The wisdom of the crowd: A case of post- to ante-mortem face matching by police super-recognisers

Josh P Davis (University of Greenwich) [a]

Andreea Maigut (University of Greenwich) [b]

Charlotte Forrest (University of Greenwich) [c]

[a] Dr Josh P Davis, Reader in Applied Psychology, Applied Psychology Research Group, University of Greenwich, London SE9 2UG, United Kingdom, j.p.davis@gre.ac.uk (corresponding author)

[b] Andreea Maigut, Department of Psychology, Social Work and Counselling, University of Greenwich, London SE9 2UG, United Kingdom, andreeamaigut@gmail.com

[c] Dr Charlotte Forrest, Applied Psychology Research Group, University of Greenwich, London SE9 2UG, United Kingdom, c.forrest@gre.ac.uk

Acknowledgements: We would like to thank Sergeant Yvonne Saunderson of the Metropolitan Police Service (MPS) River Police Unit for consulting with the authors on this case. We would also like to thank Mick Neville, now retired from the MPS for arranging police officer participation, and two anonymous reviewers of this manuscript for their positive suggestions.

Author contributions:

Josh Davis: Conceptualization, Methodology, Investigation, Writing – Original Draft, Supervision, Project Administration

Andreea Maigut: Conceptualization, Methodology, Investigation, Writing – Review & Editing

Charlotte Forrest: Methodology, Formal Analysis, Writing – Review & Editing, Visualization

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data sharing: The authors are willing to share the anonymised data upon request.

Ethics: This case study received ethical approval from the University of Greenwich Research Ethics Committee as part of an unfunded research collaboration with the Metropolitan Police Service

Conflict of interests: There are no conflicts of interest for any of the authors or for the University of Greenwich in terms of the case report described in this manuscript. None of the authors received any financial compensation, or indeed any other compensation for this work. After the report was submitted to the coroner, the first author, Davis was subsequently funded by the European Commission in a collaboration with the Metropolitan Police Service. More recently, he has provided consultancy services via the University of Greenwich for international and national police forces. However, it is unlikely that any of these projects were funded as a result of this case report, and Davis has never received any financial payment for this consultancy work.

Key words

Face matching, wisdom of the crowd, post-mortem, identification, super-recognizer, recognition

Highlights

- Police super-recognisers more likely to confidently match post- and ante-mortem photos
- Wisdom of the crowd procedure raises super-recogniser's confidence ratings for a match
- Wisdom of the crowd procedure does not raise control's confidence in face matching decisions.

Abstract

This case report describes novel methodology used to identify a 43-year-old post-mortem photo of a drowned male recovered from a London river in the 1970's. Embedded in an array of foils, police super-recognisers (*n* = 25) possessing superior simultaneous face matching ability, and police controls (*n* = 139) provided confidence ratings as to the similarity of the post-mortem photo to an ante-mortem photo of a man who went missing at about the same time Indicative of a match, compared to controls, super-recognisers provided higher ratings to the target than the foils. Effects were enhanced when drawing on the combined wisdom of super-recogniser crowds, but not control crowds. These findings supported additional case evidence allowing the coroner to rule that the deceased male and missing male were likely one and the same person. A description of how similar super-recogniser wisdom of the crowd procedures could be applied to other visual image identification cases when no other method is feasible is provided.

Introduction

Police regularly identify deceased persons. The most reliable methods are DNA, odontology, or fingerprints. However, sometimes after water-submersion, the cadaver, or a photo of it, may be directly compared with a photo of a potential match taken when living (1-3, for reviews see 4,5). Even with poor-quality images, recognition of familiar living faces is normally reliable (6,7). In contrast, with images of the deceased, "immediate post-mortem changes may be significant enough to confuse the (familiar) viewer" (p. 97) (e.g., jaw slackness, skin tone, muscle tone loss) (5). In the absence of any demands on memory, simultaneous *unfamiliar* face matching of living photos is also error prone, even with high-quality, close-up, same-pose, same-day images (8,9, for a review see 10). Performance is worse when image characteristics differ (e.g., aging, viewpoint, blurring) (11-13). Not surprisingly, accuracy of ante- to post-mortem unfamiliar face matching may be even more unreliable (3-5).

There are, however, large individual differences in unfamiliar face recognition and matching ability (14-20), with some individuals excellent at both (19,20). In London, a pool of police 'super-recognisers' who make a disproportionally high number of suspect identifications from CCTV has been established (20-23). Some possess empirically tested exceptional face recognition and simultaneous face matching ability, whereas others possess exceptional face recognition ability only. The case report described here recruited police with superior face matching ability from this pool, with the aim of identifying a post-mortem photo of a male whose drowned body was recovered in the 1970's. A comparison group of control police were also recruited.

