Cross-race correlations in the abilities to match unfamiliar faces

Eesha Kokje[a,b], Markus Bindemann[a] & Ahmed M. Megreya[c]

[a] School of Psychology, University of Kent, UK

[b] Network Aging Research, University of Heidelberg, Germany

[c] Department of Psychological Sciences, College of Education, Qatar University, Qatar

Correspondence to:

Eesha Kokje

Network Aging Research, University of Heidelberg

Bergheimer Strasse 20, 69115, Heidelberg, Germany

E-mail: kokje@nar.uni-heidelberg.de

Word count (excluding title, abstract, and references): 5495

Abstract

The other-race effect in face identification has been documented widely in memory tasks, but it persists also in identity-matching tasks, in which memory contributions are minimized. Whereas this points to a perceptual locus for this effect, it remains unresolved whether matching performance with same- and other-race faces is driven by shared cognitive mechanisms. To examine this question, this study compared Arab and Caucasian observers' ability to match faces of their own race with their ability to match faces of another race using one-to-one (Experiment 1) and one-to-many (Experiment 2) identification tasks. Across both experiments, Arab and Caucasian observers demonstrated reliable other-race effects at a group level. At an individual level, substantial variation in accuracy was found, but performance with same-race and other-race faces correlated consistently and strongly. This indicates that the abilities to match same- and other-race faces share a common cognitive mechanism.

Keywords: Face matching; Unfamiliar faces; Other-race effect; Individual differences

1. Introduction

The other-race effect (ORE) refers to a phenomenon wherein faces of an observer's own race tend to be recognized more accurately than faces of another race. This effect has been reported consistently in the face recognition and eyewitness identification literature (see, e.g., Malpass & Kravitz, 1969; Marcon, Meissner, Frueh, Susa, & Maclin, 2010), and has been replicated widely with different ethnic groups (e.g., Bothwell, Brigham, & Malpass, 1989; Chiroro & Valentine, 1995; Walker & Tanaka, 2003), pointing to a remarkably robust phenomenon. Whilst the effect has been demonstrated typically with recognition memory tasks, which require the identification of newly learned faces after an interval, it is also observed with tasks in which memory factors are minimized (Megreya & Bindemann, 2009; Megreya, White, & Burton, 2011). Such matching tasks point to a perceptual locus for the ORE, at face encoding. These tasks are also characterized by broad individual differences between observers. However, limited data is still available with regard to the consistency of these individual differences in the processing of same- and other-race faces. Consequently, the question arises of whether individual differences in the identity matching of same-race and other-race faces reflect shared or dissociable mechanisms. In this study, we therefore report two experiments to explore the correlation of individual performance across these tasks.

Face-matching tasks are now used widely in psychology (for a recent review, see Fysh & Bindemann, 2017b). In these tasks, observers typically have to compare the identities of pairs of simultaneously-presented unfamiliar faces (see, e.g., Bindemann, Avetisyan, & Blackwell, 2010; Burton, White, & McNeill, 2010; White, Kemp, Jenkins, & Burton, 2014), or compare a single target to a concurrent array of identities (Bruce et al., 1999; Megreya & Burton, 2006b, 2008). These matching tasks reveal that identification of unfamiliar faces is surprisingly difficult. For example, under highly optimized conditions, in which observers compare high-quality pairs of photographs that depict people on the same day, under similar lighting, and with neutral facial expressions, 10-20% errors are typically found (Burton et al., 2010; Megreya & Burton, 2008). Accuracy is lower still when one-to-many face comparisons are required (Megreya & Burton, 2006b; Megreya, Sandford, & Burton, 2013), and when viewing conditions are further compromised by, for example, added variation in a person's appearance (Fletcher, Butavicius, & Lee, 2008; Jenkins, White, Van Monfort, & Burton, 2011; Megreya et al., 2013; White, Kemp, Jenkins, Matheson, & Burton, 2014), reduced image-quality (Bindemann, Attard, Leach, & Johnston, 2013; Burton, Wilson, Cowan, & Bruce, 1999), or time pressure (Bindemann, Fysh, Cross, & Watts, 2016; Fysh & Bindemann, 2017a; Lee, Vast, & Butavicius, 2006).

Whereas face-matching has been studied widely with Caucasian (e.g., Burton et al., 2010; Megreya & Burton, 2006a, 2006b; White, Burton, Jenkins, & Kemp, 2014) and Arab faces (e.g., Megreya & Bindemann, 2015; Megreya & Burton, 2008; Megreya et al., 2013), only a few studies have compared the performance of Caucasian and Arab observers for these different face categories. Using same- and other-race face-matching tasks with Arab and Caucasian observers, Megreya and Bindemann (2009) revealed consistent OREs, but these effects were expressed differently in both groups of observers. Specifically, Arab observers displayed a processing advantage for the internal features of faces (i.e., the region encompassing the eyes, nose and mouth), whereas Caucasian observers relied more on external features comprising the hair and face outline. In addition, these groups of observers also exhibited different response biases during face matching. Namely, Caucasian observers were biased to classify pairs of other-race faces as depicting one person, independent of whether these depicted the same person or two different people, whereas Arab observers were generally less accurate in classifying other-race faces. This finding converges with later research with one-to-many face-matching tasks, in which Caucasians were more prone to make false positive identifications for other-race faces, whereas Arab observers were more likely to decide that a target was not present in a concurrent identity array (Megreya et al., 2011).

The differences in the expression of the ORE, both in terms of the face features that are prioritized in matching decisions (i.e., internal versus external, see Megreya & Bindemann, 2009), and the measures in which this effect is expressed (Megreya et al., 2011), suggest that different attributes might be required to match same- and other-race faces. In turn, this raises the question of whether observers who are good at matching faces of their own race are also good at processing faces of another race. It is now well established that substantial individual differences exist among observers performing pairwise (e.g., Bindemann, Avetisyan, & Rakow, 2012; Burton et al., 2010) and one-to-many face-identity comparisons (e.g., Bindemann, Brown, Koyas, & Russ, 2012; Bobak, Hancock, & Bate, 2016; Bruce et al., 1999; Megreya & Bindemann, 2013; Megreya & Burton, 2006b). In pairwise face-matching, for example, these individual differences are such that accuracy ranges from close-to-chance to perfect across participants (see, e.g., Bindemann et al., 2012; Burton et al., 2010). With regard to the ORE, these individual differences are interesting theoretically, as these may shed further light on the cognitive mechanisms governing face processing.

This question is also important practically, as broad individual differences are found in trained professionals who perform face-matching daily in occupational environments, such as security officers at passport control (White, Kemp, Jenkins, Matheson, & Burton, 2014; Wirth & Carbon, 2017). Passport officers encounter people from many different races in these real-life face-matching settings. However, as little is still known about how individual differences in a person’s face matching ability transcend across races, it is unresolved whether person identification at passport control is compromised by the ORE. An important step for investigating this problem further is to understand the relationship between same- and other-race face-matching accuracy in individual observers.

To investigate this question, the current study compared the matching of same- and other-race faces in Arab and Caucasian observers using pairwise (Experiment 1) and one-to-many (Experiment 2) identity face-matching tasks. Consistent with previous research, we expected to find a clear ORE for both groups of observers in these experiments. The question of main interest was whether performance with same- and other-race faces would also correlate across individuals.

