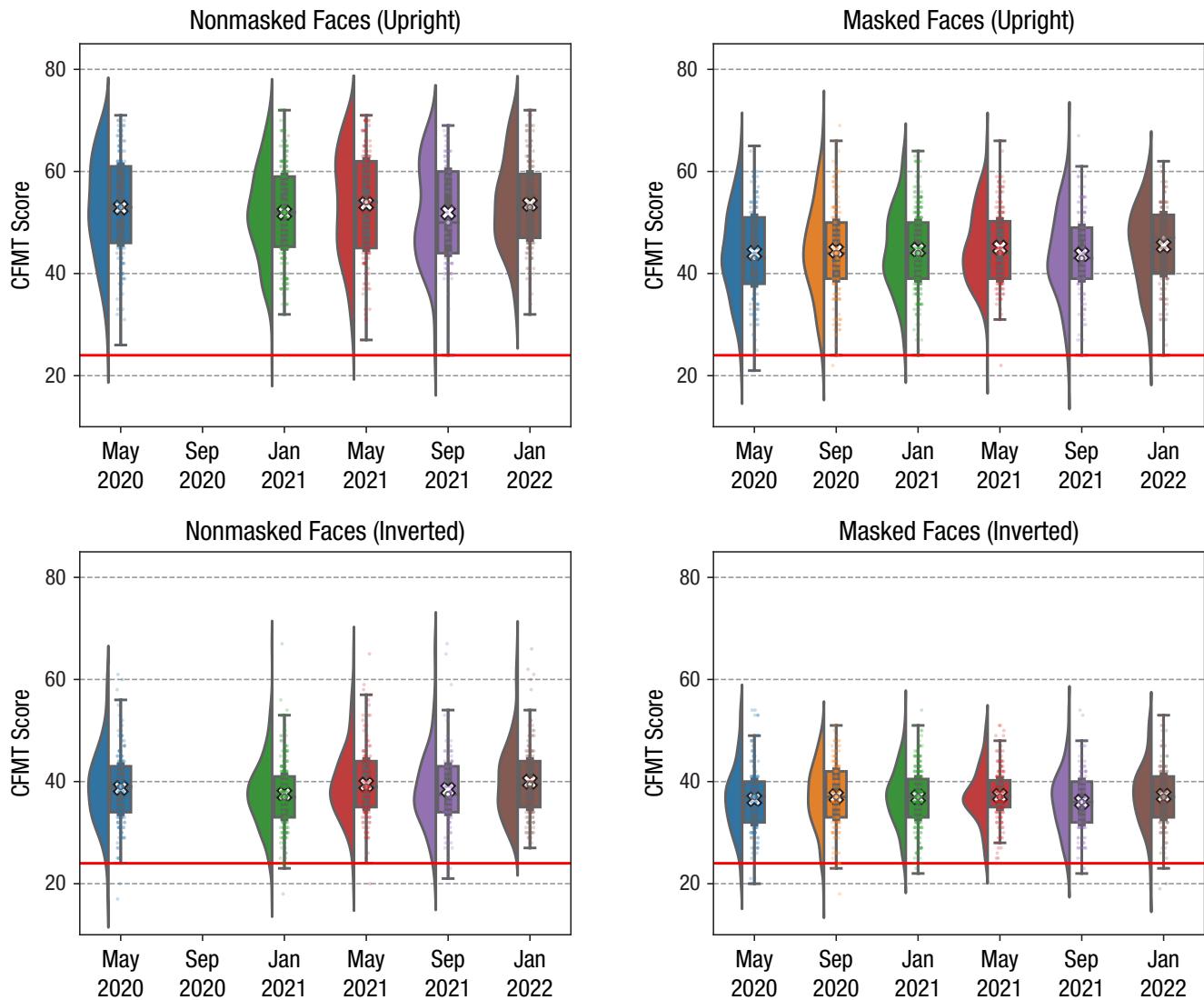


Fig. 3. Accuracy results across time points. Results of the standard Cambridge Face Memory Test (CFMT) experiment are shown for nonmasked and masked faces in both the upright and inverted conditions. The results show no improvement for masked faces over time. For all time points, performance was significantly impaired for masked faces compared with nonmasked faces. An inversion effect was found for both face types, but it was significantly smaller for masked faces. The results are similar across the different time points. Note that in September 2020, data were collected only for the masked faces. For each time point, we present the distribution (shaded region), the individual data points, and the quartiles of the data set (box plot). The whiskers depict the rest of the distribution excluding outliers, and the white Xs denote the means. The red horizontal line shows the chance-level score.

points, accompanied by a qualitative change in the processing of the masked faces.

Next, to further evaluate the consistency of recognition performance of masked faces across time points, we focused our analysis on the upright-masked-faces condition. We conducted a Bayesian ANOVA with time point (six levels) on the accuracy scores. In contrast to null-hypothesis significance testing, a Bayesian ANOVA can also provide evidence in favor of the null hypothesis (Wagenmakers et al., 2018). Accordingly, the Bayesian ANOVA decisively supported the null hypothesis (i.e., no difference between the six time points; $BF_{10} = 0.003$);

the null hypothesis was 332 times more likely than the alternative hypothesis. Similar results were observed when the Bayesian ANOVA was employed on the masked inverted faces ($BF_{10} = 0.003$).

