Feature-by-feature comparison and holistic processing in unfamiliar face matching

Ahmed M. Megreya

Department of Psychological Sciences, College of Education, Qatar University, Doha, Qatar

ABSTRACT

Identity comparisons of photographs of unfamiliar faces are prone to error but imperative for security settings, such as the verification of face identities at passport control. Therefore, finding techniques to improve face-matching accuracy is an important contemporary research topic. This study investigates whether matching accuracy can be enhanced by verbal instructions that address feature comparisons or holistic processing. Findings demonstrate that feature-by-feature comparison strategy had no effect on face matching. In contrast, verbal instructions focused on holistic processing made face matching faster, but they impaired accuracy. Given the recent evidence for the heredity of face perception and the previously reported small or no improvements of face-matching ability, it seems reasonable to suggest that improving unfamiliar face matching is *not* an easy task, but it is presumably worthwhile to explore new methods for improvement nonetheless.

INTRODUCTION

Matching a face to a photo ID is a very common procedure in security settings although a large number of experimental studies showed that face matching is rather error-prone (e.g., for a review see *Robertson, Middleton & Burton, 2015*). For example, *Bruce et al.* and colleagues (*1999*) asked observers to match a target face presented above a target-present/target-absent 10-face line-up. All images were taken on the same day, under good lighting conditions, and were presented in a full-face view. In spite of these optimal conditions, which could be never met in any real-life scenario, participants' performance was rather low, with an error rate of 30% for target-present and target-absent trials.

This low level of performance has been widely replicated (e.g., *Bindemann et al., 2012*; *Megreya & Burton, 2008*; *Megreya, White & Burton, 2011*; *Megreya & Bindemann, 2013*; *Megreya & Bindemann, 2015*; *Megreya & Bindemann, 2017*) even when the heavy demands of this 1-in-10 array methodology were remarkably reduced to a 1-in-1 face-matching task using a range of face-matching databases (e.g., *Bruce et al., 2001*; *Burton, White & McNeill, 2010*; *Henderson, Bruce & Burton, 2001*; *Megreya et al., 2012*; *Megreya & Burton, 2008*). In addition, high error rates were observed using a range of person-to-photo matching tasks, which resemble to a great extent the widely used verification procedures in security settings (*Davis & Valentine, 2009*; *Kemp, Towell & Pike, 1997*; *Megreya & Burton, 2008*; *White et al.,*

Corresponding author
Ahmed M. Megreya,
amegreya@qu.edu.qa

2014). Furthermore, the performance of police and passport officers was rather low using the photo-to-photo and person-to-photo matching tasks (*Robertson et al., 2016*; *White et al., 2014*).

In a stark contrast, it has been known that the recognition of familiar faces is robust even under very challenging circumstances. For example, people can easily recognize familiar faces using highly degraded images (*Burton et al., 1999*; *Watier & Collin, 2009*), and even after very long retention intervals (*Bahrick, Bahrick & Wittlinger, 1975*). In addition, a short familiarization procedure was found to improve face matching performance using the 1-in-10 face matching task (*Megreya & Burton, 2006*; *Megreya & Burton, 2007*).

Accordingly, it has been proposed that unfamiliar face matching relies on image-based processes, while familiar face recognition engages a more sophistic and specialized type of processing (for a review, see *Hancock, Bruce & Burton, 2000*). For example, *Megreya & Burton (2006)* found no correlation between matching upright familiarized and unfamiliar faces; but there were strong positive correlations between matching upright and inverted unfamiliar faces and between matching upright unfamiliar faces and inverted familiar faces. As face inversion impairs configural, but not featural, processing (e.g., for a review see e.g., see *Bartlett, Searcy & Abdi, 2003*), *Megreya & Burton (2006)* suggested that unfamiliar faces in the matching tasks are treated as ''images'' or ''simple visual patterns'' and matched on this basis without domain-specific expertise. Consistently, *Lobmaier & Mast (2007)* found that the recognition of new faces, compared to old ones, relies relatively more on the processing of featural information, but with the course of familiarization, the significance of holistic processing increased.

