The combined influence of the own-age, -gender, and -ethnicity biases on face recognition

Abstract

Whether the own-group (own-ethnicity, own-gender, and own-age) biases in face recognition are based on the same mechanism and whether their effects are additive or not are as yet unanswered questions. Employing a standard old/new recognition paradigm, we investigated the combined crossover effects of the own-ethnicity, own-gender, and own-age biases in a group of 160 participants. Result showed that while face recognition accuracy decreased as the number of out-group features increased, the own-ethnicity bias appeared to have more of a unique influence on face recognition than the other biases. Furthermore, we established that in a single group of participants, these biases appear to be based on different mechanisms: the own-ethnicity bias is based on individuation whereas the own-age and own-gender biases are based on motivation.

Keywords: own-ethnicity bias; own-gender bias; own-age bias; own-group bias; face recognition; additive effects

The combined influence of the own-age, -gender, and -ethnicity biases on face recognition

Face recognition an important ability vital in forming relationships (Benjamin, 2013). While face recognition is an ability that humans are extremely adept at, there are types of faces that we are not so good at recognising: those of other-groups. The own-group biases are revealed through faster and more accurate recognition of faces of one's own group relative to those of out-groups (Blaine, 2007). The three most widely researched and theorised biases are: the own-ethnicity[1] (Meissner & Brigham, 2001), own-gender (Lovén, Herlitz & Rehnman, 2011), and own-age bias (Rhodes & Anastasi, 2012). Here, we briefly review the research on these biases in an attempt to ascertain how faces that fit multiple categories might be processed.

The own-ethnicity bias is the tendency to recall faces that are of the same ethnicity as ourselves, with better accuracy than those that are not (Meissner & Brigham, 2001). The mean effect size for this bias is r=.38, Cohen's d=.82. This is the most researched of the biases and most theoretical models of them have been based on this bias. It is present across most ethnicities (Bothwell, Brigham, & Malpass, 1989) and appears to be related to the amount of (quality) contact one has with faces of another ethnicity (Brigham, Maass, Snyder, & Spaulding, 1982; Chiroro & Valentine, 1995; Cross Cross, & Daly, 1971; Stelter, Rommel, & Degner, submitted; Walker & Hewstone, 2006).

[1] These biases are sometimes known by other names. Sometimes the word "race" replaces "ethnicity" in describing the own-ethnicity bias, however, the word "ethnicity" is more appropriate since there is only one human subspecies (race). Moreover, even if “race” is used to describe the major anthropological groups, it is incorrect (as is common in the literature) to use the term “race” to refer to ethnicities such as “Hispanic” (see Valentine, Lewis, & Hills, 2016).

The own-age bias is the tendency to recognise those of the same or similar age range as oneself more accurately than other ages (Hills, 2012; Wright & Stroud, 2002). The own-age bias has be shown for older adults (Lamont, Stewart-Williams & Podd, 2005; Anastasi & Rhodes, 2005; Perfect & Harris, 2003), young adults (Anastasi & Rhodes, 2006), and children (Anastasi & Rhodes, 2005; Lindholm, 2005; Hills & Lewis 2011). The own-age bias also appears to be partially dependant on experience (Harrison & Hole, 2009). However, unlike the own-ethnicity bias, faces that were once own-age become other-age which indicates that it is recent experience that appears to be moderating the bias more so than historical experience (Hills, 2012). The average effect size for this bias is r=.18, Cohen's d=.37 (Rhodes & Anastasi, 2012).

The own-gender bias is indicated by superior recognition of faces of one's own gender relative to the other gender (McKelvie, 1987). This bias is highly asymmetrical (Herlitz & Lovén, 2013), with the bias more commonly found in women (Lovén, Herlitz, & Rehnman, 2011; Rehnman & Herlitz, 2006, 2007) - only a handful of studies have found the full crossover bias in both men and women (Ellis, Shepherd & Bruce, 1973; Man & Hills, 2016; Wright & Sladden, 2003). This bias is also not based on recent experience (since half the population is female; Hills, Pake, Dempsey, & Lewis, 2018) but may be based on early experience (Herlitz & Lovén, 2013). The own-gender bias is smaller than the other biases, with an average effect size of r=.27 for women and r=.02 for men, overall Cohen's d=.29 (Herlitz & Lovén, 2013).

Very few studies have investigated multiple biases at the same time. This research gap means that it is difficult to compare the effects in one bias to another due to potential cohort effects and individual differences in the magnitude of these biases (but see Hills et al., 2018 who showed that the own-ethnicity bias was larger than the own-age bias and the own-gender bias in a single group of participants). Thus, each bias might be based on a different and potentially unique mechanism, or subtle methodological differences across studies result in conflicting findings. Further, although the same participants can show multiple biases in the same study, it is not clear what happens to faces that fit multiple out-group categories. For example, an own-age but other-gender and -ethnicity face might be considered partially in-group or entirely out-group depending on how the own-group is classified. In other words, the biases might be all or none, additive, or the different biases might have different weightings.

