University of Reading

The effect of face masks and sunglasses on identity and expression recognition with super-recognisers and typical observers

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Noyes, E., Davis, J. P., Petrov, N., Gray, K. L.H. ORCID: https://orcid.org/0000-0002-6071-4588 and Ritchie, K. L. (2021) The effect of face masks and sunglasses on identity and expression recognition with super-recognisers and typical observers. Royal Society Open Science, 8 (3). 201169. ISSN 2054-5703 doi: 10.1098/rsos.201169 Available at https://centaur.reading.ac.uk/96696/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1098/rsos.201169

Publisher: The Royal Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

University of Reading

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

ROYAL SOCIETY OPEN SCIENCE

royalsocietypublishing.org/journal/rsos

Research

Cite this article: Noyes E, Davis JP, Petrov N, Gray KLH, Ritchie KL. 2021 The effect of face masks and sunglasses on identity and expression recognition with super-recognizers and typical observers. *R. Soc. Open Sci.* **8**: 201169. https://doi.org/10.1098/rsos.201169

Received: 7 July 2020
Accepted: 4 March 2021

Subject Category:
Psychology and cognitive neuroscience

Subject Areas:
psychology

Keywords:
face recognition, emotion recognition, masks, super-recognizers, face matching

Author for correspondence:
Eilidh Noyes
e-mail: e.noyes@hud.ac.uk

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5348151.

THE ROYAL SOCIETY PUBLISHING

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

The effect of face masks and sunglasses on identity and expression recognition with super-recognizers and typical observers

Eilidh Noyes[1], Josh P. Davis[2], Nikolay Petrov[2], Katie L. H. Gray[3] and Kay L. Ritchie[4]

[1]Department of Psychology, University of Huddersfield, Huddersfield HD1 3DH, UK
[2]School of Human Sciences, University of Greenwich, London SE10 9LS, UK
[3]School of Psychology and Clinical Language Sciences, University of Reading, Reading RG6 6UR, UK
[4]School of Psychology, University of Lincoln, Lincoln, Lincolnshire, LN6 7TS, UK

EN, 0000-0001-8709-879X; JPD, 0000-0003-0017-7159; KLHG, 0000-0002-6071-4588; KLR, 0000-0002-1348-760X

Face masks present a new challenge to face identification (here matching) and emotion recognition in Western cultures. Here, we present the results of three experiments that test the effect of masks, and also the effect of sunglasses (an occlusion that individuals tend to have more experienced with) on (i) familiar face matching, (ii) unfamiliar face matching and (iii) emotion categorization. Occlusion reduced accuracy in all three tasks, with most errors in the mask condition; however, there was little difference in performance for faces in masks compared with faces in sunglasses. Super-recognizers, people who are highly skilled at matching unconcealed faces, were impaired by occlusion, but at the group level, performed with higher accuracy than controls on all tasks. Results inform psychology theory with implications for everyday interactions, security and policing in a mask-wearing society.

1. Introduction

The human face provides us with a great deal of information about a person, and arguably the two most important pieces of information in a face are the person's identity and their emotional state [1]. Judgements of identity and emotion facilitate social interactions [2] and can inform identification procedures in policing and security contexts. In the light of the

© 2021 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

current COVID-19 pandemic, some countries are instructing that people wear face masks [3]. Face masks occlude the lower features of a face, with the potential to impact on identity and emotion perception.

Most people can recognize the faces of their friends, family and favourite celebrities with ease despite variation in the appearance of these faces across different encounters [1,4–7]. Familiar face recognition is robust against image-level changes, meaning that the face can be recognized across changes in pose, lighting, illumination, camera distance and, in some cases, deliberate disguise [4,8–14]. It is often possible to identify familiar faces when parts of the face are occluded [15–17]. By contrast, identity comparisons for faces that we have not previously encountered, or have very little experience with, is error prone [8,10,12,13,18–20]. Any differences between the comparison images (e.g. differences in pose, illumination or expression) increases task difficulty (e.g. [21]) and these effects are aggravated by natural changes to appearance such as from ageing [22]. People frequently mistake two images of the same unfamiliar face as different people [4,23], and mistake different people of similar appearance for the same person [24]. The error-prone nature of unfamiliar face identification has important consequences in applied settings.

Face comparisons inform identifications in many policing and security scenarios. A typical task at border control, for example, involves comparing the person at the border with the image on their passport. This sort of unfamiliar face matching task has been replicated in various laboratory and live settings. A 10–20% error rate is typical in standard face matching tasks which present two images side by side on a computer screen and ask participants to decide whether the images show the same person or two different people [25]. Presentation of a live person and a face photograph does not yield improved performance [7,26,27], nor does experience with face matching in passport officers [28].

Past studies show that drastic changes to facial features such as changing a person's appearance to look deliberately unlike themselves, or like someone else, severely impairs unfamiliar face identification [13,16,17,29]. When faces are altered to include competing identity information, identification may also be severely impacted [30,31]. For example, when the top half of one face (target region) is aligned with the bottom half of another (distractor region), observers' perception of the target region is biased [32]. In sequential matching paradigms, this 'composite face effect' results in observers making errors when judging the target region, often reporting that the target regions are different when they are in fact the same [33]. The composite face effect is thought to be driven by holistic processing, where local features are processed as a unified whole [34].

A fully occluded face, such as a face occluded by a stocking mask, is extremely difficult to identify [35]. However, the effect of partial occlusion is less understood. The simple addition of props to a face (e.g. sunglasses or glasses) has also been found to reduce accuracy on unfamiliar face matching tasks [13,16,17,29,36,37], where props remove information about the identity by occluding features of the face. Previous studies which have attempted to assess the importance of specific facial features for face identification have typically grouped features generically as 'external features' (e.g. hair) or 'internal features' (e.g. eyes, nose, mouth [38–40]). The relative contribution of specific internal features of the face is unknown. For unfamiliar face matching, the occlusion of any facial feature will probably reduce performance, with task difficulty increasing as more of the face is covered [20]. The eye region is argued to be the most diagnostic cue for face identification [41,42]. Glasses partially occlude, and sunglasses fully occlude the eye region of the face. Two studies have shown that unfamiliar face matching is impaired when one image in the pair is unconcealed and the other wears glasses [36,37]. The performance was further reduced when the eye region in one image in the pair was fully occluded through sunglasses, thus suggesting a role for the eye region in unfamiliar face matching [36]. It has been suggested that the mouth region may also be useful for face identification. Mileva & Burton [43] found higher face matching accuracy for pairs of images presenting open-mouthed smiles compared with neutral facial expressions. The authors argue that a person's smile provides idiosyncratic information about their identity, and therefore helps with face matching. Thus, occluding the mouth region of the face ought to reduce face matching accuracy. A recent paper, examining surgical face masks only, superimposed masks on to existing images of celebrities and also on to images from a large face database [44]. They found reduced familiar and unfamiliar face matching when one or both images in the pair was masked, compared with both images being unconcealed [44]. To date, no studies have directly compared the effect of each of these conditions (occluding the eye region compared against occluding the mouth region) on face matching performance.

While unfamiliar face matching is error prone there are some people who perform well above typical levels—referred to as 'super-recognizers' (see [45] and [46] for reviews). At the group level, super-recognizers perform with consistently high accuracy [47–51]. There have been several attempts to explore potential underpinnings of superior face recognition ability, with some studies suggesting that

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

super-recognizers use different features of the face to inform their identification. One study [52] reported that super-recognizers fixate on the nose region of the face more often relative to controls during face matching tasks. Other studies have suggested that fixations to the eyes [53], or just below the eye region, is associated with optimal face identification [54]. To date, there has been no investigation of super-recognizers' performance on face matching tests in which parts of the face are occluded. However, in two memory-based experiments, Davis & Tamonyte [55] found that super-recognizers outperformed controls at the recognition of faces wearing balaclavas (eye region only visible), hats and sunglasses (eye region covered), as well as with no facial occlusion. In their first experiment, test phase 10-person line-ups were presented immediately after the first phase single face image 8 s familiarization trial. In Experiment 2, a single 1 min video was displayed in Phase 1. The delays between Phase 1 and viewing a video line-up in Phase 2 was at least one week. In both experiments, accuracy was highest in the no occlusion condition, and lowest in the balaclava condition; although effect sizes were larger for correct rejections of previously unseen faces, than when correctly identifying those viewed before. At an individual level, fixation patterns, and performance accuracy is not always consistent with that expected of a super-recognizer [48,50,51,56]. Any study of super-recognizer performance on concealed faces should consider the performance of super-recognizers at the group level, and also consider the spread of performance at the individual level.

In addition to providing us with identity information, faces provide a highly informative cue to individuals' emotions. Seven 'basic' expressions (anger, fear, disgust, neutral, happiness, sadness and surprise) are thought to be recognized universally [57–59]; however, see [60]. These emotional expressions are processed rapidly [61], and recognized accurately in neurotypical adults [62], irrespective of face familiarity [63]. Facial expressions can be decomposed into action units [64] where each expression is described as a configuration of muscle movements. Some emotional expressions are best described by action units that are mainly in the mouth region of a face, such as happiness, and others are mainly in the eye region, such as fear [64]. To investigate whether this information is used by observers to categorize emotion, researchers have presented partially occluded faces (using 'bubbles' or other shapes) and calculated which face regions are correlated with categorization accuracy for different expressions [65–67]. Generally, mouth regions are most informative for happy, surprised and disgusted expressions, whereas eye regions are most informative for fearful and angry expressions, and both regions are informative for sad and neutral expressions [65,66]. Not only are the regions that are informative for each expression more likely to be fixated in an emotional categorization task, occluding these informative regions also disproportionately impacts accuracy [68]. Removing half of the image (i.e. only presenting the top or bottom half of an emotional face) leads to similar findings, with happy and disgust being most recognizable from the bottom half of the face and anger, fear and sadness being most recognizable from the top half of the face [69]. However, limited research has addressed the effects of naturalistic occlusion on emotional expression judgements.

One exception investigated the impact of sunglasses and masks on emotion categorization [70]. In this study, sunglasses were added onto a validated set of emotional faces using image editing software. The face mask condition, however, consisted of a non-realistic grey ellipse being added to the mouth region of the faces. Although occluding some of the same region as a realistic face mask, the ellipse did not cover the nose in most of the images presented. The authors [70] found that adults classified each emotional expression (happy, sad, surprise, fear and anger) less accurately when sunglasses were added to the images than when the images were unaltered. Accuracy was reduced further by masks than by sunglasses when all emotional expressions were combined, but the reduction in accuracy for each expression was not reported for the mask condition.