2. Experiment 1

In this experiment, the ORE was assessed in a pairwise face-matching task, in which Arab and Caucasian observers were shown two side-by-side images of unfamiliar faces. To assess the OREs, these pairs consisted either of Arab or Caucasian faces, and depicted either the same person (an identity match) or two different people (an identity mismatch). The aim of this experiment was to assess whether individuals’ matching performance correlated for same- and other-race faces, thus pointing at shared underlying processing mechanisms, or whether it was strictly dissociable.

2.1. Method

2.1.1. Participants

A total of 74 participants volunteered to take part in this study. These comprised 40 Caucasian participants (30 female) from the University of Kent with a mean age of 19.7 years (SD = 1.7), and 34 Arab observers (24 female) from Menoufia University in Egypt, with a mean of 21.6 years (SD = 4.9). Participants received course credit or a small payment for taking part in the study. None of these participants had spent over 3 months in a country with a majority population dissimilar from their own race. All participants reported normal or corrected-to-normal vision.

2.1.2. Stimuli

The stimuli consisted of 200 face pairs. Of these, 100 were Caucasian face pairs taken from the Glasgow University Face Database (GUFD; see Burton et al., 2010), and 100 were Arab face pairs from an Arab database (see Megreya & Burton, 2008). Half of each of the Caucasian and Arab face pairs depicted the same person in both the images (identity matches), and half of them depicted two different people (identity mismatches). All faces were male, as suitable comparison faces of Arab women were unavailable due to the headscarf culture. In addition, all faces were presented in greyscale on a white background, with a neutral expression, and in full-face frontal view (see Figure 1). Images of the same person were only taken a few minutes apart, but with different cameras to ensure these images did not match in their pictorial aspects (see Burton, 2013). Each face image measured maximally 350 pixels in width at a resolution of 72 ppi.

Note that the Caucasian identity mismatches were created in a previous study with a sorting technique, which was applied to generate pairwise similarity measures (see Burton et al., 2010). The face identities that were rated most similar were then paired together. For the mismatch pairs from that stimulus set that were employed in the current study, none of the identities were repeated on match trials. Moreover, all of the mismatching identities appeared in only a single mismatch pair, except for six identities that appeared in two mismatch pairs each. For Arab face pairs, the mismatch pairings were created by the experimenters of a previous study, based on their perceived similarity of these identities (see Megreya & Burton, 2008). For the face pairs from that stimulus set that were employed in the current study, eight identities appeared in both one of the match and one of the mismatch face pairs. No other identities were repeated within the match or mismatch stimuli.

FIGURE 1. Examples of Arab (left) and Caucasian (right) face pairs from the matching task in Experiment 1, depicting an identity match and a mismatch.

2.1.3. Procedure

Arab and Caucasian participants were subjected to an identical experimental procedure. Participants were tested individually in a laboratory using a standard desktop PC. The stimuli was presented on a 21” screen using PsychoPy software (Peirce, 2007), which also recorded button-press responses. Each trial began with a central fixation cross, which was presented for one second, followed by a face pair, which was displayed until a response was registered. The participants were asked to make same- or different-identity judgements about the face pairs by pressing one of two keys (S versus D) on the computer keyboard. The task was self-paced and accuracy was emphasized.

Each participant was presented 200 face pairs, in five blocks consisting of 40 trials, which were interspersed by short, self-paced breaks. Each of these blocks consisted of 20 Caucasian and 20 Arab face pairs, comprising 10 identity-match and 10 identity-mismatch trials. The order in which the blocks were presented was counterbalanced across participants, and trial order within each block was randomized individually for each participant.

2.2. Results

2.2.1. Group accuracy

For each participant, the percentage accuracy of responses was calculated for match and mismatch trials and for same-race and other-race faces. The cross-subjects means of these data are provided in Figure 2. A 2 (observer race: Arab versus Caucasian) x 2 (face race: same-race versus other-race) x 2 (trial type: match vs. mismatch) mixed-factor ANOVA revealed a main effect of observer race, $F(1,72) = 4.38$, $p < .05$, $\eta_p^2 = .06$, reflecting higher accuracy for Caucasian (86.7%, SD = 8.3) than Arab observers (83.3%, SD = 8.5). More importantly, a main effect of face race was also found, $F(1,72) = 23.44$, $p < .001$, $\eta_p^2 = .25$, due to generally higher accuracy on same-race (87.5%, SD = 7.6) than other-race trials (82.7%, SD = 8.9). In addition, a main effect of trial type, $F(1,72) = 11.42$, $p < .001$, $\eta_p^2 = .14$, and an interaction of face race and trial type were observed, $F(1,72) = 6.83$, $p < .05$, $\eta_p^2 = .09$. Analysis of simple effects revealed an effect of trial type for same-race faces, $F(1,72) = 4.65$, $p < .05$, $\eta_p^2 = .06$, and other-race faces, $F(1,72) = 16.03$, $p < .001$, $\eta_p^2 = .18$, due to higher accuracy on match than mismatch trials. More importantly, simple main effect analysis also showed that the ORE persisted on identity match trials, $F(1,72) = 5.22$, $p < .05$, $\eta_p^2 = .07$, and identity mismatch trials, $F(1,72) = 24.33$, $p < .001$, $\eta_p^2 = .25$, with attenuated accuracy for other-race faces. None of the other interactions were significant, all Fs ≤ 1.53, ps $\geq .221$.

