The super-recogniser advantage extends to the detection of hyper-realistic face masks

Article:

Robertson, David J., Davis, Josh P., Sanders, Jet and Towler, Alice (2024) The super-recogniser advantage extends to the detection of hyper-realistic face masks. Applied Cognitive Psychology, 38 (4). ISSN 0888-4080

https://doi.org/10.1002/acp.4222

Received: 6 October 2022 | Revised: 28 May 2024 | Accepted: 13 June 2024
DOI: 10.1002/acp.4222

RESEARCH ARTICLE

WILEY

The super-recogniser advantage extends to the detection of hyper-realistic face masks

David J. Robertson[1] | Josh P. Davis[2] | Jet G. Sanders[3] | Alice Towler[4]

[1]Department of Psychological Sciences and Health, University of Strathclyde, Glasgow, UK
[2]School of Human Sciences, Institute of Lifecourse Development, University of Greenwich, London, UK
[3]Department of Psychology and Behavioural Sciences, London School of Economics and Political Science, London, UK
[4]School of Psychology, The University of Queensland, Brisbane, Australia

Correspondence
David J. Robertson, Department of Psychological Sciences and Health, University of Strathclyde, Glasgow G1 1QE, UK.
Email: david.j.robertson@strath.ac.uk

Abstract

Hyper-realistic silicone masks provide a viable route to identity fraud. Over the last decade, more than 40 known criminal acts have been committed by perpetrators using this type of disguise. With the increasing availability and bespoke sophistication of these masks, research must now focus on ways to enhance their detection. In this study, we investigate whether super-recognisers (SRs), people who excel at identity recognition, are more likely to detect this type of fraud, in comparison to typical-recogniser controls. Across three tasks, we examined mask detection rates in the absence of a pre-task prompt (covert task), and again after making participants aware of their use in criminal settings (explicit task). Finally, participants were asked to indicate which aspects of the masks could support their detection (regions of interest task). The findings show an SR advantage for the detection of hyper-realistic masks across the covert and explicit mask detection tasks. In addition, the eye, mouth, and nose regions appear to be particularly indicative of the presence of a mask. The lack of natural skin texture, proportional features, expressiveness, and asymmetry are also salient cues. The theoretical and applied implications of these findings are discussed.

KEYWORDS
deception, disguise, hyper-realistic masks, identity fraud, super-recognisers

1 | INTRODUCTION

Identity verification is a critical process in policing, the criminal justice system, and at border control (Davis & Robertson, 2020; Robertson et al., 2015, 2019). It is therefore not surprising that identity fraudsters continually seek new ways to disguise their appearance to evade the authorities (Dhamecha et al., 2014; Noyes & Jenkins, 2019). While traditional approaches to facial disguise have tended to use individual items such as wigs, beards, hats, makeup and glasses (Dhamecha et al., 2014; Kramer & Ritchie, 2016; Noyes & Jenkins, 2019; Righi et al., 2012; Terry, 1994), within the last decade, a more sophisticated form of identity fraud has emerged in the form of hyper-realistic silicone masks (Sanders et al., 2017, 2019; Sanders & Jenkins, 2018, 2021; see Figure 1 for an example).

These masks first emerged for use in the entertainment industry, and typically comprise of a single layer of flexible silicone that covers the whole head, neck, and upper chest area, with advanced structural components that allow them to fit seamlessly to the wearer's face (Sanders & Jenkins, 2021). Originally prohibitively expensive, the cost of a hyper-realistic mask has decreased significantly in recent years ($400–$2000; Bernstein, 2010), increasing the number of units sold (estimated 2000–4000 per year; Sanders & Jenkins, 2021). Importantly, manufacturers are creating increasingly sophisticated bespoke products that can include hand-finished detailed paintwork and

FIGURE 1 An example of a hyper-realistic mask in use. This figure shows a hyper-realistic mask (left), an undisguised individual (middle), and the individual wearing the mask (right) (Sanders et al., 2017; reproduced with permission).

punched human hair (Sanders & Jenkins, 2021). Such 'complete' disguises provide a pathway for fraudsters to emulate the appearance of an individual in a stolen face-photo ID document, or to obscure their own identity when committing a crime. There are now several real-world examples of such fraud attacks.

In 2010, a young Asian male used a hyper-realistic mask to impersonate an elderly White man whose passport he had stolen. He passed through several identity checks at Hong Kong International Airport and successfully boarded a flight to Canada. During the flight, he removed the mask in the lavatory and on returning to his seat, only then did a fellow passenger detect the deception (Zamost, 2010). In the same year, in the USA, White male Conrad Zdzierak robbed several banks while wearing a Black male mask. CCTV footage and eyewitness accounts led to a Black male being apprehended for the crimes. The perpetrator was only caught, and the innocent man released, when his girlfriend alerted police to the existence of the mask (Gardner, 2010). More recently, a conman defrauded individuals out of an estimated $90 m by using a hyper-realistic mask to impersonate the French Minister of Defence (Schofield, 2019). Indeed, over the last decade, there have been no fewer than 41 reported cases of fraud attacks using these masks, with the majority occurring in the United States (66%), as robberies (76%), by young (94%), White (61%), male (94%) perpetrators, wearing old (70%) White (89%), male (90%) masks, with only 56% of masks being detected by witnesses during the crime and only 61% of fraudsters being caught (Sanders & Jenkins, 2021).

Such real-world examples provide a compelling account of the utility of these masks in criminal settings, and recent work in applied psychological science supports this view. Three papers have investigated mask detection rates using a *covert* procedure that examined the extent to which observers detected these masks spontaneously, and an *explicit* detection task in which observers were made aware that they may encounter this type of disguise (Sanders et al., 2017, 2019; Sanders & Jenkins, 2018). These studies show that masked face photos can go entirely undetected (0%) under covert conditions, with detection rates rising to just 1.3% when a mask wearer was present in person. Explicit detection rates were also poor. For those without prior knowledge that such masks existed, detection rates were 55% for masked face photos and 43% when asked to judge the appearance of a mask wearer in person. These findings support the real-world evidence for the effectiveness of this disguise, and research must now focus on developing ways to enhance the detectability of hyper-realistic masks in critical applied contexts. One potential route to achieving this may be through an individual differences approach that has previously revealed exceptional face processing skills in some people—super-recognisers (SRs; Bate et al., 2021; Davis, 2019, 2020; Davis et al., 2016; Ramon, 2021; Robertson et al., 2016; Russell et al., 2009).

A 'super-recogniser' is the current general term for an individual who performs at the top end of a normally distributed face recognition continuum (Bobak et al., 2016; McCaffery et al., 2018; Russell et al., 2009; Verhallen et al., 2017; Wilmer, 2017). The SR ability is likely to be hereditary (Shakeshaft & Plomin, 2015; Wilmer et al., 2010), and it has been consistently reported in face matching and memory tasks that typically use novel instances of a target face without occlusion or disguise (Davis et al., 2016; Robertson, Black, et al., 2020). However, recent research has suggested that the SR advantage for face identification extends to situations in which the target *is* disguised (e.g., with sunglasses, hats, balaclavas) including COVID-19 masks (e.g., Davis & Tamonytė, 2017; Noyes et al., 2021). This research further supports the selection and recruitment of SRs for roles in which accurate identity verification, even in the presence of traditional disguise, is the critical task. However, hyper-realistic masks do not represent a typical form of disguise, and it has not yet been established whether SRs would be more likely to detect these masks than typical recogniser controls.

One recent study did investigate the link between face identification performance and hyper-realistic mask detection in a live passport-checking task (Robertson, Sanders, et al., 2020). The study reported that participants who detected the mask in the covert detection condition showed Glasgow Face Matching Test scores (Burton et al., 2010) that were within the normal, rather than exceptional, range. At first glance, this could suggest that there may not be a link between face identification performance (i.e., the SR ability) and hyper-realistic mask detection. However, this finding was generated by a small sample size, it did not include verified SRs, and the single live mask detection trial precluded a full individual differences analysis. Therefore, it is critical, given the increasing use of hyper-realistic masks in criminal activities, that we investigate whether the SR advantage generalises to this type of full-face disguise.

