A background of bias: subtle changes in line up backgrounds increase the own race bias

HAVARD, Catriona, BREESE, Emily, THIRKETTLE, Martin
<http://orcid.org/0000-0002-6200-3130>, KASK, Kristjan and LEOL, Kris-Loreen

Citation:

Copyright and re-use policy

A Background of Bias: Subtle Changes in Lineup Backgrounds Increase the Own Race Bias

Catriona Havard[1] · Emily Breese[1] · Martin Thirkettle[2] · Kristjan Kask[3] · Kris-Loreen Leol[3] · Kaja Mädamürk[3]

Abstract

In police photo lineups, there can sometimes be small variations in shades and hues of the background images due to the faces being filmed under different lighting and cameras. Own race bias refers to a situation where people are better at remembering the faces of those who are the same race as them and find it more difficult to recognise faces from a different race. In this paper, we investigated the influence of small colour variations in backgrounds for the recognition of Black and White faces. Across 3 experiments, we found when small changes were introduced into the backgrounds of the images this increased false identifications for previously unseen Black faces, but not White faces. This finding suggests that the police need to ensure that the backgrounds of the photo lineups they use are all uniform to reduce mistaken identifications of innocent suspects.

Keywords Face recognition · Eyewitness identification · Own race bias · Biased backgrounds

Introduction

In 1982, a young White woman was brutally attacked by a Black man and during the attack he told the victim 'she wasn't special' as he had been with a White girl before. The local police instantly singled out Marvin Anderson, not because he had any police record, but because he was known to have a White girlfriend. In fact, because Marvin had no criminal record, the police did not have any mugshots of him, so they used a colour photograph taken from Marvin's work ID card and placed this in a photo lineup along with six black and white mugshots. The victim identified Marvin as her assailant, although it was discovered later that the actual perpetrator was present in the lineup. The court convicted him of two accounts of rape, sodomy, robbery and abduction, he was sentenced to 210 years. In 2002 Marvin was eventually exonerated when DNA evidence demonstrated

that it was John Otis Lincoln who had been the real perpetrator (Garrett 2011). This is one example whereby a biased police lineup was constructed, so that one innocent person stood out and was misidentified by a witness, there are many more examples that have been highlighted by the Innocence Project (Innocence Project 2021).

Psychological research has also investigated the influence of bias in lineups with the aim to reduce mistaken identifications of innocent suspects. Lineup bias can take various forms, such as the instructions given whilst viewing a lineup, ensuring the witness is informed the 'person may or may not be there' (Clark 2005; Malpass and Devine 1981), the format of the lineup presentation, i.e. sequential versus simultaneous lineups (McQuiston-Surrett et al. 2006; Wells et al. 2015) and whether the lineup administrator knows who the suspect is (Clark et al. 2009, 2013; Kovera and Evelo 2020). The aim of the current paper is to examine the less well researched area of image properties and how this influences identification decisions.

One of the few studies that looked at the image properties of a lineup investigated the influence of biased instructions and a suggestive simultaneous photo lineup. In the biased lineup, the target's image was a different angle to the rest of the filler images and had a different facial expression (Buckhout et al. 1975). Unsurprisingly, participants were more likely to choose the target in the biased lineup, however the study only employed target present lineups, so the

✉ Catriona Havard
catriona.havard@open.ac.uk

1 School of Psychology and Counselling, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK

2 Department of Psychology, Sociology and Politics, Centre for Behavioural Science and Applied Psychology, Sheffield, UK

3 School of Natural Health and Sciences, Tallinn University, Tallinn, Estonia

influence of a biased lineup could not be examined for the selection of an innocent person from a target absent lineup.

In a more recent study, Havard et al. (2019) suggested there are variations in the backgrounds of video and photo lineups due to the way faces are filmed using different lighting conditions and different cameras. Using the well-established 1 in 10 face recognition task (Bruce et al. 1999, 2001; Megreya and Burton 2006, 2007), participants were asked to identify previously seen White (own race) and Black (other race) target faces from target present (TP) and target absent (TA) face arrays. They found when the backgrounds had small variations in colour, participants were more likely to incorrectly choose a face they had not seen from a TA lineup and thereby make a false positive identification. Havard et al. (2019) suggest when background colours varied certain faces may appear to be more salient than others and this could increase the likelihood of an innocent person being chosen from a lineup and could be especially the case when a witness is trying to identify someone who is a different race to them. They also found that people were better at identifying own race faces, as compared to other race faces.

The cross-race deficit, cross race effect, or own race bias (ORB) the term we will use throughout this paper, is the widely researched phenomenon whereby people are better at remembering the faces of those who are the same race as them and find it more difficult to recognise faces that belong to those that are deemed to be from a different race. Psychological research investigating the own race bias (ORB) can be traced back to 1960s and 70s where Black and White participants in the US were asked to recognise Black and White faces. The results often found evidence for the ORB, but especially so for White participants trying to recognise Black faces (Cross et al. 1971; Malpass and Kravitz 1969). Since the twenty-first century research exploring the ORB has employed a variety of different populations around the world and demonstrated through a number of different paradigms, such as face recognition (Hayward et al. 2017; Meissner et al. 2005; Wan et al. 2015), face matching tasks (Havard 2021; Kokje et al. 2018; Meissner et al. 2013), eyewitness paradigms (Dodson & Dobolyi 2016; Havard et al. 2017; Jackiw et al. 2008; Marcon et al. 2008), and meta-analyses (Meissner and Brigham 2001; Singh et al. 2021). Most studies have found that people are more likely to correctly identify a previously seen face if it belongs to the same race as them and more likely to make a false positive response (falsely recognise a face) if it belongs to someone who is from a different race.

Several explanations have developed to understand the mechanisms behind the ORB. The contact hypothesis suggests that people gain expertise in recognising those they have more frequent social interactions with, but struggle to then recognise people from races with which they have less frequent interactions (Chiroro and Valentine 1995; Meissner and Brigham 2001; Singh et al. 2021). According to the contact hypothesis, the amount of contact people have with individuals of another race will positively correlate with recognition accuracy for faces of that race (Brigham and Malpass 1985). However, research has found mixed results in support of the contact hypothesis, some research has found a relationship between self-reported levels of interracial contact and recognition of other race faces (Hancock and Rhodes 2008; Walker and Hewstone 2006a, b), whilst others have found no relationship (Wong et al. 2020).

Another factor that can influence the ORB is a person's general face recognition ability when viewing own race faces, and there is some evidence this can influence their ability to recognise other race faces. In a face matching task, using own race faces and other race faces, performance on own race faces correlated to performance on the other race faces, i.e. those who were more accurate with own race faces were also more accurate with other race faces and the reverse pattern was found for those who were less accurate with own race faces (Kokje et al. 2018). This pattern was also confirmed by recognition studies, where those who performed poorly recognising own race faces were even less accurate at other race faces (Correll et al. 2021; Wan et al. 2017). These studies suggest that both underlying face recognition ability and interracial contact are important aspects of the ORB.