It is notable that the large sample size allowed us to compare not only the means but also the distribution of results across individuals at each of the time points. We focused this analysis on the upright orientation separately for masked and nonmasked faces. We did not observe any significant changes in the distribution of the results across time points for both the masked and the nonmasked conditions. Two-sample Kolmogorov-Smirnov

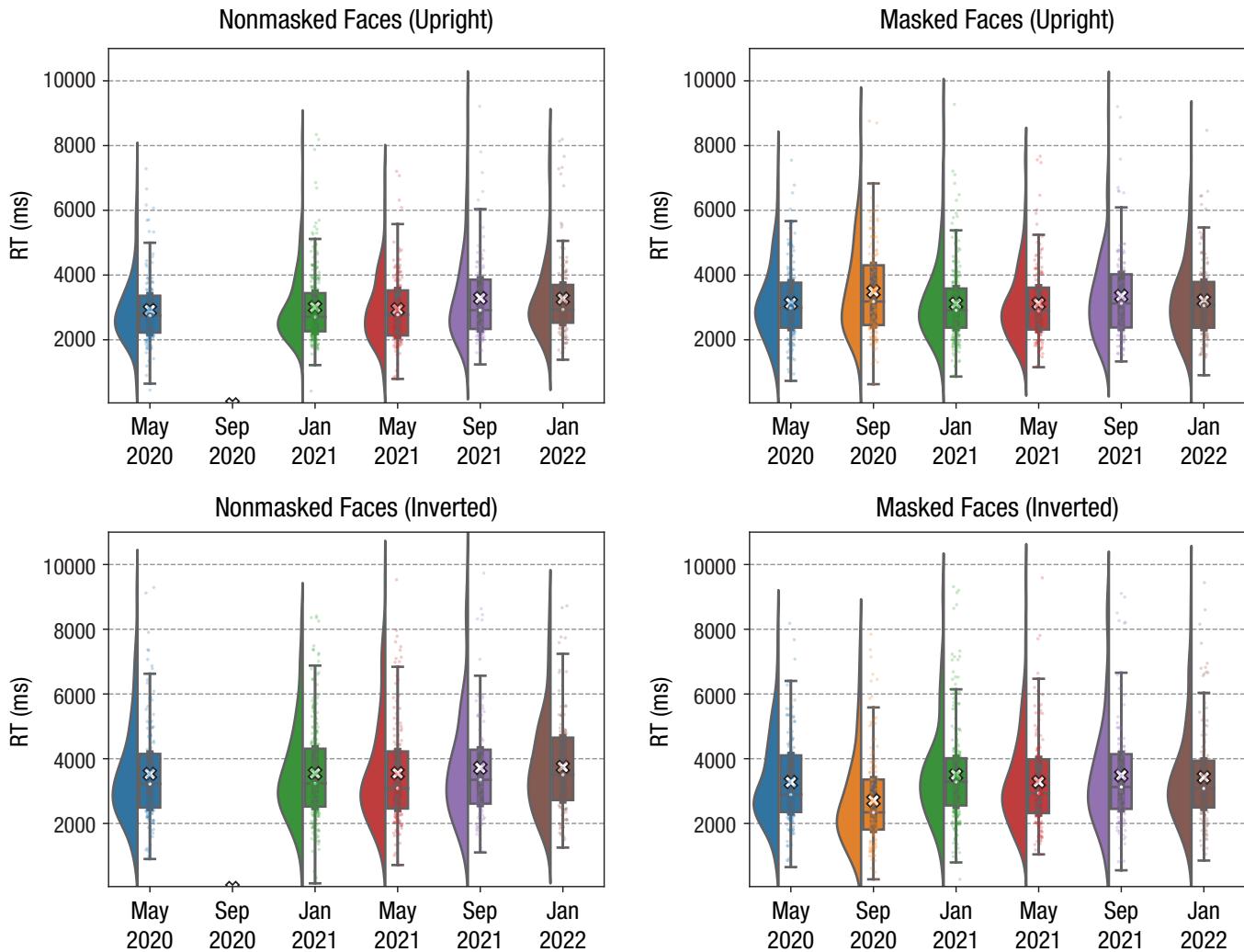


Fig. 4. Reaction time (RT) across time points. Results of the standard Cambridge Face Memory Test experiment are shown for nonmasked and masked faces in both the upright and inverted conditions. An inversion effect was found for both face types, but it was significantly smaller for masked faces. The results are similar across the different time points. Note that in September 2020, data were collected only for the masked faces. For each time point, we present the distribution (shaded region), the individual data points, and the quartiles of the data set (box plot). The whiskers depict the rest of the distribution excluding outliers, and the white Xs denote the means.

tests were employed across possible combinations for the masked and nonmasked conditions and confirmed that the distributions were not different from each other at the different time points ($D_s < 0.2$, $p > .2$). Together, the analyses of the accuracy scores provide robust evidence against changes in the processing style or processing efficiency of masked faces following extended exposure.

RTs are not commonly analyzed for the CFMT. However, we chose to analyze the RTs to provide an additional measure of face-perception abilities and to account for possible speed-accuracy trade-off effects. As shown in Figure 4, the RT results mainly mirrored the effects observed for the accuracy scores. In particular, the

ANOVA demonstrated main effects of orientation (upright faster than inverted), $F(1, 1570) = 113$, $p < .001$, $\eta_p^2 = .068$. The interaction between orientation and group was also significant, with a reduced inversion effect for masked faces, $F(1, 1570) = 22.4$, $p < .001$, $\eta_p^2 = .014$, despite the lack of a main effect for group, $F(1, 1570) < 1$. Importantly, there was no interaction between time point and group (mask status), $F(1, 1570) < 1$, suggesting that the mask effect remained constant across time points. Finally, a Bayesian ANOVA on the RT for upright faces provided support for the null hypothesis ($BF_{10} = 0.084$). Thus, the RT results support the conclusion that time and experience with masked faces did not change or improve the face-mask effect.

Longitudinal analysis

Whereas the cross-sectional analysis described above allowed us to recruit a large group of participants with no attrition, a potential disadvantage of this approach is that it does not allow examination of within-subjects changes and, therefore, might be less sensitive to subtle changes in performance. To address this concern, we invited participants who completed the experiment in January 2021 to retake the CFMT in January 2022. Two hundred nine participants (out of 495) were assigned to their original group and completed the retest experiment (mask, no mask). We note that the CFMT includes an item-specific property (presenting the same faces twice); however, such item-specific improvement in masked faces is compared against the baseline performance of participants in the nonmasked-faces condition. Therefore, to the extent that there is a general (within-subject) improvement for masked faces, it is expected that participants in the masked condition will show greater improvement in performance (item-specific improvement + general improvement) compared with participants tested twice in the nonmasked condition (for which only item-specific improvement is expected).