To that end, here we present a group of super-recognisers and typical recogniser control participants (motivated + standard undergraduate sample) with a covert mask detection task followed by an explicit mask detection task. Participants were asked to assume the role of a border control officer, and to detect anything 'suspicious' about the face of a traveller presented on each trial (*covert detection*), and then, having been briefed on hyper-realistic mask fraud, they were shown the faces again and asked to detect them (*explicit detection*). At the end of the study, in a 'regions of interest' task, participants were asked to indicate which aspects of the face were most indicative of the presence of a mask, and to provide any further qualitative responses that could support the detection of this type of disguise.

2 | METHOD

2.1 | Ethics and data availability statement

The research reported in this paper received concurrent approval from the Ethics Committee of the University of Strathclyde, and the Ethics Committee of the University of Greenwich (01/15/2018/A). A copy of the dataset that supports analyses reported in this paper is available via https://osf.io/uhb7f/.

2.2 | Participants

2.2.1 | G*power analysis

A G*Power analysis (http://www.gpower.hhu.de/) with the alpha level set at .05, and power set at .80, recommended a minimum sample size of 33 participants per group to detect a medium effect size (set at .32; based on the group effect size reported in Noyes et al., 2021). Our final sample consisted of a minimum of 157 participants per group, thereby achieving the required statistical power.

2.2.2 | Super-recognisers (SRs)

We recruited 628 White participants from an existing University of Greenwich database created by author JPD. All volunteers had previously provided consent to be invited to future online research in this field. Each had previously completed the Cambridge Face Memory Test-Long Version (CFMT+; see Russell et al., 2009) and the Glasgow Face Matching Test (GFMT; Burton et al., 2010) and they provided consent for these scores to be used in the present study. From this sample, 210 participants achieved the ≥93% CFMT+ accuracy rate required for super-recognition status (Bobak et al., 2016; Davis et al., 2016). However, 43 participants did not have their SR status confirmed by similarly exceptional scores on a second test (i.e., the GFMT; ≥95% cut-off; the initial SR mean in this dataset) and were excluded from the group. Therefore, the final super-recogniser (SR) group consisted of 157 participants. Participants recruited from this database chose to take part in these studies to test their own ability and to receive their performance scores. No further inducement or monetary reimbursement was offered for participation. Group demographics for the SRs and for each control group are presented in Table 1.

2.2.3 | Motivated controls (MCs)

The Motivated Control group were selected from the remaining 418 participants, recruited from the Greenwich database, who did not meet the CFMT+ threshold score for super-recognition. To generate discrete groups, a further 180 of these participants were excluded as although they showed typical performance on the CFMT+, their GFMT scores fell within the exceptional range. Therefore, the final motivated control (MC) group consisted of 238 participants who each scored within the typical range on both the CFMT+ and the GFMT. We label this group as 'motivated' as they are a distinct group of individuals with an interest in their face recognition ability, and achieving high scores on face tests, but whose actual scores fall below the threshold for super-recognition.

2.2.4 | Undergraduate controls (UCs)

For the Undergraduate Control group, 200 White participants were recruited separately from the SRs and MCs using the University of Strathclyde Psychology Participant Pool. There were 19 exclusions based on a failure to complete the full study, and 8 further exclusions based on CFMT+ scores which were above the typical-recogniser range. As this control group was required to complete the CFMT+, our primary grouping measure, and the three mask tasks in one session, time constraints meant that we did not require them to complete the GFMT. Therefore, the final undergraduate control (UC) group consisted of 173 participants who scored within the typical range on the CFMT+. These participants had chosen to take part in this study in return for course credit and may therefore better reflect levels of motivation and performance in the general population.

2.2.5 | Group demographics

As shown in the Table 1, there was a similar gender balance for SRs and MCs and there was no significant difference in age between these two groups, $t < 1$. The UC group, as expected, had a higher proportion of females and was significantly younger than both the SR group, $t(328) = 20.45$, $p < .001$, $d = 2.25$, and the MC group, $t(409) = 18.05$, $p < .001$, $d = 1.80$.

TABLE 1 Group demographics summary.

	Undergraduate controls $N = 173$ 88% female			Motivated controls $N = 238$ 56% female			Super-recognisers $N = 157$ 66% female		
	Mean	SD	Range	Mean	SD	Range	Mean	SD	Range
Age	21	6	18–59	39	12	19–72	38	10	20–64
CFMT+	70	13	42–92	78	14	30–92	95	2	93–100
GFMT	-	-	-	87	5	68–93	97	1	95–100

Note: This table presents the sample characteristics for the three groups. Mean percentage score is shown for the CFMT+ and the GFMT.

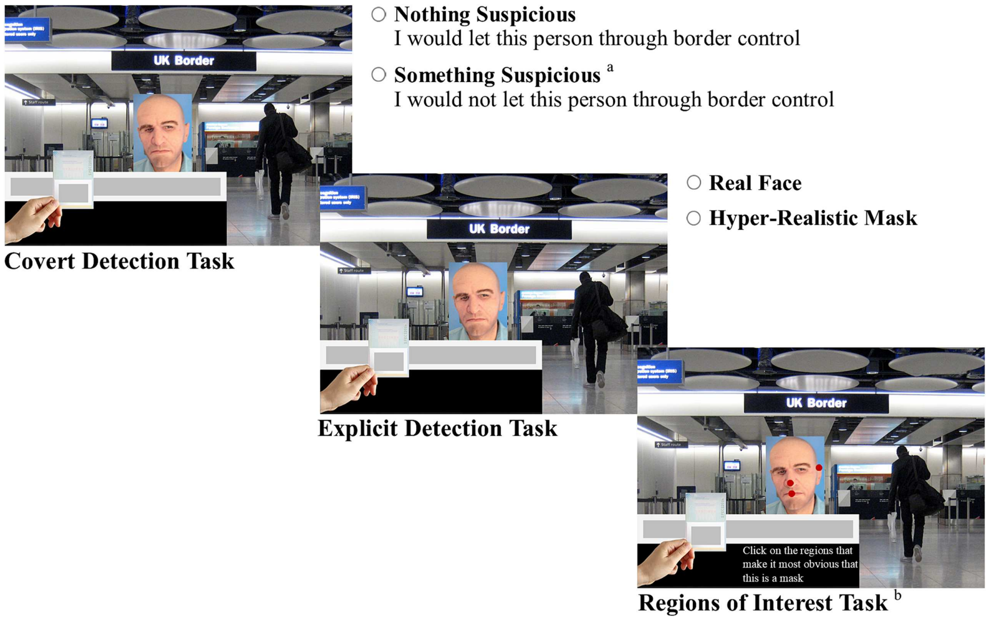

FIGURE 2 Example trials, task procedure, and response options. The 'traveller' is wearing a hyper-realistic mask in this image (authors own). Due to copyright restrictions, we cannot show the background image used in the study, however, the one presented in this figure is a close approximation of it (CC BY 2.0; attribution Danny Howard; https://tinyurl.com/3d5878x9; no changes made to original image; additional content overlaid). [a] If a 'Something Suspicious' response was selected, a text box appeared for the participants to elaborate on the source of their suspicion. [b] Each trial also contained a text box for the participants to provide any further points which might help increase the detection of this disguise.

2.3 | Stimuli and apparatus

2.3.1 | Covert detection task, explicit detection task, regions of interest task

Example stimuli are presented in Figure 2. Thirty-six face photos were selected from an existing set used by Sanders et al. (2017; 2018; 2019); each image was reduced in size to 135px × 180px and was presented in colour. Twelve of the images show a person wearing a hyper-realistic mask, 8 (67%) were male, 8 (67%) were categorised as being towards the 'older' age range in appearance (i.e., 50+; see Sanders & Jenkins, 2021), and all were classed as portraying White individuals (100%). The remaining 24 images consisted of real faces, 16 (67%) were male, 16 (67%) were categorised as being towards the older age range (67%), and all were White (100%). In this way, the demographics of our set of masked faces approximates the frequency of their reported real-world use (Sanders & Jenkins, 2021), and both the masked faces and the real faces had similar profiles in terms of age, sex, ethnic group, and presence or absence of spectacles, hair, and hats. An image of a real border control scene was downloaded from Google Images and reduced in size to 600px × 378px, upon which each face photo, and an ID checking desk set up, was overlaid. Qualtrics (https://www.qualtrics.com/) was used to present the stimuli and collect the data. Participants were instructed not to complete the experiment on a smartphone and were asked to confirm that they were using a desktop PC/Mac, laptop, or a reasonably sized tablet.