The aim of the current paper was to extend the research examining the influence of small colour variations on the background images of face arrays on the accuracy of face identification. The limited previous research has suggested that small variations in colour of backgrounds for face arrays increases false positive responses, especially when trying to identify other race faces. The current study aims to see if this finding is replicated using larger samples from different populations and to determine if this pattern is related to the own race bias. In the UK, there are two different video lineup systems that are employed by police forces, PROMAT (2010) and VIPER® (2021), both are very similar in appearance with the main differences being that the colours of the backgrounds differ, VIPER® lineups have grey backgrounds and PROMAT lineups have green backgrounds. In the first experiment, the aim was to replicate the findings from Havard et al. (2019) study using a larger sample of participants. In the second experiment the research was extended further using two different racial groups, Black and White participants and include interracial contact measures and general face recognition ability to investigate the ORB. The third experiment examined whether the ORB and influence of background variations was colour specific (hues of green found with PROMAT video lineups) and would generalise when a different colour (shades of grey found with VIPER video lineups) was employed for the background arrays.

Experiment 1

The aim of the first experiment was to investigate whether small variations in the backgrounds of target present and target absents would influence the identification of own race and other race faces. A further aim was to see if the findings from Havard et al. (2019) study conducted in the UK could be replicated with a larger sample of participants in a different geographical region in Estonia.

Method

Participants

A total of 42 participants took part in the study (21 female, ages 19–53, M = 29.55, SD = 7.21). All participants were White, spoke Estonian (93%), or Russian (7%) as a native language. Sixty percent had higher education, 40% secondary or vocational education. The data from two participants was incomplete due to a software malfunction, leaving 40 participants for the analyses.

Procedure

The experiment was conducted in the experimental psychology laboratory in Tallinn University, Estonia. Participants were tested individually and on average the experiment lasted for about 45 min. The board of research ethics at Tallinn University (Estonia) approved the study before the data collection commenced. First they were asked to sign a voluntary consent form to participate in the experiment. The experiment was conducted using laptop Lenovo ThinkPad T440 with external monitor HP EliteDisplay E232, 23", 60 Hz using E-Prime 3.0 software (Psychology Software Tools, Pittsburgh, PA). Prior to the experiment, the participants were given instructions.

The 1 in 10 face recognition paradigm was employed (see Fig. 1), using the same methodology at Havard et al. (2019). Each trial began with the presentation of an image of colour face, presented for 5 s, on a grey background. Then after a delay of 2 s, an array of 10 faces (2 rows of 5) where each face was presented on a green square and all were presented on a White background. Each array face was numbered, from 1 through to 10 and they were displayed until the participants responded. The images had been used in a previous

Fig. 1 An example of the 1 in 10 paradigm with mixed grey background on a black target present array (Experiment 3). The correct response is number 9
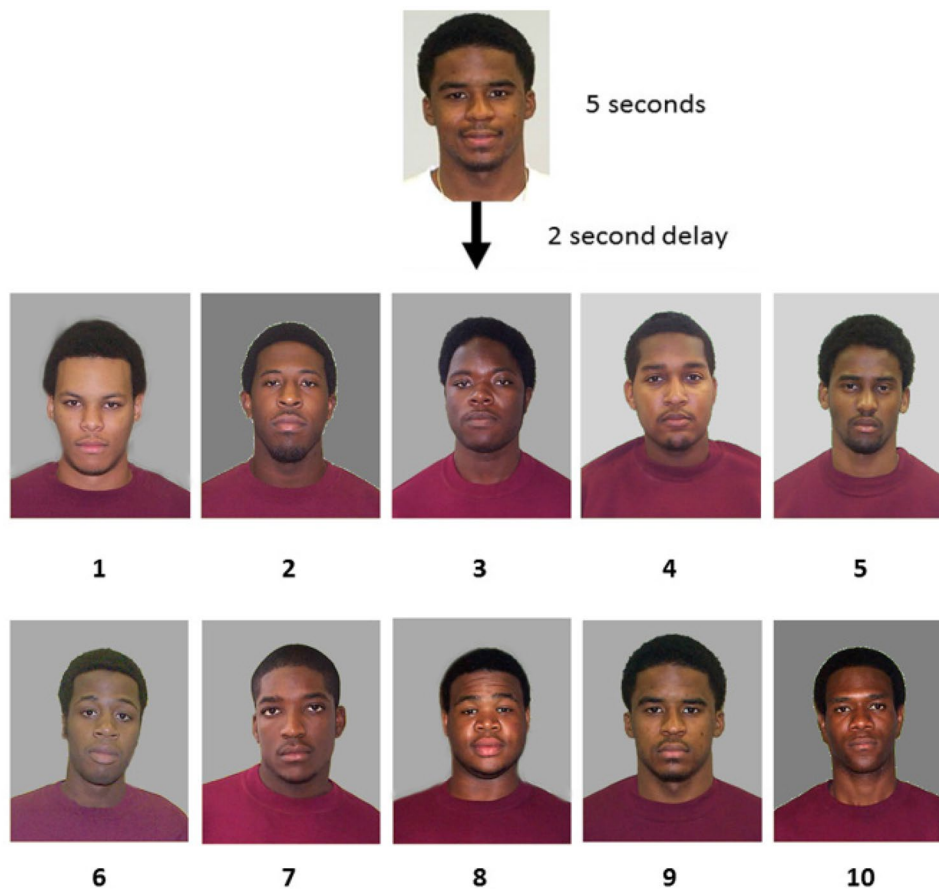
study, for details please see Havard et al. (2019), and were taken from a database that had been utilized in previous research by Meissner et al. (2005).

Participants were presented with 160 trials. 80 trials had White targets and arrays and 80 were Black targets and arrays. Half the arrays were target present (TP), where the target face was present in the array, and half were target absent (TA) arrays, where the target face was not present in the array. When the target was present in the array, a different image of the same target face was utilised. For half of the arrays all the square's backgrounds were the same colour; the mean background colour of all images in the database. For the other half of the arrays the backgrounds consisted of the 4 colour variations; colour values 2 SD from the mean in CIE colour space, as well as the mean background (See Havard et al. 2019, for details on the background processing technique). Therefore, there were 5 different coloured backgrounds that were randomly assigned. The arrays were counterbalanced so that the target appeared in every location in the array.

Results

Table 1 shows the mean performance of recognition accuracy for faces in both the mean (all mean coloured squares) and mixed (all four colours 2 SD from the mean along with the mean coloured background), for White and Black arrays.

For TP arrays, three measures were calculated: hits (correctly identifying the target), misses (incorrectly saying the target was not present) and misidentifications (MisID: incorrectly choosing someone other than the target). For TA arrays, there are two types of responses, they are False positives (FPs: choosing someone from the array) and Correct rejections (correctly saying the person is not there), as these two measures are inversely proportional only false positives are presented. Each of the measures were subjected to a repeated measures analysis of variance with face race (White vs. Black) and array background (mean vs. mixed) as factors.