A repeated measures ANOVA with group (masked, nonmasked), orientation (upright, inverted), and time point (2021, 2022) revealed a two-way interaction between time point and group, $F(1, 207) = 27.8, p < .001, \eta_p^2 = .119$; for the nonmasked group, an improvement was observed for the second time point—upright: $F(1, 207) = 63.2, p < .001$; inverted: $F(1, 207) = 13.9, p < .001$. However, no improvement was found for the masked group (upright and inverted: $Fs < 1$). This interaction was qualified by a three-way interaction with orientation, $F(1, 207) = 3.82, p = .052, \eta_p^2 = .018$, because the improvement for the nonmasked faces was slightly greater for upright faces (Fig. 5a). The RT analysis revealed faster performance in January 2022 for both masked and nonmasked faces, $F(1, 207) = 31.6, p < .001, \eta_p^2 = .13$. Notably, this effect was similar across the two categories (Time \times Group interaction; $F < 1$), with no effect of group ($F < 1$; Fig. 5b). The improved performance for the nonmasked faces (in terms of accuracy and RT) and masked faces (only RT) could be attributed to the repeated test structure and test items across the two time points (for a similar result, see Murray & Bate, 2020). The lack of improvement for the masked faces in terms of accuracy might suggest impaired processing of these stimuli. Again, the key result from this experiment is the lack of improvement for masked faces. These results extend and reinforce our conclusion from the between-subjects analysis.

Finally, this longitudinal experiment allowed us to evaluate the test-retest reliability of the CFMT. In

accordance with a previous study (Murray & Bate, 2020), we found robust correlations between the two sessions for both the upright masked faces ($r = .6, p < .001$) and nonmasked faces ($r = .68, p < .001$), further establishing the reliability of the CFMT.

Individual differences with masked faces

So far, we demonstrated that the mask effect was consistent across time points using a combination of longitudinal and cross-sectional approaches. However, it is difficult to draw strong conclusions about the absence of perceptual learning without knowing the characteristics of the perceptual experiences of the target populations. Specifically, as the pandemic continued, some people (e.g., health workers) might have progressively gained greater experience with masked faces, whereas others (e.g., students who learned from home) did not gain the same amount of experience. Thus, it is plausible that we were unable to identify significant changes in the mask effect because only a subgroup of our participants showed this type of improvement. Although this conclusion is not supported by the averaged data, which should still show a directional shift toward better performance with masked faces over time, we decided to further address this concern using an additional set of experiments and analyses aimed at testing the possible role of individual differences in exposure to masked faces.

The EMF questionnaire

First, participants who completed the experiment in September 2021 (200 out of 236) and January 2022 (longitudinal and cross-sectional samples; $n = 704$) filled out the EMF questionnaire to evaluate their experience with masked faces (see the Method section). The January 2022 participants also completed the PI20 (Shah, Sowden, et al., 2015), a validated measure of subjective face-processing abilities. To increase the objectiveness of the questionnaire in measuring exposure to masked faces, we also asked participants to indicate their work location (home, office, unemployed). As expected, participants who work from their office reported greater experience with masked faces (Items 5 and 6), $F(2, 701) = 21, p < .001, \eta_p^2 = .057$, and more encounters with masked individuals (Item 9), $F(2, 701) = 4.89, p < .01, \eta_p^2 = .01$.

We conducted a repeated measures ANOVA with group (masked, nonmasked), orientation (upright, inverted), and work location (home, office, unemployed). Whereas there was a robust main effect of masks, $F(1, 686) = 143, p < .001, \eta_p^2 = .17$, there was