It has been suggested that super-recognizers who outperform controls on face matching and memory tests, may also be superior at identifying emotional expressions. Rhodes *et al*. [71] report that a person's emotion and face recognition abilities correlate more strongly than emotion and car recognition ability. In addition, Connolly *et al*. [72] identified a positive relationship between scores on emotion and face recognition tests for all of the basic emotions other than happiness. Very recent work has linked face matching ability as scored by the Glasgow face matching task, a standardized test of face matching ability [25], with recognition ability for emotions anger, fear and happiness, and also for neutral faces [73]. No differences were observed for the recognition of disgust or sadness [73]. Links between face recognition ability and emotion recognition have also been observed at the other end of the face recognition ability spectrum—developmental prosopagnosia may also be associated with the impaired perception of emotional expressions [74]. Together these results suggest some degree of positive

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

relationship between emotion and identity recognition ability. It is unknown whether this will extend to concealed faces.

The effect of face concealment on familiar and unfamiliar face recognition, as well as emotion recognition, is increasingly relevant as many countries around the world recommend the wearing of face masks which cover the mouth and nose in an attempt to slow or stop the spread of COVID-19. Mask wearing raises new questions in face perception—how accurate will unfamiliar face matching be for masked faces? Will we still be able to recognize familiar faces and gauge how a person is feeling? Here, we provide the first comprehensive assessment of the effect of masks on familiar (Experiment 1) and unfamiliar face matching (Experiment 2), and expression categorization (Experiment 3). We compare face matching and expression recognition accuracy for unconcealed faces, faces wearing sunglasses and faces wearing mouth and nose covering face masks. We used sunglasses as a comparison for face masks for three reasons: (i) this allows for a comparison of concealment of the upper (sunglasses) and lower (face masks) parts of the face; (ii) there is some previous work using sunglasses on which we can base our predictions; and (iii) historically, sunglasses have been more commonly seen than face masks in Western countries. We also compare performance on all three tasks between control participants and super-recognizers. We predicted that both sunglasses and face masks would give rise to poorer face matching performance for unfamiliar faces. A recent paper found a detrimental effect of digitally added surgical masks on familiar face recognition [44], and so based on this we predicted that masks would give rise to poor familiar face matching accuracy. We also predicted that sunglasses and face masks would affect emotion categorization differently for different expressions. We predicted that expressions containing diagnostic information in the top half of the face, including fear and anger, would be more affected by sunglasses, whereas expressions containing diagnostic information in the bottom half of the face, including happiness and surprise, would be more affected by masks. Finally, we hypothesized that super-recognizers would outperform controls in all three tasks. In sum, here we provide the first study to directly compare face matching and emotion categorization performance for super-recognizers and typical observers for faces in no concealment, sunglasses and masks. Our study is novel in its use of real mask images (rather than computer-generated facial occlusions) in the matching tasks. Our matching task mask stimuli are representative of the types of masks worn during COVID-19.

2. Experiment 1—familiar face matching

Familiar face identification is robust against many forms of image manipulation [4,8–14]; however, drastic changes in the appearance of a familiar face can impair identification [13]. Super-recognizers are typically better at face identification tasks than controls [47,49,51,75,76]. The advantage extends to some types of concealed faces [55]. Experiment 1 tests the effect of masks and sunglasses on familiar face identification for control participants and super-recognizers.

Importantly, in all three studies presented here, we use two different groups of control participants. Our super-recognizers were recruited from a large database of participants used in previous research (e.g.[77–79]). These participants were originally attracted to take part in research after seeing media reports about a super-recognizer test. Consistent with previous studies [78–80], super-recognizers are defined as those scoring 40/40 on the Glasgow face matching test: short version (GFMT) [25] and 95+/102 (93%) on the Cambridge face memory test: extended (CFMT+) [81]. An estimated 2% of the population score 95 or above on the CFMT+ [81,82], while less than 5% achieve the maximum on the GFMT [25]. During that original database recruitment process, many participants did not meet the criteria to be classed as super-recognizers. These other participants have continued to take part in various face perception studies. Following research precedent [78,79,80], typical-ability participants invited from this second group who had previously scored within one standard deviation of the normal population mean on both the CFMT+ (i.e. 58–83: [49]) and GFMT (i.e. 28–36: [25]) were allocated to a 'practised controls' group. It is possible that practise on such tasks, and high levels of interest in participating in face recognition tasks, may themselves boost performance to a level closer to that of super-recognizers than unpractised controls [45]. In addition to the practised controls, we recruited one further control group of participants from the online recruitment platform, Prolific.co. These participants, although routinely completing surveys, do not routinely complete face processing tasks. This group had not been pre-screened using the GFMT or CFMT+, and so provide a random sample of the population [45].

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

Figure 1. Example image pairs for each concealment condition. The above images all show the same person. (Copyright restrictions prevent publication of the images used in the experiment. Images in figure 1 are illustrative of the experimental stimuli and depict someone who did not appear in the experiments but has given permission for the images to be reproduced here).

2.1. Method

2.1.1. Participants

The unpractised control group (i.e. those who were randomly recruited and who may or may not have done face processing experiments previously) were recruited from Prolific.co. We included the specification that they must be resident in the UK so as to maximize the likelihood that they would recognize all of our celebrity faces, two of whom were likely only to be recognized in the UK. Each unpractised control participant was given £2.11 to compensate them for their time. The final sample was 102 (36 male, 65 female, one other; mean age 35 years; age range 18–63 years; 87.25% Caucasian).

The practised control group (i.e. those who have participated in the previous face processing experiments) were recruited from a large database of interested participants from the UK, and were not given monetary compensation. Members of this database are practised in that all have taken the CFMT+, the GFMT and at least one other face recognition test previously, and they may have been randomly selected to be invited to up to six online face processing projects per annum. Most projects provide debriefing feedback in terms of final test scores, so participants are normally roughly aware of their own ability, although no individual trial feedback is normally provided. In addition, no cross-referencing record is kept of whether they respond to invites or not. An initial sample of 306 participants took part. All claimed never to have taken the CFMT+ and GFMT prior to their database-stored score. Three were removed due to incomplete data. In order to match the unpractised control sample, we conducted the analyses on the first 102 participants to complete the tasks (30 male, 72 female; mean age 43 years; age range 21–72 years; 88.23% Caucasian). Practised controls had a mean GFMT score of 34.12/40 (s.d. = 1.89), and a mean CFMT+ score of 72.50 (s.d. = 7.09).

The super-recognizers were recruited from the same large database as the practised control participants, and will probably have been randomly invited to a similar number of face processing projects as the practised control group, and were not given monetary compensation. An initial sample of 159 participants took part, with one removed due to incomplete data. In order to match the two control samples, we conducted the analyses on the first 102 participants to complete the tasks (25 male, 77 female; mean age 39 years; age range 21–67 years; 91.17% Caucasian). Super-recognizers all scored 40/40 on the GFMT, and a mean CFMT+ score of 97.24 (s.d. = 1.75) as assessed in a previous battery of unpublished tests.

2.1.2. Stimuli

Four images of 12 celebrities were taken from the Internet in conditions (i) reference image (unconcealed), (ii) comparison image (unconcealed), (iii) sunglasses image, and (iv) mask image (figure 1). All images were gathered via Google Image search following the procedures used in previous research (e.g. [7,83]), with the only constraints being that the image should be good quality (i.e. not blurry), and show the face in a mostly front-facing view. There were no constraints in terms of facial expression displayed. All images were cropped to show head and shoulders at 380×570 pixels. The reference image was chosen as the more front-facing, neutral expression of the two unconcealed images, so as to approximate a passport-style image. A different identity 'foil' face image was selected for each identity to serve as the reference image in non-match trials. The foil identities were chosen to match the same verbal description as the target identity, e.g. 'young woman, blonde hair'. In all trials, the

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

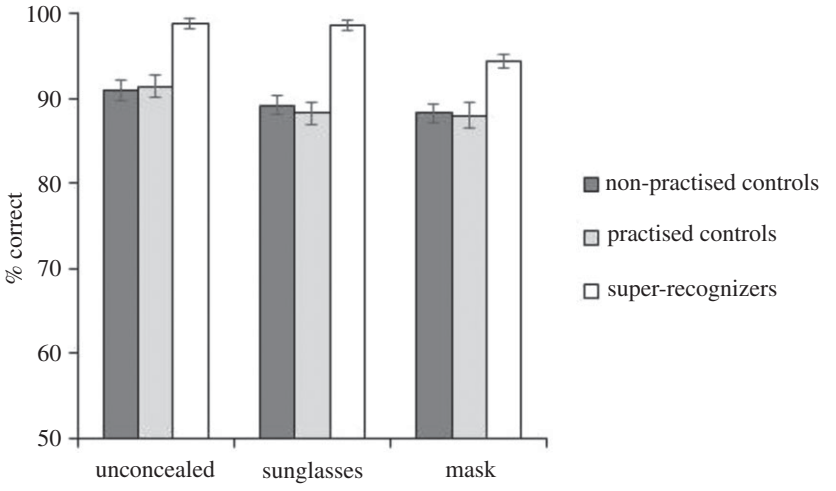

Figure 2. Data from Experiment 1, familiar face matching performance. Error bars show the within-subjects standard error [87].

image on the left was a reference image (either the reference image for the same ID trials of the celebrity, or the foil face for different identity trials). The image on the right was either the unconcealed comparison image, sunglasses image or mask image (figure 1).

2.1.3. Procedure

Participants completed the experiment online using the Qualtrics platform. Online tests of cognitive processing have become increasingly popular, and have been found to yield high-quality data that is indistinguishable from that collected in the laboratory [84–86]. In the familiar face matching task, participants were instructed that they would view pairs of face images and that their task was to decide whether the images were of the same person or two different people. Each face pair was presented side-by-side with the text 'Do you think these images depict the same or different people?' below the images. The response options presented below the text were 'Same' and 'Different'. The images and text remained on the screen until participants responded and clicked 'next' to see the next trial. On each trial, participants were also instructed to provide a confidence judgement for their same/different identity response for each trial—provided by the use of a sliding scale from 0 to 100. The confidence data are not presented in this paper. A practise trial with images not used in this experiment ensured that participants understood the paradigm. Identities were randomly assigned to conditions between participants, and each participant saw each identity only once, resulting in 12 trials. Participants saw two trials in each concealment condition (unconcealed, sunglasses, mask) for each trial type (match, non-match). The small number of trials in this experiment reflects both the difficulty of finding celebrities who would be familiar to most of our participants, and then finding images of those celebrities wearing face masks and sunglasses. At the end of the experiment, participants reviewed a list of names of celebrities and were asked to select all names for whom they would recognize the face. Participants received their accuracy score upon completion of all three experiments.