Wiese (2012) found that own-age or own-ethnicity faces were recognised significantly more accurately than other-age and other-ethnicity and combined other-age-other-ethnicity faces. All out-group faces were recognised at an equivalent level to each other. In Herlitz and Lovén's (2013) meta analysis on the own-gender bias, they indicated that the magnitude of this bias was the same whether the faces were own- or other-ethnicity. These results suggest that once a face has one out-group feature, it is considered out-group and there is no additional effects of further out-group features: there is little evidence for graded group categorisation (only that group categorisation is contextual). However, Rehnman and Herlitz (2006) found contrasting results, finding an additive effect of age and ethnicity to the effect of gender in the own-group biases. It is difficult to reconcile the differences in these studies given similarities in the methods: potentially individual difference variables might account for these differences.

Broadly speaking, there are two broad classes of explanations for the own-group biases: perceptual accounts and socio-cognitive motivational accounts. The perceptual expertise models typically suggest that contact leads to individuals having differential proficiency in processing own-ethnicity versus other-ethnicity faces (e.g., Bukach, Cottle, Ubiwa, & Miller, 2012; Valentine & Endo, 1992). Increased experience with other-group faces can then lead to more opportunities for differentiating faces with variability in different facial features. Due to the lack of contact with those of dissimilar characteristics, individuals become relatively inept at distinguishing between out-group faces (MacLin & Malpass, 2001). One perceptual account of the own-group biases suggests that we utilise expert holistic processing (processing that is selectively applied to faces based on encoding a face as a gestalt whole or processing the facial features in parallel, Richler, Palermi & Gauthier, 2012) more so for faces of one's own group compared to faces of other groups (Michel, Rossion, Han, Chung, & Caldara, 2006). Given that the deployment of holistic processing does not have to be all or none, it can be assumed that faces that are considered more of an out-group will be processed with less holistic processing than faces that are considered less of an out-group.

In Sporer's (2001) in-group/out-group model, faces are processed more deeply using effortful processing if they are considered to be an in-group. As an individual categorises a stimulus as an out-group, motivation to process them deeply is reduced which then leads to a weaker and less effective processing of individuating features (Bernstein et al., 2007; Rodin, 1987; Fiske & Neuberg, 1990). Such shallow processing leads them to be less well recognised subsequently (Bernstein, Young & Hugenberg, 2007). This model suggests that potential cognitive overload caused by processing all faces deeply is reduced by categorisation. Indeed, categorising ambiguous faces as either own-group or out-group alters the accuracy with which they will be recognised (MacLin & Malpass, 2001). Therefore, out-group faces will be processed more shallowly than in-group faces.

A recent theoretical advancement is the categorisation-individuation model (Hugenberg, Young, Bernstein, & Sacco, 2010). In this model, three factors contribute to the own-group biases: social categorisation, perceiver motivation, and perceiver experience with other-group faces. When participants encounter a face, they categorise it according to group and potentially engage in individuation for own-group faces but not other-group faces, depending on their level of motivation (see also, Young, Hugenberg, Bernstein, & Sacco, 2012). This individuation process still requires perceptual expertise to be present in order to process other-group faces. This theory allows for each own-group biases to be based on different mechanisms depending on the level of experience with processing those faces and the motivation to individuate (Hugenberg, Wilson, See, & Young, 2013). Therefore, the biases may add together, but one bias might be more prominent than the others.

In order to answer the empirical question regarding whether the effects of the biases are additive, or whether there is a simple in-group/out-group classification, we ran an old/new recognition paradigm employing the own-age, -gender, and -ethnicity biases. Faces could be own- or other- for each group, thereby creating eight groups of faces. If the biases are not additive and there is a simple categorisation process (predicted by a strict in-group/out-group model), then all faces with one or more out-group feature will be processed to the same level of performance below that of in-group faces. If the biases are additive then faces with more out-group features will be recognised less accurately than faces with less out-group features in a graded fashion (based on different levels of holistic processing being engaged in for example). Finally, if the biases are not equivalent and one is stronger than the others (for example, the own-ethnicity bias), then while the biases might add together, the effect might be larger for one bias than the others (predicted by the categorisation-individuation model, Hugenberg et al., 2010). These predictions are called the "all or none" hypothesis, the graded model, and the differential model respectively.

We used a number of dependent measures to address the mechanisms behind the own-group biases: Accuracy is used to establish the bias and whether the effects are additive, all or none, graded or differential - in this way, accuracy establishes if the biases are based on the same mechanisms; response time and encoding time measure effort engaged in for processing (Crookes & Rhodes, 2017); distinctiveness ratings establish the amount of individuation employed for faces (Valentine & Endo, 1992); and response bias used to measure participants willingness to falsely recognise faces.

Method

Participants

An opportunity sample of 160 individuals were recruited for this study from Bournemouth and London. These individuals varied on three characteristics: age (younger group=18 to 29 years, older=more than 30 years), gender (female, male), and ethnicity (Black, White) creating eight equal groups of participants (20 participants in each group). Sample size was determined based on the effect size of the own-gender bias (as it has the smallest effect size), assuming a power of 0.95. Using GPower, we established that 159 participants would be required to find a significant effect. All participants were fluent in English and understood the instructions which had been presented to them.