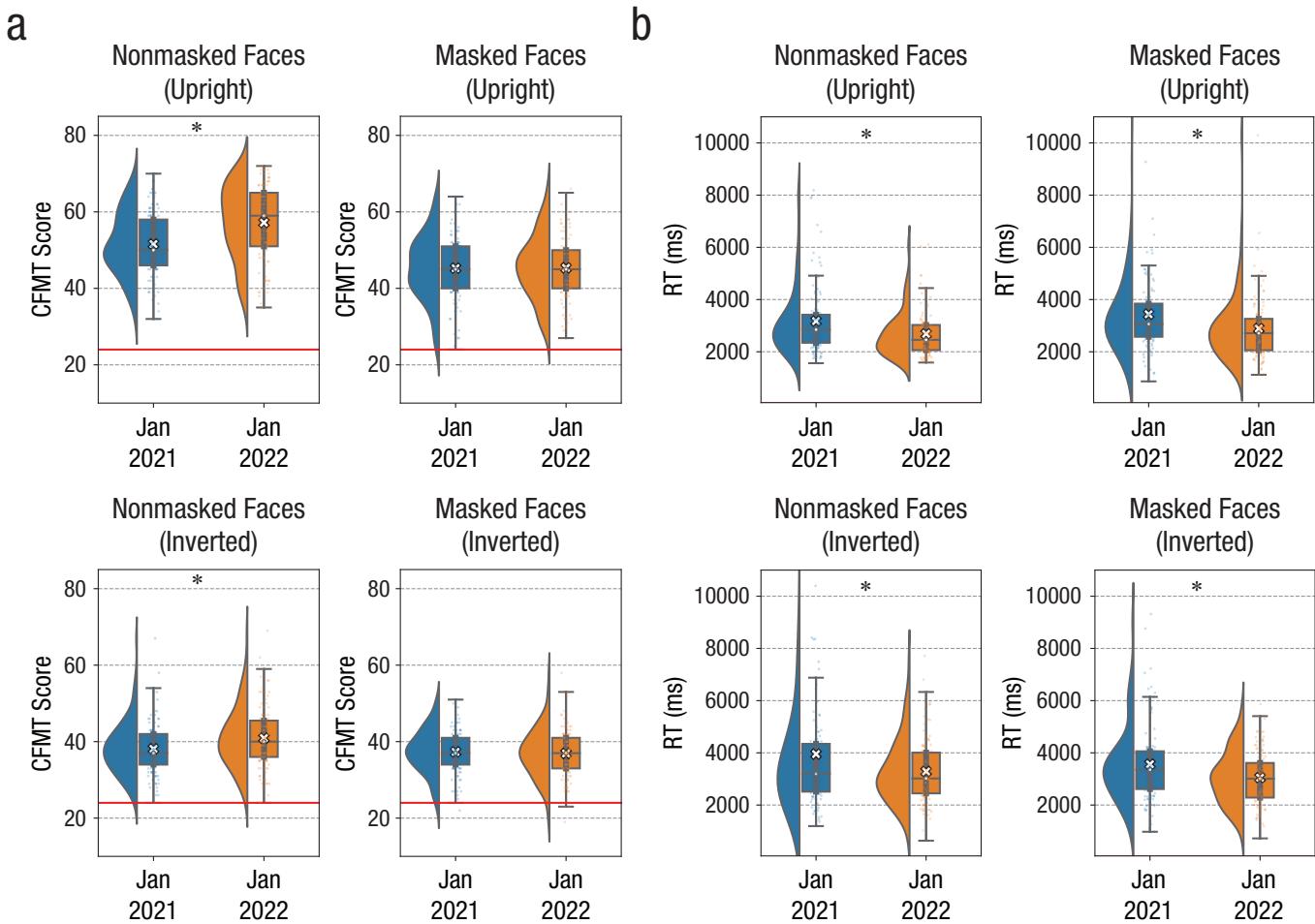


Fig. 5. Results of the longitudinal study. Results of the Cambridge Face Memory Test (CMFT) experiment are shown for nonmasked and masked faces in both the upright and inverted conditions. (a) Improved performance was observed for nonmasked faces but not for masked faces. This was true for both upright and inverted faces, with a more robust improvement for upright, nonmasked faces. The red horizontal line shows the chance-level score. (b) Faster reaction times (RTs) were observed in January 2022 for both masked and nonmasked faces. For each time point, we present the distribution (shaded region), the individual data points, and the quartiles of the data set (box plot). The whiskers depict the rest of the distribution excluding outliers, and the white Xs denote the means. A significant change between time points is denoted by an asterisk.

no interaction with work location, $F(2, 686) = 2.03, p = .132, \eta_p^2 = .005$ (note that, numerically, the mask effect was even slightly larger for the office workers). This result provides the first line of evidence that greater experience with masked faces did not modulate the mask effect. A Bayesian ANOVA on the accuracy score for masked upright faces further established the lack of work location effect ($BF_{10} = 0.119$; the null hypothesis is 7.4 times more probable than the alternative hypothesis).

The next analysis was conducted only on the participants who completed the mask condition (similar results were obtained when both groups were included). We found a robust correlation between the Experience scale (Items 1, 2, 7, and 8) and the continuous evaluation of experience with masked faces (Item 9; $r = .494,$

$p < .001$), between the Experience and Regulation scales ($r = .51, p < .001$), and between the Regulation scale and the continuous evaluation of experience with masked faces ($r = .52, p < .001$). As expected, we also found strong correlations between performance with upright and inverted faces for both accuracy and RT. Importantly, there were no correlations between face-perception abilities and experience with masked faces. This was true for accuracy scores and RTs for upright and inverted faces (all $p > .1$; Table 3). Participants from January 2022 also completed the PI20, a validated questionnaire of face-perception abilities (Shah, Sowden, et al., 2015). The PI20 score showed a significant correlation with the subjective scale of the EMF questionnaire ($r = .619, p < .001$).

Table 3. Correlations Between Face-Perception Measurements (CFMT—Mask Status as a Within-Subject Variable) and Scales of the Experience With Masked Faces (EMF) Questionnaire

Variable	Upright	Inverted	RT upright	RT inverted	Experience	Regulation	Subjective
Upright faces	—						
Inverted faces	.5*	—					
RT upright faces	.01	.04	—				
RT inverted faces	-.05	-.01	.39*	—			
EMF Experience	-.07	.03	.06	.02	—		
EMF Regulation	.04	.09	.09	.06	.51*	—	
EMF Subjective	-.09	-.08	.03	.1	.01	.01	—
Percentage of masked faces	.02	.02	.04	0	.49*	.52*	-.06

Note: CFMT = Cambridge Face Memory Test; RT = reaction time.

* $p < .05$.

These results corroborate the results of the work location analysis and indicate that the amount of reported experience with masked faces is not correlated with performance. Thus, the lack of improvement in recognition of masked faces is unlikely to reflect a reduced amount of experience with these faces.

CFMT: within-subjects design

The results of the previous analysis showed that there were no correlations between the CFMT scores and the amount of reported experience with masked faces. This was true for participants who completed the masked version of the CFMT and for those who completed the nonmasked version of the CFMT. Nevertheless, because we employed a between-subjects design in which masked faces were presented to only half of the participants, we could not directly examine whether the mask effect was modulated by the amount of experience with masked faces. Thus, we recruited a new group of participants who completed the CFMT (upright

version) twice—with masked and nonmasked faces—as well as the EMF questionnaire.

A repeated measures ANOVA with mask status as a within-subjects variable revealed a robust main effect, $F(1, 299) = 499, p < .001, \eta_p^2 = .626$, with greater accuracy for nonmasked faces (CFMT nonmasked: 54.4; CFMT masked: 44.81; mask effect = ~18%). RT analysis corroborated this finding with longer RTs for masked faces, $F(1, 299) = 19.2, p < .001, \eta_p^2 = .06$ (CFMT nonmasked: 3,631 ms; CFMT masked: 4,539 ms).