2.2. Results

We did not remove any trials based on participants' reported familiarity with the identities, and instead simply took familiarity as a group-level manipulation (as in [83] for example). Unpractised controls were familiar with a mean of 75.00% (9/12, s.d. = 2.30) celebrities, practised controls with 75.00% (9/12, s.d. = 3.52) celebrities, and super-recognizers with 91.67% (11/12, s.d. = 1.46) celebrities, therefore, we are satisfied that participants were, as groups, familiar with the celebrities presented in this experiment.

Mean accuracy for Experiment 1 is shown in figure 2. We began by analysing overall accuracy (collapsed across match and non-match conditions). A 3 (participant group: unpractised controls, practised controls, super-recognizers) × 3 (concealment: unconcealed, sunglasses, mask) mixed ANOVA showed a significant main effect of participant group $F_{2,303} = 20.89$, $p < 0.001$, $\eta_p^2 = 0.12$. Super-recognizers ($M = 97.22\%$) significantly outperformed both unpractised controls $M = 89.46\%$,

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

$t_{202} = 5.68$, $p < 0.001$, $d = 2.68$ and practised controls $M = 89.22\%$, $t_{202} = 8.01$, $p < 0.001$, $d = 2.87$, with no difference in performance accuracy between the unpractised and practised controls, $t_{202} = 0.15$, $p = 0.883$, $d = 0.07$. The ANOVA also showed a significant main effect of concealment $F_{2,606} = 5.12$, $p = 0.007$, $\eta_p^2 = 0.02$. Bonferroni corrected *t*-tests showed significantly poorer performance with pairs in which one face wore a mask ($M = 90.20\%$), compared with when both faces were unconcealed ($M = 93.71\%$), $t_{305} = 3.14$, $p = 0.006$, $d = 0.18$, and no difference between pairs in which one face was in sunglasses ($M = 91.99\%$) compared with when both were unconcealed, $p = 0.276$, and no difference in performance between the sunglasses and mask conditions, $p = 0.360$. The ANOVA showed a non-significant interaction $F_{4,606} = 0.67$, $p = 0.611$, $\eta_p^2 < 0.004$.

These results show that super-recognizers outperformed both of our control groups, with no difference between practised and unpractised controls. Signal detection analysis showed the same pattern of results (see electronic supplementary material, §S1). This result is consistent with prior research which found that super-recognizers outperformed control participants at identifying celebrities from poor quality images [75], and at matching celebrities with lookalikes in a pixellated face matching test [47]. In addition, both Davis *et al*. [75] and our present study found that super-recognizers claim to be familiar with more celebrities than control participants. This familiarity advantage may explain the difference in the performance of super-recognizers and controls.

The small overall decrease in performance for masked faces is driven predominantly by poorer performance on non-match trials (see electronic supplementary material, §S1) [88]. Our results show a subtle decrease in familiar face matching performance for masked faces compared with unconcealed faces. These results are broadly in line with a recent paper [44], although we found that masks gave rise to a smaller reduction in familiar face matching performance. When considering face identification in security settings, it is important to consider the effects of concealment on unfamiliar face recognition, therefore, we carried out a second experiment testing unfamiliar face matching with pairs of images showing unconcealed faces, as well as images wearing sunglasses or face masks.

3. Experiment 2—unfamiliar face matching

In this experiment, we investigated the effect of masks and sunglasses on unfamiliar face matching. Based on previous research, e.g. [36], we expected sunglasses to have a detrimental effect on face matching accuracy. Of particular interest was whether face masks would confer an additional detrimental effect on face matching.

3.1. Method

3.1.1. Participants, stimuli and procedure

The same participants who took part in Experiment 1 completed this experiment. Here, we used images of identities chosen to be unfamiliar to our participants. We collected images of 60 identities (30 female) which were publicly available on the Internet. Image selection and cropping were carried out in the same way as Experiment 1. As in Experiment 1, we collected two unconcealed images, one image wearing sunglasses and one image wearing a face mask which covered the mouth and nose. For each identity, we also collected one unconcealed image of a foil identity chosen to match the same verbal description as the target identity. The procedure was the same as that used in Experiment 1, but using 60 trials. Participants saw each identity once with the assignment of identities to conditions randomized across participants. Participants saw 10 trials in each concealment condition (unconcealed, sunglasses, mask) for each trial type (match, non-match).

3.2. Results

Unpractised controls reported recognizing a mean of 0 (s.d. = 0.24) identities, practised controls with 0 (s.d. = 0.24) identities and super-recognizers with 0 (s.d. = 0.29) identities, therefore, we are satisfied that the identities presented in this experiment were unfamiliar to our participants. Mean accuracy for Experiment 2 is shown in figure 3. We began by analysing overall accuracy. A 3×3 mixed ANOVA showed a significant main effect of participant group $F_{2,303} = 71.66$, $p < 0.001$, $\eta_p^2 = 0.32$, a significant main effect of concealment $F_{2,606} = 61.77$, $p < 0.001$, $\eta_p^2 = 0.17$ and a significant interaction $F_{4,606} = 3.19$, $p = 0.013$, $\eta_p^2 = 0.02$.

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

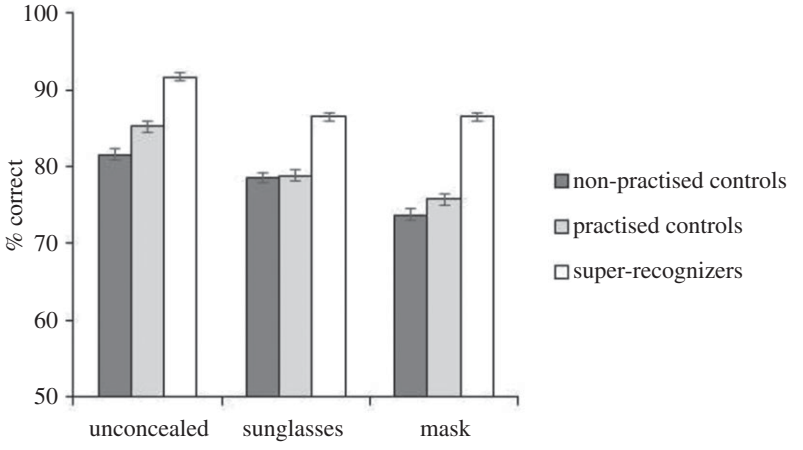

Figure 3. Data from Experiment 2, unfamiliar face matching performance. Error bars show the within-subjects standard error [74].

Bonferroni corrected *post hoc* *t*-tests revealed that super-recognizers performed significantly better than both control groups across all concealment conditions (all ps < 0.05), and that the practised control group outperformed the unpractised control group only for unconcealed faces ($p < 0.05$). The practised control group and super-recognizer groups performed with higher accuracy on unconcealed trials than on sunglasses trials (practised: unconcealed $M = 85.20$, sunglasses $M = 78.77$, $t_{101} = 5.05$, $p < 0.001$, $d = 0.50$; super-recognizers: unconcealed $M = 91.67$, sunglasses $M = 86.51$, $t_{101} = 5.08$, $p < 0.001$, $d = 0.57$). These same groups also performed with higher accuracy on unconcealed trials than mask trials (practised: mask $M = 75.74$, $t_{101} = 7.76$, $p < 0.001$, $d = 0.77$; super-recognizers: mask $M = 86.47$, $t_{101} = 5.71$, $p < 0.001$, $d = 0.57$). Unpractised controls showed no difference in their performance for unconcealed ($M = 81.52$) and sunglasses trials ($M = 78.53$), $p = 0.054$, but performed with higher accuracy for unconcealed trials than mask trials (mask $M = 73.73$), $t_{101} = 5.74$, $p < 0.001$, $d = 0.57$. Both control groups performed significantly better with sunglasses than masks (unpractised: $t_{101} = 3.62$, $p < 0.001$, $d = 0.36$, practised: $t_{101} = 2.33$, $p < 0.001$, $d = 0.23$). Super-recognizers showed no difference in performance for the sunglasses and mask conditions, corrected $p > 0.999$.

For unfamiliar face matching, super-recognizers outperformed both of our control groups, and performance was poorer with concealed faces compared with unconcealed faces. The reduction in performance for the control groups with masked faces was qualified by an increase in bias across all three participant groups for masked faces (see electronic supplementary material, §S2). For pairs of faces in which one image wore a mask, participants were more biased to respond 'non-match' or that the two images showed different people, as compared with unconcealed faces or sunglasses. These results are again broadly in line with a recent paper [37], but again our observed reduction in performance is smaller. It is important to note that performance across all groups and all conditions was consistently well above chance (50%), and so although concealment hindered performance, it did not completely destroy participants' ability to complete the task.

The super-recognizers in our study all scored 40/40 on the GFMT, and had a mean score of 97.2 on the CFMT+. Participants in the practised control group scored with lower accuracy than our super-recognizers on both of these standard tests. Some participants in the practised control group performed with very high accuracy on our tasks, and some participants in the super-recognizer group performed with accuracy levels well below the control mean. This result demonstrates noise in experimental testing [56] and speaks to the theoretical issues associated with the definition and selection of super-recognizers [45].

We had access to the GFMT scores and CFMT+ scores for both our practised control and super-recognizer group, and found significant correlations between performance on our tasks and the GFMT (figure 4), and our tasks and the CFMT+ (figure 5). This is consistent with the idea that performance on face tests tends to correlate with each other at the group level [89–91]. The scatterplots also clearly demonstrate spread in performance at the individual level for participants in each group. Our results reiterate the challenges associated with the definition of what it means to be a super-recognizer, and the importance of presenting super-recognizer data at both the group and individual level so that accurate conclusions on super-recognizer performance, and consistency of super-recognizer performance across tasks, can be drawn [45,50,51,56].