Materials

Two versions of 200 faces from the Minear and Park database (Minear & Park, 2004) and stimuli from Hills and Lewis (2013) were used in this experiment: one was presented during the learning and one was presented at test (this was counterbalanced and done to minimise pictorial recognition). These faces belonged to the same eight categories as the participants based on: age (younger=18 to 29 years, older=more than 30 years), gender (female, male), and ethnicity (Black, White). There were an equal number of each type of stimuli. The images were adjusted to have the same plain white background and cropped to mask out clothing. Faces were presented in 640 x 480 px size. The faces were presented in full frontal view and displayed either a neutral or a smiling expression. The images were presented using the software OpenSesame on an ASUS Model T5501 PC. Faces of each group were rated for distinctiveness and attractiveness by a separate group own-group participants to ensure that the stimuli were equivalent. No differences were found (all *p*s>.253).

Design

A mixed-subjects design was used with the factors of age, ethnicity and gender of the faces presented and observer. The accuracy of response which was measured using the Signal Detection Theory (SDT; Green & Swets, 1966) measure, *d'*. Response bias was measured using the SDT measure, *C.* We also measured response time during learning and during test in addition to analysing distinctiveness ratings made to faces as this can act as an index of individuation of faces. Counterbalancing was employed such that each face appeared as a target or as a distractor an equal number of times.

Procedure

Participants were tested individually in a quiet setting. After providing informed consent, participants were seated directly in front of a laptop at a distance of approximately 60 cm. Subsequently, the experiment involved the same three consecutive stages (learning, distraction and test) repeated three times (due to the large number of stimuli presented).

In the learning phase, participants were instructed they would see a set of faces that they would have to recognise later. They were shown 34 (or 32, in one version) faces, selected at random from the overall sample of faces, sequentially. The faces appeared in the centre of the screen and presented in a random order. Using the keyboard, participants were required to rate each face on a scale of 1 to 9 on distinctiveness by answering the question "how easy would this face be to spot in a crowd?" with the anchor points "difficult" and "easy" (Light, Kayra-Stuart, & Hollander, 1979). This was done to ensure that the participants paid attention to the face and can provide an index of how participants individuate faces. Participants made their responses whilst the face was on screen. The face was on screen for 2 s. Between each face there was a random noise mask presented for 150 ms.

After this phase, participants completed the social experiences questionnaire. This was adapted from Walker and Hewstone’s (2006) scale that measured experience with the other-ethnicity, to create two additional versions measuring experience with the other-age and -gender (one was presented in each cycle). Each variant consisted of 13 items; such as, "In infancy I often spent time with my mother." Questions were completed on a 5-point Likert scale, ranging from "strongly agree" to "strongly disagree". The original scale had high internal reliability

(Cronbach's alpha= 0.83), construct and face validity (Walker & Hewstone, 2006). Participants also provided their age, gender, and ethnicity. These questions lasted roughly 2-3 minutes.

The test phase followed immediately after this. Participants were shown all faces that they had previously seen in the learning face in addition to the same number of new faces, selected at random from the overall pool of faces. These were presented one-at-a-time in the middle of the screen. Using the keyboard, participants had to press either "m" (if they recalled seeing the face in the learning stage) or "z" (if they did not recall seeing the face). Between each face, there was a random-noise mask presented centrally for 150 ms.

Following the test phase, the cycle was repeated twice for a different set of faces. At the end of the final test phase, participants were thanked and debriefed.

Results

We present the response time data separately to the distinctiveness rating data, the recognition accuracy, and response bias data. All data were subjected to a 2 x 2 x 2 within-subjects ANOVA with the factors ethnicity, age, and gender: these were coded as own- and other- for each variable. These data are presented in Table 1. This analysis allowed for an assessment of the additive effects of these biases. We ran further planned comparisons to directly test the hypothesis that the degree of out-groupness would be related to coding and recognition accuracy. We coded each face according to how many out-group features it had (0, 1, 2, or 3) and ran a one-way ANOVA on the accuracy data, shown in Figure 1. The full data set is available at University data repository service (BoRDaR).

Recognition Accuracy

Recognition responses were converted into the SDT measure of stimulus discriminability, *d'*, using the (Macmillan & Creelman, 2005) method. *d'* combines the hit rate (accurately recalled faces), the false alarm rate (recalled an inaccurate face) and ranges from 0 (chance recognition) to 3.92 (perfect recognition for the number of stimuli used in the present experiment).

This analysis revealed a significant own-ethnicity bias, $F(1, 159)=123.31$, $MSE=0.73$, $p<.001$, $\eta_p^2=.44$, a significant own-gender bias, $F(1, 159)=26.66$, $MSE=0.36$, $p<.001$, $\eta_p^2=.14$, and a significant own-age bias, $F(1, 159)=110.05$, $MSE=0.44$, $p<.001$, $\eta_p^2=.41$. The effect sizes for these interactions indicate that the own-ethnicity and own-age biases were larger than the own-gender bias.