2.4 | Procedure

The mask tasks were presented in a fixed order to each participant (covert detection, explicit detection, regions of interest). The UC group also completed the CFMT+ at the start of their testing session. For the *covert mask detection task*, participants were instructed to assume the role of a border control officer trying to detect suspicious travellers making their way through the airport. At no point prior to or during the instruction phase was any reference made to hyper-realistic masks. They were then presented with 36 face photos (12 mask/24 real face), see Figure 2 for example trials, and were asked to decide whether the individual looked 'suspicious', 'is there anything not quite right about their appearance', 'are they trying to deceive you in anyway'. The response options were: Nothing Suspicious—I WOULD let this person through Border Control; Something Suspicious—I WOULD NOT let this person through border control. A 'something suspicious' response would generate an onscreen text box for the participant to elaborate on their suspicions, providing the opportunity for spontaneous mask detection to be recorded.

For the *explicit mask detection task*, participants were made aware that some of the faces they had seen in the previous block were people wearing hyper-realistic masks. They were then provided with an onscreen information sheet which outlined what a hyper-realistic mask was, how they have been used in identity fraud, alongside example images. The 36 face photos were then shown again, this time with the explicit instruction to detect whether the individual was wearing a mask or not. The response options were 'Real Face' or 'Hyper-Realistic Mask'.

For the final *regions of interest task*, participants were presented with the 12 trials that contained a masked face, and on each trial, they were asked to click on the regions of the face that 'make it most obvious that this is a hyper-realistic mask'. There were 12 clickable regions of the face (hair/hairline, forehead, left eye, right eye, midpoint between the eyes, left cheek, right cheek, nose, left ear (when visible), right ear (when visible), mouth, chin/neck), these region 'boxes' were not visible to participants on screen, and they could make up to 10 clicks. For each of the 12 trials, a text entry box was also present, and participants were asked to enter anything else they would like to note that might help a border control officer detect this type of disguise.

Prior to the experimental debrief, participants were asked whether they had *any prior knowledge* of hyper-realistic masks (had they seen them before, did they know of their use in fraud, had they taken part in other mask related research studies; 56% of UCs; 44% of MCs; and 47% of SRs answered 'yes' to at least one of these questions). In addition, for *quality control* purposes, participants were asked to confirm that the study presentation (i.e., clarity of text and stimuli) allowed them to complete the study to the best of their ability. Three participants reported internet server issues but this did not prevent them from submitting a full response set and so they were retained within the dataset.

The order of presentation of the trials within each block was randomised across participants. Each task was self-paced, but participants were instructed that they should expect to complete the study within 30–35 min for the SR and MC group, and 50–55 min for the UC group (additional time for completion of the CFMT+), and in one sitting.

3 | RESULTS

3.1 | Covert detection task

Performance on the covert detection task reflects the extent to which participants correctly attribute suspicion to a masked face trial. Using signal detection analysis (Stanislaw & Todorov, 1999), a hit was categorised as a 'something suspicious' response to a masked face trial, while a false alarm was categorised as a 'something suspicious' response to a real face trial. These values were used to calculate overall detection sensitivity (d prime; d') and response bias (criterion; c). In this analysis, and throughout the results section, when the assumption of homogeneity of variance between the groups has not been violated, we use ANOVA followed by Tukey–Kramer post hoc tests. When the assumption of homogeneity of variance between the groups has been violated, we report Welch's ANOVA (Delacre et al., 2019), and use Games-Howell post hoc tests.

One-way between subjects ANOVAs with group as the factor (UC, MC, SR) revealed a main effect for hits, $F_{Welch}(2, 353) = 45.51$, $p < .001$, $\omega^2 = .073$, false alarms, $F_{Welch}(2, 345) = 3.05$, $p = .049$, $\omega^2 = .004$, detection sensitivity (d'), $F(2, 567) = 27.17$, $p < .001$, $\eta^2 = .088$, and response bias (c), $F(2, 567) = 32.23$, $p < .001$, $\eta^2 = .102$. As seen in Figure 3, Games-Howell post hoc tests revealed that hit rates were highest for the SRs ($M = 72\%$, $SD = 22\%$) followed by the MCs ($M = 61\%$, $SD = 24\%$), $p < .001$, 95% CI [6.03, .17.20], and then the UC group ($M = 47\%$, $SD = 26\%$) in turn, $p < .001$, 95% CI [7.94, .19.92]. For false alarms, only the difference between the SRs ($M = 14\%$, $SD = 13\%$) and the UCs ($M = 11\%$, $SD = 10\%$) was significant, $p = .047$, 95% CI [.03, 6.01].

Importantly, as seen in Figure 3, Tukey–Kramer post hoc tests revealed that overall detection sensitivity for correctly attributing a 'something suspicious' response to a masked face was greatest for the SRs ($M = 1.94$, $SD = 0.83$) followed by the MCs ($M = 1.61$, $SD = 0.84$), $p < .001$, 95% CI [.13, .52], $d = 0.389$, and then the UCs ($M = 1.28$, $SD = 0.74$) in turn, $p < .001$, 95% CI [.14, .52], $d = 0.414$. The same pattern of results was found for response bias, with SRs ($M = 0.28$, $SD = 0.50$) demonstrating a more liberal response pattern in comparison to MCs ($M = 0.49$, $SD = 0.51$), $p < .001$, 95% CI [.09, .34], $d = 0.421$, with UCs ($M = 0.74$, $SD = 0.56$) producing the most conservative response bias, $p < .001$, 95% CI [.13, .37], $d = 0.469$. Overall, for the covert detection task, SRs showed a significant advantage, in comparison to controls, in correctly attributing suspicion to the mask wearers present in the image set.

3.2 | Covert detection task: Prior mask awareness

Participants were not aware, prior to completing the covert detection task, that mask wearers would be present in the image set. However, at the end of the full experiment they were asked whether they had any pre-study knowledge of these masks. There was a relatively even split for each group in relation to pre-study awareness; 56% of UCs; 44% of MCs; and 47% of SRs answered 'yes' to at least one of the

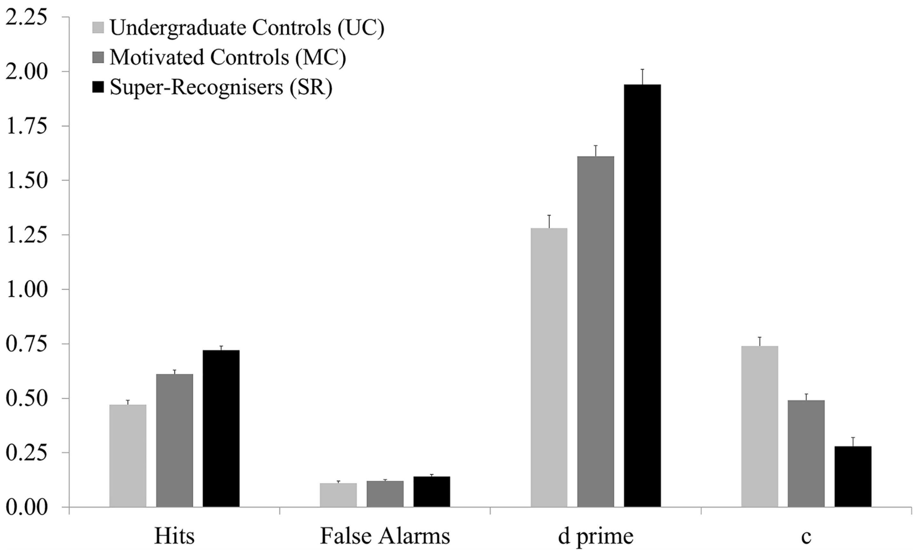

FIGURE 3 Covert detection task performance. This figure shows the mean task performance across each of the signal detection measures for the covert mask detection task. Percentage hits and false alarms are shown here as decimal numbers (error bars denote standard error of the mean).