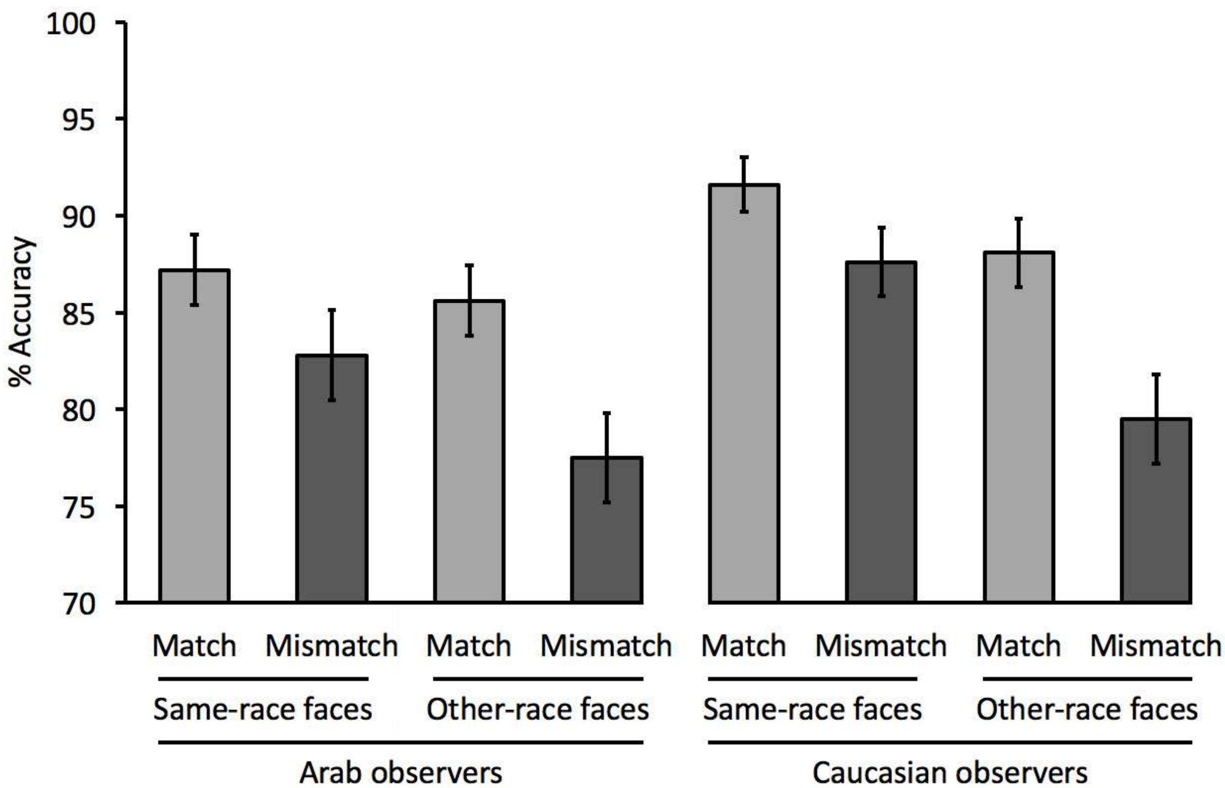

FIGURE 2. Percentage of correct responses by Arab and Caucasian observers for same-race and other-race face pairs on match and mismatch trials in Experiment 1. Error bars show standard error of the means.

2.2.2. Individual differences in accuracy

The group analysis confirms the presence of the ORE in face matching in both Arab and Caucasian observers. However, an inspection of individual data, which are illustrated in Figure 3, reveals broad individual differences in accuracy for both groups. For example, these individual differences were such that performance on same-race match trials ranged from 56% to 100% in Arab observers, and from 60% to 100% in Caucasian observers. Despite these individual differences, strong Pearson's correlations between the matching of same-race and other-race faces were found in Arab observers for identity match trials, $r(32) = .41$, $p < .05$, and identity mismatch trials, $r(32) = .58$, $p < .001$. This occurred in a context in which accuracy for match and mismatch trials was not correlated in these observers for same-race faces, $r(32) = -.22$, $p = .22$, and other-race faces, $r(32) = .13$, $p = .48$. Correspondingly, correlations of race were present in Caucasian observers for identity matches, $r(38) = .75$, $p < .001$, and mismatches, $r(38) = .67$, $p < .001$, but not between match and mismatch trials for same-race and other-race stimuli, $r(38) = -.07$, $p = .66$ and $r(38) = -.11$, $p = .51$, respectively.

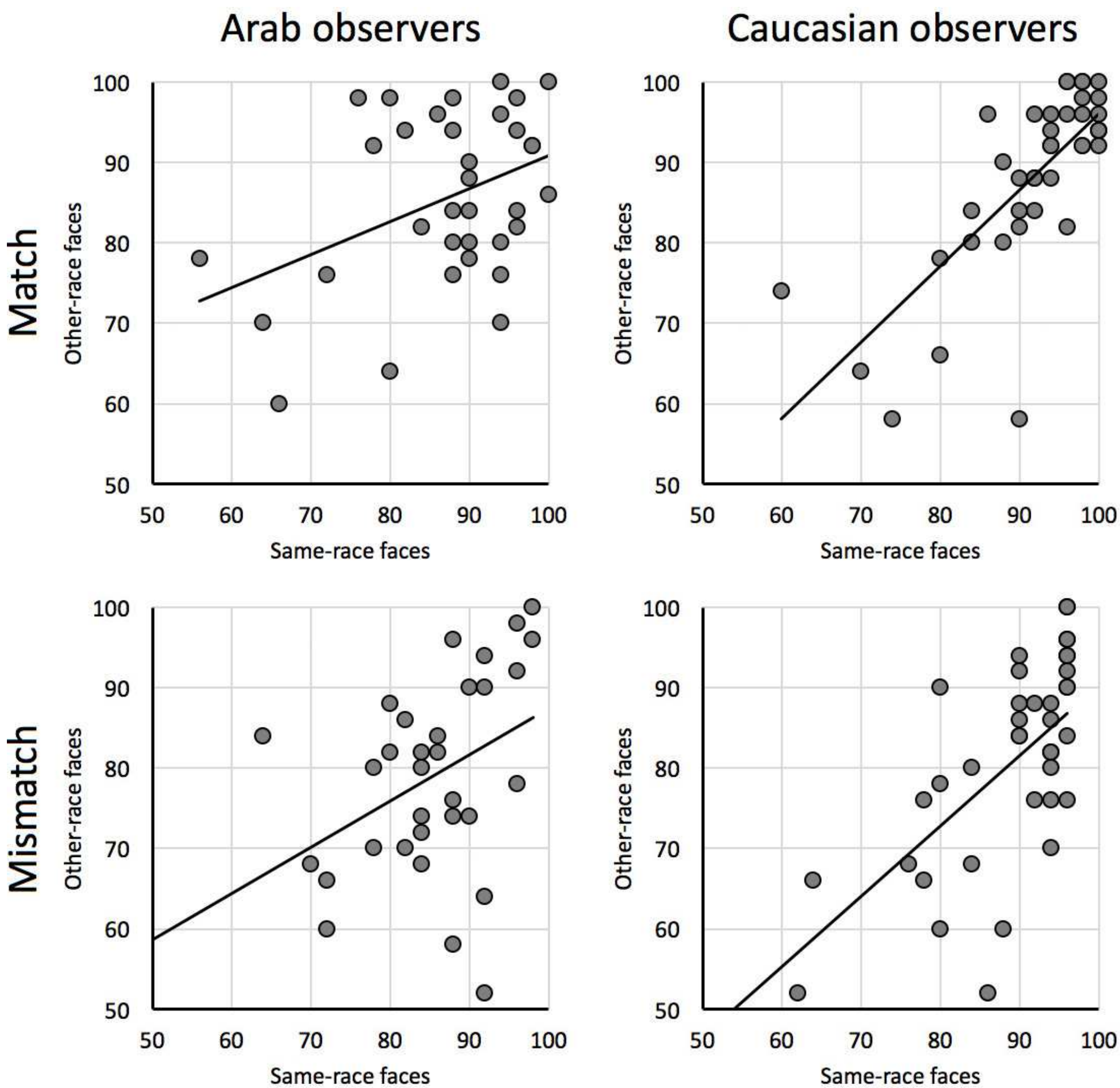

FIGURE 3. Individual face-matching accuracy (% accuracy) of Arab and Caucasian observers for same- and other-race faces on match and mismatch trials.

2.3. Discussion

This experiment reveals an ORE in face matching in Arab and Caucasian observers, which converges with previous face-matching studies (Megreya & Bindemann, 2009; Megreya et al., 2011). The current experiment adds to these findings by revealing strong positive associations in individual performance in the matching of same-race and other-race faces. These correlations were present in both groups of observers, and for identity match and mismatch trials, pointing to a robust effect. These results therefore indicate that observers who are good at matching same-race faces are also good at matching other-race faces, and suggest a common underlying mechanism for accurate performance in this task.

It is notable that these effects were obtained in a context in which correlations between conditions were not observed universally. Previous studies indicate that face-matching performance on identity match and mismatch trials differs qualitatively, as correlations for these trial types are not found (Megreya & Burton, 2007). The current experiment provides further support for these findings, by demonstrating that match and mismatch performance was fully dissociable even though same- and other-race processing is not. In turn, the absence of such correlations for identity match and mismatch trials serves to underline the presence of correlations in matching performance for same- and other-race faces.