The own-age bias interacted with the own-gender bias, $F(1, 159)=11.58$, $MSE=0.37$, $p=.001$, $\eta_p^2=.07$. This interaction was revealed through a larger effect of age when for own-gender faces $t(159)=9.80$, $p<.001$[2], Cohen's $d=0.77$, than for other-gender faces, $t(159)=5.58$, $p<.001$, Cohen's $d=0.44$. Similarly, the own-gender bias was larger for own-age faces, $t(159)=6.04$, $p<.001$, Cohen's $d=0.46$, than for other-gender faces, $t(159)=1.22$, $p=.225$, Cohen's $d=0.09$. Neither two-way interaction involving the factor ethnicity were significant: with age, $F(1, 159)=0.01$, $MSE=0.65$, $p=.909$, $\eta_p^2<.01$, and with gender, $F(1, 159)=0.03$, $MSE=0.43$, $p=.862$, $\eta_p^2<.01$.

[2] All post-hoc *t*-tests throughout this manuscript were Bonferroni- Šidák corrected for multiple comparisons.

Table 1.
*Mean (and standard error) recognition accuracy (*d'*), response bias (*C*), response time (*ms*), response time during learning (*ms*), and distinctiveness ratings.*

		Own-Age Faces		**Other-Age Faces**	
		Own-Gender Faces	Other-Gender Faces	Own-Gender Faces	Other-Gender Faces
Recognition Accuracy (*d'*)	Own-Ethnicity Faces	2.34 (0.06)	1.96 (0.06)	1.76 (0.07)	1.778 (0.07)
	Other-Ethnicity Faces	1.73 (0.07)	1.53 (0.06)	1.30 (0.07)	1.17 (0.06)
Response Bias (*C*)	Own-Ethnicity Faces	.18 (.04)	.18 (.04)	.13 (.04)	.17 (.04)
	Other-Ethnicity Faces	.19 (.04)	.15 (.05)	.24 (.04)	.18 (.04)
Recognition Response Time (*ms*)	Own-Ethnicity Faces	1573 (74)	1545 (68)	1467 (53)	1485 (57)
	Other-Ethnicity Faces	1553 (70)	1505 (76)	1490 (60)	1575 (74)
Learning Response Time (*ms*)	Own-Ethnicity Faces	3101 (226)	2826 (123)	2764 (152)	2917 (150)
	Other-Ethnicity Faces	2943 (147)	3055 (176)	3036 (259)	3076 (221)
Distinctiveness Ratings	Own-Ethnicity Faces	5.13 (0.11)	5.28 (0.10)	5.20 (0.11)	5.17 (0.11)
	Other-Ethnicity Faces	4.89 (0.11)	5.03 (0.11)	4.92 (0.12)	4.95 (0.11)

Finally, the three-way interaction was significant, $F(1, 159)=6.81$, $MSE=0.32$, $p=.010$, $\eta_p^2=.04$. A series of *t*-tests were conducted to explore this interaction, revealing that the magnitude of the own-ethnicity bias was consistent and significant across all other conditions with *t* values ranging from 5.06 to 7.94 (all *p*s<.001). Similarly, the own-age bias was significant across all other conditions (all *p*s<.016). However, the magnitude of the own-gender bias was not significant for other-age faces (*p*s>.05) but was for own-age faces (*p*s<.003).

Further planned comparisons were run to directly test the hypothesis that the degree of out-groupness would be related to recognition accuracy. We coded each face according to how many out-group features it had (0, 1, 2, or 3) and ran a one-way ANOVA of the accuracy data, shown in Figure 1. This analysis revealed a significant effect of number of out-group features, $F(2.23, 354.44$[3]$)=39.15$, $MSE=0.47$, $p<.001$, $\eta_p^2=.18$. Bonferoni-Šidák corrected pairwise comparisons revealed that face recognition accuracy was significantly different for all categories (all *p*s<.001) except when there were 2 or 3 out-group features ($p=.535$).

*Response Bias (*C*)*

A parallel analysis was run on response criterion data, measured using the Signal Detection Theory measure *C* (calculated using the Macmillan & Creelman, 2010, method). The own-ethnicity bias was not significant, $F(1, 159)=0.45$, $MSE=0.33$, $p=.504$, $\eta_p^2<.01$, nor was the own-gender bias, $F(1, 159)=0.43$, $MSE=0.14$, $p=.511$, $\eta_p^2<.01$, nor the own-age bias, $F(1, 159)=0.08$, $MSE=0.18$, $p=.785$, $\eta_p^2<.01$. No interactions were significant, largest $F(1, 159)=2.83$, smallest $p=.094$, largest $\eta_p^2=.02$.

[3] The Huynh-Feldt correction was applied to the degrees of freedom because Mauchley's test of sphericity was significant, $W(5)=.51$, $p<.001$, and the epsilon was .74.