Next, we examined the relationship between the mask effect (in terms of accuracy and RT) and the scores obtained from the EMF questionnaire for the longitudinal data (Table 4). As expected, we found a robust correlation between the Experience scale (Items 1, 2, 7, and 8) and the continuous evaluation of experience with masked faces (Item 9; $r = .64, p < .001$). We also found strong correlations between the Experience and Regulation (Items 3 and 4) scales ($r = .41, p < .001$) and between the Regulation scale and the continuous evaluation of experience with masked faces ($r = .46, p < .001$).

Table 4. Correlations Between Face-Perception Measurements (CFMT—Mask Status as a Within-Subject Variable) and Scales of the Experience With Masked Faces (EMF) Questionnaire

Variable	Mask effect	Mask effect RT	Experience	Regulation	Subjective
Mask effect	—				
Mask effect RT	0	—			
EMF Experience	-.03	.02	—		
EMF Regulation	.1	.03	.41*	—	
EMF Subjective	-.07	-.09	.02	.09	—
Percentage of masked faces	-.08	-.04	.64*	.46*	.03

Note: CFMT = Cambridge Face Memory Test.

* $p < .05$.

Table 5. Correlations Between Face-Perception Measurements (GFMT) and Scales of the Experience With Masked Faces (EMF) Questionnaire

Variable	<i>d'</i>	RT	Experience	Regulation	Subjective
<i>d'</i>	—				
RT	.22*	—			
EMF Experience	-.06	.06	—		
EMF Regulation	-.01	.2*	.51*	—	
EMF Subjective	-.06	-.01	-.01	.04	—
Percentage of masked faces	-.04	.16*	.65*	.54*	-.02

Note: GFMT = Glasgow Face Matching Test (short version).

* $p < .05$.

Importantly, we did not find any correlations between the mask effect (in terms of either accuracy or RT) and experience with masked faces (all $ps > .1$).

Finally, we collected information about the occupational profile of our participants. We then focused our analysis on occupations that are more likely to have an increased amount (health and retail workers; $n = 46$) or reduced amount (students and unemployed; $n = 124$) of masked-face experiences. First, as expected, we found that health and retail workers had more experience with masked faces—Experience scale: $F(3, 166) = 8.33, p < .001, \eta_p^2 = .13$; percentage-of-masked-faces scale: $F(3, 166) = 8.54, p < .001, \eta_p^2 = .13$. Critically, an additional ANOVA with workplace and mask status showed a robust main effect for mask status, $F(1, 166) = 22.3, p < .001, \eta_p^2 = .119$, but no main effect for workplace ($F < 1$) nor an interaction between these factors ($F < 1$), suggesting that the mask effect was not modulated by the participants' occupational profile.

GFMT: masked faces

Last, we collected data from an additional established test—the GFMT (Burton et al., 2010). In the short GFMT, participants were shown 40 pairs of faces, photographed in full-face view but with different cameras, and were asked to make same/different judgments. One advantage of the GFMT over the CFMT is that external face cues (e.g., hair) that might assist in identification of masked faces are not removed. Previous research has already demonstrated the existence of the mask effect for this task (Carragher & Hancock, 2020), but it is still unclear whether experience with masked faces is correlated with performance.

To test the possible role of experience, we examined the relationship between face-perception performance (*d'* and RT) and the scores obtained from the EMF questionnaire (Table 5). Similar to the results observed for the CFMT (see above), we found a robust correlation between the Experience scale (Items 1, 2, 7, and 8) and

the continuous evaluation of experience with masked faces (Item 9; $r = .65, p < .001$), between the Experience and Regulation scales ($r = .51, p < .001$), and between the Regulation scale and the continuous evaluation of experience with masked faces ($r = .54, p < .001$). Importantly, once again, we did not find any correlations between face-perception abilities and experience with masked faces (all $ps > .1$). There were weak, yet significant correlations between RT and the Regulation scale ($r = .2$) as well as for RT and the continuous evaluation of experience with masked faces ($r = .16$); participants who reported greater experience with masked faces had slower RTs. Note that a reversed pattern is predicted if greater exposure to masked faces improves processing efficiency.

Discussion

Face masks were an important tool in the effort to minimize COVID-19 virus transmission (Cheng et al., 2020). Accordingly, the years 2020 to 2022 provided an unprecedented opportunity to examine the effects of prolonged and frequent exposure to occluded faces on recognition abilities. Here, we have documented persistent quantitative and qualitative alterations in face-processing abilities for masked versus nonmasked faces, with no evidence of improvement in the processing of masked faces over time. Using a combined cross-sectional and longitudinal approach, we found that the CFMT scores for upright faces decreased by approximately 15% when masks were added to the faces. This reduction remained statistically constant across 20 months, a period of extensive exposure to masked faces. This finding suggests that the matured face-processing system did not benefit from the prolonged exposure. Additional experiments and analyses confirmed and extended this conclusion and showed that the consistent decrement in face processing of masked faces was evident even when individual differences in exposure to these faces were considered.

Another key finding is the consistent and robust reduction of the face-inversion effect for masked faces across all time points. In particular, the inversion effect was roughly 43% smaller for masked faces. The inversion effect is suggested to reflect difficulties extracting the configural relationships between face parts (Farah et al., 1995; Freire et al., 2000). Hence, the smaller inversion effect for masked faces may be taken as evidence that holistic processing is largely reduced (although not entirely abolished). This qualitative change in the processing of masked faces was consistent across time points, providing additional evidence for the rigidity of the matured face-processing system.

Why is there no improvement in masked-face recognition?

The consistent effect of masks across time points could reflect the rigidity of the matured face-processing system. In particular, face perception rapidly develops in infancy but is then subject to a prolonged developmental trajectory (Pascalis et al., 2011, 2020). In early childhood, face processing is shaped by experience with other faces (Bate et al., 2020). One of the best examples of this malleability comes from the other-race effect, which is evident early in life (Kelly et al., 2009) but could be reversed or disappear if a child is regularly exposed to other-race faces (De Heering et al., 2010; Sangrigoli et al., 2005). In contrast, in adulthood, face-processing mechanisms are already in place and are less likely to be affected by experience (Pascalis et al., 2020; White, Kemp, Jenkins, Matheson, & Burton, 2014; Yovel et al., 2012). Here, we show that even extensive, naturalistic exposure to masked faces is not sufficient to facilitate the recognition of these faces, even though the eyes region, which is disproportionately critical for face recognition (Butler et al., 2010; Caldara et al., 2005; Royer et al., 2018; Tardif et al., 2019), remains uncovered.