3. Experiment 2

To replicate and extend the findings of Experiment 1, we conducted a further face-matching experiment. In contrast to pairwise comparisons, this task was based on one-to-many comparisons by asking observers to detect the presence of a target in a concurrent ten-face array. To provide an analogy to identity match and mismatch trials of the pairwise matching tasks, the target could be present or absent from these arrays. If the results of Experiment 1 are robust, then an ORE should be present again at a group level in both Arab and Caucasian observers. More importantly, identification of same- and other-race faces should correlate in individual accuracy.

3.1. Method

3.1.1. Participants

A total of 60 participants volunteered to take part in this study. These comprised 31 Caucasian participants (22 female) from the University of Glasgow with a mean age of 21.2 years (SD = 2.3) and 30 Arab observers (15 female) from Menoufia University in Egypt, with a mean age of 19.4 years (SD = 1.7). Participants received course credit or a small payment for taking part, and all reported normal or corrected-to-normal vision. None of the participants had taken part in the previous experiment.

3.1.2. Stimuli

A total of 200 stimulus arrays were created from separate image databases of Caucasian and Arab faces. Each of these arrays consisted of a target face, which was displayed centrally above a ten-face lineup (for an illustration, see Figure 4). Half of the arrays comprised target-present lineups, in which the target was present among the ten faces. The other half comprised target-absent lineups, in which the target was not shown in the concurrent ten-face display. For the Caucasian lineups, the faces for these displays were taken from the UK Home Office (PITO) database and comprised images of young males (18 to 35 years old). Note that the non-target faces for the stimulus arrays were chosen as the identities that were rated most similar to the target face by independent raters in a previous study (see Bruce et al., 1999). Arab arrays were created in similar fashion, with male student volunteers of comparable age (20-22 years old; see Megreya & Burton, 2008). For both face sets, the target faces at the top of the arrays and the lineup faces below were captured with different devices (camcorder and digital camera). All faces were shown without facial hair, jewelry, or distinguishing marks.

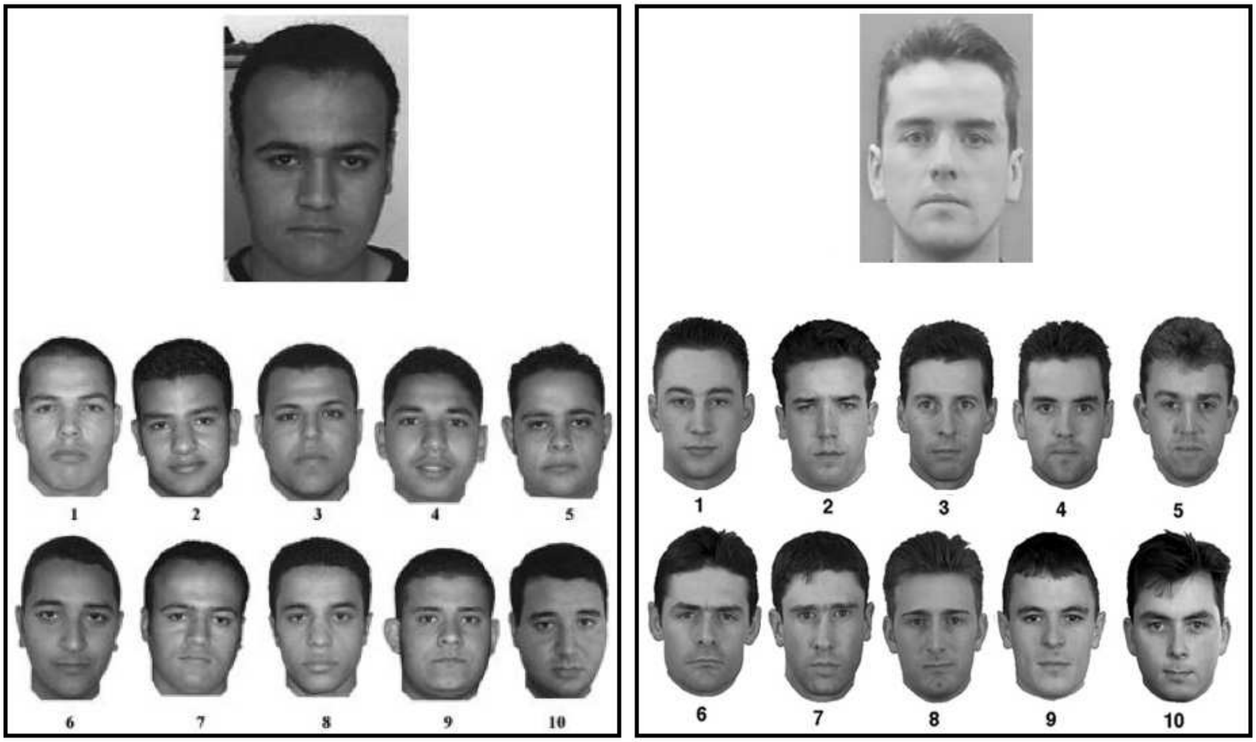

FIGURE 4. Examples of Arab (left) and Caucasian (right) face arrays used in the matching task in Experiment 2, depicting two target-present arrays (the targets are lineup face number 7 and face number 4, respectively).

3.1.3. Procedure

Arab and Caucasian participants were subjected to an identical experimental procedure. In the experiment, participants were seated in a laboratory equipped with a standard desktop computer. Each trial consisted of a fixation cross, which was shown for one second, followed by a stimulus array, which remained onscreen until a response was registered. Participants were asked to decide whether the person depicted at the top of the stimulus array was present in the concurrent lineup, and if so, to indicate who it is, by pressing the corresponding number key on computer keyboard (with '0' for face number 10) or by pressing '+' for target-absent. In this way, each participant completed 100 trials comprising 25 target-present and 25 target-absent trials each of the Arab and Caucasian face displays. The presence /absence of a target in a given lineup was counterbalanced across observers and trial order was randomized individually. In addition, Arab and Caucasian face trials were randomly intermixed. As in Experiment 1, accuracy of response was emphasized.

3.2. Results

Face-matching accuracy data was broken down into five measures of performance. For target-present trials, these comprised the correct identification of a target from a concurrent lineup (hits), the identification of an incorrect face as the target (misidentifications), or the incorrect decision that the target is absent (misses). For target-absent trials, responses reflected either the correct response that a target is not in the face lineup (a correct rejection) or the false identification of an incorrect face (a false positive). These measures are inversely proportionate, so only false positives are reported. Overall accuracy was also calculated by averaging hits and correct rejections. These measures are illustrated in Figure 5 for Arab and Caucasian observers and same- and other-race faces.