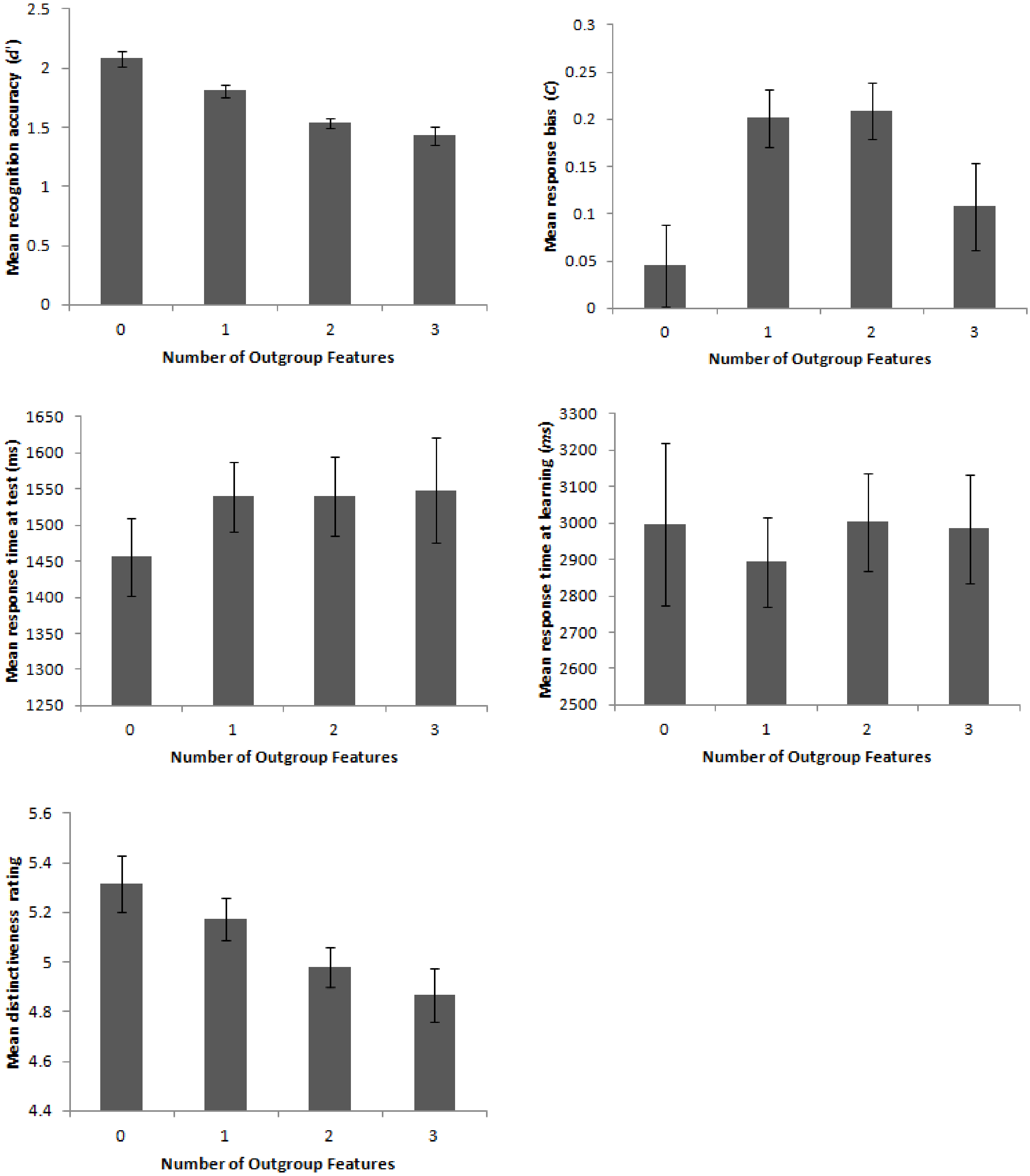

Figure 1. Mean face recognition accuracy (*d'*), response bias (*C*), response time at test (*ms*), response time at learning (*ms*), and distinctiveness ratings split by the number of out-group features. Error bars represent standard error of the mean.

Shown in Figure 1, response bias did depend on number of out-group features, $F(2.34, 372.41^{4})=6.93$, *MSE*=0.18, $p=.001$, $\eta_p^2=.04$. Pairwise comparisons revealed bias was lower (less of a tendency to respond with a "new" response) when there were no out-group features compared to when there were one, two (both *p*s=.001), or three ($p=.093$) out-group features. There were no significant differences in bias for the different values of out-group features (all *p*s>.22).

*Response Time (*ms*)*

Parallel analyses were run on the response time data. No significant effects were observed for response time at learning, nor at test, largest $F(1, 159)=1.57$, smallest $p=.212$, largest $\eta_p^2=.01$, Figure 1 highlights there was no effect of number of out-group features on response time.

Distinctiveness Ratings

Finally, we ran a parallel set of analyses on the distinctiveness ratings data as this might give an indication of depth of processing and individualisation processes being engaged in. Own-ethnicity faces were rated as more distinctive than other-ethnicity faces, $F(1, 159)=5.93$, *MSE*=3.30, $p=.016$, $\eta_p^2=.04$. The own-age bias, $F(1, 159)=0.20$, *MSE*=1.05, $p=.660$, $\eta_p^2<.01$, nor the own-gender bias, $F(1, 159)=1.15$, *MSE*=1.31, $p=.286$, $\eta_p^2<.01$, were significant. No interactions were significant in this analysis, largest $F(1, 159)=1.82$, smallest $p=.179$, $\eta_p^2=.01$.