The suggestion that unfamiliar face matching is a feature-based process, which relies on comparison of individual facial parts such as the eyes, nose or mouth, indicates that face matching could be improved by directing observers to the diagnostic features that are particularly beneficial for face matching. In fact, police and passport officers usually receive feature-by-feature comparison training (for reviews, see, e.g., *Robertson, Middleton & Burton, 2015*; *Towler, Kemp & White, 2017*), though the results of experiments that examined the effectiveness of this strategy are inconsistent (*Berman & Cutler, 1998*; *Towler, White & Kemp, 2014*; *Towler, White & Kemp, 2017*; *Woodhead, Baddeley & Simmonds, 1979*). Specifically, some experiments found that feature-by-feature comparison have no benefits for face recognition (*Woodhead, Baddeley & Simmonds, 1979*) and even impaired it (*Berman & Cutler, 1998*). Nevertheless, a more recent study by *Towler, White & Kemp (2017)* found that a feature similarity rating task could improve unfamiliar face matching. Specifically, a pair of faces was presented above a list of eleven facial features (ears, jawline, chin, check area, face shape, eyes, mouth area, forehead, mouth, nose, and scars/blemishes), and observers were asked to rate the similarity of those features. They were then asked to decide whether the face pair showed the same person or different people. The results demonstrated that this rating task improved matching accuracy on identity match, but not on non-match, trials. In addition, *Towler, Kemp & White (2017)* found no benefits for rating the similarity of personality traits that have been thought to rely on holistic processing (for a review, see, e.g., *Coin & Tiberghien, 1997*). Consistently, *White et al. (2015)* found that facial forensic examiners who received training on feature-by-feature analysis of

facial images outperformed students in face-matching task. This superiority was especially strong when faces were turned upside down, suggesting that the feature-by-feature strategy employed by forensic examiners might be the locus of their expertise (*White et al., 2015*).

The same suggestion might also indicate that increasing the reliance of holistic processing during matching unfamiliar faces may improve participants' performance. Holistic processing refers to the ability to recognize a face as a gestalt, rather than a collection of distinguishable features. This type of processing has been thought to underlie the robustness of familiar face recognition (e.g., for reviews see *Maurer, Le Grand & Mondloch, 2002*; *Tanaka & Gordon, 2011*). However, the results of previous experiments that examined the association between face recognition accuracy and holistic processing were mixed. For example, *Wang et al. (2012)* found that face recognition accuracy correlated with the extent to which observers processed the faces holistically (as indexed by the composite-face effect and the whole-part effect). Nevertheless, *Konar, Bennett & Sekuler (2010)* found a close to zero correlation between unfamiliar face matching accuracy and the magnitude of the composite face effect, which is a well-established measure of holistic processing.

The aim of this study was to examine the effectiveness of feature-by-feature comparison versus holistic processing on matching unfamiliar faces. Participants were instructed to use either a feature-by-feature matching strategy (where they could compare the features of two faces piece-by-piece) or a holistic matching strategy (where they could encode each face as a whole piece, rather than a collection of separable features) during matching pairs of unfamiliar faces. Because of the importance of face matching for security settings such as the country borders and the well-documented low levels performance on face matching, exploring new techniques for improving unfamiliar face matching is of particular importance and one of the central questions in the ongoing face identification literature (see, e.g., *Robertson, Middleton & Burton, 2015*).

METHOD

Participant

Forty under-graduate students from Qatar University volunteered to participate in this experiment ($M_{age} = 20.5$, $SD_{age} = 0.7$; 65% female). All participants had normal or corrected to normal vision. Ethical approval for participation in this study was provided by Qatar University's institutional review board (QU-IRB 576-EA/16) and all methods were performed in accordance with the QU-IRB guidelines and regulations.

Stimuli

A total of 120 Arab match/mismatch pairs of faces were used as stimuli in this experiment that were taken from an Egyptian face-matching database (*Megreya & Burton, 2008*). Each pair consisted of a target video still (taken by a video camera) and a test photograph (taken by a digital camera), which depicted the same face or two different persons. Pairing in mismatch pairs was made according to subjective overall similarity between the faces. All target and test images were taken at the same time and under the same lighting conditions. All faces were shown in a frontal view, with a neutral expression, and any extraneous

background was removed. The size of each image was approximately 5 cm × 7 cm and all images were presented in grey-scale (for full details of these stimuli see (*Megreya & Burton, 2008*), where examples of face pairs can be found).