Case background

1970's police records describe the deceased's body as probably water-immersed for 4-5 days, although a medical doctor reported minimal decomposition. One grainy facial photo of the deceased exists (hereafter 'post-mortem photo'). Approximately 43 years later, members of the family of a man who had been reported missing during the same month, viewed the post-mortem photo, stating that although there were similarities with their relative, they could not be certain of identity. The family supplied the police with two photos (hereafter 'ante-mortem' photos), taken 'some years' before his disappearance. A comparison of the missing man's medical records, family remembrances, and a medical doctor's descriptions of the deceased suggested a match based on medical, physical and circumstantial characteristics. There was also some evidence that pointed against a match, as details differed in the medical records and the doctor's description. However, the missing man's more recent medical records had been destroyed, while some features (i.e. old scars), may have been missed by a medical doctor inspecting a river recovered body.

The investigating officer in the case had first submitted the images described in this report to the Metropolitan Police Service's (MPS) forensic facial comparison specialists. However, the images were deemed to be of insufficient of quality for facial comparison techniques, and so the officer approached the current authors for assistance via the Central Forensic Image Circulation Unit (CFIT) of the MPS.

Current investigation

In the investigation described here, a line-up array was constructed containing one of the post-mortem photos of the missing man together with same-era similarly-faded photos of physically similar foils selected following best practice procedures used for eyewitness identification procedures (see below). Police officers provided a confidence rating to each array face as to the likelihood that it depicted the person in the post-mortem photo. Collecting confidence ratings in this manner increases eyewitness identification and face matching accuracy in contrast to the standard procedure of selecting a single line-up member (e.g. 24). The 'wisdom of the crowd' paradigm was also employed (25). With this, combined group performance at cognitive estimation tasks, including forensic facial examiner face matching is better than that of individuals making up the group (e.g., 25-28). The phenomenon has been applied to a number of paradigms including category learning (26). In this, participants exposed to repeating trials of stimuli (e.g. shapes, nonsense words), require several attempts in order to achieve close to perfect categorisation performance. In contrast, combining those decisions using the Wisdom of the Crowd results in near perfect performance far earlier, suggesting a faster route to achieving reliable judgements. In a similar vein, face matching decision making accuracy can be improved by the fusion of human and computerised face recognition decisions (29). Indeed, the fusion of the decisions of multiple super-recognisers and/or expert forensic facial examiners with those from high-performing commercial face recognition algorithms can enhance face matching accuracy over and above that achieved by the same algorithm and controls, or by algorithms or humans alone (30). The empirical evidence clearly provides theoretical evidence that the wisdom of the crowd of police super-recognisers could assist in cases of the type described here. However, the current case study may be the first application in a real police investigation.

Wisdom of the crowd theory suggests that amalgamated decisions are more accurate as a consequence of *signal amplification*, while the impact of noise on decisions is reduced (26). In a face matching task, aggravated identity decisions from the crowd should exceed those from individuals making up the crowd, but only as long as a signal can be detected. In other words, there should be sufficient idiosyncratic features available in both images to allow an identification, over and above random guessing or noise. Super-recognisers may more efficiently extract the signal – controls may be more likely to guess, suggesting the intriguing possibility that a super-recogniser crowd may be wiser than a control crowd.

Surowiecki (26) suggests a wise crowd requires three components. First, all decisions should be *independent*. In the current study, police were asked not to confer. Second, the crowd should employ *diverse* strategies. No crowd advantage will ensue if all use the exact same method. Third, the crowd should be *de-centralised* so that each member draws on different information sources when making decisions. These may be based on training or past experiences. Although no details of police job roles were collected during this procedure, later research on many of the same super-recogniser pool revealed that only a few had received facial identification training (23). Most were front-line police officers deployed across many different police departments. As such, strategy unity was unlikely, optimising crowd success likelihood.

Two groups of police were recruited. Super-recognisers had previously scored substantially higher than normalised data on the standardised *Glasgow Face Matching Test* (GFMT) (16). The remaining police were designated as 'controls,' although in some cases no data were available as to face matching ability. Based on the hypothesis that post- and ante-mortem photos depicted one and the same person, it was predicted that higher confidence ratings would be provided to the ante-mortem photo than the foils. Super-recognisers' ratings to the ante-mortem photo were also expected to be higher than controls. Amalgamated ratings provided to the ante-mortem photo by 'the crowd', particularly the super-recogniser crowd, were also predicted to be higher than those by individual participants.

2. Methods

2.1 Participants

Police, warned of image nature, were invited to contribute by the Metropolitan Police Service (MPS) Central Forensic Image Circulation Unit (CFIT) based at New Scotland Yard in London via an intranet advert. They were asked to click a tick box if they had previously taken part in the first author's research in order to provide permission to access data attached to an anonymised code.