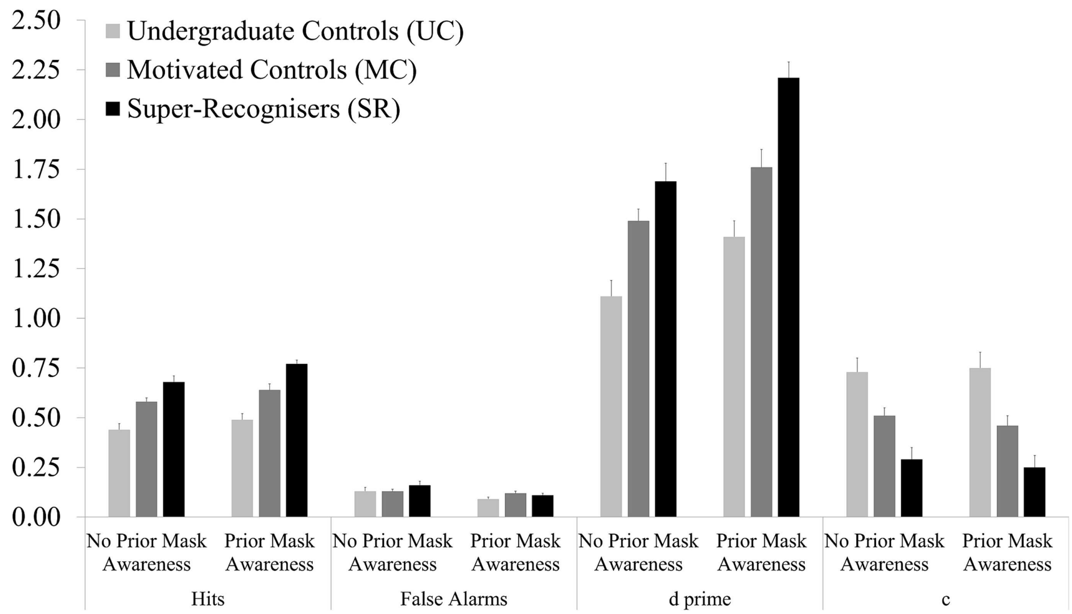

FIGURE 4 Covert detection task performance: Pre-study mask awareness. Covert detection task performance presented as a function of participants pre-study awareness of hyper-realistic masks. Percentage hits and false alarms are shown here as decimal numbers (error bars denote standard error of the mean).

related questions. Therefore, we ran an additional analysis, this time splitting the sample into six groups (UCs, MCs, SRs × pre-study mask awareness, no pre-study mask awareness), to assess whether pre-study mask knowledge affected covert task performance.

As seen in Figure 4, the data for hits, false alarms, and response bias for each group largely mirrored the overall pattern reported above for the full sample. Note that attribution errors remained evident even in those with prior awareness of masks ($M = 15\%$ not responding with suspicion to mask trials, on average, across the groups; $M = 8\%$ responding with suspicion to real faces, on average, across the groups). However, the ANOVA on overall detection sensitivity revealed an interesting finding. There was a main effect of group, $F_{Welch}(5,\ 247) = 20.84$, $p < .001$, $\omega^2 = .028$, and as seen in Figure 4, SRs with pre-study mask knowledge (PA = prior awareness, NPA = no prior awareness; SRs-PA, $M = 2.21$, $SD = 0.73$) significantly outperformed each of the control groups and SRs without pre-study awareness (MCs-PA, $M = 1.76$, $SD = 0.96$; UCs-PA, $M = 1.41$, $SD = 0.78$; SRs-NPA, $M = 1.69$, $SD = 0.83$; MCs-NPA, $M = 1.49$, $SD = 0.70$; UCs-NPA $M = 1.11$, $SD = 0.66$), p's ≤ .007, 95% CI's [.09, 1.44], for the differences). SRs with no prior mask awareness outperformed the UCs, $p < .001$, 95% CI [.23, .94], but not the MCs, $p = .440$. This finding shows that the super-recognition ability allied with prior knowledge of hyper-realistic masks produced the most accurate performance in terms of correctly attributing a 'something suspicious' response to a masked face. A lack of prior mask awareness, as expected, was detrimental to performance, with statistically similar performance between SRs and motivated controls.

3.3 | Covert detection task: Spontaneous detection, qualitative descriptors

Using a word frequency counter (https://tinyurl.com/2fce8sfy), we analysed the extent to which a correct response to a mask trial was justified with the word 'mask' as a descriptor. In other words, how likely was it that spontaneous mask detection occurred in the covert task. The analysis showed that spontaneous detection rates were low, with just 16% of the qualitative descriptor content referring to the word 'mask' to elaborate upon their (correctly attributed) suspicions. SRs showed the greatest use of this descriptor (29% for SRs with prior mask awareness; 13% for SRs without prior awareness), followed by the motivated controls (20% for MCs with prior mask awareness; 9% for MCs without prior awareness) and then the undergraduate controls (15% for UCs with prior mask awareness; 5% for UCs without prior awareness). The next most frequently cited terms were 'disguise' (7%) and 'fake' (7%). As spontaneous mask/disguise detection rates were low, there was a range of other justifications provided for attribution of suspicion, including: 'looks untrustworthy', 'looks strange', 'looks angry', 'unusual face shape/bone structure', 'skin texture does not match the eyes', 'something seems off', 'the smirk makes them look suspicious', 'looks like they have taken drugs'.

Taken together, the findings from the covert detection task show that SRs, with pre-study mask knowledge, were better than controls in detecting that something was amiss in the facial appearance of a mask wearer. However, voluntarily attributing that suspicion to the presence of a hyper-realistic mask in that *un-prompted context* was less likely. Next, we examine whether the covert detection advantage found for the SRs with prior mask knowledge holds for explicit mask detection, in which participants *have been prompted* to detect mask wearers.

3.4 | Explicit detection task

Performance on the explicit detection task reflects the extent to which participants correctly detect a mask wearer in the image set. For signal detection purposes a hit was categorised as a 'mask' response to a masked face trial, while a false alarm was categorised as a 'mask' response to a real face trial. One-way between subjects ANOVAs with group as the factor (UC, MC, SR) revealed a main effect for hits, $F_{Welch}(2, 363) = 22.70$, $p < .001$, $\omega^2 = .030$, false alarms, $F_{Welch}(2, 357) = 24.44$, $p < .001$, $\omega^2 = .036$, and overall detection sensitivity (d'), F_{Welch} $(2, 360) = 60.96$, $p < .001$, $\omega^2 = .083$. Although there was a trend for a more conservative response bias for the SRs, the main effect of group was not significant $F(2, 567) = .538$, $p = .584$, $\eta^2 = .002$. As seen in Figure 5, Games-Howell post hoc tests revealed that the SR group produced the greatest proportion of hits ($M = 89\%$, $SD = 12\%$) and lowest proportion of false alarms ($M = 5\%$, $SD = 10\%$) in comparison to both control groups (For hits; $M = 83\%$, $SD = 16\%$ for MCs; $M = 80\%$, $SD = 16\%$ for UCs; For false alarms; $M = 11\%$, $SD = 11\%$ for MCs; $M = 13\%$, $SD = 11\%$ for UCs), $p's \leq .001$, 95% CI's [2.96, 13.36] for the differences. While there were trends for greater hit rates/lower false alarm rates for the MC group compared to the UC group, post hoc comparisons using the Games-Howell procedure did not reach significance, $p's \geq .115$. Importantly, overall detection sensitivity was greatest for the SRs ($M = 3.02$, $SD = 0.64$) followed by the MCs ($M = 2.45$, $SD = 0.76$), $p < .001$, 95% CI [.41, .74], and then the UCs ($M = 2.20$, $SD = 0.78$) in turn, $p = .005$, 95% CI [.06, .43], showing that the SR advantage reported in the covert task extends to explicit mask detection.