Procedure

All participants were tested individually using an Apple laptop. Experimental software was used to present face-matching stimuli and to record participants' responses. In two successive sessions, the participants' task was to make match/mismatch decisions for the face pairs, which were presented randomly. Each pair was presented until a participant's response was made by pressing two labelled keys in the standard computer keyboard. All participants were presented with 30 face pairs (15 matches and 15 mismatches) while half of the participants were given instructions that encouraged them to use either a feature-by-feature matching strategy or a holistic matching strategy. In the feature-by-feature matching instruction, participants were asked to match the faces by comparing the faces feature-by-feature. In the holistic matching instruction, participants were asked to look at the faces globally and not to focus on their parts. Immediately after these instructions, all participants were presented with 30 new face pairs (15 matches and 15 mismatches). Notably, match/mismatch pairs were counter-balanced across participants. Therefore, each face target was equally seen in match and mismatch pairs across the experiment, and there was no repetition of face identities across all trials. The face-matching tasks were also counter-balanced across the pre- and post-training sessions.

RESULTS

The accuracy of face matching was calculated using three main measures. Correct Identification refers to the correct positive decision that the two faces are of the same person. Correct Rejection refers to the correct negative decision that the two faces depict two different people. Overall Accuracy refers to the total correct responses in match and mismatch pairs that were calculated by combining correct identification and correct rejections. Figure 1 shows the percentage averages for participants' performances in this experiment. These data were subjected to three 2 (instructions: holistic processing versus feature-by-feature comparison) × 2 (training: before versus After) mixed-design, where the instructions variable was treated as a between-participant factor while training was treated as a within-participant factor. There were significant interactions between instructions and training for overall accuracy, $F(1, 38) = 8.87$, $p < 0.01$, $\eta_p^2 = 0.19$ and correct rejections, $F(1, 38) = 14.50$, $p < 0.01$, $\eta_p^2 = 0.27$, but not for correct identifications, $F < 1$. A series of paired-sample t-tests between participants' performances before and after training (with alpha adjusted to $0.05/2 = 0.05$ for two comparisons) revealed no significant effects for feature instructions on overall accuracy, $t(19) = 0.40$, $p = 0.69$, Cohen' $d = 0.09$, and correct rejections, $t(19) = 0.83$, $p = 0.42$, Cohen' $d = 0.27$. However, holistic instructions had detrimental effects on overall accuracy, $t(19) = 3.80$, $p = 0.001$, Cohen' $d = 1.17$, and correct rejections, $t(19) = 4.21$, $p < 0.001$, Cohen' $d = 1.18$.

In addition, the averages of medians of the response times were calculated for correct identification and correct rejection (see Fig. 2). These data were also subjected to two
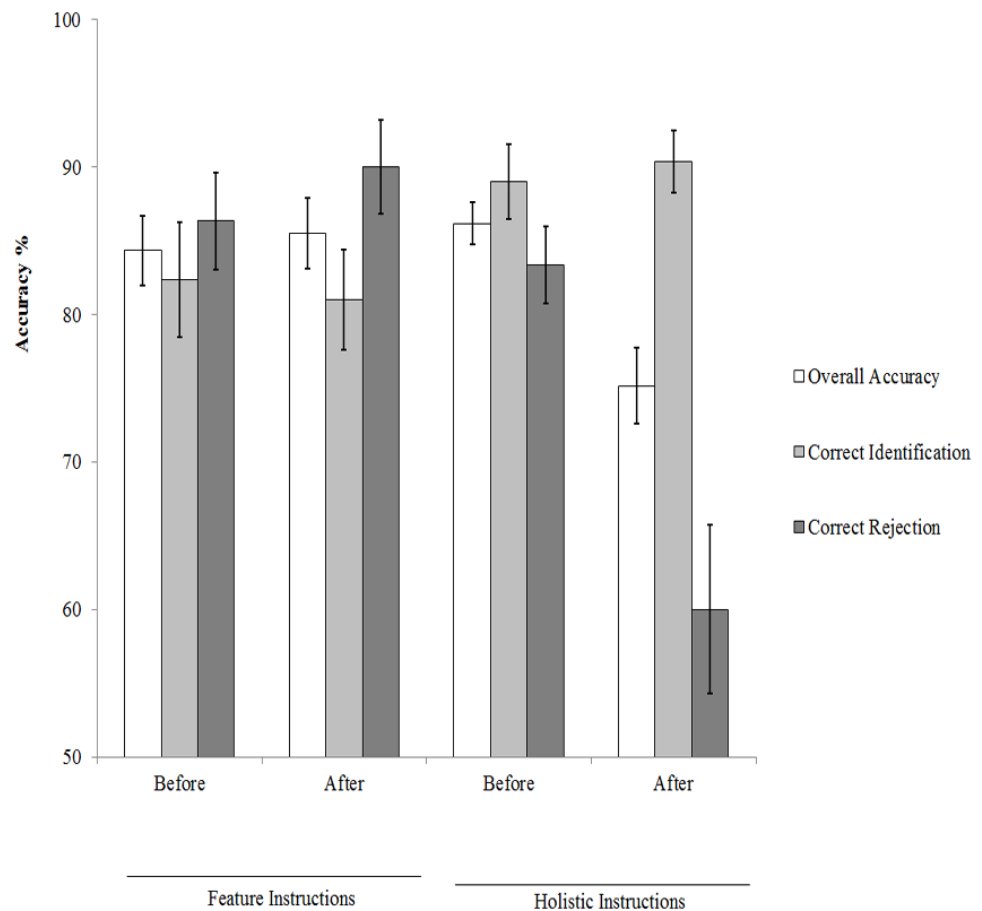
Figure 1 **Participants' performance in this experiment: accuracy measures.**