Super-recognisers (n = 25) had achieved ≥ 37 (92.5%) out of 40 on the *Glasgow Face Matching Test* (GFMT) (short version) (16) at an MPS event. A one-sample t-test demonstrated that with large effect sizes, super-recognisers scored significantly higher (M = 94.7%, SD = 0.8%) than published GFMT normalised data (n = 194) (M = 81.3%, SD = 9.7%), $t(24) = 260.73$, $p < .001$, Cohen's d = 1.95. The remaining police designated as controls (n = 139), had scored below 92.5% on the GFMT (n = 36), had not contributed to the previous research, or could not remember their anonymised code and therefore no ability data were available (n = 103). Additional pilot participants (n = 76) assisted in foil selection and array fairness checks. No demographic data were collected, or information as to current job roles, or experience of, or training in facial identification or comparison procedures.

2.2 Materials

To protect the privacy of the person(s) of interest and family, as much as possible individuating data and other information are not reported.

Target photographs: The black-and-white post-mortem photograph depicted the deceased male lying down with eyes closed from a close-to-profile view. Two black-and-white ante-mortem photos of the missing male were available. However, only one was used here, as this matched the foils' full-face viewpoints. Images were cropped to display head-and-shoulders only.

Array construction and fairness checks: A three-stage strategy ensured foils met England and Wales police guidelines that they should resemble the suspect (e.g., age, gender, ethnicity, position in life) (31), which in this case meant they matched that of the ante-mortem photo on these features.

First, as the ante-mortem photo was over 40-years-old and had slightly faded, the second author and the investigating police sergeant located 21 potential foils by searching police archives of 1960's/1970's photographic mug-shots of males possessing a similar appearance.

Second, two randomly ordered and counterbalanced arrays (3 x 7) containing the 21 shortlisted foils were simultaneously displayed below the ante-mortem photo. Pilot participants (n = 24) each selected ten foils they believed most closely resembled the ante-mortem photo. Along with the ante-mortem photo, the seven foils selected most often were included in the final counterbalanced 2 x 4 line-up array.

Third, to ensure the ante-mortem photo did not stand out from the foils which would make the line-up unfair, the mock witness paradigm was employed (32). With this, five pilot assessors provided written descriptions of the post-mortem photo. Consistent descriptors were retained and amalgamated into a single 'modal' description. Mock witnesses (n = 38) selected the face they believed most closely matched the modal description from the 2 x 4 eight-person line-up array of foils and ante-mortem photo. All foils were selected at least twice (range: 2 to 10); while six mock witnesses (15.8%) selected the ante-mortem photo (95% CI: 4.2% – 27.4%) and as this was not significantly above chance (12.5%, p > .05) there was no evidence of bias. Tredoux's E (32) which provides an estimate of how many line-up members sufficiently meet the modal description = 5.73 (95% CI = 4.41 – 8.18). This additionally suggests that most foils were plausible matches to the amalgamated description of the post-mortem photo. Together, these steps provide evidence that the final array was a fair test of participants' ability.

2.3 Design and Procedure

Police were invited to contribute by the MPS, and if interested clicked on a Qualtrics survey system link (33) and if relevant, provided permission to access past GFMT (16) data attached to an anonymised personal code. The GFMT consists of 40 pairs of simultaneously presented white-Caucasian faces in greyscale. Participants respond 'same' (20 faces) or 'different' (20 faces) to each pair. The test is scored out of 40.

Police were asked not to collaborate on the task. After providing informed consent they viewed the post-mortem photo presented above the randomly counterbalanced 2 x 4 eight-photo numbered array which always included the anti-mortem photo. Using a between-participants simultaneous face-matching design, they provided confidence ratings (0% = definitely not the same person to 100% = definitely the same person) as to whether the post-mortem photo depicted each array member. Adapted standardised eyewitness instructions, that the 'culprit may or may not be present' in a line-up, which reduces innocent suspect identification risk (34) read, 'it might be possible, but it is not definite that the deceased male photo and the missing male photo depict the same person'. To encourage confidence ratings to all images, they were informed that "unlike in a normal police investigation, we are also interested in 'not sure' answers". The entire procedure took less than 5 minutes.

3. Results

Participants' confidence ratings of the similarity between the post-mortem photo and each array image were compiled and the mean ratings are presented in Figure 1 as a function of group. As expected, the ante-mortem photo received the highest mean ratings (*M* = 53.96, *SD* = 33.08) in comparison to the combined mean ratings provided to the foils (*M* = 21.19, *SD* = 12.17), with super-recognisers' ratings to the target image higher than ratings from controls.