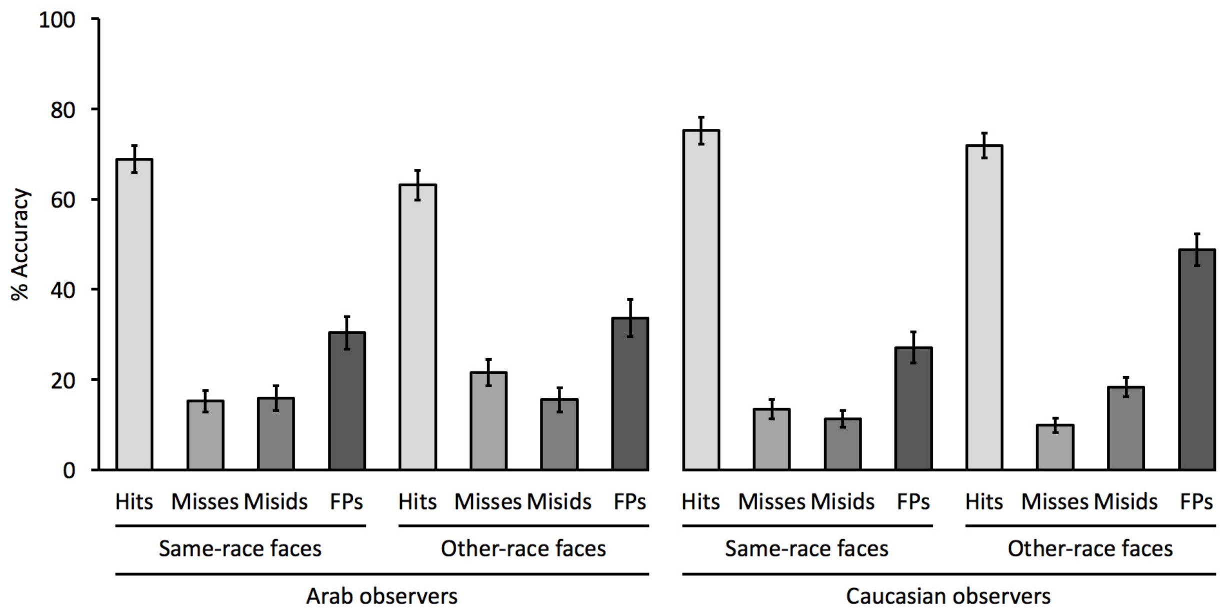

FIGURE 5. Performance of Arab and Caucasian observers for same-race and other-race face arrays on target-present trials (hits, misses, misidentifications) and target-absent trials (false positives / FPs) in Experiment 2. Error bars show standard error of the means.

To analyze this data, separate 2 (observer race: Arab versus Caucasian) x 2 (face race: same-race versus other-race) mixed-factor ANOVAs were conducted for each of the five measures. For overall accuracy, ANOVA did not find a main effect of observer race, $F(1,58) = 0.00$, $p = .97$, $\eta_p^2 = .00$, but revealed a main effect of face race, $F(1,58) = 40.25$, $p < .001$, $\eta_p^2 = .41$, and an interaction between factors, $F(1,58) = 6.55$, $p < .05$, $\eta_p^2 = .10$. Analysis of simple main effects did not show a difference in accuracy between Arab and Caucasian observers in the processing of same-race faces, $F(1,58) = 0.68$, $p = .41$, $\eta_p^2 = .01$, or other-race faces, $F(1,58) = 0.70$, $p = .41$, $\eta_p^2 = .01$. However, an ORE was found for Arab observers, $F(1,58) = 7.16$, $p = .01$,

η_p^2 = .11, and Caucasian observers, $F(1,58) = 39.64$, $p < .001$, η_p^2 = .41, with both groups displaying higher accuracy in the identification of same-race compared to other-race face targets.

A similar overall pattern was evident in the analysis of the individual performance measures, though some variation in these measures was observed. For example, for hits a main effect of observer race, $F(1,58) = 3.54$, $p = .07$, η_p^2 = .06, and an interaction of observer race and face race were not found, $F(1,58) = 0.67$, $p = .42$, η_p^2 = .01. However, a main effect of face race was present, $F(1,58) = 8.79$, $p < .01$, η_p^2 = .13, reflecting more correct target identifications for same-race than other-race face arrays.

For misses, a main effect of face race was not present, $F(1,58) = 1.05$, $p = .31$, η_p^2 = .02, but a main effect of observer race, $F(1,58) = 4.95$, $p < .05$, η_p^2 = .08, and an interaction between factors, $F(1,58) = 14.30$, $p < .001$, η_p^2 = .20, was found. Analysis of simple main effects showed that Arab and Caucasian observers were matched in their accuracy for same-race faces, $F(1,58)$ = 0.30, $p = .59$, η_p^2 = .01, but Arab observers were more likely than Caucasian observers to miss targets in other-race arrays, $F(1,58) = 12.03$, $p < .001$, η_p^2 = .17. Arab observers were also more likely to miss other-race than same-race faces, $F(1,58) = 11.54$, $p < .001$, η_p^2 = .17. Caucasian observers demonstrated the reverse trend, but this did not reach significance, $F(1,58) = 3.81$, p = .06, η_p^2 = .06.

For misidentifications, a main effect of observer race was not found, $F(1,58) = 0.08$, p = .78, η_p^2 = .00, but a main effect of face race, $F(1,58) = 6.66$, $p < .05$, η_p^2 = .10, and an interaction were found, $F(1,58) = 8.39$, $p < .01$, $\eta_p^2 = .13$. This interaction reflects a pattern whereby Arab and Caucasian observers made a comparable number of misidentifications for same-race faces, $F(1,58) = 1.88$, $p = .18$, $\eta_p^2 = .03$, and other-race faces, $F(1,58) = 0.68$, $p = .41$, $\eta_p^2 = .01$. However, whereas Arab observers also committed a comparable percentage of misidentifications on same-race and other-race trials, $F(1,58) = 0.05$, $p = .82$, $\eta_p^2 = .00$, Caucasian observers were more likely to misidentify non-target faces as the target in the other-race than the same-race condition, $F(1,58) = 15.00$, $p < .001$, $\eta_p^2 = .21$.

Finally, analysis of false positives did not show a main effect of observer race, $F(1,58) = 1.47$, $p = 0.23$, $\eta_p^2 = .03$, but revealed a main effect of face race, $F(1,58) = 45.42$, $p < .001$, $\eta_p^2 = .44$, due to more false identifications of other-race than same-race faces on target-absent trials, and an interaction between factors, $F(1,58) = 25.10$, $p < .001$, $\eta_p^2 = .30$. Analysis of simple main effects revealed that Caucasian observers committed more false positives with other-race than same-race faces, $F(1,58) = 69.03$, $p < .001$, $\eta_p^2 = .54$, whereas Arab observers produced a comparable percentage of false positives for same- and other-race faces, $F(1,58) = 1.50$, $p = .23$, $\eta_p^2 = .03$. In addition, Caucasian observers recorded more false positives than Arab observers with other-race faces, $F(1,58) = 7.71$, $p < .01$, $\eta_p^2 = .12$, but not with same-race faces, $F(1,58) = 0.45$, $p = .50$, $\eta_p^2 = .01$.

3.2.1. Individual differences in accuracy

The group data confirms the presence of the ORE in face matching for both groups of observers in this experiment. However, as in Experiment 1, broad individual differences were observed between observers, with overall accuracy ranging from 30% to 94% in Arab observers, and from 32% to 98% in Caucasian observers (see Figure 6). To explore whether accuracy with same- and other-race faces was associated at an individual level, Pearson's correlations were performed for all measures. These revealed that performance for same- and other-race face arrays was correlated highly in Arab observers in overall accuracy, $r(28) = .84$, $p < .001$, hits, $r(28) = .83$, $p < .001$, misidentifications, $r(28) = .76$, $p < .001$, misses, $r(28) = .71$, $p < .001$, and false positives, $r(28) = .78$, $p < .001$. Similarly, Caucasian observers presented strong correlations for same- and other-race face arrays in overall accuracy, $r(28) = .74$, $p < .001$, hits, $r(28) = .64$, $p < .001$, misidentifications, $r(28) = .64$, $p < .001$, misses, $r(28) = .69$, $p < .001$, and false positives, $r(28) = .72$, $p < .001$. Similar to Experiment 1, performance was also dissociable for target-present and target-absent trials, as evident from the absence of correlations in hits and false positives in Arab observers on same-race, $r(28) = -.28$, $p = .13$, and other-race trials, $r(28) = -.14$, $p = .45$, and in Caucasian observers, $r(28) = -.07$, $p = .71$ and $r(28) = -.29$, $p = .12$, respectively.