An additional account for the lack of improvement in recognizing masked faces relates to the nature of the interaction. One can argue that mere exposure to masked unfamiliar faces may not suffice to revamp face-processing mechanisms. However, we note that daily encounters with masked people typically include more than just passive viewing. For example, in the grocery store, a person may need to identify their neighbor or their preferred cashier. An office worker needs to recognize peers and customers. Parents who pick up their children from school interact with other parents, children, and teachers. Hence, daily experiences provide a rich arena of exposures and the need to recognize masked faces. Yet our data suggest that such naturalistic exposures and interactions might be insufficient in eliciting adaptation of the face-processing system. A more

refined view is that improvement in face-processing abilities in adulthood depends on deliberate, systematic training programs and does not rely on naturalistic exposure. This view is supported by recent studies that show effects of systematic training programs that include individuation tasks (McGugin et al., 2011; Yovel et al., 2012) and ongoing feedback (White, Kemp, Jenkins, & Burton, 2014). Note, however, that even these systematic training programs bring only very moderate improvement in face recognition.

The results could also be attributed to another intriguing possible mechanism; the current situation may be part of a vicious circle, one that reduces the chances to improve. On the one hand, there is massive exposure to masked faces, which, in many cases, require effective recognition. On the other hand, however, people have the chance to meet and to encounter nonmasked people in the privacy of their homes or via electronic media. It is possible, therefore, that such a hybrid state of affairs provides the system with a convenient escape from effectively dealing with masked faces. In other words, the current situation may limit the system's ability to adapt, even in the face of a clear need to do so. This proposed mechanism could account for the lack of improvement that we report (almost) 2 years into the pandemic. An intriguing question is for how long such lack of improvement could persist. This, of course, depends on the extent and length of the pandemic.

Finally, the observed limited malleability of the matured face-processing system raises important questions about the ability of children to improve in recognizing masked faces. A recent study reported that in school-age children, masks hinder face-processing ability to a similar or even greater extent compared with adults (Stajduhar et al., 2022). Whether children exhibit improved masked-face recognition following prolonged exposure to masked faces in everyday life remains to be determined.

Limitations

The current investigation is timely and unique and benefits from the large sample size and combination of approaches. However, there are still important limitations that should be addressed in future studies. First, although the CFMT is a reliable test that has been used extensively over the past two decades (Bobak et al., 2016; Russell et al., 2009), the faces included in this test are all Caucasian men. Given the gender effect observed in our data as well as by other groups (Bobak et al., 2016), it is important to examine the reported effects using other, more diverse tests (Scherf et al., 2017). Another concern regards the ecological validity of the CFMT. Specifically, external face cues, which are

important for real-life face recognition, are not available in this test. This concern might be more detrimental in the case of masked faces. However, it is important to note that previous studies reported correlations between CFMT scores and subjective reports of face-recognition abilities (Shah, Gaule, et al., 2015), between the CFMT and other measurements of face-processing abilities (DeGutis et al., 2013; Russell et al., 2009), and, most importantly, between CFMT scores and naturalistic assessments of face-perception abilities (Balas & Saville, 2017). It is also worth noting that previous studies demonstrated the existence of the mask effect for other test and image sets, including the GFMT (Carragher & Hancock, 2020; see also the control experiment described above) and the Karolinska Directed Emotional Faces (Marini et al., 2021), in which external face cues are preserved.

The concern regarding ecological validity also applies to the absence of other cues that might facilitate person recognition, such as motion, voice, and body shape. Importantly, however, it is established that faces play a superior role in person recognition even when other cues are available (Hahn et al., 2016). This is demonstrated in cases of prosopagnosia, which is experienced in daily life even when all cues are available.

Another limitation of the current image set (as well as other image sets used in previous studies) is that the masks were added to existing pictures in an artificial manner. This might lead to an omission of face shape cues that are normally available and plausibly critical for recognizing masked faces in naturalistic settings. Although we cannot rule out the detrimental effect of the artificial mask on face perception, a recent study by Marini and colleagues (2021) demonstrated the existence of a mask effect even for transparent masks that reveal important cues from the lower part of the face. Hence, it is unlikely that the mask effect observed here, especially the lack of improvement in face perception for masked faces, is solely due to the nature of the stimuli.

Conclusion

The current study provides evidence for a persistent deficit in recognizing masked faces despite extensive, naturalistic exposure to these faces. Deficient performance in recognizing masked faces, along with the qualitative difference in processing those faces, may have long-term implications for daily activities, especially social interactions. Continuous exposure to masked faces could also lead to long-lasting effects in other domains, such as processing facial emotion, propagating feelings of anxiety and alienation in a world heavily occupied by masked people (Calbi et al., 2021).

Transparency

Action Editor: Vladimir Sloutsky

Editor: Patricia J. Bauer

Author Contributions

E. Freud, T. Ganel, R. S. Rosenbaum, and G. Avidan developed the study concept and contributed to the study design. A. Stajduhar and D. Di Giamarino collected the data. E. Freud analyzed the data and drafted the first version of the manuscript. T. Ganel, R. S. Rosenbaum, and G. Avidan edited the manuscript. All the authors approved the final manuscript for submission.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada and Canada First Research Excellence Fund Grant CFREF, 2016–2023.

Open Practices

All data and analysis code have been made publicly available via the Open Science Framework and can be accessed at <https://osf.io/tq92h/>. This article has received the badge for Open Data. More information about the Open Practices badges can be found at <http://www.psychologicalscience.org/publications/badges>.