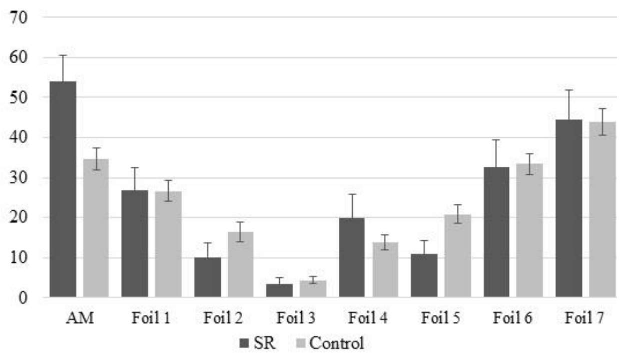

Figure 1: Mean (and SEM) confidence ratings provided to the ante-mortem photo (AM) and the seven foils as a function of group

Using SPSS statistical software (35), a 2 (group: super-recognisers, controls) x 2 (image: ante-mortem photo, foil mean) mixed ANOVA was conducted on the Table 1 data. The main effect of group was significant, $F(1, 162) = 5.80$, $p = .017$, $\eta^2 = .035$. Super-recognisers provided higher ratings than controls. Supporting the first hypothesis, the main effect of image was also significant, $F(1, 162) = 37.46$, $p < .001$, $\eta^2 = .188$. With strong effect sizes, the ante-mortem photo received higher ratings than the mean of foils. The interaction was also significant, $F(1, 162) = 8.22$, $p = .005$, $\eta^2 = .048$. A Bonferroni-corrected paired t-test found that the mean ratings to foils provided by the super-recognisers and the controls did not differ, $t(162) = 0.59$, $p > .2$, Cohen's $d = 0.13$; whereas in support of the second hypothesis, with a medium effect size, super-recogniser ratings to the ante-mortem photo were significantly higher than those from controls, $t(162) = 2.82$, $p < .05$, Cohen's $d = 0.60$. [1]

Additional analyses measured the advantage of drawing on the wisdom of the crowd using RStudio (36) with the pROC package (37). On the assumption that the ante- and post-mortem photos depicted one and the same person, the confidence ratings provided by each participant to the ante-mortem photo were treated as hits, those to the foils as false alarms. This meant that for the analysis, the target was labelled as 1 and all other foils were labelled as 0 in the target part of the AUC function in the pROC package. The *Area under the Curve* (AUC) was calculated for each participant. An independent-measures t-test confirmed that the mean *Area under the Curve* (AUC) (calculated for each participant separately) was significantly higher for super recognisers ($M = 0.75$, $SD = 0.25$) than controls ($M = 0.63$, $SD = 0.27$), $t(162) = 2.00$, $p = .048$, Cohen's d = 0.46.

We then ran a wisdom of the crowd analysis to look at whether crowds of super recognisers were more accurate than crowds of controls. For this we examined the effect of increasing crowds sizes. Following procedures described in [27,28] separate random sampling of super recognisers and controls was conducted on the confidence ratings. The selected participants confidence ratings were then averaged for each sample crowd.. This sampling procedure was conducted 25 times for each crowd size (the maximum number of super recognisers and so the maximum number of iterations possible in the crowd size of 1), this is also enough for the main conclusions to be consistent if the analysis were to be rerun, even if the exact figures would vary. For each iteration average the AUC was calculated.

---

[1] Note: The GFMT threshold for super-recognisers reported here was that used in the report to the coroner (minimum of 37 out of 40: 92.5%). Subsequent analyses were conducted with the GFMT threshold raised to 38 (95%) ($n = 15$ super-recognisers; $n = 149$ controls). The critical main effects of image, $F(1, 162) = 27.63$, $p < .001$, $\eta^2 = .146$, and the interaction, $F(1, 162) = 5.81$, $p = .017$, $\eta^2 = .035$ were significant and consistent with the conclusions reported here. A Bonferroni-corrected post-hoc t-test comparing super-recogniser and control ratings to the ante-mortem photo was marginally significant, $t(162) = 2.21$, $p < .1$, Cohen's d = 0.56, but with a similar effect size to that reported in the main text.

This was then averaged across the iterations for a crowd size. The results are reported for computed group sizes (crowds) of 1, 2, 4, 8 and 16. The upper limit was 16 as there were only 25 participants in the super-recogniser group.