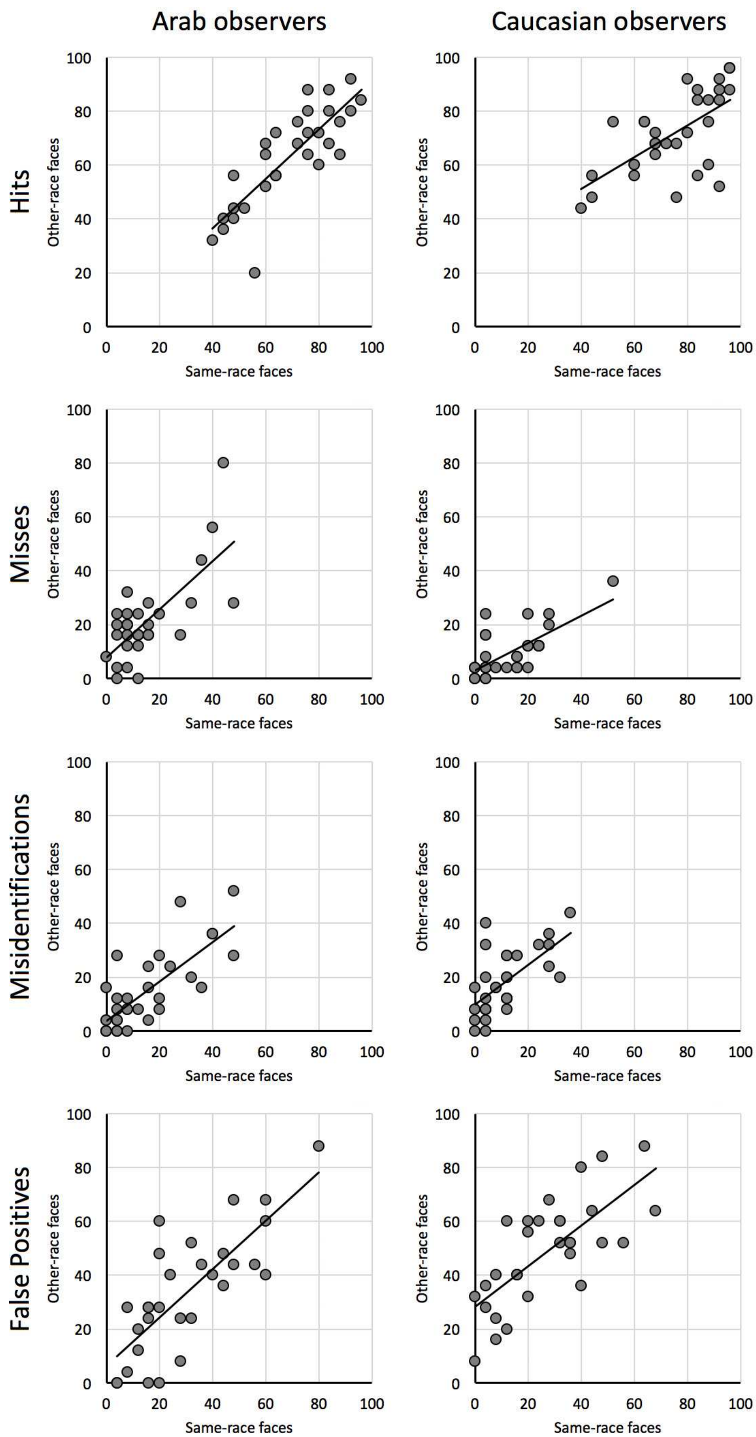

FIGURE 6. Individual face-matching accuracy of Arab and Caucasian observers for same-race and other-race face arrays on target-present trials (hits, misses, misidentifications) and target-absent trials (false positives / FPs) in Experiment 2.

3.3. Discussion

This experiment sought to replicate the correlation in same- and other-race matching that was observed in Experiment 1, and to extend this to a scenario in which one face target is compared to a lineup of ten faces. As in Experiment 1, both Arab and Caucasian observers demonstrated a clear ORE, whereby the identification of target faces from a lineup, as well as the identification of absence of targets, was more accurate for same- than other-race faces. Correspondingly, observers were also less likely to incorrectly identify a non-target face as the target in the same-race than the other-race conditions. Crucially, however, individuals' performance correlated strongly for same- and other-race faces in all of the measures, thus adding further evidence that observers who are good at matching same-race faces are also good at matching other-race faces. Similar to Experiment 1, this was observed in a context in which performance on target-present and target-absent trials was not correlated in either group of participants, thus providing further evidence that the ability to match same-race and other-race faces is associated, but the ability to detect the presence of a target or its absence is not (see Megreya & Burton, 2007). We return to a fuller discussion of these findings in the General Discussion.

4. General Discussion

This study examined how an individual's ability to match unfamiliar faces of their own race is related to their ability to match unfamiliar faces from another race. For this purpose, Arab and Caucasian observers were presented with same-race and other-race faces using one-to-one (Experiment 1) and one-to-many (Experiment 2) face identification tasks. At a group level, these experiments demonstrated robust OREs in both Arab and Caucasian observers, whereby identification accuracy was better with same-race than other-race faces. At an individual level, a broad range in performance was observed. In Experiment 1, for example, individual performance on same-race trials ranged from 56% to 100% in Arab observers, and from 60% to 100% in Caucasian observers. Similarly, overall accuracy ranged from 36% to 92% and from 42% to 96% in these observer groups in Experiment 2. Importantly, despite these individual differences, strong associations were consistently found for the processing of same- and other-race faces across all of the measures here. This indicates that observers with a high ability to match same-race faces are also likely to perform with higher accuracy in the matching of other-race faces. In turn, this suggests that the abilities to match same- and other-race faces share a common cognitive mechanism.

This is an interesting finding considering evidence that same- and other-race faces engage different perceptual processes. Several studies demonstrate, for example, that same-race faces are processed more holistically or configurally, whereby individual facial features (such as the eyes, nose, and mouth) are integrated into a Gestalt-like percept, whereas other-race faces are processed in a more piece-meal fashion that is based on an individual analysis of features (see, e.g., Michel, Caldara, & Rossion, 2006; Michel, Corneille, & Rossion, 2007; Michel, Rossion, Han, Chung, & Caldara, 2006; Rhodes, Brake, Taylor, & Tan , 1989). For example, employing the parts-wholes paradigm, Tanaka, Kiefer, and Bukach (2004) demonstrated that Caucasian observers, who had reported very little previous contact with Asian people, recognized Caucasian face parts more accurately when these were presented in the context of whole faces than when these were presented in isolation. By contrast, similar parts of Asian faces were recognized with similar accuracy across these conditions. On the other hand, Asian observers who reported having more contact with Caucasians than Asians (as a result of longer experience living with

Caucasian people), recognized both Caucasian and Asian face parts more accurately when they were presented in a whole-faces context than when they were shown in isolation. Of course, holistic and featural face processing mechanism may not be engaged in an all-or-nothing manner in such studies, but may contribute differentially to same- and other-race processing (Mondloch et al., 2010; Rhodes, Hayward, & Winkler, 2006). This is corroborated further by an ERP study in which the N170 component, which is sensitive to face inversion, and thereby to holistic/configural processing, demonstrated an increased amplitude for inverted same- as well as other-race faces (Wiese, Stahl, & Schweinberger, 2009). This indicates that identification of other-race faces also involves holistic processing mechanisms, which in turn suggests that more than one processing mechanism may simultaneously be involved during face identification.