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

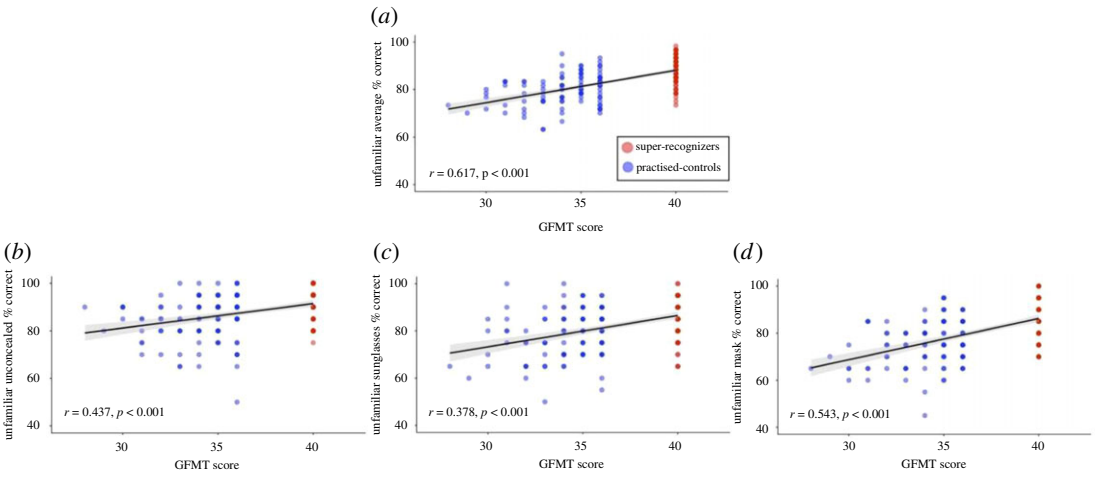

Figure 4. Scatterplots showing the relationship between GFMT performance and unfamiliar matching accuracy for (*a*) the average across all conditions, (*b*) unconcealed, (*c*) sunglasses and (*d*) mask conditions. Super-recognizers are given in red, and practised controls given in blue; individual data points are semi-transparent. 95% CI is given in grey.

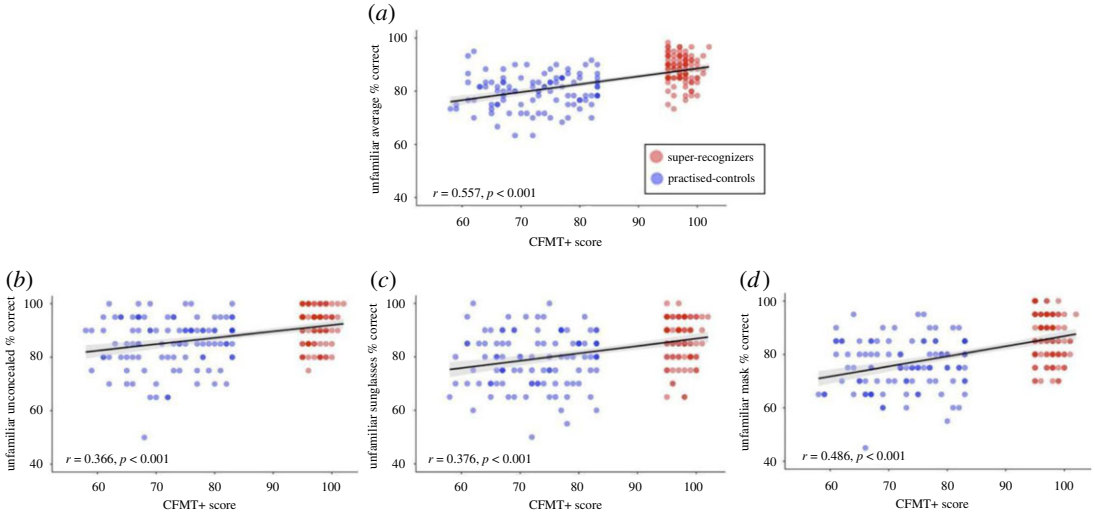

Figure 5. Scatterplots showing the relationship between CFMT+ performance and unfamiliar matching accuracy for (*a*) the average across all conditions, (*b*) unconcealed, (*c*) sunglasses and (*d*) mask conditions. Super-recognizers are given in red, and practised controls given in blue; individual data points are semi-transparent. 95% CI is given in grey.

In addition to posing a problem for identity recognition, face masks may also impair facial expression recognition. Our third experiment examined expression categorization with unconcealed faces, sunglasses and face masks.

4. Experiment 3—expression recognition

Typically, the whole face is used to categorize emotions [69]. However, different face regions are relatively more or less important for categorizing different emotional expressions; whereas some expressions contain critical diagnostic information in the eye region, others contain diagnostic information in the mouth region [64–67,69]. As sunglasses and masks block information from eye and mouth regions, respectively, the pattern of categorization impairment we find with these forms of concealment is likely to depend on the location of diagnostic information for each expression. Research on individual differences suggests that face identity recognition ability and expression recognition accuracy are related [71,72], indicating super-recognizers may have better emotion categorization than control participants.

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

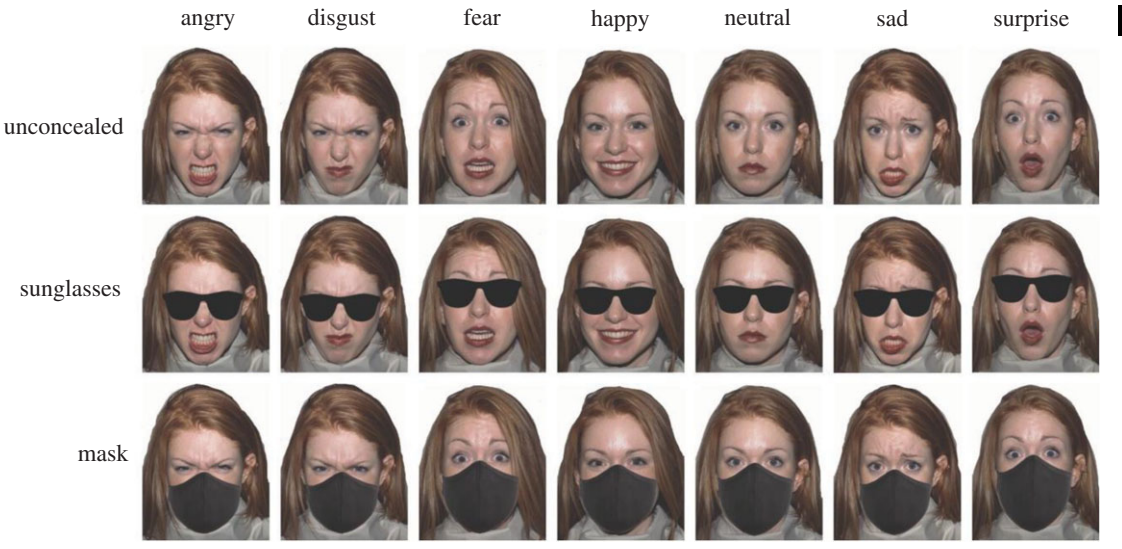

Figure 6. Example stimuli from Experiment 3 (emotional expression). An example from one identity is given across the different expressions and concealment conditions.

4.1. Method

4.1.1. Participants

Participants were taken from the same pool of participants as Experiments 1 and 2. Two participants from the unpractised control group (recruited via Prolific.co) did not have complete data and were removed, leaving 100 participants in this sample. To match the unpractised control group, the first 100 responders in the practised control group and super-recognizers were selected. One of the super-recognizers in this sample also did not have complete data, and so was replaced.

4.1.2. Stimuli

Images of 18 identities (nine female) displaying angry, disgust, fear, happy, neutral, sad and surprise expressions were selected from the NimStim face database [92] for their high emotional validity. Examples of sunglasses and face masks were sourced from the Internet, and added onto the images using Adobe Photoshop (figure 6). We chose to add sunglasses and masks to existing stimuli here (as opposed to using images of faces actually wearing sunglasses and masks as in Experiments 1 and 2) in order to ensure that the intensity and validity of the underlying expression was consistent across concealment conditions.

4.1.3. Procedure

At the beginning of the task, participants were instructed that they would see a face image and were asked to determine the emotion portrayed by the person in the image. Faces were presented as single images on the screen for 1000 ms, immediately followed by a choice of seven emotion response buttons. Participants chose when to proceed by pressing the next key. Seven practise trials, with unconcealed emotional expressions from identities not used in this experiment, ensured that participants understood the task. Each participant completed 63 trials (7 expressions × 3 concealment conditions × 3 repetitions). The three repetitions per condition consisted of different model identities. For each participant, and each emotion condition, nine identities were randomly selected from the possible 18 and were then randomly allocated to concealment condition. This enabled the restriction that the same identity could not be presented expressing the same emotion in different concealment conditions.

4.2. Results

The accuracy data for Experiment 3 is shown in figure 7 and table 1. We conducted a 3 (concealment: unconcealed, mask, sunglasses) × 7 (emotion: angry, fear, disgust, happy, neutral, sad, surprise) × 3