Figure 1 shows the crowd effect as a function of sample size (one, two, four, eight, sixteen), and demonstrates that combining identity confidence ratings in this manner enhanced super-recognisers' AUC scores only. A 2 (group: super-recognisers, Controls) x 5 (crowd size: 1, 2, 4, 8, 16) ANOVA revealed a significant group main effect, $F(1, 240) = 187.49$, $p < .001$, $\eta^2 = .409$ with super-recognisers being more confident (90.7%) than controls (70.3%). There was also a significant group x crowd size interaction, $F(1, 240) = 6.37$, $p < .001$, $\eta^2 = .055$, driven with strong effect sizes by a significant increase with increasing crowd size for super-recognisers, $F(1, 120) = 6.65$, $p < .001$, $\eta^2 = .181$, but not controls, $F(1, 120) = 1.25$, $p = .292$, $\eta^2 = .040$. Indeed, the ratings provided at a single super-recogniser level were higher than the mean ratings given by the largest crowd of 16 controls, $t(39) = 5.60$, $p < .001$, $\eta^2 = .395$.

These analyses support the proposal that the post- and ante-mortem photo depict the same person, as super-recognisers' confidence ratings to the ante-mortem photo were as hypothesied higher than those from controls. In addition, the amalgamated confidence levels of the crowd of super-recognisers were as predicted higher than the individual decisions. *Figure 2* illustratrating the wisdom of the crowd scores supports this interpretation as for the super-recognisers but not controls, the scores for a group of 16 are almost optimal, as these crowds had an average AUC of 1.

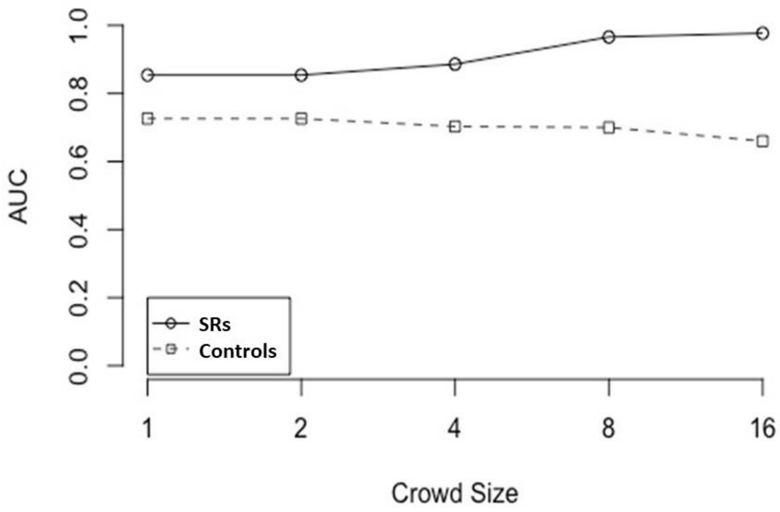

Figure 2: Accuracy judgements as assessed using AUC analyses for increasing crowd sizes separately for super-recognisers *(SRs) and controls on the assumption that the post-mortem and ante-mortem photograph depicted one and the same person.*