ORCID iDs

Erez Freud <https://orcid.org/0000-0003-3758-3855>

R. Shayna Rosenbaum <https://orcid.org/0000-0001-5328-8675>

References

- Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia*, 49(9), 2541–2552. <https://doi.org/10.1016/j.neuropsychologia.2011.05.002>
- Balas, B., & Saville, A. (2017). Hometown size affects the processing of naturalistic face variability. *Vision Research*, 141, 228–236.
- Bate, S., Adams, A., & Bennetts, R. J. (2020). Guess who? Facial identity discrimination training improves face memory in typically developing children. *Journal of Experimental Psychology: General*, 149(5), 901–913.
- Bentin, S., Golland, Y., Flevaris, A., Robertson, L. C., & Moscovitch, M. (2006). Processing the trees and the forest during initial stages of face perception: Electrophysiological evidence. *Journal of Cognitive Neuroscience*, 18(8), 1406–1421. <https://doi.org/10.1162/jocn.2006.18.8.1406>
- Bobak, A. K., Pampoulov, P., & Bate, S. (2016). Detecting superior face recognition skills in a large sample of young British adults. *Frontiers in Psychology*, 7, Article 1378. <https://doi.org/10.3389/fpsyg.2016.01378>

- Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., Rivolta, D., Wilson, C. E., & Yovel, G. (2009). Diagnosing prosopagnosia: Effects of ageing, sex, and participant–stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. *Cognitive Neuropsychology*, 26(5), 423–455.
- Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow Face Matching Test. *Behavior Research Methods*, 42(1), 286–291. <https://doi.org/10.3758/BRM.42.1.286>
- Butler, S., Blais, C., Gosselin, F., Bub, D., & Fiset, D. (2010). Recognizing famous people. *Attention, Perception, & Psychophysics*, 72(6), 1444–1449.
- Calbi, M., Langiulli, N., Ferroni, F., Montalti, M., Kolesnikov, A., Gallese, V., & Umiltà, M. A. (2021). The consequences of COVID-19 on social interactions: An online study on face covering. *Scientific Reports*, 11(1), Article 2601. <https://doi.org/10.1038/s41598-021-81780-w>
- Caldara, R., Schyns, P., Mayer, E., Smith, M. L., Gosselin, F., & Rossion, B. (2005). Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *Journal of Cognitive Neuroscience*, 17(10), 1652–1666.
- Carragher, D. J., & Hancock, P. J. B. (2020). Surgical face masks impair human face matching performance for familiar and unfamiliar faces. *Cognitive Research: Principles and Implications*, 5(1), Article 59. <https://doi.org/10.1186/s41235-020-00258-x>
- Cheng, V. C.-C., Wong, S.-C., Chuang, V. W.-M., So, S. Y.-C., Chen, J. H.-K., Sridhar, S., To, K. K.-W., Chan, J. F.-W., Hung, I. F.-N., & Ho, P.-L. (2020). The role of community-wide wearing of face mask for control of coronavirus disease 2019 (COVID-19) epidemic due to SARS-CoV-2. *Journal of Infection*, 81(1), 107–114.
- Corrow, S. L., Davies-Thompson, J., Fletcher, K., Hills, C., Corrow, J. C., & Barton, J. J. (2019). Training face perception in developmental prosopagnosia through perceptual learning. *Neuropsychologia*, 134, Article 107196. <https://doi.org/10.1016/j.neuropsychologia.2019.107196>
- DeGutis, J. M., Chiu, C., Grossi, M. E., & Cohan, S. (2014). Face processing improvements in prosopagnosia: Successes and failures over the last 50 years. *Frontiers in Human Neuroscience*, 8, Article 561. <https://doi.org/10.3389/fnhum.2014.00561>
- DeGutis, J., Cohan, S., Mercado, R. J., Wilmer, J., & Nakayama, K. (2012). Holistic processing of the mouth but not the eyes in developmental prosopagnosia. *Cognitive Neuropsychology*, 29(5–6), 419–446. <https://doi.org/10.1080/02643294.2012.754745>
- DeGutis, J. M., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition*, 126(1), 87–100. <https://doi.org/10.1016/j.cognition.2012.09.004>
- De Heering, A., De Liedekerke, C., Deboni, M., & Rossion, B. (2010). The role of experience during childhood in shaping the other-race effect. *Developmental Science*, 13(1), 181–187.
- Dhamecha, T. I., Singh, R., Vatsa, M., & Kumar, A. (2014). Recognizing disguised faces: Human and machine evaluation. *PLOS ONE*, 9(7), Article e99212. <https://doi.org/10.1371/journal.pone.0099212>
- Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, 44(4), 576–585.
- Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance*, 21, 628–634.
- Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception*, 29(2), 159–170.
- Freud, E., Stajduhar, A., Rosenbaum, R. S., Avidan, G., & Ganel, T. (2020). The COVID-19 pandemic masks the way people perceive faces. *Scientific Reports*, 10(1), Article 22344. <https://doi.org/10.1038/s41598-020-78986-9>
- Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research*, 41(17), 2261–2271. [https://doi.org/10.1016/S0042-6989\(01\)00097-9](https://doi.org/10.1016/S0042-6989(01)00097-9)
- Grundmann, F., Epstude, K., & Scheibe, S. (2021). Face masks reduce emotion-recognition accuracy and perceived closeness. *PLOS ONE*, 16(4), Article e0249792. <https://doi.org/10.1371/journal.pone.0249792>
- Hahn, C. A., O'Toole, A. J., & Phillips, P. J. (2016). Dissecting the time course of person recognition in natural viewing environments. *British Journal of Psychology*, 107(1), 117–134.
- JASP Team. (2020). *JASP* (Version 0.14) [Computer software]. <https://jasp-stats.org>
- Kelly, D. J., Liu, S., Lee, K., Quinn, P. C., Pascalis, O., Slater, A. M., & Ge, L. (2009). Development of the other-race effect during infancy: Evidence toward universality? *Journal of Experimental Child Psychology*, 104(1), 105–114.