Although all participants had an awareness of these masks before beginning the explicit task, a re-analysis of the explicit response data

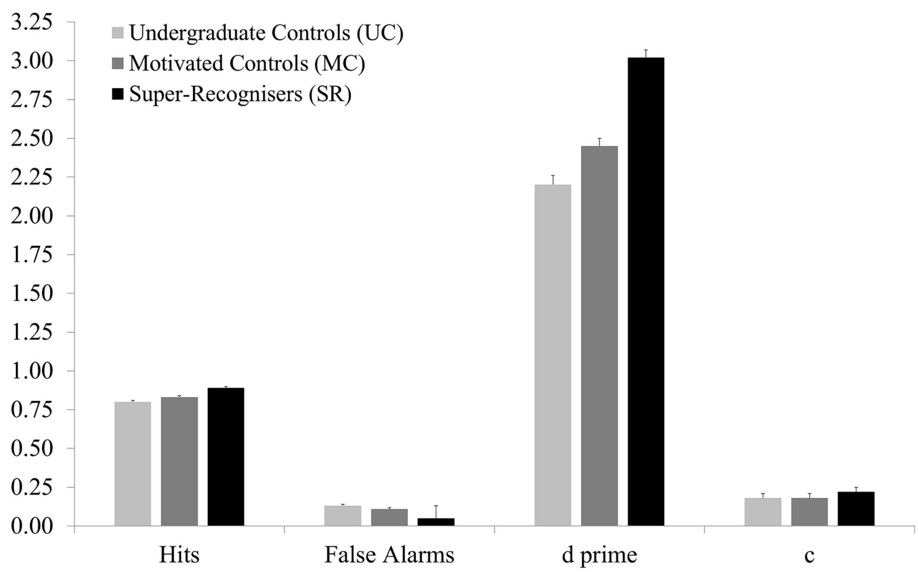

FIGURE 5 Explicit mask detection task performance. This figure shows the mean task performance across each of the signal detection measures for the explicit mask detection task. Percentage hits and false alarms are shown here as decimal numbers (error bars denote standard error of the mean).

with the sample split into the six groups with pre-study mask awareness as a factor (UCs, MCs, SRs × pre-study mask awareness, no pre-study mask awareness) showed that there was now no difference in performance between SRs with and without pre-study mask knowledge, $p = .475$, and that both SR groups outperformed each of the other group combinations, p's ≤ .005, 95% CI's [.08, 1.46]. This finding shows that mask awareness, either pre-study or within-session, combined with the super-recogniser ability, produced the most accurate explicit mask detection performance.

3.5 | Covert and explicit detection performance

To assess consistency in performance across the two tasks, participants covert and explicit detection sensitivity scores were entered into a Pearson's correlation analysis. An outlier check led to the removal of 14 datapoints and this accounts for the differences in the degrees of freedom reported below. The analysis showed significant positive correlations across the groups both for those with pre-study mask awareness (SRs $r(70) = .357$, $p = .002$, 95% CI [.133, .546]; MCs $r(105) = .702$, $p < .001$, 95% CI [.589, .788]; UCs $r(95) = .454$, $p < .001$, 95% CI [.277, .601]) and those without (SRs $r(80) = .479$, $p < .001$, 95% CI [.289, .632]; MCs $r(132) = .249$, $p = .004$, 95% CI [.082, .403]; UCs $r(72) = .265$, $p = .024$, 95% CI [.036, .468]). This finding shows that although spontaneous mask detection rates in the covert task were low (i.e., correct response + 'mask' descriptor; $M = 16\%$ on average), individuals within each group, who were better at correctly attributing suspicion to a masked face, for any reason, also tended to perform accurately at explicit mask detection.

3.6 | Region of interest task: Quantitative response data

Here we examine whether some regions of a masked face were more likely than others to indicate the presence of a mask. Due to violations of homogeneity of variance across the groups for most face regions, we report individual one-way between groups Welch's ANOVAs for each region of interest, and due to the number of statistical tests, we apply the Bonferroni correction to our alpha level (new $\alpha = .004$). As this task was based on selecting the salient features from the masked faces presented, we report findings from the three groups without the pre-study mask awareness distinction. Figure 6 shows the mean percentage click rate for each region of the masked faces, with each of the groups placing particular importance on the eyes, lips, and nose as salient cues for mask detection. The ANOVAs revealed significant group effects for the left eye, $F_{Welch}(2, 353) = 8.37$, $p < .001$, $\omega^2 = .014$, and nose, $F_{Welch}(2, 337) = 7.35$, $p = .001$, $\omega^2 = .011$. While, as seen in Figure 6, there were trends for group differences for the right eye, mouth, forehead, and right ear, these did not reach significance at our adjusted alpha level, p's ≥ .027. The group effects for the remaining regions were not significant, p's ≥ .057.

Games-Howell post hoc tests showed that for the left eye region, SRs and MCs were equally likely to click on this region ($M = 15\%$, $SD = 7\%$ for SRs; $M = 15\%$, $SD = 8\%$ for MCs), $p = .968$, and to a significantly greater extent than the UC group ($M = 12\%$, $SD = 8\%$), p's = .001, 95% CI's [1.04, 5.17]. For the nose region, the SRs and UCs were equally likely to click on this region ($M = 11\%$, $SD = 7\%$ for SRs; $M = 11\%$, $SD = 9\%$ for UCs), $p = .586$, and to a greater extent than the MC group ($M = 9\%$, $SD = 7\%$), p's ≤ .020, 95% CI's [.24, 4.52]. These findings support previous work which has shown that

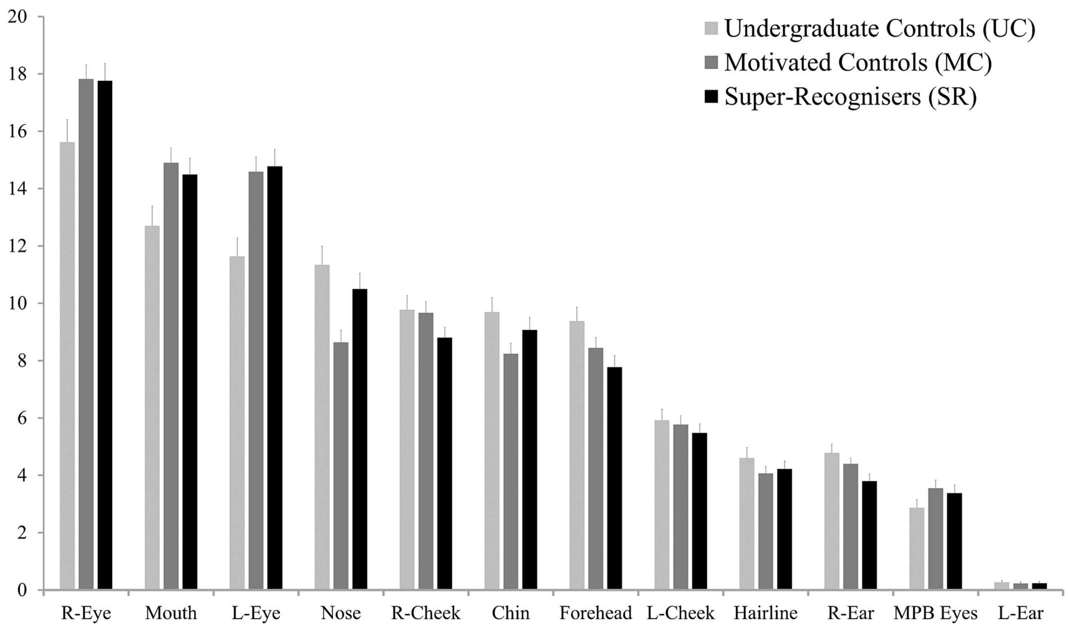

FIGURE 6 Mean percentage clicks for masked face regions that could support detection. This figure shows the mean percentage clicks for each of the 12 pre-defined masked face regions of interest (MPB Eyes; Midpoint between the eyes; error bars denote standard error of the mean).

the eye-regions may provide a key area to help differentiate between a mask and a real face (Sanders & Jenkins, 2018). The lack of consistent group effects suggests that the SR advantage is unlikely to be underpinned by any qualitatively different face processing strategies (Nador et al., 2021).