In a similar vein, it has also been shown that part-based and space-based visual information may in itself be processed by associated mechanisms during face processing (Yovel & Kanwisher, 2008). Thus, although some differences may clearly exist in the processing of same- and other-race faces, these might reflect quantitative differences in the engagement of cognitive processes, rather than profound qualitative differences (DeGutis, Mercado, Wilmer, & Rosenblatt, 2013; Harrison, Gauthier, Hayward, & Richler, 2014; Hayward, Crookes, & Rhodes, 2013; Hayward, Rhodes, & Schwaninger, 2008).

Previous face *matching* studies also reveal some differences in the processing of same- and other-race faces, and these appear to be modulated further by the race of the observer. For example, Arab observers display a processing advantage for the internal features of faces in pairwise matching tasks compared to Caucasian observers (Megreya & Bindemann, 2009; Megreya, Memon, & Havard, 2012), and may also be generally less accurate in the identification of other-race faces, whereas Caucasian observers are more likely to commit false match decisions and false positives in one-to-one and one-to-many matching (Megreya & Bindemann, 2009; Megreya et al., 2011). In the current study, a similar effect was observed in false positives in Experiment 2. Despite this, false positives correlated strongly for same- and other-race faces. We therefore suggest that some nuanced differences exist between Arab and Caucasian observers in face matching tasks, but abilities to process these same- and other-race faces share a common cognitive mechanism.

Some questions arise from these findings that require further investigation. Firstly, we note that our study comprised of only Caucasian and Arab participants, and face stimuli of these races. It therefore remains to be seen whether similar correlations are found with other races, such as African and Asian faces. Such correlations have been observed in recognition *memory* paradigms with Caucasian, Asian and African faces (see, e.g., Brown, Uncapher, Chow, Eberhardt, & Wagner, 2017; Wan, Crookes, Dawel, Pidcock, Hall, & McKone, 2017), which suggests that the current pattern in a *matching* task could hold more generally.

Secondly, it remains unclear at present whether any shared processes for the matching of same- and other-race faces are face-specific or might reflect general cognitive abilities. In the recognition domain, face processing ability appears to be dissociable from general intelligence and cognitive ability, as well as the processing of non-face stimuli, such as houses, animals, and cars (see, e.g., McKone, Kanwisher, & Duchaine, 2007; Wilmer et al., 2010; Zhu et al., 2010). Studies of people with prosopagnosia also show that while these individuals are impaired on face recognition, similar processing of other visual stimuli can remain intact (Duchaine, Yovel, Butterworth, & Nakayama, 2006; Farah, 1991, 1996; Farah, Levinson, & Klein, 1994). On the other hand, unfamiliar face *matching* performance correlates with measures of visual short-term memory, perceptual speed, and the matching of non-face figures (see Burton et al., 2010;

Megreya & Burton, 2006b). Thus, it is possible that the cross-race correlations that were observed here reflect either face-specific or more general cognitive mechanisms.

We close by noting that the current findings may have some practical implications for occupational environments, in which face matching is performed routinely for security reasons. Passport officers at borders and airports, for example, routinely have to match unfamiliar faces, but they also demonstrate similar broad individual differences to lay participants in psychology experiments (White et al., 2014). Passport officers also encounter people from many different races in these real-life face-matching settings, but little is still known about how their face matching ability transcends across different races. The current study suggests that professionals with high identification ability for one race are likely to show similar ability for faces of other races.

Conflicts of interest

None.

Funding

This work did not receive funding from agencies in the public, commercial, or not-for-profit sectors.

References

Bindemann, M., Attard, J., Leach, A., & Johnston, R. A. (2013). The effect of image pixelation on unfamiliar-face matching. *Applied Cognitive Psychology*, *27*(6), 707-717. doi:10.1002/acp.2970

Bindemann, M., Avetisyan, M., & Blackwell, K. A. (2010). Finding needles in haystacks: Identity mismatch frequency and facial identity verification. *Journal of Experimental Psychology: Applied*, *16*(4), 378-386. doi:10.1037/a0021893

Bindemann, M., Avetisyan, M., & Rakow, T. (2012). Who can recognize unfamiliar faces? Individual differences and observer consistency in person identification. *Journal Of Experimental Psychology: Applied*, *18*(3), 277-291. doi:10.1037/a0029635

Bindemann, M., Brown, C., Koyas, T., & Russ, A. (2012). Individual differences in face identification postdict eyewitness accuracy. *Journal of Applied Research in Memory and Cognition*, *1*(2), 96-103. doi:10.1016/j.jarmac.2012.02.001

Bindemann, M., Fysh, M., Cross, K., & Watts, R. (2016). Matching faces against the clock. *i-Perception*, *7*(5), 2041669516672219. doi:10.1177/2041669516672219

Bobak, A., Hancock, P., & Bate, S. (2015). Super-recognisers in action: Evidence from face-matching and face memory tasks. *Applied Cognitive Psychology*, *30*(1), 81-91. doi:10.1002/acp.3170

Bothwell, R. K., Brigham, J. C., & Malpass, R. S. (1989). Cross-racial identification. *Personality and Social Psychology Bulletin*, *15*(1), 19-25. doi:10.1177/0146167289151002

Brown, T. I., Uncapher, M. R., Chow, T. E., Eberhardt, J. L., & Wagner, A. D. (2017). Cognitive control, attention, and the other-race effect in memory. *PLoS ONE, 12*(3): e0173579. doi:10.1371/journal.pone.0173579

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J. B., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*, *5*(4), 339. doi:10.1037/1076-898x.5.4.339

Burton, A. M., (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology*, *66*(8), 1467-1485. doi:10.1080/17470218.2013.800125

Burton, A. M., White, M., & McNeill, D. (2010). The Glasgow Face Matching Test. *Behavior Research Methods, 42*(1), 286-291. doi:10.3758/brm.42.1.286

Burton, A. M., 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

Chiroro, P., & Valentine, T. (1995). An investigation of the contact hypothesis of the own-race bias in face recognition. *Quarterly Journal of Experimental Psychology*, *48*(4), 879-894. doi:10.1080/14640749508401421

DeGutis, J., Mercado, R. J., Wilmer, J., & Rosenblatt, A. (2013). Individual differences in holistic processing predict the own-race advantage in recognition memory. *PLoS One*, *8*(4), e58253. doi:10.1371/journal.pone.0058253

Duchaine, B. C., Yovel, G., Butterworth, E. J., & Nakayama, K. (2006). Prosopagnosia as an impairment to face-specific mechanisms: Elimination of the alternative hypotheses in a developmental case. *Cognitive Neuropsychology*, *23*(5), 714-747. doi:10.1080/02643290500441296

Farah, M. J. (1991). Patterns of co-occurrence among the associative agnosias: Implications for visual object representation. *Cognitive Neuropsychology*, *8*(1), 1-19. doi:10.1080/02643299108253364

Farah, M. J. (1996). Is face recognition 'special'? Evidence from neuropsychology. *Behavioural Brain Research*, *76*(1), 181-189. doi:10.1016/0166-4328(95)00198-0

Farah, M. J., Levinson, K. L., & Klein, K. L. (1995). Face perception and within-category discrimination in prosopagnosia. *Neuropsychologia*, *33*(6), 661-674. doi:10.1016/0028-3932(95)00002-K

Fletcher, K. I., Butavicius, M. A., & Lee, M. D. (2008). Attention to internal face features in unfamiliar face matching. *British Journal of Psychology*, *99*(3), 379-394. doi:10.1348/000712607X235872

Fysh, M. C., & Bindemann, M. (2017a). Effects of time pressure and time passage on face-matching accuracy. *Royal Society Open Science*, *4*(6), 170249. doi:10.1098/rsos.170249

Fysh, M. C., & Bindemann, M. (2017b). Forensic face matching: A review. In M. Bindemann & A. M. Megreya (Eds.), *Face processing: Systems, Disorders and Cultural Differences* (pp. 1-20). New York, NY: Nova Science Publishing, Inc.