Please note the data from this study precludes the calculation of signal detection measures of sensitivity and criterion

Table 1 Percentage of responses for White and Black Targets on mean and mixed arrays (standard deviations in parentheses)

Array	Hits	MisIDs	Misses	FPs
White Mean	76.5 (13.31)	7.38 (9.34)	16.75 (10.59)	21.88 (21.53)
White Mixed	67.38 (16.64)	11.50 (10.33)	19.25 (14.21)	20.50 (20.71)
Black Mean	64.13 (13.77)	16 (12.05)	18.38 (12.00)	35.38 (21.97)
Black mixed	68.88 (12.63)	13.59 (11.82)	15.16 (7.56)	40.50 (23.53)

because participants sometimes make misidentifications in TP trials.

For Hits there was a main effect of Face Race, $F(1, 39) = 10.99$, $p = 0.002$, $\eta p^2 = 0.220$, with more hits for White faces than Black faces, but no reliable main effects of Array background, $F(1, 39) = 1.91$, $p = 0.175$, $\eta p^2 = 0.047$. There was a significant interaction for Face race and array background, $F(1, 39) = 19.47$, $p = 0.001$, $\eta p^2 = 0.333$, revealing more hits for White faces on mean backgrounds, as compared to Black faces on mean backgrounds ($p < 0.001$), but no significant differences for hits for both races for the mixed background ($p = 0.522$). For White faces there were more hits for the mean backgrounds, compared to mixed backgrounds ($p = 0.001$), whereas for Black faces there were more hits for mixed backgrounds ($p = 0.02$).

For MisIDs, there was a main effect of Race, $F(1, 39) = 26.03$, $p = 0.001$, $\eta p^2 = 0.400$, with more MisIDs for Black faces compared to White faces, but no reliable main effects of array background, $F(1, 39) = 0.097$, $p = 0.758$, $\eta p^2 = 0.002$. There was a significant interaction for race and array background, $F(1, 39) = 19.29$, $p = 0.001$, $\eta p^2 = 0.331$, with more MisIDs for Black faces as compared to White faces for mean backgrounds ($p < 0.001$). However, for the mixed backgrounds there were no significant differences in MisIDs for either race ($p = 0.382$). For the White faces there were more MisIDs for mixed backgrounds, as compared to mean backgrounds ($p = 0.008$). Whilst for Black faces there were more MisIDs for mean backgrounds as compared to the mixed backgrounds ($p = 0.029$).

For Misses, there was a main effect of Race, $F(1, 39) = 4.16$, $p = 0.048$, $\eta p^2 = 0.096$, with more Misses for White compared to Black faces, but no reliable main effects of array background, $F(1, 39) = 3.30$, $p = 0.077$, $\eta p^2 = 0.078$. There was a significant interaction for race and array background, $F(1, 39) = 12.03$, $p = 0.001$, $\eta p^2 = 0.236$. There were more misses for the White faces as compared to Black faces for mixed backgrounds ($p < 0.001$), whereas for the mean backgrounds there were no significant differences in misses for White or Black faces ($p = 0.433$). For the Black faces, there were significantly more misses for the mean backgrounds compared to the mixed backgrounds (< 0.001), whilst there were no significant differences in misses between the backgrounds for the White faces ($p = 0.179$).

For the false positives (FPs), there was a main effect of race, $F(1, 39) = 52.72$, $p = 0.001$, $\eta p^2 = 0.575$, with more false positives for Black faces compared with White faces, but no reliable main effects of array background, $F(1, 39) = 1.57$, $p = 0.217$, $\eta p^2 = 0.039$. There was a significant interaction for race and array background, $F(1, 39) = 6.65$, $p = 0.014$, $\eta p^2 = 0.146$. There were more FPs for Black faces on both the mixed and mean backgrounds as compared to White faces (all ps > 0.001). There were more FPs for Black faces on mixed backgrounds as compared to mean backgrounds

($p = 0.006$), whilst for the White faces the FPs for mean backgrounds and were very similar mixed backgrounds ($p = 0.524$).

Discussion

The aim for the first experiment was to investigate whether the variability in the hues of backgrounds of face arrays would influence the identification of own race and other races faces and if Havard et al. (2019) findings of an increase in false positives responses for mixed backgrounds could be replicated using a larger sample from a different geographical location. We found mixed arrays increased false positive responses replicating Havard et al. (2019) study, however only for the Black faces and not for the White faces. Interestingly, the false positives for White faces were very similar for mean backgrounds as compared to the mixed backgrounds. Looking at Havard's data, the influence of the mixed arrays did appear to increase FPs to a greater magnitude for Black faces as compared to White faces, however this was not found to be statistically significant. The differences between the current study and that of Havard and colleagues could be due to differences in samples, Havard's sample was relatively small, and all were from the UK, whilst the current study had a larger sample, nearly twice that of Havard et al. (2019) and all were from Estonia, which has a more homogenous population in relation to different racial groups (World Population Review n.d.). The current study suggests that the mixed backgrounds are especially detrimental for correctly rejecting target absent arrays with Black faces, and less so for White faces.

When it came to correctly identifying a face from an array, the background had different effects depending on the race of the face. Replicating Havard et al.'s findings there was there was an increase in correct identifications for mean backgrounds for White faces, whereas for Black faces there were more correct identifications for mixed backgrounds. Unlike Havard's study there were also influences of the backgrounds for miss responses and misidentifications. The current study found more misidentifications for Black faces on mean backgrounds, whereas the opposite pattern was found for White faces. Furthermore, participants were more likely to incorrectly say the target was not present in the array if the face was White on a mixed background, whilst the opposite pattern was found for Black faces.

Participants demonstrated an ORB for both TP and TA lineups, being more accurate with own race (White) faces. This finding replicates many other studies that have found that people are more accurate at identifying faces that are the same race as them (Havard 2021; Havard et al. 2017, 2019; Hayward et al. 2017; Meissner et al. 2005; Wan et al. 2015). There was an ORB for correct identifications for

mean backgrounds, compared to mixed backgrounds, which was diminished for mixed backgrounds, which is the same pattern as Havard et al. (2019) found. Replicating previous research there were more misidentification responses for other race faces as compared to own race faces (Havard et al. 2019). Interestingly the current study also found a significant influence of race and backgrounds for the miss responses, unlike Havard et al. (2019) where no significant effects were found. In the current study there were more miss responses for own race (White) faces as compared to other race (Black) faces and mean arrays appeared to increase miss responses for other race faces, but not for own race faces. There were significantly more false positive responses for other race faces as compared to own race faces, replicating the findings that people are more likely to identify an innocent person if they are from another race from a TA lineup (Havard et al. 2017, 2019; Jackiw et al. 2008; Wylie et al. 2015).

Both the current study and that of Havard et al. (2019) replicated the own race bias; however, one criticism of both studies is that they only employed White participants, so it was not possible to determine whether the own race bias could also be demonstrated for Black participants trying to identify White faces and if the influence of mixed backgrounds for increasing false positives is a factor linked to the ORB.