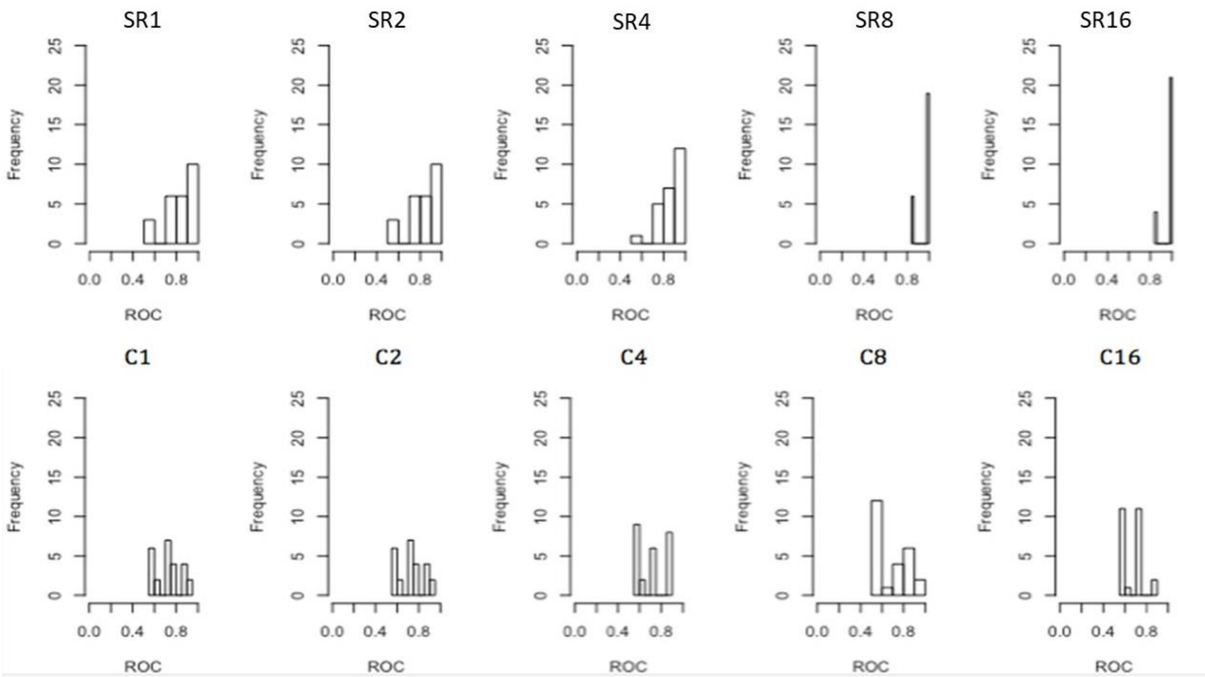

Figure 3. Frequency histograms showing the distribution of area under AUC values for each super-recogniser (SR) and control (C) size. Group AUC values were calculated separately for all possible combinations of participants at the initial solo and four levels of group size.

4. Discussion

This paper describes a novel methodology used to identify a post-mortem photograph of a drowned male taken more than 40 years previously. Super-recogniser and control police independently viewed the post-mortem photo and a simultaneously presented eight-person array. The array contained an ante-mortem photo of a man who went missing in the same month as the deceased's body was recovered. Supporting the hypotheses that the two images depicted one and the same person, mean confidence ratings were higher when matching the post-mortem and ante-mortem photo, than to the post-mortem photo and foils. Super-recognisers also provided higher confidence ratings to this pair of photos than controls. However, increasing the size of the randomly selected control crowds had no effect on confidence judgments. In contrast, successively increasing numbers of super-recognisers (1-2-4-8-16) in the crowd significantly increased confidence judgments, supporting the proposal that post- and ante-mortem photos depicted the same person.

The procedures for selecting foils for the array were designed to meet best practice recommendations for eyewitness procedures (for a review see 38), and to protect against cognitive bias (see 39), the researchers were provided with few case facts beforehand. All police participants were instructed not to confer, and they were blind to the randomly counterbalanced location of the ante-mortem photo in the line-up. On competition of the research, the first author of this paper provided a report to the investigating officer in the case, and only then were the researchers provided with full case details. This report primarily consisted of a description of the procedures listed in the current method and results, with some brief additional background information about super-recognition, individual differences in face processing ability, and best practice methods for the selection of foils. Based on these findings, and other evidence in the case, in 2015, a coroner ruled that the deceased and the missing male were "more likely than not" one and the same person, providing "sufficient proof' to support an application to add the missing man's name to the death certificate. It is not possible to determine how much weight was given by the coroner to the identification evidence described in the current report.

It should be acknowledged that simultaneous unfamiliar face matching, even with high-quality images is error prone, and although police super-recognisers may be superior at such tasks (e.g. 21), other techniques might be more suitable. Indeed, if the family of the missing man had confidently identified the post-mortem image as that of their relative, the procedures described here would not have been required. Furthermore, the international Facial Identification Scientific Working Group (FISWG) (40) recommends the use of morphological face analyses for facial comparison purposes. With this, facial features are compared and classified. This was considered unfeasible by the MPS Facial Image Comparison team. As such, the method described here could be attempted in future cases if no other traditional methodology is possible. It could also be performed when rapid decisions of identity are required perhaps during the initial stages of an investigation when police may wish to determine whether a visual image matches that of a suspect on a database. Creating an array of foils here required accessing archive photographs as quality had faded. It should normally be far quicker to assemble an array of contemporary foils.

Although a growing body of empirical evidence suggests that the wisdom of the crowd technique provides enhanced identification accuracy (28,29), to the authors knowledge this is the first police investigation to employ the technique. Additional validation is required to ensure reliability in future investigations. For this, empirical evaluations of specialist super-recognisers could generate data of error rates across a wide range of face matching tasks. Rigorous assessment using a variety of images in which ground truth of guilt is known would be necessary.

If applied to additional investigations in future, super-recognisers could be conceived of as a ‘crowdsourcing tool’, for forensic analysts to draw on in order to produce statistical probabilities of a match, in the manner of other forensic techniques. Conceptually, the data produced would be analogous to that produced by typical scientific measuring equipment, only the data would be generated from humans rather than a machine. In cases in which forensic facial examiners are unable to provide any more than weak support for a mismatch or match, it could provide additional case evidence alongside traditional forensic facial comparison techniques. Regular checks using ground truth images could also be randomly inputted into case work in order to ensure maintenance of the reliability of each super-recogniser as well as the crowd.