2×2 mixed-design ANOVAs, which revealed interactions between instructions and training for correct identification, $F\ (1,38) = 5.86$, $p = 0.02$, $\eta_p^2 = 0.26$ and correct rejections $F\ (1,38) = 10.69$, $p < 0.001$, $\eta_p^2 = 0.28$. A series of paired-sample t-tests between participants' performances before and after training (with alpha adjusted to $0.05/2 = 0.05$ for two comparisons) revealed no significant effects for feature instructions on correct identification, $t\ (19) = 0.86$, $p = 0.40$, Cohen' $d = 0.20$, and correct rejections, $t\ (19) = 0.49$, $p = 0.63$, Cohen' $d = 0.09$. However, participants were quicker to make the face-matching decisions after receiving holistic instructions using correct identification, $t\ (19) = 4.63$, $p \leq 0.001$, Cohen' $d = 1.60$, and correct rejections, $t\ (19) = 4.44$, $p < 0.001$, Cohen' $d = 1.24$.

DISCUSSION

This experiment examined the effects of verbal instructions focused on feature-by-feature comparisons and holistic processing on unfamiliar face matching. The results showed that feature-by-feature comparison strategy did not benefit face-matching accuracy and latency (see Figs. 1 and 2). This finding converges with the results of an early study (*Woodhead,*
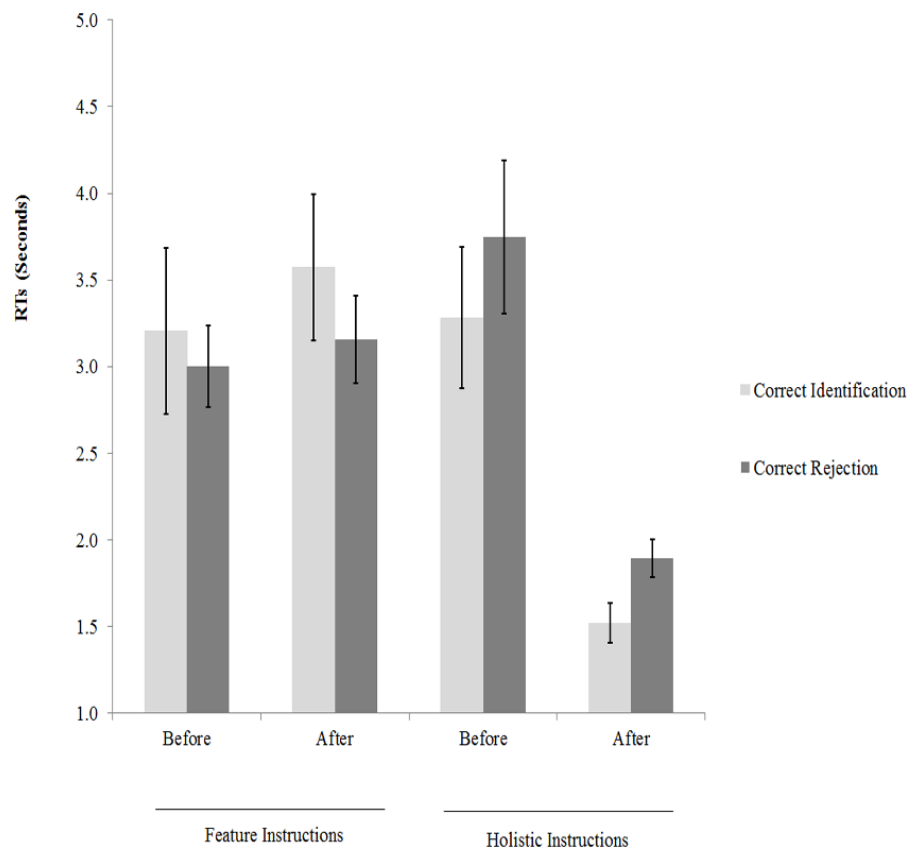
Figure 2 Participants' performance in this experiment: response time measures.