Experiment 2

The aim of Experiment 2 was to investigate if the influence of the mixed backgrounds and own race bias could be replicated using both Black and White participants. Another aim was to measure levels of interracial contact and examine whether performance identifying own race faces was related to ability to recognise other race faces.

Methodology

Ethical approval was gained from the Open University's Human Research Ethics Committee (HREC), prior to commencing the research.

Participants

Ninety-four participants took part in the study (45 female, 49 males ages – 18–54, M = 28.5). 48 identified as White (24 UK, 6 Portugal, 6 Italy, 3 Poland, 2 Czech Republic, 2 Slovenia, 1 Estonia, 1 Latvia, 1 Austria, 1 USA) and 46 identified as Black (24 UK, 18 USA, 2 Ireland, 1 Canada, 2 Spain). They were recruited through Prolific and reimbursed with a voucher for their time. Ninety-seven of participants took part, but 3 failed the online attention checks.

Stimuli and Procedure

The 1 in 10 face recognition paradigm was employed, using the same stimuli as for Experiment 1. The experiment was presented online using Gorilla Experiment Builder (www.gorilla.sc), on participants own computers or tablets. Five attention checks were included, four within the face matching trials, where participants were asked to respond a specific way, and one within the questionnaire, where they had to respond selecting a certain answer.

After completing the face memory task, participants were asked to complete a questionnaire regarding their demographical data and an interracial contact questionnaire (Walker and Hewstone 2006a). The questionnaire consists of questions that measure social contact and individuating experience. The social contact scale consisted of 5 items. Item 1 asked 'How many Black (White) people do you know very well?' with answer choices: Up to 2, Up to 5, Up to 8, Up to 12 and more than 12. Items 2–5 were worded as follows: 'I often spend time with Black (White) people', 'I spend a lot of my free time doing things with Black (White) people', 'I often go round to the houses of Black (White) people', and '(White) people often come around to my house' (all using the following scale: strongly agree, sort of agree, not sure, sort of disagree, strongly disagree).

In the individuating experience scale, participants were asked how often they engaged in activities with other-race individuals using the following scale: very often, quite often, sometimes, hardly ever and never. The five items were: 1) 'I have looked after or helped a Black (White) friend when someone was causing them trouble or being mean to them', 2) 'A Black (White) person has looked after me or helped me when someone was causing me trouble or being mean to me', 3) 'I have comforted a Black (White) friend when they have been feeling sad', 4) 'A Black (White) person has comforted me when I have been feeling sad', and 5) 'I have asked an Black (White) person to be on my team or in my group during sports or activities'.

Results

Table 2 shows the mean performance for White and Black participants accuracy for faces in both the mean (all mean coloured squares) and mixed (all 4 colours 2 SD from the mean along with the mean coloured background), for White and Black arrays. The same measures for the TP (hits, missIDs, Miss) and TA (FP) arrays were calculated. For each measure a mixed factor ANOVA was performed: Between subject factor Participant Race (Black vs. White) and within-group factors Face Race (White vs. Black) and Array Background (Mean vs. Mixed).

For Hits there was a significant main effect for Face Race $F(1,92) = 13.73$, $p < 0.001$, ηp2 = 0.13, and Background $F(1,92) = 8.61$, $p = 0.004$, ηp2 = 0.94). There were more Hits for White faces (68.67) as compared to Black faces (63.76) and more hits for mixed backgrounds (67.55) as compared to mean backgrounds (64.88). In addition, there was a significant interaction for Face Race and Ethnicity $F(1,92) = 11.97$, $p < 0.001$, ηp2 = 0.12, and for Face Race and Background, $F(1,92) = 4.49$, $p = 0.037$, ηp2 = 0.05. White participants made significantly more hits for White faces as compared to Black faces (68.54 vs 59.38, $p < 0.001$), whilst Black participants had a similar hit rate for Black or White faces (68.15 vs 68.48 $p = 0.86$). There were more hits for the White targets ($M = 68.39$) as compared to Black targets ($M = 61.36$) for the mean backgrounds ($p < 0.001$), but no significant differences in hits for mixed backgrounds ($M = 68.95$ vs. 66.16, $p = 0.11$). For White faces, there were no differences in hits for either background array ($p = 0.68$). In contrast, for the Black faces there were more hits for the mixed arrays as compared to the mean arrays ($p < 0.001$).

For Misses there were no significant main effects (all ps > 0.1), but there was an interaction for Face Race and Background, $F(1,92) = 7.68$, $p = 0.007$, ηp2 = 0.08. There were more Misses for Black faces on mean backgrounds compared to mixed backgrounds ($M = 18.53$ vs. 15.16, $p = 0.003$), but no significant differences for Misses for White faces on the mean or mixed backgrounds ($M = 16.16$ vs. 17.09, $p = 0.39$).

Table 2 Percentage of responses by Black and White participants for White and Black Targets on mean and mixed green arrays (standard deviations in parentheses)

Participants		Hits	Miss	MissID	FPs
White	Own-race Mix background	68.54 (15.61)	16.56 (13.69)	14.90 (14.86)	35.00 (25.24)
	Own-race Mean background	69.17 (15.86)	14.06 (10.93)	16.77 (15.73)	33.85 (23.84)
	Other-race Mix background	61.46 (17.89)	15.10 (13.07)	23.44 (15.92)	49.90 (25.78)
	Other-race Mean background	57.29 (19.71)	17.81 (14.76)	24.90 (17.24)	46.25 (24.13)
Black	Own-race Mix background	70.87 (21.02)	15.22 (18.35)	13.91 (19.00)	33.80 (27.66)
	Own-race Mean background	65.43 (21.26)	19.24 (20.71)	15.33 (17.78)	26.30 (25.70)
	Other-race Mix background	69.35 (21.02)	17.61 (18.10)	13.04 (18.00)	29.46 (29.12)
	Other-race Mean background	67.61 (19.71)	18.26 (18.33)	14.13 (17.77)	32.39 (26.79)

For MissIDs there was a significant effect of Background $F(1,92) = 4.26$, $p = 0.042$, $\eta p2 = 0.044$, and Face Race $F(1,92) = 21.15$, $p < 0.001$, $\eta p2 = 0.19$. There were more MissIDs for faces on mean backgrounds (19.39) as compared to mixed backgrounds (14.71) and more MissIDs for Black faces (17.78) as compared to White faces (16.32). There was an interaction for Face Race and Participant Race, $F(1,92) = 12.85$, $p = 0.001$, $\eta p2 = 0.12$, which revealed significantly more MissIDs made by White participants for Black faces as compared to White faces ($M = 24.17$ vs 15.83, $p < 0.001$), but no significant differences in MissIDs made by Black participants as a function of face race ($M = 14.60$ vs 13.59, $p = 0.48$). White participants made significantly more MisIDs for Black for faces than Black participants ($p = 0.007$), but there were no differences in MisIDs for White faces by either group of participants ($p = 0.495$).