- Kret, M. E., & de Gelder, B. (2012). Islamic headdress influences how emotion is recognized from the eyes. *Frontiers in Psychology*, 3, Article 110. <https://doi.org/10.3389/fpsyg.2012.00110>
- 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.
- Marini, M., Ansani, A., Paglieri, F., Caruana, F., & Viola, M. (2021). The impact of facemasks on emotion recognition, trust attribution and re-identification. *Scientific Reports*, 11(1), Article 5577. <https://doi.org/10.1038/s41598-021-84806-5>
- McGugin, R. W., Tanaka, J. W., Lebrecht, S., Tarr, M. J., & Gauthier, I. (2011). Race-specific perceptual discrimination improvement following short individuation training with faces. *Cognitive Science*, 35(2), 330–347. <https://doi.org/10.1111/j.1551-6709.2010.01148.x>
- 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), Article 12820. <https://doi.org/10.1038/s41598-019-49202-0>
- Michel, C., Rossion, B., Han, J., Chung, C.-S., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science*, 17(7), 608–615.
- Murray, E., & Bate, S. (2020). Diagnosing developmental prosopagnosia: Repeat assessment using the Cambridge Face Memory Test. *Royal Society Open Science*, 7(9), Article 200884. <https://doi.org/10.1098/rsos.200884>
- Pascalis, O., Fort, M., & Quinn, P. C. (2020). Development of face processing: Are there critical or sensitive periods? *Current Opinion in Behavioral Sciences*, 36, 7–12. <https://doi.org/10.1016/j.cobeha.2020.05.005>
- Pascalis, O., de Viviés, X. D., Anzures, G., Quinn, P. C., Slater, A. M., Tanaka, J. W., & Lee, K. (2011). Development of face processing. *WIREs Cognitive Science*, 2(6), 666–675. <https://doi.org/10.1002/wcs.146>
- Royer, J., Blais, C., Charbonneau, I., Déry, K., Tardif, J., Duchaine, B., Gosselin, F., & Fiset, D. (2018). Greater reliance on the eye region predicts better face recognition ability. *Cognition*, 181, 12–20. <https://doi.org/10.1016/j.cognition.2018.08.004>
- Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin & Review*, 16(2), 252–257.
- Sangrigoli, S., Pallier, C., Argenti, A.-M., Ventureyra, V. A. G., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science*, 16(6), 440–444.
- Scherf, K. S., Elbich, D. B., & Motta-Mena, N. V. (2017). Investigating the influence of biological sex on the behavioral and neural basis of face recognition. *eNeuro*, 4(3). <https://doi.org/10.1523/ENEURO.0104-17.2017>
- Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, 14(5), 391–396.
- Shah, P., Gaule, A., Sowden, S., Bird, G., & Cook, R. (2015). The 20-item prosopagnosia index (PI20): A self-report instrument for identifying developmental prosopagnosia. *Royal Society Open Science*, 2(6), Article 140343. <https://doi.org/10.1098/rsos.140343>
- Shah, P., Sowden, S., Gaule, A., Catmur, C., & Bird, G. (2015). The 20-item prosopagnosia index (PI20): Relationship with the Glasgow Face-Matching Test. *Royal Society Open Science*, 2(11), Article 150305. <https://doi.org/10.1098/rsos.150305>
- Stacchi, L., Huguenin-Elie, E., Caldara, R., & Ramon, M. (2020). Normative data for two challenging tests of face matching under ecological conditions. *Cognitive Research: Principles and Implications*, 5(1), Article 8. <https://doi.org/10.1186/s41235-019-0205-0>
- Stajduhar, A., Ganel, T., Avidan, G., Rosenbaum, R. S., & Freud, E. (2022). Face masks disrupt holistic processing and face perception in school-age children. *Cognitive Research: Principles and Implications*, 7(1), Article 9. <https://doi.org/10.1186/s41235-022-00360-2>
- Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, 93(1), B1–B9. <https://doi.org/10.1016/j.cognition.2003.09.011>
- Tanaka, J. W., & Pierce, L. J. (2009). The neural plasticity of other-race face recognition. *Cognitive, Affective, & Behavioral Neuroscience*, 9(1), 122–131.
- Tardif, J., Morin Duchesne, X., Cohan, S., Royer, J., Blais, C., Fiset, D., Duchaine, B., & Gosselin, F. (2019). Use of face information varies systematically from developmental prosopagnosics to super-recognizers. *Psychological Science*, 30(2), 300–308. <https://doi.org/10.1177/0956797618811338>
- Vinette, C., Gosselin, F., & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: It's in the eyes. *Cognitive Science*, 28(2), 289–301. https://doi.org/10.1207/s15516709cog2802_8
- Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., & Boutin, B. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, 25(1), 58–76.
- White, D., Kemp, R. I., Jenkins, R., & Burton, A. M. (2014). Feedback training for facial image comparison. *Psychonomic Bulletin & Review*, 21(1), 100–106.
- White, D., Kemp, R. I., Jenkins, R., Matheson, M., & Burton, A. M. (2014). Passport officers' errors in face matching. *PLOS ONE*, 9(8), Article e103510. <https://doi.org/10.1371/journal.pone.0103510>
- Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., Nakayama, K., & Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences, USA*, 107(11), 5238–5241. <https://doi.org/10.1073/pnas.0913053107>
- Wong, H. K., Estudillo, A. J., Stephen, I. D., & Keeble, D. R. T. (2021). The other-race effect and holistic processing across racial groups. *Scientific Reports*, 11(1), Article 8507. <https://doi.org/10.1038/s41598-021-87933-1>
- Yovel, G., Halsband, K., Pelleg, M., Farkash, N., Gal, B., & Goshen-Gottstein, Y. (2012). Can massive but passive exposure to faces contribute to face recognition abilities? *Journal of Experimental Psychology: Human Perception and Performance*, 38(2), 285–289. <https://doi.org/10.1037/a0027077>