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

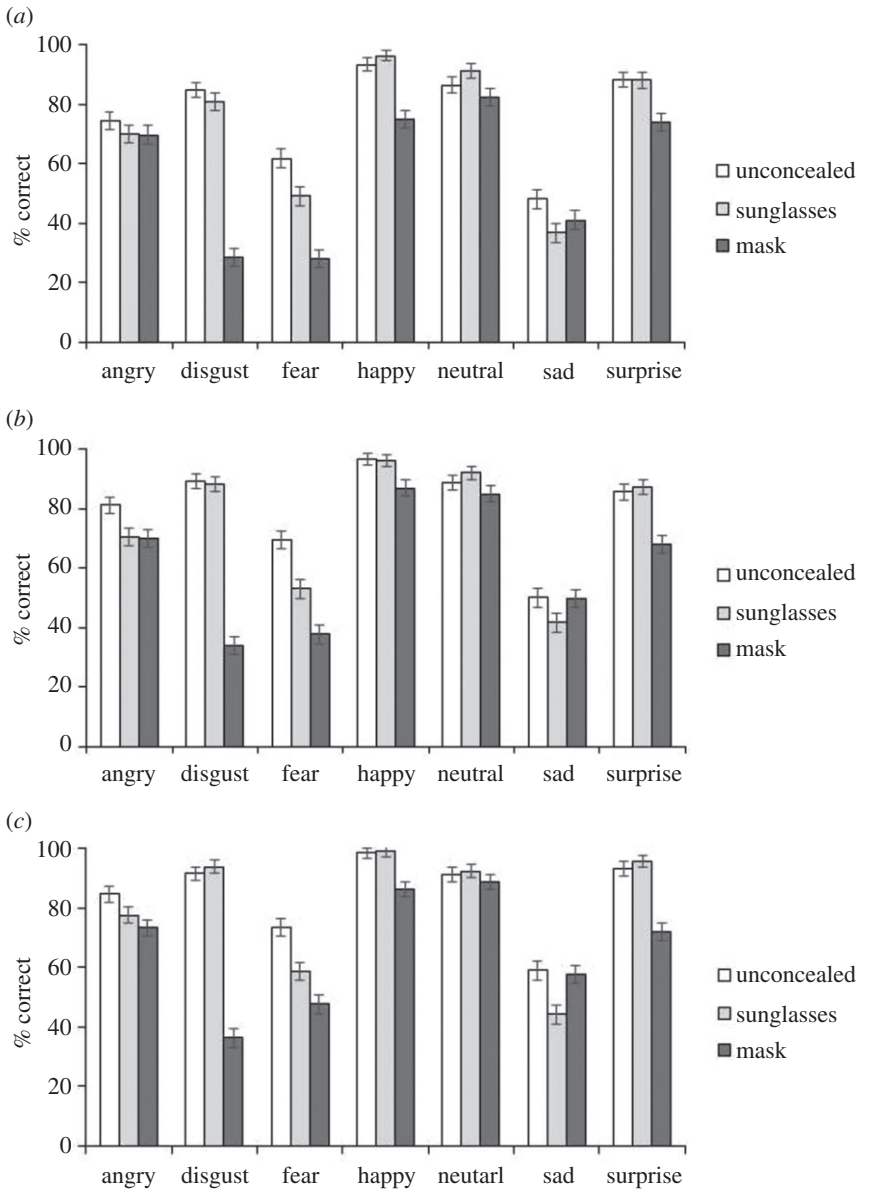

Figure 7. Data for Experiment 3, expression recognition. (*a*) Unpractised controls. (*b*) Practised controls. (*c*) Super-recognizers. Error bars show the within-subjects standard error [74].

(group: unpractised controls, practised controls, super-recognizers) mixed ANOVA. Where sphericity was violated, Greenhouse–Geisser adjustments are reported; *post hoc t*-tests are Bonferroni corrected. We found a significant main effect of group, $F_{2,297} = 26.09$, $p < 0.001$, $\eta_p^2 = 0.15$, whereby super-recognizers ($M = 76.95\%$) outperformed both the practised controls ($M = 72.46\%$) and unpractised controls ($M = 68.89\%$), and the practised controls also performed better than the unpractised controls (all ps < 0.015). There was also a significant main effect of emotion, $F_{5.1,1503.4} = 386.22$, $p < 0.001$, $\eta_p^2 = 0.57$, whereby all emotions differed from each other (ps < 0.015). Happy expressions were recognized best ($M = 92.07\%$), followed by neutral ($M = 88.63\%$), surprise ($M = 83.56\%$), angry ($M = 74.63\%$), disgust ($M = 69.74\%$), fear ($M = 53.19\%$) and then sad ($M = 47.56\%$) expressions. The ANOVA also showed a main effect of concealment $F_{1.9,577.5} = 381.66$, $p < 0.001$, $\eta_p^2 = 0.56$. Unconcealed expressions were recognized most accurately ($M = 80.48\%$), followed by emotions with sunglasses ($M = 76.32\%$), with expressions in the mask condition being most poorly recognized ($M = 61.51\%$; all ps < 0.001).

The pattern of errors made in emotional categorization tasks can be informative (e.g. [62,93]). These data are given in electronic supplementary material, §S3.

These main effects were subsumed within two significant interactions: an emotion by concealment interaction $F_{9.6,2856.2} = 74.98$, $p < 0.001$, $\eta_p^2 = 0.20$, and a small emotion by group interaction $F_{12,3636} = 2.24$, $p = 0.008$, $\eta_p^2 = 0.02$.

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

Table 1. Mean categorization performance (% correct) in Experiment 3. SRs, super-recognizers.

	unconcealed % correct (s.d.)			masks % correct (s.d.)			sunglasses % correct (s.d.)		
	SRs	practised controls	non-practised controls	SRs	practised controls	non-practised controls	SRs	practised controls	non-practised controls
angry	84.67 (20.88)	81.33 (23.36)	74.33 (28.76)	73.33 (21.19)	70.00 (28.23)	69.67 (29.24)	77.67 (24.64)	70.67 (26.50)	70.00 (27.83)
disgust	91.67 (15.98)	89.33 (16.33)	84.67 (20.88)	36.33 (32.17)	34.00 (29.12)	28.67 (26.39)	94.00 (13.72)	88.33 (20.31)	80.67 (26.03)
fear	73.67 (26.08)	69.33 (28.30)	61.67 (31.91)	47.67 (29.30)	37.67 (32.36)	28.00 (26.68)	58.67 (28.87)	53.00 (33.20)	49.00 (29.38)
happy	98.67 (6.56)	96.67 (12.08)	93.33 (18.35)	86.33 (22.27)	87.00 (21.13)	75.00 (29.35)	99.00 (5.71)	96.33 (10.48)	96.33 (14.13)
neutral	91.33 (16.15)	88.67 (21.30)	86.33 (23.73)	88.67 (19.08)	85.00 (22.91)	82.33 (26.15)	92.33 (15.61)	92.00 (17.16)	91.00 (21.11)
sad	59.00 (32.77)	50.00 (31.25)	48.00 (30.45)	57.67 (28.37)	49.67 (30.88)	41.00 (32.77)	44.33 (32.50)	41.67 (33.63)	36.67 (34.33)
surprise	93.33 (15.71)	85.67 (21.32)	88.33 (19.17)	72.00 (30.60)	68.00 (26.77)	73.67 (26.92)	95.67 (13.12)	87.33 (18.82)	88.00 (19.26)

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

In the emotion by concealment interaction, we ran one-way ANOVAs to probe the effect of concealment within each emotion. There was a significant effect of concealment in each of the expressions (angry: $F_{2,585.5} = 12.39$, $p < 0.001$, $\eta_p^2 = 0.04$; disgust: $F_{1.8,535.6} = 532.4$, $p < 0.001$, $\eta_p^2 = 0.67$; fear: $F_{1.9,576.5} = 92.46$, $p < 0.001$, $\eta_p^2 = 0.24$; happy: $F_{1.3,387.8} = 79.87$, $p < 0.001$, $\eta_p^2 = 0.21$; neutral: $F_{2,598} = 9.17$, $p < 0.001$, $\eta_p^2 = 0.03$; sad: $F_{2,598} = 12.06$, $p < 0.001$, $\eta_p^2 = 0.04$ and surprise: $F_{1.9,553.1} = 75.47$, $p < 0.001$, $\eta_p^2 = 0.20$).

For most expressions, performance was reduced on mask trials compared with unconcealed trials (angry: $t_{299} = 4.71$, $p < 0.001$, $d = 0.27$; disgust: $t_{299} = 28.98$, $p < 0.001$, $d = 1.67$; fear: $t_{299} = 14.22$, $p < 0.001$, $d = 0.90$; happy: $t_{299} = 8.95$, $p < 0.001$, $d = 0.52$ and surprise: $t_{299} = 9.84$, $p < 0.001$, $d = 0.57$). This was not the case for neutral ($p = 0.069$) or sad ($p = 0.696$) expressions, where there was no significant difference between unconcealed and mask conditions. Sunglasses significantly affected performance compared with the unconcealed condition, for some of the expressions (angry: $t_{299} = 4.03$, $p < 0.001$, $d = 0.23$; fearful: $t_{299} = 6.93$, $p < 0.001$, $d = 0.40$; sad: $t_{299} = 4.84$, $p < 0.001$, $d = 0.28$); sunglasses had no significant effect on performance for the other expressions ($ps > 0.108$). Expressive faces with sunglasses tended to be more accurately categorized than those with masks (disgust: $t_{299} = 26.83$, $p < 0.001$, $d = 1.55$; fear: $t_{299} = 6.45$, $p < 0.001$, $d = 0.37$; happy: $t_{299} = 9.74$, $p < 0.001$, $d = 0.56$; neutral: $t_{299} = 4.09$, $p < 0.001$, $d = 0.24$; surprise: $t_{299} = 10.06$, $p < 0.001$, $d = 0.58$), with two exceptions. The mask and sunglasses conditions did not significantly differ in the angry expressions ($p = 0.389$), and sad expressions were categorized more accurately in the mask than the sunglasses condition ($t_{299} = 3.43$, $p = 0.003$, $d = 0.20$).

To explore the emotion by group interaction, we ran one-way ANOVAs to probe the effect of group within each emotion. There was a significant effect of group in most of the expressions (angry: $F_{2,297} = 4.62$, $p = 0.011$, $\eta_p^2 = 0.03$; disgust: $F_{2,297} = 11.12$, $p < 0.001$, $\eta_p^2 = 0.07$; fear: $F_{2,297} = 12.56$, $p < 0.001$, $\eta_p^2 = 0.08$; happy: $F_{2,297} = 8.42$, $p < 0.001$, $\eta_p^2 = 0.05$; sad: $F_{2,297} = 8.07$, $p < 0.001$, $\eta_p^2 = 0.05$ and surprise: $F_{2,297} = 6.20$, $p = 0.002$, $\eta_p^2 = 0.04$), with no effect of group in the neutral expression ($p > 0.1$). We followed up these analyses with *t*-tests in all expressions except for neutral.

Super-recognizers outperformed non-practised controls in each of the expressions (angry: $t_{198} = 3.00$, $p = 0.012$; disgust: $t_{198} = 4.42$, $p < 0.001$; fear: $t_{198} = 5.42$, $p < 0.001$; happy: $t_{198} = 3.53$, $p = 0.002$; sad: $t_{198} = 4.12$, $p < 0.001$), except surprise ($p = 0.06$) and neutral. However, performance in the super-recognizer group was not significantly better than that of practised controls for most expressions (all $ps > 0.060$), except fear ($t_{198} = 2.41$, $p = 0.034$) and surprise ($t_{198} = 3.68$, $p < 0.001$), where super-recognizers did outperform practised controls. Practised controls tended to perform better than non-practised controls for some expressions (disgust: $t_{198} = 2.85$, $p = 0.01$; fear: $t_{198} = 2.43$, $p = 0.032$; happy: $t_{198} = 2.73$, $p = 0.014$); whereas for angry, sad and surprise expressions the two control groups did not statistically differ ($ps > 0.081$).

The pattern of errors made in emotional categorization tasks can be informative (e.g. [62,93]). These data are given in electronic supplementary material, §S3.