For the false positives (FPs), there were main effects of Participant Race $F(1, 92) = 4.57$, $p = 0.35$, $\eta p2 = 0.69$, Face Race $F(1, 92) = 22.73$, $p < 0.001$, $\eta p2 = 0.198$, and Array background $F(1, 92) = 5.378$, $p = 0.023$, $\eta p2 = 0.055$. White participants made more FPs as compared to Black participants ($M = 41.25$ vs 30.49), there were more FPs for Black faces as compared to White faces ($M = 39.06$ vs. 32.68) and more FPs for mixed backgrounds as compared to mean backgrounds ($M = 37.04$ vs. 34.70). In addition there was a significant three way interaction for Participant Race, Face Race and Array background $F(1, 92) = 5.107$, $p < 0.026$, $\eta p2 = 0.053$. Black participants made significantly more FPs for White faces as compared to Black faces for mean backgrounds ($M = 32.39$ vs. 26.30, $p = 0.007$). Whilst for mixed backgrounds there were marginally more FPs for Black faces as compared to White faces ($M = 33.80$ vs 29.46, $p = 0.068$). White participants made significantly more FPs for Black faces as compared to White faces for the mixed backgrounds ($M = 49.90$ vs 35.00, $p < .001$) and for the mean background ($M = 46.25$ vs 33.85, $p < .001$). For the Black faces there were more FPs for mixed backgrounds as compared to mean backgrounds and this was the case for Black participants (33.80 vs 26.30, $p < 0.001$) and White participants (49.90 vs 46.25, $p < 0.001$). There was no significant influence of background on FP for the White faces (all ps > 0.05).

Inter-racial Contact

Following (Walker and Hewstone 2006a, b), the social and individuating contact measures were calculated, by collapsing scores across the questions to calculate mean scores for each participant (please see Table 3). Black participants had a significantly higher level of social contact with Whites, than White participants did with Blacks ($t(92) = 4.29$, $p < 0.001$) and individuating experiences than White participants ($t(92) = 6.12$, $p < 0.001$). Correlations for the Black and White participants revealed no significant relationships

Table 3 Mean responses of social and individual inter-racial contact for the Black and White participants

	Participant ethnicity	Mean	Std. Deviation
Social contact	Black	2.54	1.260
	White	1.56	.965
Individual contact	Black	3.49	.866
	White	2.20	1.126

between measures for individuating experience and accuracy for other race faces for either the mean or mixed background (all ps > 0.1).

Own Race Versus Other Race Performance

The response data were collapsed across the TP and TA mixed and mean arrays and an overall accuracy means were calculated for White and Black arrays for the White and Black participants. Separate correlations performed on the different groups found that there were significant positive relationships between accuracy on own race faces and accuracy for other race faces for the Black participants ($r(48) = .891$, $p < .001$) and also for White participants ($r(48) = .850$, $p < 0.001$).

Discussion

The aim of Experiment 2 was to investigate whether the own race bias and the influence of background array that was found in Experiment 1 could be replicated with a larger sample employing both White and Black participants. The findings revealed that mixed backgrounds increased false positive responses for target absent arrays for Black faces, but not White faces, replicating the findings of Experiment 1. The increased false positive responses for mixed arrays were produced by both Black and White participants suggesting this was not a result of the ORB. Replicating the findings from Experiment 1 the mixed backgrounds increased correct identifications for Black faces on target present arrays, however unlike Experiment 1 there were no differences for correct identifications for White faces for either background condition. Interestingly there were more misidentifications for mixed backgrounds as compared to mean backgrounds, and no influence of the race of the target face. This replicates the findings from experiment 1 for White faces, however the opposite was found for Black faces. When it came to incorrectly saying the person was not there for a target absent lineup there were fewer miss responses for Black faces on mean backgrounds, replicating the finding from experiment 1. However, the current experiment found no

differences for miss responses for White faces a function of array background, whilst experiment 1 found that mixed arrays increased misses for White faces.

An ORB was found for White participants who made significantly more correct identifications for own race faces and more misidentifications and false positives for other race faces. This replicates the finding from experiment 1 and previous research that has found White people demonstrating an ORB by being more accurate at identifying faces that are White and less accurate with faces that are of a different race (Kokje et al. 2018; Marcon et al. 2008; Singh et al. 2021). Interesting, Black participants also demonstrated an ORB by making fewer false positives for Black faces, but only for target absent arrays with mean coloured backgrounds, for all other conditions there was no ORB demonstrated by Black participants.

Contact measures revealed that Black participants had more inter-racial contact with White individuals than White participants had with Black individuals, which aligns with the countries that the participants were living in as all were countries where the most commonly encountered faces would be White (e.g. Europe, UK and US). There were no significant correlations between measures of contact and accuracy for other race faces. There was a significant relationship between accuracy for own race faces and other race faces.

The first two experiments and that conducted by Havard et al. (2019) have found using small variations in the hue of the background colours of the face arrays could influence responses. All the arrays used green backgrounds derived from the PROMAT lineup backgrounds, and the third experiment sought to investigate if the effects could be replicated using changes in lightness, rather than hue.

Experiment 3

Currently in the UK, there are two video lineups systems, PROMAT which uses green backgrounds and VIPER® which uses grey backgrounds. The current experiment aimed to investigate whether the influence of small variations in the hues of green coloured backgrounds for face recognition arrays could be replicated for different shades of grey backgrounds.

Methodology

Ethical approval was gained from the Open University's Human Research Ethics Committee (HREC), prior to commencing the research.

Participants

80 participants took part in the study (44 female, 36 males ages – 18–54, $M = 32.5$). 40 identified as White (7 from the UK, 8 Poland, 8 Portugal, 1 Estonia, 1 Finland, 1 Germany, 2 Greece, 2 Hungary, 2 Ireland, 4 Italy, 1 Mexico, 1 Netherlands, 1 Spain, and 1 Switzerland) and 40 identified as Black (16 from the UK, 15 US, 2 Canada, 1 Poland, 2 Portugal and 4 South Africa). They were recruited through Prolific and reimbursed with a voucher for their time. 84 participants took part, but 4 failed the online attention checks.

Stimuli and Procedure

The same procedure and stimuli were employed from the first experiment. However, the test 10-face arrays were placed on grey backgrounds, rather than green backgrounds (please see Fig. 1). For the test stimuli, 3 different grey backgrounds were chosen using the 8-bit grayscale of RGB images. In the absence of a database of images to approximate, as with the colour backgrounds, we selected a mid-grey (RGB 170), and two grey values equally lighter and darker than this midpoint (RGB 128 and 212 respectively).

Results

The same measures and analyses were used as for experiment 2. For hits, there were no significant main effects (all *ps* were > 0.05), there was an interaction for Participant Race and Face Race, $F(1,78) = 10.46$, $p = 0.002$, ηp2 = 0.12. Follow up analyses found that there were more hits made by Black participants for Black faces ($M = 68.50$) as compared to White faces ($M = 63.81$; $p = 0.008$), whilst the opposite pattern was found for White participants, who made more hits for White faces, however this did not reach statistical significance ($M = 64.38$ *vs.* 67.56, $p = 0.068$).