Harrison, S. A., Gauthier, I., Hayward, W. G., & Richler, J. J. (2014). Other-race effects manifest in overall performance, not qualitative processing style. *Visual Cognition*, *22*(6), 843-864. doi:10.1080/13506285.2014.918912

Hayward, W. G., Crookes, K., & Rhodes, G. (2013). The other-race effect: Holistic coding differences and beyond. *Visual Cognition*, *21*(9-10), 1224-1247. doi:10.1080/13506285.2013.824530

Hayward, W. G., Rhodes, G., & Schwaninger, A. (2008). An own-race advantage for components as well as configurations in face recognition. *Cognition*, *106*(2), 1017-1027. doi:10.1016/j.cognition.2007.04.002

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition*, *121*(3), 313-323. doi:10.1016/j.cognition.2011.08.001

Lee, M. D., Vast, R. L., & Butavicius, M. A. (2006). Face matching under time pressure and task demands. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society, Vancouver, Canada* (pp. 1675-1680).

Malpass, R., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology, 13*(4), 330-334. doi:10.1037/h0028434

Marcon, J., Meissner, C., Frueh, M., Susa, K., & MacLin, O. (2010). Perceptual identification and the cross-race effect. *Visual Cognition, 18*(5), 767-779. doi:10.1080/13506280903178622

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences*, *11*(1), 8-15. doi:10.1016/j.tics.2006.11.002

Megreya, A. M., & Bindemann, M. (2009). Revisiting the processing of internal and external features of unfamiliar faces: The headscarf effect. *Perception*, 38(12), 1831-1848. doi:10.1068/p6385

Megreya, A. M., & Bindemann, M. (2013). Individual differences in personality and face identification. *Journal of Cognitive Psychology*, *25*(1), 30-37. doi:10.1080/20445911.2012.739153

Megreya, A. M., & Bindemann, M. (2015). Developmental improvement and age-related decline in unfamiliar face matching. *Perception*, *44*(1), 5-22. doi:10.1068/p7825

Megreya, A. M., & Burton, A. M. (2006a). Recognising faces seen alone or with others: When two heads are worse than one. *Applied Cognitive Psychology*, *20*(7), 957-972. doi:10.1002/acp.1243

Megreya, A. M., & Burton, A. M. (2006b). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition*, *34*(4), 865-876. doi:10.3758/BF03193433

Megreya, A. M., & Burton, A. M. (2007). Hits and false positives in face matching: A familiarity-based dissociation. *Perception & Psychophysics*, 69(7), 1175-1184. doi:10.3758/BF03193954

Megreya, A. M., & Burton, A. M. (2008). Matching faces to photographs: Poor performance in eyewitness memory (without the memory). *Journal of Experimental Psychology: Applied*, *14*(4), 364. doi:10.1037/a0013464

Megreya, A. M., Memon, A., & Havard, C. (2012). The headscarf effect: Direct evidence from the eyewitness identification paradigm. *Applied Cognitive Psychology, 26*(2), 308-315. doi:10.1002/acp.1826

Megreya, A. M., Sandford, A., & Burton, A. M. (2013). Matching face images taken on the same day or months apart: The limitations of photo ID. *Applied Cognitive Psychology, 27*(6), 700–706. doi:10.1002/acp.2965

Megreya, A., White, D., & Burton, A. M. (2011). The other-race effect does not rely on memory: Evidence from a matching task. *Quarterly Journal of Experimental Psychology, 64*(8), 1473-1483. doi:10.1080/17470218.2011.575228

Michel, C., Caldara, R., & Rossion, B. (2006). Same-race faces are perceived more holistically than other-race faces. *Visual Cognition, 14*(1), 55-73. doi:10.1167/4.8.425

Michel, C., Corneille, O., & Rossion, B. (2007). Race categorization modulates holistic face encoding. *Cognitive Science*, *31*(5), 911-924. doi:10.1080/03640210701530805

Michel, C., Rossion, B., Han, J., Chung, C., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science*, *17*(7), 608-615. doi:10.1111/j.1467-9280.2006.01752.x

Mondloch, C. J., Elms, N., Maurer, D., Rhodes, G., Hayward, W. G., Tanaka, J. W., & Zhou, G. (2010). Processes underlying the cross-race effect: An investigation of holistic, featural, and relational processing of own-race versus other-race faces. *Perception*, *39*(8), 1065-1085. doi:10.1068/p6608

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1), 8-13. doi:10.1016/j.jneumeth.2006.11.017

Rhodes, G., Brake, S., Taylor, K., & Tan, S. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology*, *80*(3), 313-331. doi:10.1111/j.2044-8295.1989.tb02323.x

Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review*, *13*(3), 499-505. doi:10.3758/BF03193876

Tanaka, J., Kiefer, M., & Bukach, C. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, *93*(1), B1-B9. doi10.1016/j.cognition.2003.09.011

Walker, P. M., & Tanaka, J. W. (2003). An encoding advantage for own-race versus other-race faces. *Perception*, *32*(9), 1117-1125. doi:10.1068/p5098

White, D., Burton, A. M., Jenkins, R., & Kemp, R. I. (2014). Redesigning photo-ID to improve unfamiliar face matching performance. *Journal of Experimental Psychology: Applied*, *20*(2), 166-173. doi:10.1037/xap0000009

White, D., Kemp, R. I., Jenkins, R., & Burton, A. M. (2014). Feedback training for facial image comparison. *Psychonomic Bulletin & Review*, *21*(1), 100-106. doi:10.3758

White, D., Kemp, R., Jenkins, R., Matheson, M., & Burton, A. M. (2014). Passport officers' errors in face matching. *PLoS ONE, 9*(8), e103510. doi:10.1371/journal.pone.0103510

Wiese, H., Stahl, J., & Schweinberger, S. R. (2009). Configural processing of other-race faces is delayed but not decreased. *Biological Psychology*, *81*(2), 103-109. doi:10.1016/j.biopsycho.2009.03.002

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. doi:10.1073/pnas.0913053107

Wirth, B. E., & Carbon, C. C. (2017). An easy game for frauds? Effects of professional experience and time pressure on passport-matching performance. *Journal of Experimental Psychology: Applied*, *23*(2), 138-157. doi:10.1037/xap0000114

Yovel, G., & Kanwisher, N. (2008). The representations of spacing and part-based information are associated for upright faces but dissociated for objects: Evidence from individual differences. *Psychonomic Bulletin & Review*, *15*(5), 933-939. doi:10.3758

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*(2), 137-142. doi:10.1016/j.cub.2009.11.067