There are a number of limitations of the current research that should be acknowledged. The GFMT (16) used to attribute superior-face-matching ability has been criticised for its ceiling effects (e.g. 20), and the super-recogniser threshold employed here was lower than that associated with the highest levels of ability. This was partly due to necessity, as only a minority of MPS officers from the super-recogniser pool of over 100 who had taken the GFMT test at the time contributed to this investigation. A more robust test of ability would be recommended if similar procedures were to be used in other cases. Furthermore, the face matching ability of most police controls was not assessed. Although this did not impact the results reported here or case conclusions, it limits theoretical understanding of the Wisdom of the Crowd effect. Most relevant research has found improvements in cognitive estimation tasks from this procedure by typical participants (25-28), and it is not clear why null effects were found here. Instead, positive effects from the Wisdom of the Crowd procedure were only significant for the super-recognisers. In terms of signal amplification (26), it might suggest super-recognisers could detect an idiosyncratic identification signal from the ante- and post-mortem photos, so that the super-recogniser crowd performed at a higher level than individual super-recognisers. In contrast, controls were unable to extract any idiosyncratic identification signal, and were effectively inducing guessing noise into the process. However, without knowing control's face matching skills, it is not possible to determine what level of ability is required to extract a signal.

An additional limitation is that the mean confidence rating by super-recognisers of just over 50% for a match between the ante- and post-mortem, even if higher than to any individual foil may not appear conclusive. However, low ratings may be a consequence of instructions stating that unlike in normal police investigations, 'not sure' responses were encouraged. This instruction may have implicitly encouraged lower ratings. Adapted instructions might prove to be more suitable and should be evaluated in future research.

Conclusions: In the case study described here, super-recognisers were more likely than controls to confidently match a post-mortem photo with an ante-mortem photo than they were to match the post-mortem photo to foils; with larger crowds of super-recognisers even more likely to do so than a single super-recogniser. Although the wisdom of the crowd method described in this paper does not guarantee identity, it did support alternative case evidence. As such, assuming the weight of evidence is correct, the results suggest that a super-recogniser crowd may provide 'wiser' identity-decisions than a crowd made up of those with typical face recognition ability. Such a method could be applied to generate forensically reliable matching decisions in alternative cases in which photographic identification is a key element in a police investigation. The potentially quick-to-arrange technique might best provide the first investigatory evidence of identification in a case for later evidential corroboration.

References

1. Soomer H, Ranta H, Penttila A. Identification of victims from the M/S Estonia. Int J Legal Med 2001;114:259–62.
2. Padmitilaka K. Tsunami disaster 2004. Emergency management and aftercare in General Hospital Matara. Matara: Matara General Hospital, 2005.
3. Caplova Z, Obertova Z, Gibelli DM, Mazzarelli D, Fracasso T, Vanezis P et al. The reliability of facial recognition of deceased persons on photographs. J Forensic Sci 2017;62(5):1286-91.
4. Caplova Z, Obertova Z, Gibelli DM, De Angelis D, Mazzarelli D, Sforza C. et al. Personal identification of deceased persons: An overview of the current methods based on physical appearance. J Forensic Sci 2017: doi: 10.1111/1556-4029.13643
5. Wilkinson C. Craniofacial analysis and identification. In: Valentine T, Davis JP, editors. Forensic facial identification: theory and practice of identification from eyewitnesses, composites and CCTV. Chichester: Wiley-Blackwell, 2015:93-126.
6. Bruce V, Henderson Z, Newman C, Burton AM. Matching identities of familiar and unfamiliar faces caught on CCTV images, J. Exp. Psychol: Appl. 2001;7(3):207-18.
7. Burton AM, Wilson S, Cowan M, Bruce V. Face recognition in poor-quality video: evidence from security surveillance. Psy Sci 1999;10(3):243–8.
8. Bruce V, Henderson Z, Greenwood K, Hancock P, Burton AM, Miller P. Verification of face identities from images captured on video. J Exp Psych App 1999;5(4):339-60.
9. Jenkins R, White D, Van Montfort X, Burton AM. Variability in photos of the same face, Cognition 2011;121(3):313-23.
10. Davis JP, Valentine T. Human verification of identity from photographic images. In: Valentine T, Davis JP, editors. Forensic facial identification: theory and practice of identification from eyewitnesses, composites and CCTV. Chichester: Wiley-Blackwell, 2015:211-38.

11. Megreya AM, Sandford A, Burton AM. Matching face images taken on the same day or months apart: the limitations of photo ID. App Cog Psych 2013;27(6):700-6.

12. Estudillo A, Bindemann M. Generalization across view in face memory and face matching. i-Per 2014;5(7):589-601.

13. Bindemann M, Attard J, Leach A, Johnston R. The effect of image pixelation on unfamiliar face matching. App Cog Psych 2013;27(6):707-17.

14. Bindemann M, Avetisyan M, Rakow T. Who can recognize unfamiliar faces? Individual differences and observer consistency in person identification. J Exp Psych App 2012;18(3):277-91.