For Misses, there was a significant a main effect for Array Background $F(1,78) = 7.46$, $p = 0.008$, ηp2 = 0.09, and Face Race $F(1,78) = 9.76$, $p = 0.002$, ηp2 = 0.11). There were more misses for White faces (M = 17.41) as compared to Black faces (M = 15.03) and there were more misses for mean backgrounds (17.72) as compared to mix backgrounds (M = 14.72). There was also a significant interaction for face race and background, $F(1,78) = 4.91$, $p = 0.03$, ηp2 = 0.06, which revealed higher misses for Black faces on mean backgrounds compared to mixed backgrounds (17.63 vs 12.44; $p < 0.001$), whilst no differences for White faces on either background (17.81 vs 17.00, $p = 0.58$).

For MissIDs, there were no significant main effects (all *ps* > 0.05), there was a significant interaction for Participant Race and Face race, $F(1,78) = 17.22$, $p < 0.001$, ηp2 = 0.18. There were more MissIDs made by White participants for

Black faces as compared to White faces ($M = 19.44$ vs 12.44, $p < 0.001$), but no significant differences in MissIDs made by Black participants as a function of face race ($M = 18.25$ vs 20.75, $p = 0.13$). For White faces there were more MissIDs made by Black participants than by White participants (20.75 vs 12.44, $p = 0.015$). Whereas for Black faces there were no significant differences in MissIDs as a function of participant race ($p = 0.70$).

For the false positives (FPs), there were main effects of Face Race $F(1, 78) = 37.19$, $p < 0.001$, $\eta p2 = 0.32$, and Array Background $F(1, 78) = 39.25$ (24.40), $p < 0.001$, $\eta p2 = 0.16$. There were more FPs for Black faces as compared to White faces ($M = 42.40$ vs. 34.31) and more FPs for mixed backgrounds as compared to mean backgrounds ($M = 40.5$ vs. 36.22). There was also a significant interaction for significant interaction for Face Race and Participant Race $F(1,78) = 36.61$, $p < 0.001$, $\eta p2 = 0.32$. White participants made more FPs for Black faces (42.94 *vs.* 26.81, $p < 0.001$), but Black participants made no more FPs for Black or White faces (41.88 vs 41.81, p = 0.97). For White faces, there were significantly more FPs made by Black participants as compared to White participants (41.81 vs 26.81, p = 0.005). For Black faces there were no significant differences in FPs between the Black and White participants (41.88, vs. 42.94, p = 0.83). There was also a significant Face Race and Background, $F(1,92) = 12.13$, $p = 0.001$, $\eta p2 = 0.13$. interaction, revealing more FPs for Black faces on mixed backgrounds as compared to mean backgrounds (46.56 *vs.* 38.25, $p < 0.001$), but no significant differences in FPs for White faces on either background (34.44 vs. 34.19, $p = 0.87$) (Table 4).

Inter-racial Contact

Following Walker and Hewstone (2006a, b), the social and individuating contact measures were calculated, by collapsing scores across the questions to calculate mean scores for each participant (please see Table 5). Black participants had significantly higher levels social contact ($t(78) = 5.96$, $p > 0.001$) and individuating experiences than White participants ($t(78) = 4.39$, $p > 0.001$). There were no significant correlations between levels of interracial contact and accuracy

Table 5 Mean responses of social and individual inter-racial contact for the Black and White participants

	Participant ethnicity	Mean	Std. Deviation
Social contact	Black	3.66	1.21
	White	2.14	1.17
Individual contact	Black	3.45	.898
	White	2.44	1.14

for other race arrays for either Black or White participants (all *ps* < 0.05).

Own Race Versus Other Race Performance

The response data were collapsed across the TP and TA mixed and mean arrays and an overall mean accuracy was calculated for White and Black arrays for the White and Black participants. Separate correlations performed on the different groups found a significant positive relationship between accuracy on own race faces and accuracy for other race faces for the Black participants ($r(40) = 0.891$, $p < 0.001$) and also for White participants ($r(40) = 0.892$, $p < 0.001$).

Discussion

The aim of this experiment was to investigate whether the findings from Experiment 2 would be replicated using grey, rather than green backgrounds. For target present lineups, there was no influence of background variations for correct identifications and misidentifications for target present lineups unlike experiments 1 and 2. This could be because the shades of grey were more similar to one another than the different hues of green used in experiments 1 and 2. However, there was an influence of background for incorrectly saying the face was not there, with more misses for Black faces on mean backgrounds, replicating the findings from Experiment 2. There were also more false positives for Black faces on

Table 4 Percentage of responses by Black and White participants for White and Black Targets on mean and mixed grey arrays (standard deviations in parentheses)

Participants		Hits	Miss	MissID	FPs
White	Own-race Mix background	67.62 (15.36)	20.75 (15.96)	11.63 (11.0)	26.75 (21.11)
	Own-race Mean background	67.50 (18.98)	19.25 (14.74)	13.25 (15.17)	26.87 (19.99)
	Other-race Mix background	66.75 (17.23)	13.50 (11.391)	19.75 (13.82)	42.13 (27.13)
	Other-race Mean background	62.00 (16.94)	18.88 (13.75)	19.12 912.86)	37.25 (24.42)
Black	Own-race Mix background	70.13 (20.83)	11.38 (10.13)	18.50 (16.06)	44.50 (24.70)
	Own-race Mean background	66.87 (18.73)	15.12 (12.93)	18.0 (16.13)	39.25 (24.40)
	Other-race Mix background	63.63 (21.73)	14.50 (12.70)	21.87 (19.04)	42.13 (27.13)
	Other-race Mean background	64.00(26.02)	16.37 (15.06)	19.63 (17.37)	41.50 (26.49)

mixed arrays, and no differences for White faces, replicating the findings from experiments 1 and 2.

An ORB was demonstrated for both Black and White participants, with more hits for own race faces, replicating findings from previous experiments. There were more false positive responses and misidentifications made by White participants for Black faces, but Black participants did not demonstrate an ORB for false positives or misidentification responses, replicating previous reach that has found that not all races demonstrate an ORB, especially if they are in a country where they are the minority (Havard et al. 2017; Sporer and Horry 2011).

There was no relationship between levels of interracial contact in the self-reported questionnaire and accuracy for other race faces, replicating the findings from experiment 2 and this will be discussed in more detail in the next section. There was a positive relationship between accuracy for identifying own race faces and for other race faces, again replicating the results from experiment 2.

General Discussion

The aim of the current paper was to investigate whether small variations in the colours of background images influenced face identifications, using a paradigm akin to the task of an eyewitness trying to identify a culprit from a photo lineup. The consistent finding across all three experiments was that small colour variations (mixed) backgrounds increased false positive responses for target absent arrays, which also replicated findings from Havard et al. (2019). This was consistently found to be the case for Black faces, but not White faces. The increase in false positive responses for Black faces on mixed backgrounds was demonstrated by both White and Black participants, suggesting that this effect was not a result of the ORB, but perhaps something related to the image properties themselves. This finding has serious implications for the police using photo lineups where the background images might vary due to suspects and foils being filmed under different lighting conditions and cameras as this could result in false identifications of innocent suspects, especially when those suspects are Black.