3.7 | Region of interest task: Qualitative response data

In addition to providing quantitative data in the form of clickable regions of masked faces, participants could also provide additional qualitative response data to highlight further aspects of the masks that could aid in their detection. To provide the most useful descriptors of the cues which may best support mask detection, we report those provided by the SR group. Qualitative responses were not a mandatory part of this task, however, 1119 responses out of a possible 1884 (59%) were provided. Figure 7 shows the most frequently cited descriptors that participants suggested an observer should focus on to help detect the presence of a hyper-realistic mask (percentage descriptor = word frequency/N responses; the percentages do not sum to 100% as several descriptors were often included within one text response). In line with the quantitative region of interest data, the qualitative data supports the importance that the participants placed on the eyes, nose, and mouth.

However, the qualitative data also provided additional cues in relation to skin, wrinkles, hair, beards, and facial expression. In Table 2 we provide a sample of the text responses that are representative of the most frequently cited aspects of the face. One notable common theme, in addition to comments relating to the lack of natural texture, proportionality of the features, and expressiveness, is the lack of asymmetry in the hyper-realistic masks. That is, in real faces participants appear to expect to see a nose, a hairline, or a wrinkle, for example, that is not perfectly symmetrical or set in a straight line, and the unnatural symmetry which can be typical of some of these masks could therefore assist in their detection.

4 | DISCUSSION

Hyper-realistic masks present an emerging, sophisticated, and viable route to identity fraud. In both the real world (Sanders & Jenkins, 2021) and experimental settings (Sanders et al., 2017, 2019; Sanders & Jenkins, 2018), the detection of these masks is a challenging task and one that is prone to error. Having established the viability of this 'complete' method of disguise, focus must now turn to

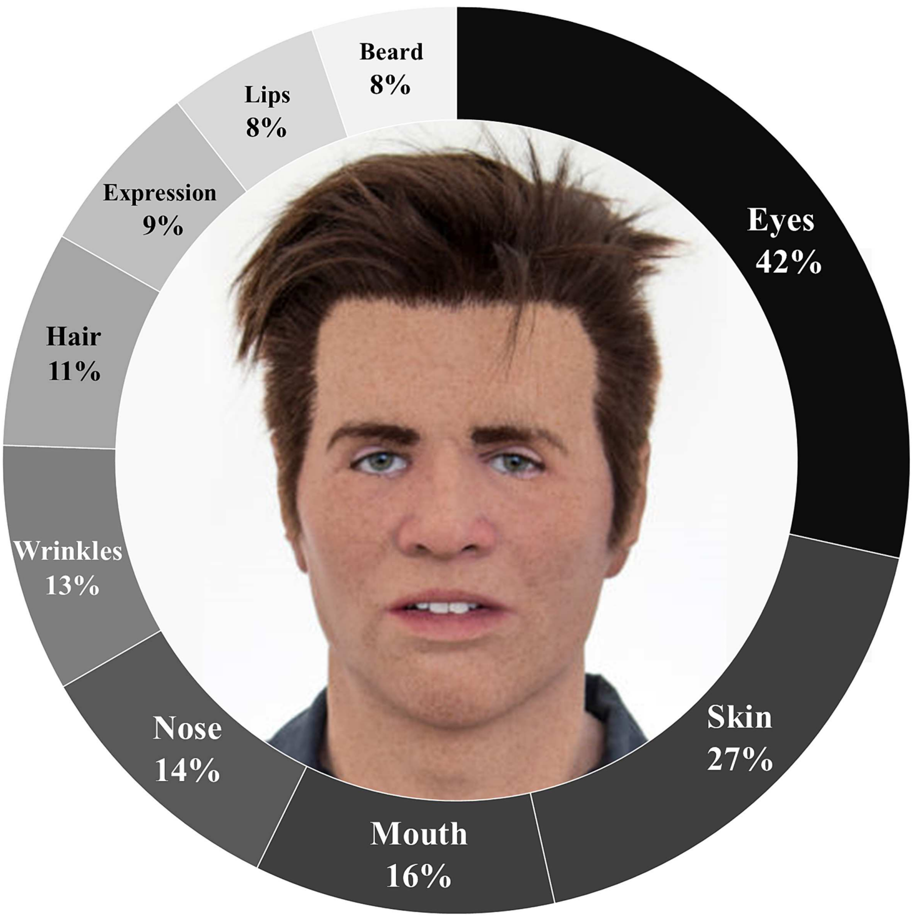

FIGURE 7 Percentage qualitative descriptors for masked face regions that could support detection. This figure shows the most frequent descriptors that participants used to highlight which aspects an observer should focus on to help detect the presence of a hyper-realistic mask (mask image is the authors own; percentage descriptor = word frequency/N responses; the percentages do not sum to 100% as several descriptors were often included within one text response).

TABLE 2 Representative text responses for the masked face regions that could support detection.

	Representative text responses
Eyes	Eyes appear 'set too deep', 'set back', 'sunken' within the face
	Eyes don't 'fit' with the overall expression on the rest of the face
Skin	Skin doesn't look 'natural', too 'consistent in colour across the face'
	Skin is 'too smooth', with no 'texture', looks 'waxy' or 'shiny'
Mouth	Mouth is 'inexpressive', 'not alive', 'unnaturally open', 'too large'
	Mouth seems 'detached' from other areas of the face
Nose	Nose is 'too big', 'too shiny', 'too smooth'
	Nose is 'too symmetrical', 'too wide for face'
Wrinkles	Wrinkles are 'too symmetric', 'there are too many', 'excessive'
	Wrinkles make the face 'look too strained and unmoveable'
Hair	Hair looks 'like a wig', is 'artificial', 'too thick'
	The hairline is 'unnatural', 'too straight', 'too perfect'
Expression	'Expressionless especially around the mouth and eyes'
	'Expression is 'dull', 'empty'
Lips	Lips 'lack structure', 'are too big', 'should be thinner'
	Lips 'look disconnected with inside of mouth'
Beard	Beard 'looks fake', 'is too straight', 'is too perfect'
	Beard 'hair texture is unrealistic'

Note: In these text responses, participants were attempting to provide border control officers with descriptors that may help them to detect the presence of a traveller wearing a hyper-realistic mask.

developing ways to enhance the detection of these masks. In the present study, we took an individual differences approach to assess whether a group of people, known as super-recognisers (SRs), who excel at the recognition of *identity* from facial information, would be any better at detecting hyper-realistic masks than two groups of control participants. We utilised a *covert mask detection task*, which invited unprompted/spontaneous recognition of this deception, and an *explicit mask detection task* in which participants sought to detect this disguise having received information and examples of their use.

Our findings from the covert detection task showed that SRs, with pre-study mask knowledge, outperformed controls in correctly attributing suspicion to mask wearers. That is, without any initial prompting that the face photos would contain anything other than real faces, the SR group showed a greater propensity to detect that something was amiss with the facial appearance of the masked individuals. While rates of 'spontaneous' mask detection (i.e., including 'mask' as a qualitative descriptor to a suspicious response) were relatively low, the proportion of specific 'mask' responses or more general disguise-related descriptors was also numerically greater for the SRs (with prior mask awareness) compared to the other groupings. Importantly, and regardless of pre-study mask knowledge, there was a clear SR advantage for explicit mask detection, after all participants had received the same information about this type of identity fraud, and there was a link between initial covert suspicions and explicit detection rates.

These findings show that the SR ability for identity recognition (Davis et al., 2016) and the detection of traditional item-based disguise (Noyes et al., 2021), does appear to generalise to hyper-realistic masks. This finding is of theoretical interest as the detection of these masks is quite different to the identity-based focus of SR research to date. In this context, there is no person identification component, and the task is not to see through a disguise, rather it is to decide whether an entire face is artificial or not. Therefore, at first glance, it may have been logical to assume that there would be no connection between the SR face identification ability and their aptitude at detecting masks. However, recent research reported by Faghel-Soubeyrand et al. (2021) and Nador et al. (2021) point to SRs having more proficient low-level and mid-level visual face processing abilities than typical-recogniser controls. This makes sense given the contribution that the processing of both texture, and shape, make to face recognition (Troje & Bülthoff, 1996), and it may explain why the SR advantage for face identity generalises to the detection of these synthetic masks.