Baddeley & Simmonds, 1979), in which a three-day training course emphasized on "isolated" facial features failed to make improvements on face recognition memory as well as face matching across different views and expressions. However, *Towler, White & Kemp (2017)* found that the similarity ratings of a list of specific facial features, prior to matching decisions, improved face-matching accuracy. Notably, there are several methodological differences that can explain the contrast between the results of the present experiment and those of *Towler, White & Kemp (2017)* study. For example, the present experiment used verbal instructions, whereas *Towler, White & Kemp (2017)* used a feature similarity rating task. Therefore, those verbal instructions might not cause participants to spend longer comparing individual features to the same extent as *Towler, White & Kemp (2017)*.

The present experiment also found that verbal instructions focusing on holistic processing made participants' performances faster (see Fig. 2), but they harmed accuracy (see Fig. 1). It has been commonly thought that faces are processed holistically rather than as a corpus of individual parts (e.g., for reviews see *Maurer, Le Grand & Mondloch, 2002*; *Piepers & Robbins, 2012*; *Tanaka & Gordon, 2011*; but for a different view see *Burton et al., 2015*). This holistic processing of faces is assumed to emerge very rapidly (*Goffaux & Rossion, 2006*; *Richler et al., 2009*). Consistently, encouraging observers to adopt a holistic processing strategy during matching faces significantly speeded performance on both match

and mismatch trials. Importantly however, this response time advantage was associated with a lower level of face matching accuracy, especially in mismatch trials (see Fig. 1), suggesting that holistic processing strategy induced a speed-accuracy tradeoff.

CONCLUSION

Previous studies reported that the face-matching performance of highly experienced passport and security officers was error-prone and similar to naïve participants (*White et al., 2014*; *Wirth & Carbon, 2017*). Existing data indicates that methods for improving matching accuracy in security settings are required (see, e.g., *Robertson, Middleton & Burton, 2015*; *White et al., 2014*). The current study indicates that verbal instructions focusing on feature-by-feature comparisons could not improve face matching. Conversely, verbal instructions focusing holistic processing could improve face matching latency but significantly harm accuracy. Given the genetic basis of face perception (*Wilmer et al., 2010*; *Zhu et al., 2010*) and the small or no improvements of face matching ability reported in this experiment and in previous studies (e.g., for a recent review see, *Young & Burton, 2018*), it seems reasonable to suggest that improving unfamiliar face matching is *not* an easy task, but it is presumably worthwhile to explore new methods for improvement nonetheless.

ACKNOWLEDGEMENTS

The statements made herein are solely the responsibility of the author.

Funding
This study was made possible by NPRP grant #NPRP9-272-5-033 from the Qatar National Research Fund (a member of Qatar foundation). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures
The following grant information was disclosed by the author:
Qatar National Research Fund: #NPRP9-272-5-033.

Competing Interests
The author declares there are no competing interests.

Author Contributions
- Ahmed M. Megreya conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.

Human Ethics
The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

Ethical approval for participation in this study was provided by Qatar University's institutional review board (QU-IRB) and all methods were performed in accordance with the QU-IRB guidelines and regulations.

Data Availability

The raw data has been provided as a Supplemental File.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.4437#supplemental-information.

REFERENCES

Bahrick HP, Bahrick PO, Wittlinger RP. 1975. Fifty years of memory for names and faces: a cross-sectional approach. *Journal of Experimental Psychology: General* **104(1)**:54–75 DOI 10.1037/0096-3445.104.1.54.