4.3. Associations across tasks

To explore whether performance on one task was associated with performance on the others, we ran between-task correlations (figure 8). We found a moderate correlation between performance on familiar and unfamiliar face matching ($r = 0.400$, $p < 0.001$), a small correlation between familiar face matching and emotion categorization ($r = 0.200$, $p = 0.001$) and a moderate correlation between unfamiliar face matching and emotion categorization ($r = 0.350$, $p < 0.001$). Please see the electronic supplementary material, §S4 for correlations broken down by condition, with and without the inclusion of the super-recognizer group.

4.4. Discussion

In several countries, the general public have been advised or instructed to wear face masks that cover the nose and mouth regions of the face to help reduce the transmission of COVID-19. Here, we tested the effect of masks on familiar face matching, unfamiliar face matching and emotion categorization as performed by humans. The matching task image pairs always consisted of one unconcealed face image, which was paired with an unconcealed, sunglasses or mask image. Matching accuracy was lower for the mask condition than for two unconcealed faces, regardless of face familiarity. On average performance reduced by 3.5% for masks compared with performance for unconcealed images when the faces were familiar, and by 7.5% for masks compared with unconcealed images when the faces were unfamiliar. There was no difference in accuracy for the mask and sunglasses conditions for

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

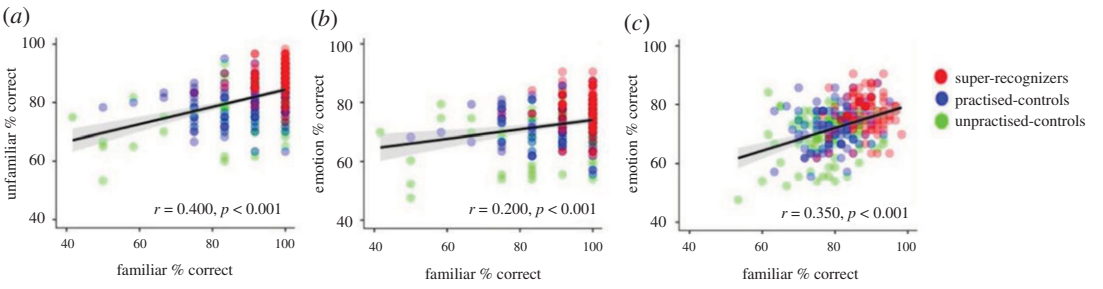

Figure 8. Correlations between average % correct scores on the familiar and unfamiliar face matching tasks (*a*), the familiar matching and emotion categorization tasks (*b*) and the unfamiliar matching and emotion categorization tasks (*c*). Individual data points are semi-transparent; 95% CI is given in grey. There is visible spread in performance at the individual level on all tasks (e.g. [50,94]).

familiar faces; however, for unfamiliar faces, this difference was 2.6% and the effect size was small. Our study has shown a smaller reduction in performance due to masks than was found in a recent paper [44]. This could be an artefact of the stimuli. Carragher & Hancock [44] superimposed surgical masks onto existing images, whereas we used real images of people wearing masks (and sunglasses), perhaps maintaining more of the structure of the face and aiding matching.

A different pattern of results was observed for familiar and unfamiliar faces. While both the sunglasses and mask conditions reduced matching performance for unfamiliar faces, occlusion of the eyes (sunglasses condition) did not impair performance on the familiar face matching task, whereas occlusion of the nose and mouth (mask condition) did reduce accuracy. Despite fewer trials in the familiar compared with unfamiliar experiment, this is unlikely to have driven differences between experiments. We present two possible explanations for the different pattern of results for familiar and unfamiliar faces. First, masks cover more features of a face than sunglasses. The eye region has previously been found to be the most diagnostic cue for face identification [41,42]; therefore, it is possible that the larger area concealed by masks leads to this relative drop in performance. Second, our participants might be more familiar with viewing sunglasses on a face than viewing masked faces in general, and may even have experience with viewing some of our famous faces in sunglasses. It will be interesting to track changes in face identification as the Western world becomes more familiar with viewing faces in masks.

The relatively small reduction in matching accuracy caused by masks or sunglasses is somewhat surprising given larger reductions in performance that have been observed for other forms of image manipulation. For example, Noyes & Jenkins [13] found that matching accuracy dropped by 35% when participants were presented with unfamiliar faces disguised to look unlike themselves, compared with performance for the same faces when they were presented without disguise. We argue that occlusion of facial features is less disruptive to identification than alteration of facial features, such as through make-up to create a disguise [13], contrast negation of facial features [95] or composite images [30,31]. Face images which have been altered contain information which can actively derail an identification by providing erroneous information about the appearance of features. For example, in the case of familiar face matching, an altered face may not match with the expected appearance for the identity [13]. When a face is unfamiliar, altered features may disrupt image-level comparisons and cause confusion for identifications. People may be better equipped to deal with occlusion, in which information from a face is removed rather than altered. This is perhaps achieved by 'filling in the gaps' created by occlusion with stored knowledge about the true appearance of features for a known identity, or by making use of the parts of the face which are visible.

The pattern of results was more complex in the emotion categorization task. In line with previous research that has linked different facial features as diagnostic for different emotions, the effect of masks and sunglasses varied across emotions. The emotional expressions that tend to have diagnostic information in the mouth region, such as disgust, happy and surprise [65,66], were most affected by the masks, showing a large reduction in categorization accuracy for disgust expressions in particular. The angry and fear expressions which tend to have diagnostic information in the eye region [65,66], were found to be affected by both sunglasses and masks; masks especially had a relatively large effect on fear categorization. When only the top half of a disgust face, or only the bottom half of a fear face is presented, there is a similar drop in performance [69]. The sad and neutral expressions, which have

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

diagnostic information across both the eyes and mouth regions [65,66] were differentially affected. Sad expressions were only disrupted in the sunglasses condition, suggesting the eye region is more critical for the accurate identification of this expression. The neutral expression accuracy actually improved in the sunglasses condition, suggesting that the information in the eye region is occasionally over-interpreted as emotional when neutral faces are unconcealed [96]. Although super-recognizers generally outperformed the unpractised control participants, and their mean scores were highest across most expressions, our analyses showed that their performance did not significantly differ from the practised controls' performance, with two exceptions. The super-recognizers had higher accuracy for fear and surprise expressions only.

Our study compared the performance of two groups of control participants and super-recognizers on each task. At the group level, super-recognizers outperformed controls on all tasks. This finding is consistent with previous work that shows that at a group level, super-recognizers consistently outperform controls on a range of face identification tasks [47–49,51,55,73,76]. However, also consistent with past work [48,49,51,55,75,76], there was a large spread in super-recognizers' performance at the individual level. All participants in our super-recognizer participant group had previously scored with 40/40 on the GFMT and with 97% accuracy on the CFMT+. High scores on these standard face recognition ability tests did not guarantee superior performance on our matching tasks. These results highlight the complexities associated with the definition of super-recognizers, and that caution must be exerted when interpreting group-level results, as these results will not necessarily transfer to the level of the individual. In Experiment 1, super-recognizers were slightly more likely to endorse being familiar with a larger number of identities than either control group. It is possible that this had an impact on the results in Experiment 1; however, this cannot be a factor in either of the other experiments. We did not record reaction times for Experiments 1–3, and so cannot comment on whether super-recognizers took longer to respond than control participants, benefiting from a speed–accuracy trade-off. Previous research has not found a difference in reaction times between super-recognizers and controls on face memory tasks [97], but comparing reaction times for these groups on a face matching task could be an interesting avenue for future research. The patterns of results from our two control groups were broadly similar. As we did not have a measure of how many face processing tasks each individual had previously participated in, we cannot say whether practise has or has not influenced the results of these experiments.

Here, we have considered face perception as performed by humans. Interestingly, a recent paper found that an automatic face recognition system, which uses a deep neural network to make similarity comparisons for face images, matched faces with lower accuracy when a face was presented in a mask than no mask, but only for some image types [44]. Specifically, ambient images of the face (similar to the images used in this experiment) with a superimposed mask, led to more errors than more controlled images with a superimposed mask [44]. It is unclear how the same algorithm would perform on a matching task that involved our ambient images for genuine mask wearing, and sunglasses wearing faces. This area requires more systematic investigation.

When we encounter a face in our everyday interactions, the face is typically accompanied by information from the person's body, voice and gait, which can all help inform our perceptions of the person's identity and current emotional state [98–100]. The results of our study are based on the presentation of still images. The additional cues that are available in live viewing environments will probably help overcome the effects of occlusion on identification and emotion perception. Masks do, however, remain an issue for unfamiliar face matching in situations where still images of the face are all that is available to inform an identification. While a new area of investigation, face identification and emotion recognition appear relatively robust against occlusion when considered against the effects of other forms of image manipulation which alter, rather than occlude facial features.

Ethics. This research was granted ethical approval from the University of Huddersfield School of Human and Health Sciences Ethics Committee (SREIC/2020/048).

Data accessibility. The datafiles for this paper are available from the online data depository site Dryad [101] (https://doi.org/10.5061/dryad.2280gb5pq).

Authors' contributions. E.N., J.P.D., N.P., K.L.H.G and K.L.R. designed the study. N.P. programmed the experiment. E.N., K.L.H.G. and K.L.R. analysed the data. E.N., J.P.D., N.P., K.L.H.G. and K.L.R. wrote the manuscript. All authors gave final approval for publication.

Competing interests. We declare we have no competing interests.

Funding. We received no funding for this study.