The findings from the present study should also be of interest to applied organisations for two reasons. First, we show that simple mask awareness substantially improved mask detection across each group. This means that a clear and quick step can be made to enhance the detection of these masks, regardless of the current level of face recognition ability of the officials, simply by making them aware of them, their use in identity fraud, and by providing examples for reference. Second, for agencies who are looking to implement the SR advantage in their own identity verification and fraud detection operations, the present study shows that the SR advantage extends to mask detection and that SRs can be recruited for increasingly multipurpose identification tasks. Indeed, research has recently shown that super-face-recognisers may also be super-voice recognisers (Jenkins et al., 2021), and here we add another tool to their armoury by demonstrating their proficiency with masks.

As stated above, while the SRs outperformed the controls, their explicit detection performance was not perfect, and we now plan to build on the qualitative data collected to develop a training program using a similar approach to Towler et al. (2021) to enhance their performance further. Training has been shown to improve the detection of other identity fraud techniques, such as face morphs (Robertson et al., 2018), but only when there are clear cues (e.g., visual artifacts from the morphing process) that can be detected by the human visual system (Kramer et al., 2019). Developing a training paradigm that draws a super-recognisers attention to the salient aspects of a masked face that cue the presence of this disguise may further enhance their performance. Such training programmes should also consider that these masks will likely become ever more sophisticated and life-like. The masked images used in this paper were sourced in 2016, so further research should also source more recent examples, both to incorporate into a stimulus set and to judge the rate at which the sophistication of this disguise is improving over time.

While these findings provide support for the view that super-recognisers may also be super-mask-detectors, our study, while representative of *known* mask fraud attempts (Sanders & Jenkins, 2021), used mainly White, older, male masks. It may be the case that there are a wider variety of masks (e.g., young, female, ethnically diverse) being used in criminal situations that are under-detected/under-reported, and so future research should look to use a more diverse set of stimuli which mirror the wider population to a greater extent. In addition, our tasks used face photos, and while findings within the face recognition field tend to be consistent regardless of whether the task involves images or live faces (Davis & Valentine, 2009), ecological validity is key in applied psychological science, and SRs should now be tested using live mask detection tasks, ideally in operational contexts. Moreover, we only presented masked faces in our ROI task, and while we believe the descriptors provided in Table 2, which unpack the reasons for the region selections, do provide a better understanding of how this disguise could be detected, future research should also incorporate real faces into this task to further refine the mask-specific attributes that can support detection.

While super-recognisers may provide a route to enhance the detection of hyper-realistic masks, in line with face-based identity recognition, peak performance is likely to occur when we pair our best SRs with our best identification and fraud detection algorithms (see Towler et al., 2023; White et al., 2015), and there is some recent work from computer science which has started to focus on mask fraud attacks (Jia et al., 2020; and see Carragher & Hancock, 2023; Fysh & Bindemann, 2018; Howard et al., 2020; and White et al., 2015 for a broader discussion on operator-algorithm interactions). Finally, here we partitioned our groups using the well-established CFMT+ in conjunction with one further confirmatory measure. However, we acknowledge that new tests are regularly being published which may in the future, if the field can agree upon a common assessment framework (see Ramon, 2021), become the new standardized SR tests, and we would support work towards that end.

In summary, over the last decade, more than 40 criminal acts have been committed by perpetrators wearing a hyper-realistic mask. These masks provide a viable route to identity fraud. This study shows that the super-recogniser advantage for identity verification generalises to the detection of this sophisticated synthetic disguise, providing a promising route to maximise their detection and minimise their impact on society.

AUTHOR CONTRIBUTIONS

David J. Robertson: Conceptualization; data curation; formal analysis; investigation; methodology; project administration; software; writing – original draft; writing – review and editing. **Josh P. Davis:** Conceptualization; data curation; formal analysis; investigation; methodology; project administration; resources; writing – review and editing. **Jet G. Sanders:** Conceptualization; data curation; formal analysis; investigation; methodology; project administration; resources; writing – review and editing. **Alice Towler:** Conceptualization; data curation; formal analysis; investigation; methodology; project administration; writing – review and editing.

ACKNOWLEDGEMENTS

We thank the participants who were recruited from the University of Greenwich face recognition database for their ongoing engagement with research.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

A copy of the dataset that supports analyses reported in this paper is available via https://osf.io/uhb7f/. This study was not preregistered.

ORCID

David J. Robertson https://orcid.org/0000-0002-8393-951X
Josh P. Davis https://orcid.org/0000-0003-0017-7159
Alice Towler https://orcid.org/0000-0003-4092-8703

REFERENCES

Bate, S., Portch, E., & Mestry, N. (2021). When two fields collide: Identifying “super-recognisers” for neuropsychological and forensic face recognition research. *Quarterly Journal of Experimental Psychology*, *74*(12), 2154–2164. https://doi.org/10.1177/17470218211027695

Bernstein, S. (2010). Masks so realistic they're arresting the wrong guy. Retrieved from http://articles.latimes.com/2010/dec/08/business/la-fi-mask-20101209

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*, 1378. https://doi.org/10.3389/fpsyg.2016.01378

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

Carragher, D. J., & Hancock, P. J. (2023). Simulated automated facial recognition systems as decision-aids in forensic face matching tasks. *Journal of Experimental Psychology: General*, *152*(5), 1286–1304. https://doi.org/10.1037/xge0001310

Davis, J. P. (2019). The worldwide impact of identifying super-recognisers in police and business. *The Cognitive Psychology Bulletin*, *4*, 17–22. http://gala.gre.ac.uk/id/eprint/24017

Davis, J. P. (2020). CCTV and the super-recognisers. In C. Stott, B. Bradford, M. Radburn, & L. Savigar-Shaw (Eds.), *Making an impact on policing and crime: Psychological research, policy and practice* (pp. 34–67). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9780429326592

Davis, J. P., Lander, K., Evans, R., & Jansari, A. (2016). Investigating predictors of superior face recognition ability in police super-recognisers. *Applied Cognitive Psychology*, *30*(6), 827–840. https://doi.org/10.1002/acp.3260

Davis, J. P., & Robertson, D. J. (2020). Capitalizing on the super-recognition advantage: A powerful, but underutilized, tool for policing and national security agencies. *The Journal of the United States Homeland Defence and Security Information Analysis Center (HDIAC)*, *7*(1), 20–25. https://hdiac.org/articles/capitalizing-on-the-super-recognition-advantage-a-powerful-but-underutilized-tool-for-policing-and-national-security-agencies/

Davis, J. P., & Tamonytė, D. (2017). Masters of disguise: Super-recognisers' superior memory for concealed unfamiliar faces. In *2017 seventh international conference on emerging security technologies (EST)* (pp. 44–49). IEEE. https://doi.org/10.1109/EST.2017.8090397

Davis, J. P., & Valentine, T. (2009). CCTV on trial: Matching video images with the defendant in the dock. *Applied Cognitive Psychology*, *23*(4), 482–505. https://doi.org/10.1002/acp.1490

Delacre, M., Leys, C., Mora, Y. L., & Lakens, D. (2019). Taking parametric assumptions seriously: Arguments for the use of Welch's F-test instead of the classical F-test in one-way ANOVA. *International Review of Social Psychology*, *32*(1), 13. https://doi.org/10.5334/irsp.198

Dhamecha, T. I., Singh, R., Vatsa, M., & Kumar, A. (2014). Recognizing disguised faces: Human and machine evaluation. *PLoS One*, *9*(7), e99212. https://doi.org/10.1371/journal.pone.0099212

Faghel-Soubeyrand, S., Ramon, M., Bamps, E., Zoia, M., Woodhams, J., Alink, A., Gosselin, F., & Charest, I. (2021). Decoding real-world visual recognition abilities in the human brain. *Journal of Vision*, *21*(9), 2604. https://doi.org/10.1167/jov.21.9.2604

Fysh, M. C., & Bindemann, M. (2018). Human–computer interaction in face matching. *Cognitive Science*, *42*(5), 1714–1732. https://doi.org/10.1111/cogs.12633

Gardner, D. (2010). The bank rubber … Raider uses movie- style mask to disguise himself. Daily Mail. https://www.dailymail.co.uk/news/article-1268215/White-robber-fools-police-weeks-elaborate-African-American-Hollywood-mask.html