When it came to correct identifications from target present lineups, the background colour manipulation appeared to have more influence when the backgrounds were green (experiments 1 & 2) producing more correct identifications for mean backgrounds for white faces, and more correct identifications for Black faces on mixed backgrounds. When the backgrounds were grey there was little influence of small changes in the background shades for correct identifications or misidentifications. The different responses for the green and grey backgrounds could be due to there being a greater perceptible difference between the green backgrounds than

the grey backgrounds. We have not investigated the variance in background colour required to be reliably detected—our focus here was on employing real-world stimuli and approximating real-world variation—future studies may want to look at systematically determining the level of background variability required for simple detection and that required to produce an impact on line-up performance, both in colour space and in greyscale.

The own race bias was demonstrated across all three experiments, especially for White participants who were consistently more accurate at correctly identifying own race faces and more likely to misidentify or falsely recognise other race faces. These findings replicate a wealth of research that has found that people are better at identifying those who are the same race as them, and are less accurate as identifying faces that belong to a different race (Cross et al. 1971; Havard 2021; Hayward et al. 2017; Kokje et al. 2018; Malpass and Kravitz 1969; Meissner et al. 2005; Wan et al. 2015). The ORB was especially pronounced in the White participants and less so for the Black participants. The reduced ORB for Black participants could relate to the countries where the participants were living, the Black participants who took part in the research were living in countries where the most commonly encountered faces were White (e.g. UK, Ireland, and USA). Research has found some populations do not exhibit an ORB when they live in a country where they are the minority race, and the most commonly seen faces are of a different race (Corenblum and Meissner 2006; Havard et al. 2017; Sporer et al. 2007; Sporer and Horry 2011).

Although the experiments confirmed the ORB, there was no relationship between the self-reported measures of interracial contact and accuracy on identifying other race faces. The lack of relationship between the levels of contact as measured by the contact questionnaire and performance on face recognition tasks could be due to the contact measures used. Although some research has found a relationship between self-reported levels of contact and recognition of other race faces (Hancock and Rhodes 2008; Walker and Hewstone 2006a), other research has not found any relationship (Wong et al. 2020). Furthermore, a meta-analysis of face recognition studies investigating the ORB found self-reported contact measures only accounted for just over 2 percent of the variation for the ORB. This finding suggests that currently used self-reported measures may not be reliable indicators for measuring interracial contact (Singh et al. 2021). One of the reasons that some studies may have found a relationship whilst others have not might be the way that the interracial contact is measured, most research has measured contact using self-report questionnaires (Hancock and Rhodes 2008; Walker and Hewstone 2006a, b). Self-reported contact measures do not usually distinguish between individual's current

inter-racial contact and past interracial contact (e.g. childhood). There is some evidence that for interracial contact to be effective at reducing the ORB it needs to be from an early age, rather than in adulthood (Zhou et al. 2019). Therefore, any future research should examine whether inter-racial contact occurred during childhood in addition to adulthood.

Another factor that can influence a person's ability to recognise faces of another race, is their ability to recognise faces in general. Across two of the experiments there was a positive relationship between people's ability to recognise own race faces and other race faces, this replicates previous research that has found that a persons' general face recognition ability can influence how accurate they are at recognising other race faces (Correll et al. 2021; Kokje et al. 2018; Wan et al. 2017). This finding suggests that in our study, own race recognition accuracy was a better predictor for performance on other race faces than self-reported levels of contact.

There was a clear pattern of results found across three experiments, however one of the limitations of the current study is that we still not fully understand why small variations in background colours increase false identifications of previous unseen Black faces. Future research should focus on understanding the mechanisms behind this increased false positive bias, for example using eye tracking methodologies to examine if there are any eye gaze patterns that might shed light on this phenomenon. Confidence measures could be taken at the time the identification to see if there is any relationship between be levels of confidence and responses. A further limitation was that we did not record whether any of the participants had any colour vision deficiencies (e.g. red/green) as this may have influenced their perception of backgrounds for experiments 1 and 2. Future research investigating variations in background colours should record these data.

To summarise, small variations in the background colours of face arrays that are similar in appearance to photo lineups can bias responses leading to an increase in false positive responses for Black faces. This has serious implications for police administered photo lineups that might have small variations in background colours due to suspects being filmed under different lighting conditions and different cameras. There are also implications for eyewitness researchers who use photo lineups to investigate eyewitness identification. Anyone administering photo lineups should ensure that images have uniform background colours to ensure that innocent persons are not falsely identified, potentially leading to wrongful convictions.

Declarations

References

Brigham JC, Malpass RS (1985) The role of experience and contact in the recognition of faces of own- and other-race persons. J Soc Issues 41(3):139–155. https://doi.org/10.1111/J.1540-4560.1985.TB01133.X

Bruce V, Henderson Z, Greenwood K, Hancock PJB, Burton AM, Miller P (1999) Verification of face identities from images captured on video. J Exp Psychol Appl 5(4):339–360. https://doi.org/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. J Exp Psychol Appl 7(3):207–218. https://doi.org/10.1037/1076-898X.7.3.207

Buckhout R, Figueroa D, Hoff E (1975) Eyewitness identification: effects of suggestion and bias in identification from photographs. Bull Psychon Soc 6(1):71–74. https://doi.org/10.3758/BF03333151

Chiroro P, Valentine T (1995) An Investigation of the Contact Hypothesis of the Own-race Bias in Face Recognition. Q J Exp Psychol Sect A 48(4):879–894. https://doi.org/10.1080/14640749508401421

Clark SE (2005) A re-examination of the effects of biased lineup instructions in eyewitness identification. Law Hum Behav 29(5):575–604. https://doi.org/10.1007/s10979-005-7121-1

Clark SE, Brower GL, Rosenthal R, Hicks JM, Moreland MB (2013) Lineup administrator influences on eyewitness identification and eyewitness confidence. J Appl Res Mem Cogn 2(3):158–165. https://doi.org/10.1016/J.JARMAC.2013.06.003

Clark SE, Marshall TE, Rosenthal R (2009) Lineup administrator influences on eyewitness identification decisions. J Exp Psychol Appl 15(1):63–75. https://doi.org/10.1037/A0015185

Correll J, Ma DS, Davis JP (2021) Perceptual tuning through contact? Contact interacts with perceptual (not memory-based) face-processing ability to predict cross-race recognition. J Exp Soc Psychol 92(January 2020):104058. https://doi.org/10.1016/j.jesp.2020.104058

Corenblum B, Meissner CA (2006) Recognition of faces of ingroup and outgroup children and adults. J Exp Child Psychol 93(3):187–206. https://doi.org/10.1016/j.jecp.2005.09.001