---

[4] The Huynh-Feldt correction was applied to the degrees of freedom because Mauchley's test of sphericity was significant, $W(5)=.57$, $p<.001$, and the epsilon was .78.

Shown in Figure 1, we found that there was an effect of number of out-group features on face recognition accuracy, $F(2.27, 360.48^5)=7.11$, $MSE=1.18$, $p=.001$, $\eta_p^2=.04$. A trend analysis showed a significant linear trend, $F(1, 159)=11.78$, $MSE=0.67$, $p=.001$, $\eta_p^2=.07$, further evidence by significantly higher distinctiveness rating for faces with no out-group feature than those having 2 ($p=.024$) or 3 ($p=.009$) out-group features and higher distinctive ratings for faces with 1 out-group feature compared to those with 3 ($p=.028$) out-group features. No other pairwise comparisons were significant (*p*s>.079).

Discussion

We have found that our participants showed the own-ethnicity, own-age, and own-gender biases consistent with various previous studies that have examined biases (Katz & Kofkin, 1997; Bernstein, Young & Hugenberg, 2007; MacLin & Malpass, 2001; Lovén, et al., 2011; Wright & Sladen, 2003; Anastasi & Rhodes, 2005; Meissner & Brigham, 2001; Hills & Lewis, 2011). While we found that the magnitude of the each bias was larger than found in previous studies (Herlitz & Lovén, 2013; Meissner & Brigham, 2001; Rhodes & Anastasi, 2011), the magnitude of the own-age bias was significantly larger than expected. We found that the own-gender bias was smaller than the other two biases, consistent with the notion that it is typically stronger in women than in men (Lovén et al., 2011). In this study, we tested the biases the same group of participants, rather than comparing them across studies, so this study might better reflect that the relative differences in sizes of the biases at least for the population tested.

[5] The Huynh-Feldt correction was applied to the degrees of freedom because Mauchley's test of sphericity was significant, $W(5)=.60$, $p<.001$, and the epsilon was .76.

We did not find that the own-group biases consistently added together as revealed through our graded analysis. Faces that contained two or three out-group features were recognised to a similar degree, whereas faces with one out-group feature were recognised better than those with two or three and less well than those with no out-group features. Such results are inconsistent with a simple categorisation account of the own-group biases. This result is also not consistent with a simple additive account of the own-group biases. The main analysis reveal that the own-ethnicity bias is much more robust than the other biases and is relatively independent of the other biases: The magnitude of the own-age and own-gender biases differs depending on whether the faces were also own- or other-gender/age. This suggests that, in our participants, the own-ethnicity bias is relatively unique. In other words, one reason we found that there was little difference between having two or three out-group characteristics is due to the fact that the own-age and own-gender bias interact. These results can only be interpreted within a flexible framework of the own-group biases such as the categorisation-individualisation model (Hugenberg et al., 2010).

We analysed the distinctive ratings to see if participants were encoding faces of other-groups in a different manner to faces of their own-group. It must be noted that there were no own-group effects in our response time data, inconsistent with a number of previous studies (Meissner & Brigham, 2001). We surmise that the encoding process (i.e., the rating participants made) caused participants to respond in a consistent time for all faces. Nevertheless, the distinctiveness ratings indicate that participants find faces of their own-group more distinctive than faces of other groups. This is consistent with the notion that participants are more likely to individuate faces of their own group (Hugenberg et al., 2007). While there were some similar patterns in the distinctiveness rating data and the recognition accuracy data, the patterns were not identical. Specifically, there were no interactions between the biases in the distinctiveness data. In other words, as faces become more of an out-group they are individuated less. While this is not directly related to recognition accuracy, it suggests that participants find it harder to distinguish between faces that are more distant to themselves in terms of outgroup features.

The distinctiveness rating data is entirely consistent with Valentine's (1991) face-space model of face memory. In this model, faces are stored in a multidimensional space where each dimension of the space represents a physiognomic feature used to differentiate faces. Because the space develops as a result of the faces encountered during one's lifetime (Valentine & Endo, 1992; Hills & Lewis, 2018), the dimensions best distinguish between features of those faces. This creates a situation in which out-group faces are stored further from the centre of the space and clustered together because the dimensions are not appropriate. Faces that have more outgroup features will be even less well coded in the space than faces with fewer outgroup features. Valentine (1991) hypothesised that this would lead to faces being less well recognised. We have shown that there might be limits in the link of distinctiveness and recognition accuracy.