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

References

1. Bruce V, Young A. 1986 Understanding face recognition. *Br. J. Psychiatry* **77**, 305–327. (doi:10.1111/j.2044-8295.1986.tb02199.x)
2. Stins JF, Roelofs K, Villan J, Kooijman K, Hagenaars MA, Beek PJ. 2011 Walk to me when I smile, step back when I'm angry: emotional faces modulate whole-body approach–avoidance behaviors. *Exp. Brain Res.* **212**, 603–611. (doi:10.1007/s00221-011-2767-z)
3. World Health Organization. 2020 In WHO. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/when-and-how-to-use-masks.
4. Jenkins R, White D, Van Montfort X, Burton AM. 2011 Variability in photos of the same face. *Cognition* **121**, 313–323. (doi:10.1016/j.cognition.2011.08.001)
5. Young AW, Burton AM. 2017 Recognizing faces. *Curr. Dir. Psychol. Sci.* **26**, 212–217. (doi:10.1177/0963721416688114)
6. Ritchie K, Burton AM. 2017 Learning faces from variability. *Q. J. Exp. Psychol. (Colchester)* **70**, 897–905. (doi:10.1080/17470218.2015.1136656)
7. Ritchie KL, Mireku MO, Kramer RSS. 2020 Face averages and multiple images in a live matching task. *Br. J. Psychol.* **111**, 92–102. (doi:10.1111/bjop.12388)
8. O'Toole AJ, Edelman S, Bülthoff HH. 1998 Stimulus-specific effects in face recognition over changes in viewpoint. *Vis. Res.* **38**, 2351–2363. (doi:10.1016/S0042-6989(98)00042-X)
9. Hill H. 1997 Information and viewpoint dependence in face recognition. *Cognition* **62**, 201–222. (doi:10.1016/S0010-0277(96)00785-8)
10. Hancock PJB, Bruce V, Burton AM. 2000 Recognition of unfamiliar faces. *Trends Cogn. Sci.* **4**, 330–337. (doi:10.1016/S1364-6613(00)01519-9)
11. Johnston A, Hill H, Carman N. 1992 Recognising faces: effects of lighting direction, inversion, and brightness reversal. *Perception* **21**, 365–375. (doi:10.1068/p210365)
12. Noyes E, Jenkins R. 2017 Camera-to-subject distance affects face configuration and perceived identity. *Cognition* **165**, 97–104. (doi:10.1016/j.cognition.2017.05.012)
13. Noyes E, Jenkins R. 2019 Deliberate disguise in face identification. *J. Exp. Psychol. Appl.* **25**, 280–290. (doi:10.1037/xap0000213)
14. Bruce V, Henderson Z, Newman C, Burton AM. 2001 Matching identities of familiar and unfamiliar faces caught on CCTV images. *J. Exp. Psychol. Appl.* **7**, 207–218. (doi:10.1037/1076-898X.7.3.207)
15. Dhamecha TI, Singh R, Vatsa M, Kumar A. 2014 Recognizing disguised faces: human and machine evaluation. *PLoS ONE* **9**, e99212. (doi:10.1371/journal.pone.0099212)
16. Terry RL. 1994 Effects of facial transformations on accuracy of recognition. *J. Soc. Psychol.* **134**, 483–492. (doi:10.1080/00224545.1994.9712199)
17. Terry RL. 1993 How wearing eyeglasses affects facial recognition. *Curr. Psychol.* **12**, 151–162. (doi:10.1007/BF02686820)
18. Burton AM, Wilson S, Cowan M, Bruce V. 1999 Face recognition in poor-quality video: evidence from security surveillance. *Psychol. Sci.* **10**, 243–248. (doi:10.1111/1467-9280.00144)
19. Hancock PJB, Bruce V, Burton AM. 2001 Unfamiliar faces: memory or coding? *Trends Cogn. Sci.* **5**, 9. (doi:10.1016/S1364-6613(00)01569-2)
20. Megreya AM, Burton AM. 2006 Unfamiliar faces are not faces: evidence from a matching task. *Mem. Cognit.* **34**, 865–876. (doi:10.3758/BF03193433)
21. Bruce V, Henderson Z, Greenwood K, Hancock PJB, Burton AM, Miller P. 1999 Verification of face identities from images captured on video. *J. Exp. Psychol.: Appl.* **5**, 339–360. (doi:10.1037/1076-898X.5.4.339)
22. Megreya AM, Sandford A, Burton AM. 2013 Matching face images taken on the same day or months apart: the limitations of photo ID. *Appl. Cogn. Psychol.* **27**, 700–706. (doi:10.1002/acp.2965)
23. Andrews S, Jenkins R, Cursiter H, Burton AM. 2015 Telling faces together: learning new faces through exposure to multiple instances. *Q. J. Exp. Psychol.* **68**, 2041–2050. (doi:10.1080/17470218.2014.1003949)
24. Cavazos JG, Noyes E, O'Toole AJ. 2019 Learning context and the other-race effect: strategies for improving face recognition. *Vis. Res.* **157**, 169–183. (doi:10.1016/j.visres.2018.03.003)
25. Burton AM, White D, McNeill A. 2010 The Glasgow face matching test. *Behav. Res. Methods* **42**, 286–291. (doi:10.3758/BRM.42.1.286)
26. Kemp R, Towell N, Pike G. 1997 When seeing should not be believing: photographs, credit cards and fraud. *Appl. Cogn. Psychol.* **11**, 211–222. (doi:10.1002/(SICI)1099-0720(199706)11:3<211::AID-ACP430>3.0.CO;2-O)
27. Megreya AM, Burton AM. 2008 Matching faces to photographs: poor performance in eyewitness memory (without the memory). *J. Exp. Psychol. Appl.* **14**, 364–372. (doi:10.1037/a0013464)
28. White D, Kemp RI, Jenkins R, Matheson M, Burton AM. 2014 Passport officers' errors in face matching. *PLoS ONE* **9**, e103510. (doi:10.1371/journal.pone.0103510)
29. Patterson KE, Baddeley AD. 1977 When face recognition fails. *J. Exp. Psychol. Hum. Learn.* **3**, 406–417. (doi:10.1037/0278-7393.3.4.406)
30. Young AW, Hellawell D, Hay DC. 2013 Configurational information in face perception. *Perception* **42**, 1166–1178. (doi:10.1068/p160747n)
31. Murphy J, Gray KLH, Cook R. 2017 The composite face illusion. *Psychon. Bull. Rev.* **24**, 245–261. (doi:10.3758/s13423-016-1131-5)
32. Gray KLH, Guillemin Y, Cenac Z, Gibbons S, Vestner T, Cook R. 2020 Are the facial gender and facial age variants of the composite face illusion products of a common mechanism? *Psychon. Bull. Rev.* **27**, 62–69. (doi:10.3758/s13423-019-01684-9)
33. Rossion B. 2013 The composite face illusion: a whole window into our understanding of holistic face perception. *Vis. Cogn.* **21**, 139–253. (doi:10.1080/13506285.2013.772929)
34. Farah MJ, Wilson KD, Drain M, Tanaka JN. 1998 What is 'special' about face perception? *Psychol. Rev.* **105**, 482. (doi:10.1037/0033-295X.105.3.482)
35. Davies GM, Flin R. 1984 The man behind the mask – disguise and face recognition. *Hum. Learn.: J. Pract. Res. Appl.* **3**, 83–95.
36. Graham DL, Ritchie KL. 2019 Making a spectacle of yourself: the effect of glasses and sunglasses on face perception. *Perception* **48**, 461–470. (doi:10.1177/0301006619844680)
37. Kramer RSS, Ritchie KL. 2016 Disguising superman: how glasses affect unfamiliar face matching. *Appl. Cogn. Psychol.* **30**, 841–845. (doi:10.1002/acp.3261)
38. Jarudi IN, Sinha P. 2003 Relative contributions of internal and external features to face recognition. Defense Technical Information Center. ADA459650. https://doi.org/10.21236/ada459650.
39. Haig ND. 1985 How faces differ—a new comparative technique. *Perception* **14**, 601–615. (doi:10.1068/p140601)
40. Young AW, Hay DC, McWeeny KH, Flude BM, Ellis AW. 1985 Matching familiar and unfamiliar faces on internal and external features. *Perception* **14**, 737–746. (doi:10.1068/p140737)
41. Vinette C, Gosselin F, Schyns PG. 2004 Spatio-temporal dynamics of face recognition in a flash: it's in the eyes. *Cogn. Sci.* **28**, 289–301.
42. 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. (doi:10.1016/j.cognition.2018.08.004)
43. Mileva M, Burton AM. 2018 Smiles in face matching: idiosyncratic information revealed through a smile improves unfamiliar face matching performance. *Br. J. Psychol.* **109**, 799–811. (doi:10.1111/bjop.12318)
44. Carragher DJ, Hancock P. 2020 Surgical face masks impair human face matching performance for familiar and unfamiliar faces. *Cogn. Res.: Princ. Implic.* **5**, 1–15. (doi:10.1186/s41235-020-00258-x)
45. Noyes E, Phillips PJ, O'Toole AJ. 2017 What is a super-recogniser? In *Face processing: systems, disorders, and cultural differences* (eds M Bindemann, AM Megreya), pp. 173–202. New York, NY: Nova.
46. Ramon M, Bobak AK, White D. 2019 Super-recognizers: from the lab to the world and back again. *Br. J. Psychol.* **110**, 461–479. (doi:10.1111/bjop.12368)
47. Robertson DJ, Noyes E, Dowsett AJ, Jenkins R, Burton AM. 2016 Face recognition by Metropolitan Police super-recognisers. *PLoS ONE* **11**, e0150036. (doi:10.1371/journal.pone.0150036)
48. Davis JP, Bretfelean D, Belanova E, Thompson T. 2020 Super-recognisers: face recognition performance after variable delay intervals. *Appl. Cogn. Psychol.* **34**, 1350–1368. (doi:10.1002/acp.3712)