15. Bobak AK, Bennetts RJ, Parris BA, Jansari A, Bate S. An in-depth cognitive examination of individuals with superior face recognition skills. Cortex 2016;82:48-62.

16. Burton AM, White D, McNeill A. The Glasgow face matching test. Behav Res Meth 2010;42(1):286-91.

17. Megreya AM, Burton AM. Unfamiliar faces are not faces: evidence from a matching task. Mem & Cog 2006;34(4):865-76.

18. Russell R, Duchaine B, Nakayama K. Super-recognizers: people with extraordinary face recognition ability. Psych B Rev 2009;16(2):252-257.

19. Bobak AK, Hancock PJ, Bate S. Super-recognizers in action: evidence from face matching and face memory tasks. App Cog Psych 2016;30(1)81-91.

20. Davis JP, Lander K, Evans R, Jansari A. Investigating predictors of superior face recognition ability in police super-recognisers. App Cog Psych 2016;30(6):827-40.

21. Robertson DJ, Noyes E, Dowsett AJ, Jenkins R, Burton AM. Face recognition by Metropolitan Police super-recognisers. PloS one 2016;11(2):e0150036–8.

22. Davis JP, Lander K, Jansari A, I never forget a face! The Psychologist 2013;26:726-29.
23. Davis JP, Treml F, Forrest C, Jansari A, Identification from CCTV: Assessing police super-recognisers ability to spot faces in a crowd and susceptibility to change blindness. App Cog Psych 2018;*32:*337-353.
24. Brewer N, Weber N, Wootton D, Lindsay DS. Identifying the bad guy in a lineup using confidence judgments under deadline pressure. Psych Sci 2012;23:1208-14.
25. Surowiecki J. The wisdom of crowds. New York: Random House 2004.
26. Danileiko I, Lee MD. A Model‐Based Approach to the Wisdom of the Crowd in Category Learning. Cog Sci;42, 861-83.
27. White D, Burton AM, Kemp RI, Jenkins R. Crowd effects in unfamiliar face matching. App Cog Psych 2013;27(6):769–77.
28. White D, Phillips PJ, Hahn CA, Hill M, O'Toole AJ. Perceptual expertise in forensic facial image comparison. Proc R Soc B: Biological Sci 2015;282:1814-22.
29. O'Toole AJ, Abdi H, Jiang F, Phillips PJ. Fusing face-verification algorithms and humans. IEEE Trans Syst, Man & Cyb B 2007;37(5), 1149-55.
30. Phillips PJ, Yates AN, Hu Y, Hahn CA, Noyes E, Jackson K, Cavazos JG, Jeckeln G, Ranjan R, Sankaranarayanan S, Chen J, Castillo CD, Chellappa R, White D, O'Toole A. Face recognition accuracy in forensic examiners, super-recognisers and algorithms. Pro Nat Acad Sci 2018;115(24), 6171-77.
31. Police and Criminal Evidence Act 1984: Code of Practice for the identification of persons by Police Officers. London: TSO (2017).
32. Tredoux CG. Statistical inference on measures of lineup fairness. Law Hum Behav 1998; 22(2):217-37.
33. Qualtrics, USA, http://www.qualtrics.com (accessed 9.01.17) 2017.

34. Clark SE. Costs and benefits of eyewitness identification reform: psychological science and public policy. Persp Psych Sci 2012;7(3):238-59.

35. IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.

36. RS RStudio Team. RStudio: Integrated Development for R. Boston, MA: RStudio Inc. 2015.

37. Robin X, Turck N, Hainard A, Tiberti N, Lisacek N, Sanchez JC, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bio 2011;12, p. 77.

38. Valentine T, Davis JP, editors. Forensic facial identification: theory and practice of identification from eyewitnesses, composites and CCTV. Chichester: Wiley-Blackwell, 2015.

39. Kassin SM, Dror IE, Kukucka J. The forensic confirmation bias: problems, perspectives, and proposed solutions. J App Res Mem Cog 2013;2(1):42-52.

40. Facial Identification Scientific Working Group (FISWC): Facial image comparison feature list for morphological analysis (Version 1.0 2013.11.22), FISWG, https://fiswg.org/FISWG_1to1_Checklist_v1.0_2013_11_22.pdf (accessed 9.01.17) 2013.

41. Fussey P, Murray D. Independent report on the London Metropolitan Police Service's trial of live facial recognition technology. University of Essex: Human Rights Centre, https://48ba3m4eh2bf2sksp43rq8kk-wpengine.netdna-ssl.com/wp-content/uploads/2019/07/London-Met-Police-Trial-of-Facial-Recognition-Tech-Report.pdf (accessed 7.07.19) 2019.