Cross JF, Cross J, Daly J (1971) Sex, race, age and beauty as factors in recognition of faces. Perception 10(6):393–396

Dodson CS, Dobolyi DG (2016) Confidence and eyewitness identifications: the cross-race effect, decision time and accuracy. Appl Cogn Psychol 30(1):113–125. https://doi.org/10.1002/acp.3178

Garrett B (2011) Convicting the innocent. where criminal prosecutions go wrong. (pp. 45–83). Harvard University Press

Hancock KJ, Rhodes G (2008) Contact, configural coding and the other-race effect in face recognition. Br J Psychol 99(1):45–56. https://doi.org/10.1348/000712607X199981

Havard C (2021). The importance of internal and external features in matching own and other race faces. https://doi.org/10.1177/03010066211043464

Havard C, Memon A, Humphries JE (2017) The own-race bias in child and adolescent witnesses. Int J Police Sci Manag 146135571717313577. https://doi.org/10.1177/1461355717731579

Havard C, Richter S, Thirkettle M (2019) Effects of changes in background colour on the identification of own- and other-race faces. I-Perception 10(2). https://doi.org/10.1177/2041669519843539

Hayward WG, Favelle SK, Oxner M, Chu MH, Lam SM (2017) The other-race effect in face learning: Using naturalistic images to investigate face ethnicity effects in a learning paradigm. Q J Exp Psychol 70(5):890–896. https://doi.org/10.1080/17470218.2016.1146781

Innocence Project (2021) https://innocenceproject.org/

Jackiw LB, Arbuthnott KD, Pfeifer JE, Marcon JL, Meissner CA (2008) Running head: Cross-race Effect. Can J Behav Sci 40(1):52–57

Kokje E, Bindemann M, Megreya AM (2018) Cross-race correlations in the abilities to match unfamiliar faces. Acta Physiol (oxf) 185(January):13–21. https://doi.org/10.1016/j.actpsy.2018.01.006

Kovera MB, Evelo AJ (2020) Improving eyewitness-identification evidence through double-blind lineup administration. Curr Dir Psychol Sci 29(6):563–568. https://doi.org/10.1177/0963721420969366

Malpass RS, Devine PG (1981) Eyewitness identification: Lineup instructions and the absence of the offender. J Appl Psychol 66(4):482–489. https://doi.org/10.1037/0021-9010.66.4.482

Malpass RS, Kravitz J (1969) Recognition for faces of own and other race. J Pers Soc Psychol 13(4):330–334. https://doi.org/10.1037/h0028434

Marcon JL, Meissner CA, Jackiw LB, Arbuthnott KD, Pfeifer JE (2008) Examining the cross-race effect in lineup identification using caucasian and first nations samples. Can J Behav Sci 40(1):52–57. https://doi.org/10.1037/0008-400x.40.1.52

McQuiston-Surrett D, Malpass RS, Tredoux CG (2006) Sequential vs. simultaneous lineups: a review of methods, data, and theory. Psychol Pub Policy Law 12(2):137–169. https://doi.org/10.1037/1076-8971.12.2.137

Meissner CA, Brigham JC (2001) Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review. Psychol Public Policy Law 7(1):3–35. https://doi.org/10.1037//1076-8971.7.1.3

Meissner CA, Brigham JC, Butz DA (2005) Memory for own- and other-race faces: A dual-process approach. Appl Cogn Psychol 19(5):545–567. https://doi.org/10.1002/acp.1097

Meissner CA, Susa KJ, Ross AB (2013) Can I see your passport please? Perceptual discrimination of own- and other-race faces. Vis Cogn 21(9–10):1287–1305. https://doi.org/10.1080/13506285.2013.832451

Megreya AM, Burton AM (2006) Unfamiliar faces are not faces: evidence from a matching task. Mem Cognit 34(4):865–876. http://www.ncbi.nlm.nih.gov/pubmed/17063917

Megreya AM, Burton AM (2007) Hits and false positives in face matching: A familiarity-based dissociation. Percept Psychophys 69(7):1175–1184

PROMAT (2010) Promat video ID system: Image capture [Computer software]. Lancashire, England: Promat Envision International

Singh B, Mellinger C, Earls HA, Tran J, Bardsley B, Correll J (2021). Does cross-race contact improve cross-race face perception? A meta-analysis of the cross-race deficit and contact. Pers Soc Psychol Bull 014616722110244. https://doi.org/10.1177/01461672211024463

Sporer SL, Horry R (2011) Recognizing faces from ethnic in-groups and out-groups: Importance of outer face features and effects of retention interval. Appl Cogn Psychol 25(3):424–431. https://doi.org/10.1002/acp.1709

Sporer SL, Trinkl B, Guberova E (2007) Matching faces: Differences in processing speed of out-group faces by different ethnic groups. J Cross Cult Psychol 38(4):398–412. https://doi.org/10.1177/0022022107302310

VIPER® (2021) National VIPER Bureau https://www.viper.police.uk/

Walker PM, Hewstone M (2006a) A developmental investigation of other-race contact and the own-race face effect. Br J Dev Psychol 24(3):451–463. https://doi.org/10.1348/026151005X51239

Walker PM, Hewstone M (2006b) A perceptual discrimination investigation of the own-race effect and intergroup experience. Appl Cogn Psychol 20(4):461–475. https://doi.org/10.1002/acp.1191

Wan L, Crookes K, Dawel A, Pidcock M, Hall A, McKone E (2017) Face-blind for other-race faces: Individual differences in other-race recognition impairments. J Exp Psychol Gen 146(1):102–122. https://doi.org/10.1037/xge0000249

Wan L, Crookes K, Reynolds KJ, Irons JL, McKone E (2015) A cultural setting where the other-race effect on face recognition has no social-motivational component and derives entirely from lifetime perceptual experience. Cognition 144(0010):91–115. https://doi.org/10.1016/j.cognition.2015.07.011

Wells GL, Steblay NK, Dysart JE (2015) Double-blind photo lineups using actual eyewitnesses: An experimental test of a sequential versus simultaneous lineup procedure. Law Hum Behav 39(1):1–14. https://doi.org/10.1037/lhb0000096

Wong HK, Stephen ID, Keeble DRT (2020) The own-race bias for face recognition in a multiracial society. Front Psychol 11(March). https://doi.org/10.3389/fpsyg.2020.00208

World Population Review (n.d.) Estonia Population 2022 (Demographics, Maps, Graphs). Retrieved March 8, 2022, from https://worldpopulationreview.com/countries/estonia-population

Wylie LE, Bergt S, Haby J, Brank EM, Bornstein BH (2015) Age and lineup type differences in the own-race bias. Psychology, Crime and Law 21(5):490–506. https://doi.org/10.1080/1068316X.2014.989173

Zhou X, Elshiekh A, Moulson MC (2019) Lifetime perceptual experience shapes face memory for own- and other-race faces. Vis Cogn 27(9–10):687–700. https://doi.org/10.1080/13506285.2019.1638478