Bartlett JC, Searcy JH, Abdi H. 2003. What are the routes to face recognition? In: Peterson A, Rhodes G, eds. *Perception of faces, objects, and scenes: analytical and holistic processes.* New York: Oxford University Press, 21–52.

Berman GL, Cutler BL. 1998. The influence of processing instructions at encoding and retrieval on face recognition accuracy. *Psychology, Crime & Law* **4**:89–106 DOI 10.1080/10683169808401751.

Bindemann M, Gillattm K, Avetisyan M, Megreya AM. 2012. Recognizing faces seen alone or with others: why are two heads worse than one? *Perception* **41**:415–435 DOI 10.1068/p6922.

Bruce V, Henderson Z, Greenwood K, Hancock PJB, Burton AM, Miller P. 1999. Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied* **5**:339–360 DOI 10.1037//1076-898X.5.4.339.

Bruce V, Henderson Z, Newman C, Burton AM. 2001. Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied* **7**:207–218.

Burton AM, Schweinberger S, Jenkins R, Kaufmann J. 2015. Arguments against a 'configural processing' account of familiar face recognition. *Perspectives on Psychological Science* **10(4)**:482–496 DOI 10.1177/1745691615583129.

Burton AM, White D, McNeill A. 2010. The Glasgow face matching test. *Behavior Research Methods* **42(1)**:286–291 DOI 10.3758/BRM.42.1.286.

Burton AM, Wilson S, Cowan M, Bruce V. 1999. Face recognition in poor-quality video: evidence from security surveillance. *Psychological Science* **10(3)**:243–248 DOI 10.1111/1467-9280.00144.

Coin C, Tiberghien G. 1997. Encoding activity and face recognition. *Memory* **5**:545–568 DOI 10.1080/741941479.

Davis JP, Valentine T. 2009. CCTV on trial: matching video images with the defendant in the dock. *Applied Cognitive Psychology* **23**:482–505 DOI 10.1002/acp.1490.

Goffaux V, Rossion B. 2006. Faces are "spatial": holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception & Performance* **32**:1023–1039.

Hancock PJB, Bruce V, Burton AM. 2000. Recognition of unfamiliar faces. *Trends in Cognitive Sciences* **4**:330–337 DOI 10.1016/S1364-6613(00)01519-9.

Henderson Z, Bruce V, Burton AM. 2001. Matching the faces of robbers captured on video. *Applied Cognitive Psychology* **15**:445–464 DOI 10.1002/acp.718 .

Kemp R, Towell N, Pike G. 1997. When seeing should not be believing: photographs, credit cards and fraud. *Applied Cognitive Psychology* **11**:211–222 DOI 10.1002/(SICI)1099-0720(199706)11:3<211::AID-ACP430>3.0.CO;2-O.

Konar Y, Bennett PJ, Sekuler AB. 2010. Holistic processing is not correlated with face-identification accuracy. *Psychological Science* **21**:38–43 DOI 10.1177/0956797609356508.

Lobmaier JS, Mast FW. 2007. Perception of novel faces: the parts have it! *Perception* **36**:1660–1673 DOI 10.1068/p5642.

Maurer D, Le Grand R, Mondloch CJ. 2002. The many faces of configural processing. *Trends in Cognitive Science* **6**:255–260 DOI 10.1016/S1364-6613(02)01903-4.

Megreya AM, Bindemann M. 2013. Individual differences in personality and face identification. *Journal of Cognitive Psychology* **25**:30–37 DOI 10.1080/20445911.2012.739153.

Megreya AM, Bindemann M. 2015. Developmental improvement and age-related decline in unfamiliar face matching. *Perception* **44**:2–44 DOI 10.1068/p7825.

Megreya AM, Bindemann M. 2017. A visual processing advantage for deaf observers in face and object matching. *Scientific Reports* **7**:41133 DOI 10.1038/srep41133.

Megreya AM, Bindemann M, Havard C, Burton AM. 2012. Identity-lineup location influences target selection: evidence from eye movements. *Journal of Police & Criminal Psychology* **27**:167–178 DOI 10.1007/s11896-011-9098-7.

Megreya AM, Burton AM. 2006. Unfamiliar faces are not faces: evidence from a matching task. *Memory & Cognition* **34**:865–876 DOI 10.3758/BF03193433.