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021

49. Bobak AK, Hancock PJB, Bate S. 2016 Super-recognisers in action: evidence from face-matching and face memory tasks. *Appl. Cogn. Psychol.* **30**, 81–91. (doi:10.1002/acp.3170)
50. Noyes E, Hill MQ, O'Toole AJ. 2018 Face recognition ability does not predict person identification performance: using individual data in the interpretation of group results. *Cogn. Res. Princ. Implic.* **3**, 1–3. (doi:10.1186/s41235-018-0117-4)
51. Phillips PJ *et al.* 2018 Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. *Proc. Natl Acad. Sci. USA* **115**, 6171–6176. (doi:10.1073/pnas.1721355115)
52. Bobak AK, Parris BA, Gregory NJ, Bennetts RJ, Bate S. 2017 Eye-movement strategies in developmental prosopagnosia and 'super' face recognition. *Q. J. Exp. Psychol. (Colchester)* **70**, 201–217. (doi:10.1080/17470218.2016.1161059)
53. Sekiguchi T. 2011 Individual differences in face memory and eye fixation patterns during face learning. *Acta Psychol.* **137**, 1–9. (doi:10.1016/j.actpsy.2011.01.014)
54. Peterson MF, Eckstein MP. 2012 Looking just below the eyes is optimal across face recognition tasks. *Proc. Natl Acad. Sci. USA* **109**, E3314–E3323. (doi:10.1073/pnas.1214269109)
55. Davis JP, Tamonyte D. 2017 Masters of disguise: super-recognisers' superior memory for concealed unfamiliar faces. In *7th Int. Conf. on Emerging Security Technologies (EST)*. https://doi.org/10.1109/est.2017.8090397.
56. Young AW, Noyes E. 2019 We need to talk about super-recognizers Invited commentary on: Ramon, M., Bobak, A. K., & White, D. Super-recognizers: from the lab to the world and back again. *Br. J. Psychiatry* **110**, 492–494. (doi:10.1111/bjop.12395)
57. Ekman P, Friesen WV. 1971 Constants across cultures in the face and emotion. *J. Pers. Soc. Psychol.* **17**, 124–129. (doi:10.1037/h0030377)
58. Ekman P. 1992 An argument for basic emotions. *Cogn. Emot.* **6**, 169–200. (doi:10.1080/02699939208411068)
59. Ekman P. 1992 Are there basic emotions? *Psychol. Rev.* **99**, 550–553. (doi:10.1037/0033-295X.99.3.550)
60. Jack RE, Garrod OGB, Yu H, Caldara R, Schyns PG. 2012 Facial expressions of emotion are not culturally universal. *Proc. Natl Acad. Sci. USA* **109**, 7241–7244. (doi:10.1073/pnas.1200155109)
61. Batty M, Taylor MJ. 2003 Early processing of the six basic facial emotional expressions. *Brain Res. Cogn. Brain Res.* **17**, 613–620. (doi:10.1016/S0926-6410(03)00174-5)
62. Prkachin GC. 2003 The effects of orientation on detection and identification of facial expressions of emotion. *Br. J. Psychol.* **94**, 45–62. (doi:10.1348/000712603762842093)
63. Young AW, McWeeny KH, Hay DC, Ellis AW. 1986 Matching familiar and unfamiliar faces on identity and expression. *Psychol. Res.* **48**, 63–68. (doi:10.1007/BF00309318)
64. Ekman P, Friesen WV. 1976 Measuring facial movement. *Environ. Psychol. Nonverbal Behav.* **1**, 56–75. (doi:10.1007/BF01115465)
65. Smith ML, Cottrell GW, Gosselin F, Schyns PG. 2005 Transmitting and decoding facial expressions. *Psychol. Sci.* **16**, 184–189. (doi:10.1111/j.0956-7976.2005.00801.x)
66. 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**, e0177239. (doi:10.1371/journal.pone.0177239)
67. Schyns PG, Bonnar L, Gosselin F. 2002 Show me the features! Understanding recognition from the use of visual information. *Psychol. Sci.* **13**, 402–409. (doi:10.1111/1467-9280.00472)
68. Schurgin MW, Nelson J, Iida S, Ohira H, Chiao JY, Franconeri SL. 2014 Eye movements during emotion recognition in faces. *J. Vis.* **14**, 14. (doi:10.1167/14.13.14)
69. Calder AJ, Young AW, Keane J, Dean M. 2000 Configural information in facial expression perception. *J. Exp. Psychol. Hum. Percept. Perform.* **26**, 527–551. (doi:10.1037/0096-1523.26.2.527)
70. Roberson D, Kikutani M, Döge P, Whitaker L, Majid A. 2012 Shades of emotion: what the addition of sunglasses or masks to faces reveals about the development of facial expression processing. *Cognition* **125**, 195–206. (doi:10.1016/j.cognition.2012.06.018)
71. Rhodes G, Pond S, Burton N, Kloth N, Jeffery L, Bell J, Ewing L, Calder AJ, Palermo R. 2015 How distinct is the coding of face identity and expression? Evidence for some common dimensions in face space. *Cognition* **142**, 123–137. (doi:10.1016/j.cognition.2015.05.012)
72. Connolly HL, Young AW, Lewis GJ. 2019 Recognition of facial expression and identity in part reflects a common ability, independent of general intelligence and visual short-term memory. *Cogn. Emot.* **33**, 1119–1128. (doi:10.1080/02699931.2018.1535425)
73. Ali L, Davis JP. 2020 The positive relationship between face identification and facial emotion perception ability extends to super-recognisers. *PsyArXiv*. (doi:10.31234/osf.io/4ktrj)
74. Biotti F, Cook R. 2016 Impaired perception of facial emotion in developmental prosopagnosia. *Cortex* **81**, 126–136. (doi:10.1016/j.cortex.2016.04.008)
75. Davis JP, Lander K, Evans R, Jansari A. 2016 Investigating predictors of superior face recognition ability in police super-recognisers. *Appl. Cogn. Psychol.* **30**, 827–840. (doi:10.1002/acp.3260)
76. Bobak AK, Dowsett AJ, Bate S. 2016 Solving the Border control problem: evidence of enhanced face matching in individuals with extraordinary face recognition skills. *PLoS ONE* **11**, e0148148. (doi:10.1371/journal.pone.0148148)
77. Belanova E, Davis JP, Thompson T. 2019 The part-whole effect in super-recognisers. *PsyArXiv*. (doi:10.31234/osf.io/mkhnd)
78. Satchell L, Davis JP, Julle-Danière E, Tupper N, Marshman P. 2019 Recognising faces but not traits: accurate personality judgment from faces is unrelated to superior face memory. *J. Res. Pers.* **79**, 49–58. (doi:10.1016/j.jrp.2019.02.002)
79. Davis JP, Bretfelean LD, Belanova E, Thompson T. 2020 Super-recognisers: face recognition performance after variable delay intervals. *Appl. Cogn. Psychol.* **34**, 1350–1368. (doi:10.1002/acp.3712)
80. Correll J, Ma DS, Davis JP. 2020 Perceptual tuning through contact? Contact interacts with perceptual (not memory-based) face-processing ability to predict cross-race recognition. *J. Exp. Soc. Psychol.* **92**, 104058. (doi:10.1016/j.jesp.2020.104058)
81. Russell R, Duchaine B, Nakayama K. 2009 Super-recognizers: people with extraordinary face recognition ability. *Psychon. Bull. Rev.* **16**, 252–257. (doi:10.3758/PBR.16.2.252)
82. Bobak AK, Pampoulov P, Bate S. 2016 Detecting superior face recognition skills in a large sample of young British adults. *Front. Psychol.* **7**, 1378. (doi:10.3389/fpsyg.2016.01378)
83. Ritchie KL, Smith FG, Jenkins R, Bindemann M, White D, Burton AM. 2015 Viewers base estimates of face matching accuracy on their own familiarity: explaining the photo-ID paradox. *Cognition* **141**, 161–169. (doi:10.1016/j.cognition.2015.05.002)
84. Crump MJC, McDonnell JV, Gureckis TM. 2013 Evaluating Amazon's mechanical turk as a tool for experimental behavioral research. *PLoS ONE* **8**, e57410. (doi:10.1371/journal.pone.0057410)
85. Germine L, Nakayama K, Duchaine BC, Chabris CF, Chatterjee G, Wilmer JB. 2012 Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychon. Bull. Rev.* **19**, 847–857. (doi:10.3758/s13423-012-0296-9)
86. Woods AT, Velasco C, Levitan CA, Wan X, Spence C. 2015 Conducting perception research over the internet: a tutorial review. *PeerJ* **3**, e1058. (doi:10.7717/peerj.1058)
87. Cousineau D. 2005 Confidence intervals in within-subject designs: a simpler solution to Loftus and Masson's method. *Tutor. Quant. Methods Psychol.* **1**, 42–45. (doi:10.20982/tqmp.01.1.p042)
88. Megreya AM, Burton AM. 2007 Hits and false positives in face matching: a familiarity-based dissociation. *Percept. Psychophys.* **69**, 1175–1184. (doi:10.3758/BF03193954)
89. Dunn JD, Summersby S, Towler A, Davis JP, White D. 2020 UNSW Face Test: a screening tool for super-recognizers. *PLoS ONE* **15**, e0241747. (doi:10.1371/journal.pone.0241747)
90. McCaffery JM, Robertson DJ, Young AW, Burton AM. 2018 Individual differences in face identity processing. *Cogn. Res. Princ. Implic.* **3**, 21. (doi:10.1186/s41235-018-0112-9)
91. Yovel G, Wilmer JB, Duchaine B. 2014 What can individual differences reveal about face processing? *Front. Hum. Neurosci.* **8**, 562. (doi:10.3389/fnhum.2014.00562)
92. Tottenham N *et al.* 2009 The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res.* **168**, 242–249. (doi:10.1016/j.psychres.2008.05.006)
93. Derntl B, Seidel E-M, Kainz E, Carbon C-C. 2009 Recognition of emotional expressions is affected by inversion and presentation time. *Perception* **38**, 1849–1862. (doi:10.1068/p6448)
94. Fysh MC, Stacchi L, Ramon M. 2020 Differences between and within individuals, and subprocesses of face cognition: implications for theory, research and personnel selection. *R. Soc. Open Sci.* **7**, 200233. (doi:10.1098/rsos.200233)

95. Gilad S, Meng M, Sinha P. 2009 Role of ordinal contrast relationships in face encoding. *Proc. Natl Acad. Sci. USA* **106**, 5353–5358. (doi:10.1073/pnas.0812396106)
96. Lee E, Kang JI, Park IH, Kim J-J, An SK. 2008 Is a neutral face really evaluated as being emotionally neutral? *Psychiatry Res.* **157**, 77–85. (doi:10.1016/j.psychres.2007.02.005)
97. Belanova E, Davis JP, Thompson T. 2018 Cognitive and neural markers of super-recognisers' face processing superiority and enhanced cross-age effect. *Cortex* **98**, 91–101. (doi:10.1016/j.cortex.2018.07.008)
98. Aviezer H, Trope Y, Todorov A. 2012 Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science* **338**, 1225–1229. (doi:10.1126/science.1224313)
99. Aviezer H, Hassin RR, Ryan J, Grady C, Susskind J, Anderson A, Moscovitch M, Bentin S. 2008 Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychol. Sci.* **19**, 724–732. (doi:10.1111/j.1467-9280.2008.02148.x)
100. Rice A, Phillips PJ, Natu V, An X, O'Toole AJ. 2013 Unaware person recognition from the body when face identification fails. *Psychol. Sci.* **24**, 2235–2243. (doi:10.1177/0956797613492986)
101. Noyes E, Davis JP, Petrov N, Gray KLH, Ritchie KL. 2021 Data from: The effect of face masks and sunglasses on identity and expression recognition with super-recognizers and typical observers. Dryad Digital Repository. (doi:10.5061/dryad.2280gb5pq).

Downloaded from https://royalsocietypublishing.org/ on 25 March 2021