The own-ethnicity bias in the distinctiveness ratings data highlight that participants were not individuating faces of other-ethnicities. In other words, participants found that faces of the other-ethnicity were more similar to each other than faces of their own-ethnicity. This pattern was not observed for the other biases: No significant own-age nor own-gender bias was observed for the distinctiveness data. If we accept the notion that rating faces for distinctiveness is a metric for how faces are stored in face-space and reflect the amount that they are individuated, then this suggests the own-age and own-gender biases are not based on the same individuation mechanism as the own-ethnicity bias. We are presenting evidence that the three biases are not based on the same mechanism or that contextual factors not considered in face recognition studies cause them to be displayed differentially. This is consistent with the categorisation-individuation model (Hugenberg et al., 2010). In other words, there are reasons why a bias might be displayed by a participant, but the experimental context will alter its magnitude. Overall, however, the own-ethnicity bias is more likely to be displayed than the other biases and the magnitude of it is more robust across faces of own- and other-age and gender.

We have implied that the own-group biases tested here might be based on different mechanisms because of the different metrics we tested. Given this, we can indicate that the own-ethnicity bias is primarily based on individuation mechanisms. Such mechanisms supersede other mechanisms responsible for the biases since this bias was not affected by the presence of the other biases. The own-age and own-gender biases interact with each other but not with the own-ethnicity bias. The mechanisms for these biases are likely to be due to a motivation to process an in-group more deeply than an out-group (Hills et al., 2018; Man & Hills, 2016). In other words, we suggest that the own-ethnicity bias is based on individuation mechanisms whereas the own-age and -gender biases are based on motivation to encode faces deeply. While the overall pattern of results could be the result of the differential use holistic processing for the different groups of faces (Tanaka et al., 2007), it does not seem likely that such a simple mechanism could explain the complexity of the present results. While there are some additive effects of these biases (given that as faces become more out-group the recognition of them diminishes), this additive effect is limited: Faces with two or more out-group features are recognised equivalently probably because the own-ethnicity bias is based on a different mechanism. In other words, these results fit with the individuation-categorisation model of Hugenberg et al. (2010), but provides some concrete context for how and when these biases might be present, at least in this group of participants.

One strength of this study is that we have looked into three own-group biases combined. We found that the biases appear to be based on different mechanisms, given the subtle differences in the patterns of significance across different dependent variables. Nevertheless, future work should aim to explore the precise mechanisms of these biases in the same participants to better understand whether the biases are really based on similar mechanisms. Overall, the magnitude of the own-gender bias was smaller than the own-ethnicity and own-age biases. Furthermore, the own-ethnicity bias was more robust than the other biases. Further, we have shown that the biases appear to be based on different mechanisms.

Acknowledgements

The authors would like to thank Amy Crowe and Evie Mathis for some of the data collection for this work.

References

Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic bulletin & review*, *12*(6), 1043-1047.

Anastasi, J. S., & Rhodes, M. G. (2006). Evidence for an own-age bias in face recognition. *North American Journal of Psychology*, *8*(2).

Benjamin, J. (2013). *The bonds of love: Psychoanalysis, feminism, & the problem of domination*. Pantheon.

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The Cross-Category Effect: Mere Social Categorization Is Sufficient to Elicit an Own-Group Bias in Face Recognition. *Psychological Science (0956-7976*), 18(8), 706-712.

Blaine, B. E. (2007). *Understanding the psychology of Diversity*. Los Angeles: SAGE Publications Ltd.

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

Brigham, J. C., Maass, A., Snyder, L. D., & Spaulding, K. (1982). Accuracy of eyewitness identification in a field setting. *Journal of Personality and Social Psychology*, *42*(4), 673.

Bukach, C. M., Cottle, J., Ubiwa, J., & Miller, J. (2012). Individuation experience predicts other-race effects in holistic processing for both Caucasian and Black participants. *Cognition*, *123*(2), 319-324.

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

Crookes, K. & Rhodes, G. (2017). Poor recognition of other-race faces cannot always be explained by a lack of effort. *Visual Cognition,* 1-12.

Cross, J. F., Cross, J., & Daly, J. (1971). Sex, race, age, and beauty as factors in recognition of faces. *Perception & psychophysics*, *10*(6), 393-396.

Ellis, H., Shepherd, J., & Bruce, A. (1973). The effects of age and sex upon adolescents' recognition of faces. *The Journal of Genetic Psychology*, *123*(1), 173-174.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In *Advances in experimental social psychology* (Vol. 23, pp. 1-74). Academic Press.

Green, D. M., & Swets, J . A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Harrison, V., & Hole, G. J. (2009). Evidence for a contact-based explanation of the own-age bias in face recognition. Psychonomic Bulletin & Review, 16, 264 –269.

Herlitz, A., & Lovén, J. (2013). Sex differences and the own-gender bias in face recognition: a meta-analytic review. *Visual Cognition*, *21*(9-10), 1306-1336.

Hills, P. J. (2012). A developmental study of the own-age face recognition bias in children. *Developmental psychology*, *48*(2), 499.

Hills, P. J., & Lewis, M. B. (2011). Rapid communication: The own-age face recognition bias in children and adults. *Quarterly Journal of Experimental Psychology*, *64*(1), 17-23.