Megreya AM, Burton AM. 2007. Hits and false positives in face matching: a familiarity based dissociation. *Perception & Psychophysics* **69**:1175–1184 DOI 10.3758/BF03193954.

Megreya AM, Burton AM. 2008. Matching faces to photographs: poor performance in eyewitness memory (without the memory). *Journal of Experimental Psychology: Applied* **14**:364–372 DOI 10.1037/a0013464.

Megreya AM, White D, Burton AM. 2011. The other race effect does not rely on memory: evidence from a matching task. *Quarterly Journal of Experimental Psychology* **64**:1473–1483 DOI 10.1080/17470218.2011.575228.

Piepers DW, Robbins RA. 2012. A review and clarification of the terms "holistic," "configural," and "relational" in the face perception literature. *Frontiers in Psychology* **3**:1–11 DOI 10.3389/fpsyg.2012.00559.

Richler JJ, Mack ML, Gauthier I, Palmeri TJ. 2009. Holistic processing of faces happens at a glance. *Vision Research* **49**:2856–2861 DOI 10.1016/j.visres.2009.08.025.

Robertson DJ, Middleton R, Burton AM. 2015. From policing to passport control. The limitations of photo ID. *Keesing: The Journal of Documents and Identity* **46**:3–8.

Robertson DJ, Noyes E, Dowsett A, Jenkins R, Mike Burton A. 2016. Face recognition by metropolitan police super-recognisers. *PLOS ONE* **11(2)**:e0150036 DOI 10.1371/journal.pone.0150036.

Tanaka JW, Gordon I. 2011. Features, configuration and holistic face processing. In: Calder AJ, Rhodes G, Johnson MH, Haxby JV, eds. *The Oxford handbook of face perception*. Oxford: OUP, 15–30.

Towler A, Kemp RI, White D. 2017. Unfamiliar face matching systems in applied settings. In: Bindemann M, Megreya AM, eds. *Face processing: systems, disorders and cultural differences*. New York: Nova Science Publishing, Inc.

Towler A, White D, Kemp RI. 2014. Evaluating training methods for facial image comparison: the face shape strategy does not work. *Perception* **43**:214–218 DOI 10.1068/p7676.

Towler A, White D, Kemp RI. 2017. Evaluating the feature comparison strategy for forensic face identification. *Journal of Experimental Psychology: Applied* **23**:47–58 DOI 10.1037/xap0000108.

Wang R, Li J, Fang H, Tian M, Liu J. 2012. Individual differences in holistic processing predict face recognition ability. *Psychological Science* **23**:169–177 DOI 10.1177/0956797611420575.

Watier NN, Collin CA. 2009. Effects of familiarity on spatial frequency thresholds for face matching. *Perception* **38**:1497–1507 DOI 10.1068/p6396.

White D, Kemp RI, Jenkins R, Matheson M, Burton AM. 2014. Passport officers' errors in face matching. *PLOS ONE* **9(8)**:103510 DOI 10.1371/journal.pone.0103510.

White D, Phillips PJ, Hahn CA, Hill M, O'Toole AJ. 2015. Perceptual expertise in forensic facial image comparison. *Proceedings of the Royal Society B: Biological Sciences* **282**:1814–1822.

Wilmer JB, Germine L, Chabris CF, Chatterjee G, Williams M, Loken E, Nakayama K, Duchaine B. 2010. Human face recognition ability is specific and highly heritable. *Proceeding of the National Academy of the Science of the United States of America* **107**:5238–5241 DOI 10.1073/pnas.0913053107.

Wirth BE, Carbon CC. 2017. An easy game for frauds? Effects of professional experience and time pressure on passport-matching performance. *Journal of Experimental Psychology: Applied* **23**:138–157 DOI 10.1037/xap000014.

Woodhead MM, Baddeley AD, Simmonds DCV. 1979. On training people to recognize faces. *Ergonomics* **22(3)**:333–343 DOI 10.1080/00140137908924617.

Young AW, Burton AM. 2018. Are we face experts? *Trends in Cognitive Psychology* **22**:100–110 DOI 10.1016/j.tics.2017.11.007.

Zhu Q, Song Y, Hu S, Li X, Tian M, Zhen Z, Dong Q, Kanwisher N, Liu J. 2010. Heritability of the specific cognitive ability of face perception. *Current Biology* **20**:137–142 DOI 10.1016/j.cub.2009.11.067.