Howard, J. J., Rabbitt, L. R., & Sirotin, Y. B. (2020). Human-algorithm teaming in face recognition: How algorithm outcomes cognitively bias human decision-making. *PLoS One*, *15*(8), e0237855. https://doi.org/10.1371/journal.pone.0237855

Jenkins, R. E., Tsermentseli, S., Monks, C. P., Robertson, D. J., Stevenage, S. V., Symons, A. E., & Davis, J. P. (2021). Are super-face-recognisers also super-voice-recognisers? Evidence from cross-modal identification tasks. *Applied Cognitive Psychology*, *35*(3), 590–605. https://doi.org/10.1002/acp.3813

Jia, S., Guo, G., & Xu, Z. (2020). A survey on 3D mask presentation attack detection and countermeasures. *Pattern Recognition*, *98*, 107032. https://doi.org/10.1016/j.patcog.2019.107032

Kramer, R. S., Mireku, M. O., Flack, T. R., & Ritchie, K. L. (2019). Face morphing attacks: Investigating detection with humans and computers. *Cognitive Research: Principles and Implications*, *4*(1), 1–15. https://doi.org/10.1186/s41235-019-0181-4

Kramer, R. S. S., & Ritchie, K. L. (2016). Disguising superman: How glasses affect unfamiliar face matching. *Applied Cognitive Psychology*, *30*(6), 841–845. https://doi.org/10.1186/s41235-019-0181-4

McCaffery, J. M., Robertson, D. J., Young, A. W., & Burton, A. M. (2018). Individual differences in face identity processing. *Cognitive Research: Principles and Implications*, *3*(1), 1–15. https://doi.org/10.1186/s41235-018-0112-9

Nador, J. D., Zoia, M., Pachai, M. V., & Ramon, M. (2021). Psychophysical profiles in super-recognizers. *Scientific Reports*, *11*(1), 1–11. https://doi.org/10.1038/s41598-021-92549-6 Psychophysical profiles in super-recognizers.

Noyes, E., Davis, J. P., Petrov, N., Gray, K. L., & Ritchie, K. L. (2021). The effect of face masks and sunglasses on identity and expression recognition with super-recognizers and typical observers. *Royal Society Open Science*, *8*(3), 201169. https://doi.org/10.1098/rsos.201169

Noyes, E., & Jenkins, R. (2019). Deliberate disguise in face identification. *Journal of Experimental Psychology: Applied*, *25*(2), 280–290. https://doi.org/10.1037/xap0000213

Ramon, M. (2021). Super-recognizers–a novel diagnostic framework, 70 cases, and guidelines for future work. *Neuropsychologia*, *158*, 1–11. https://doi.org/10.1016/j.neuropsychologia.2021.107809

Righi, G., Peissig, J. J., & Tarr, M. J. (2012). Recognizing disguised faces. *Visual Cognition*, *20*(2), 143–169. https://doi.org/10.1080/13506285.2012.654624

Robertson, D., Middleton, R., & Burton, A. M. (2015). From policing to passport control. *Keesing Journal of Documents & Identity*, *46*, 3–8. https://platform.keesingtechnologies.com/from-policing-to-passport-control/

Robertson, D. J., Black, J., Chamberlain, B., Megreya, A. M., & Davis, J. P. (2020). Super-recognisers show an advantage for other race face identification. *Applied Cognitive Psychology*, *34*(1), 205–216. https://doi.org/10.1002/acp.3608

Robertson, D. J., Fysh, M. C., & Bindemann, M. (2019). Face identity verification: Five challenges facing practitioners. *Keesing Journal of Documents & Identity*, *59*, 3–8. https://platform.keesingtechnologies.com/keesing-journal-59/

Robertson, D. J., Mungall, A., Watson, D. G., Wade, K. A., Nightingale, S. J., & Butler, S. (2018). Detecting morphed passport photos: A training and individual differences approach. *Cognitive Research: Principles and Implications*, *3*(1), 1–11. https://doi.org/10.1186/s41235-018-0113-8

Robertson, D. J., Noyes, E., Dowsett, A. J., Jenkins, R., & Burton, A. M. (2016). Face recognition by metropolitan police super-recognisers. *PLoS One*, *11*(2), e0150036. https://doi.org/10.1371/journal.pone.0150036

Robertson, D. J., Sanders, J. G., Towler, A., Kramer, R. S., Spowage, J., Byrne, A., Burton, A. M., & Jenkins, R. (2020). Hyper-realistic face masks in a live passport-checking task. *Perception*, *49*(3), 298–309. https://doi.org/10.1177/0301006620904614

Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin & Review*, *16*(2), 252–257. https://doi.org/10.3758/PBR.16.2.252

Sanders, J. G., & Jenkins, R. (2018). Individual differences in hyper-realistic mask detection. *Cognitive Research: Principles and Implications*, *3*(1), 1–10. https://doi.org/10.1186/s41235-018-0118-3

Sanders, J. G., & Jenkins, R. (2021). Realistic masks in the real world. In M. Bindemann (Ed.), *Forensic face matching: Research and practice* (pp. 216–236). Oxford Academic Press. https://doi.org/10.1093/oso/9780198837749.001.0001

Sanders, J. G., Ueda, Y., Minemoto, K., Noyes, E., Yoshikawa, S., & Jenkins, R. (2017). Hyper-realistic face masks: A new challenge in person identification. *Cognitive Research: Principles and Implications*, *2*(1), 1–12. https://doi.org/10.1186/s41235-017-0079-y

Sanders, J. G., Ueda, Y., Yoshikawa, S., & Jenkins, R. (2019). More human than human: A Turing test for photographed faces. *Cognitive Research: Principles and Implications*, *4*(1), 1–10. https://doi.org/10.1186/s41235-019-0197-9

Schofield, H. (2019). The fake French minister in a silicone mask who stole millions. https://www.bbc.co.uk/news/world-europe-48510027

Shakeshaft, N. G., & Plomin, R. (2015). Genetic specificity of face recognition. *Proceedings of the National Academy of Sciences*, *112*(41), 12887–12892. https://doi.org/10.1073/pnas.1421881112

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Terry, R. L. (1994). Effects of facial transformations on accuracy of recognition. *The Journal of Social Psychology*, *134*(4), 483–492. https://doi.org/10.1080/00224545.1994.9712199

Towler, A., Dunn, J. D., Castro Martínez, S., Moreton, R., Eklöf, F., Ruifrok, A., Kemp, R. I., & White, D. (2023). Diverse types of expertise in facial recognition. *Scientific Reports*, *13*(1), 11396. https://doi.org/10.1038/s41598-023-28632-x

Towler, A., Keshwa, M., Ton, B., Kemp, R. I., & White, D. (2021). Diagnostic feature training improves face matching accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(8), 1288–1298. https://doi.org/10.1037/xlm0000972

Troje, N. F., & Bülthoff, H. H. (1996). Face recognition under varying poses: The role of texture and shape. *Vision Research*, *36*(12), 1761–1771. https://doi.org/10.1016/0042-6989(95)00230-8

Verhallen, R. J., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Bargary, G., & Mollon, J. D. (2017). General and specific factors in the processing of faces. *Vision Research*, *141*, 217–227. https://doi.org/10.1016/j.visres.2016.12.014

White, D., Dunn, J. D., Schmid, A. C., & Kemp, R. I. (2015). Error rates in users of automatic face recognition software. *PLoS One*, *10*(10), e0139827. https://doi.org/10.1371/journal.pone.0139827

Wilmer, J. B. (2017). Individual differences in face recognition: A decade of discovery. *Current Directions in Psychological Science*, *26*(3), 225–230. https://doi.org/10.1177/0963721417710693

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*, *107*(11), 5238–5241. https://doi.org/10.1073/pnas.0913053107

Zamost, S. (2010). Exclusive: Man in disguise boards international flight. https://edition.cnn.com/2010/WORLD/americas/11/04/canada.disguised.passenger/index.html

How to cite this article: Robertson, D. J., Davis, J. P., Sanders, J. G., & Towler, A. (2024). The super-recogniser advantage extends to the detection of hyper-realistic face masks. *Applied Cognitive Psychology*, *38*(4), e4222. https://doi.org/10.1002/acp.4222