Hills, P. J. & Lewis, M. B. (2018). The development of face expertise: Statistical evidence for a qualitative change in processing. *Cognitive Development, 48,* 1-18*.*

Hills, P. J., & Pake, J. M. (2013). Eye-tracking the own-race bias in face recognition: Revealing the perceptual and socio-cognitive mechanisms. *Cognition*, *129*(3), 586-597.

Hills, P., Pake, J. M., Dempsey, J. R., & Lewis, M. B. (2018). Exploring the contribution of motivation and experience in the post-pubescent own-gender bias in face recognition. *Journal of Experimental Psychology: Human Perception and Performance*.

Hugenberg, K., Miller, J., & Claypool, H. M. (2007). Categorization and individuation in the cross-race recognition deficit: Toward a solution to an insidious problem. *Journal of Experimental Social Psychology*, *43*(2), 334-340.

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological review*, *117*(4), 1168.

Hugenberg, K., Wilson, J. P., See, P. E., & Young, S. G. (2013). Towards a synthetic model of own group biases in face memory. *Visual Cognition*, *21*(9-10), 1392-1417.

Katz, P. A., & Kofkin, J. A. (1997). Race, gender, and young children. *Developmental psychopathology: Perspectives on adjustment, risk, and disorder*, *21*, 51-74.

Lamont, A. C., Stewart-Williams, S., & Podd, J. (2005). Face recognition and aging: Effects of target age and memory load. *Memory & Cognition*, *33*(6), 1017-1024.

Lewin, C., & Herlitz, A. (2002). Sex differences in face recognition—Women's faces make the difference. *Brain and cognition*, *50*(1), 121-128.

Levin, D. T. (2000). Race as a visual feature: using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General*, *129*(4), 559.

Light, L. L., Kayra-Stuart, F., & Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(3), 212.

Lindholm, T. (2005). Own-age biases in verbal person memory. *Memory*, *13*(1), 21-30.

Lovén, J., Herlitz, A., & Rehnman, J. (2011). Women's own-gender bias in face recognition memory. *Experimental psychology*.

MacLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces: The ambiguous race face effect. *Psychology, Public Policy, and Law*, *7*(1), 98.

Macmillan, N. A., & Creelman, C. D. (2005). Detection theory: A user's guide. Cambridge, England: Cambridge University Press.

Man, T. W., & Hills, P. J. (2016). Eye-tracking the own-gender bias in face recognition: Other-gender faces are viewed differently to own-gender faces. *Visual Cognition*, *24*(9-10), 447-458.

McKelvie, S. J. (1987). Sex differences, lateral reversal, and pose as factors in recognition memory for photographs of faces. *The Journal of general psychology*, *114*(1), 13-37.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, *7*(1), 3.

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

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 630-633.

Perfect, T. J., & Harris, L. J. (2003). Adult age differences in unconscious transference: Source confusion or identity blending?. *Memory & cognition*, *31*(4), 570-580.

Rehnman, J., & Herlitz, A. (2006). Higher face recognition ability in girls: Magnified by own-sex and own-ethnicity bias. *Memory*, *14*(3), 289-296.

Rehnman, J., & Herlitz, A. (2007). Women remember more faces than men do. *Acta psychologica*, *124*(3), 344-355.

Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: a meta-analytic and theoretical review. *Psychological bulletin*, *138*(1), 146.

Rodin, M. J. (1987). Who is memorable to whom: A study of cognitive disregard. *Social Cognition*, *5*(2), 144-165.

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law*, *7*(1), 36.

Stelter, M., Rommel, M., & Degner, J. (submitted). (Eye) Tracking the Other-Race Effect: Comparison of Eye Movements during Encoding and Recognition of Proximal and Distal Outgroup Faces. *Social Cognition*.

Walker, P. M., & Hewstone, M. (2006). A perceptual discrimination investigation of the own-race effect and intergroup experience. *Applied Cognitive Psychology*, *20*(4), 461-475.

Wiese, H. (2012). The role of age and ethnic group in face recognition memory: ERP evidence from a combined own-age and own-race bias study. *Biological psychology*, *89*(1), 137-147.

Wright, D. B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta psychologica, 114*(1), 101-114. Doi:10.1016/S0001-6918(03)00052-0

Wright, D. B., & Stroud, J. N. (2002). Age differences in lineup identification accuracy: People are better with their own age. *Law and human behavior*, *26*(6), 641-654.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology Section A*, *43*(2), 161-204.

Valentine, T., & Endo, M. (1992). Towards an exemplar model of face processing: The effects of race and distinctiveness. *The Quarterly Journal of Experimental Psychology*, *44*(4), 671-703.

Valentine, T., Lewis, M. B., & Hills, P. J. (2016). Face-space: A unifying concept in face recognition research. *The Quarterly Journal of Experimental Psychology*, *69*(10), 1996-2019.

Young, S. G., Bernstein, M. J., & Hugenberg, K. (2010). When do own-group biases in face recognition occur? Encoding versus post-encoding. *Social Cognition*, *28*(2), 240-250.

Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and motivation in face recognition: A critical review of theories of the cross-race effect. *Personality and Social Psychology Review*, *16*(2), 116-142.