Social Cognition, Vol. 39, No. 3, 2021, pp. 366–395

(EYE-) TRACKING THE OTHER-RACE EFFECT: COMPARISON OF EYE MOVEMENTS DURING ENCODING AND RECOGNITION OF INGROUP FACES WITH PROXIMAL AND DISTANT OUTGROUP FACES

Marleen Stelter, Marc Rommel, and Juliane Degner
Hamburg University

People experience difficulties recognizing faces of ethnic outgroups, known as the other-race effect. The present eye-tracking study investigates if this effect is related to differences in visual attention to ingroup and outgroup faces. We measured gaze fixations to specific facial features and overall eye-movement activity level during an old/new recognition task comparing ingroup faces with proximal and distal ethnic outgroup faces. Recognition was best for ingroup faces and decreased gradually for proximal and distal outgroup faces. Participants attended more to the eyes of ingroup faces than outgroup faces, but this effect was unrelated to recognition performance. Ingroup-outgroup differences in eye-movement activity level did not emerge during the study phase, but during the recognition phase, with ingroup-outgroup differences varying as a function of recognition accuracy and old/new effects. Overall, ingroup-outgroup effects on recognition performance and eye movements were more pronounced for recognition of new items, emphasizing the role of retrieval processes.

Keywords: face recognition, other-race effect, eye movements

People are better at recognizing faces from their racial or ethnic ingroup compared to faces from outgroups—an effect known as the other-race effect (ORE[1]).

1. This effect is sometimes also referred to as "other-ethnicity effect" or "own group bias." We decided to use the term "other-race effect" to relate to the majority of previous publications. Considering critical views regarding the use of the term "race" (see Blum, 2010), we use the broader terms "ingroup" and "outgroup" (rather than "own-race" and "other-race") to refer to faces of the respective racial(ized) groups. By using these terms, we do not mean to imply that reported effects generalize to outgroups in other social domains (see Hugenberg, Wilson, See, & Young, 2013).

Supplementary materials are available online at https://osf.io/9x2qs/.

We thank Stella Rebholz for assistance with image editing.

Correspondence concerning this article should be addressed to Marleen Stelter, Department of Social Psychology, Universität Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany, E-mail: marleen.stelter@uni-hamburg.de.

This effect is one of the most robust findings in the face perception literature and has been replicated across different cultural settings (see Meissner & Brigham, 2001, for a meta-analysis). Theoretical accounts focus broadly on two origins of the ORE: perceptual experience and social-cognitive factors (see Young, Hugenberg, Bernstein, & Sacco, 2012, for a comprehensive review of theories). Theories of perceptual expertise propose that the ORE occurs due to reduced exposure with infrequently encountered racial or ethnic outgroup faces in daily life, leading to *less efficient processing* of outgroup faces (e.g., Hancock & Rhodes, 2008; Rhodes, Hayward, & Winkler, 2006; Valentine, 1991). Social-cognitive theories propose that, based on social categorization processes, people are not as motivated to individuate outgroup faces as they are to individuate ingroup faces, leading to *less effortful processing* of outgroup faces (e.g., Hugenberg, Young, Bernstein, & Sacco, 2010; Levin, 1996, 2000; Rodin, 1987; Sporer, 2001). Albeit researchers tend to agree that the ORE is most likely caused by an interplay of both, visual expertise and motivation to individuate (e.g., Hugenberg et al., 2013; Hugenberg et al., 2010; Sporer, 2001; Wan, Crookes, Reynolds, Irons, & McKone, 2015), there is an ongoing debate about the specific cognitive processes underlying the differential effort and efficiency when processing ingroup and outgroup faces (e.g., Hayward, Crookes, & Rhodes, 2013; Levin, 2000; Meissner, Brigham & Butz, 2005; Sporer, 2001; Valentine, 1991; Walker & Tanaka, 2003). The current research examines eye movements as indicators of visual attention in order to investigate if differences in visual attention to ingroup and outgroup faces contribute to the emergence of the ORE.

EYE-TRACKING THE OTHER-RACE EFFECT

A recently growing number of studies have employed eye-tracking methodology to investigate if and how eye movements differ for ingroup versus outgroup faces, in order to better understand the underlying mechanism of the ORE (see also Kawakami, Friesen, & Vingilis-Jaremko, 2018, for a recent review). Eye-tracking offers a straightforward approach to investigating processing differences of ingroup and outgroup faces, using fixations (i.e., states when the eyes remain fixated over a period of time) to derive a person's current focus of overt visual attention (Deubel & Schneider, 1996; Hoffman & Subramaniam, 1995). It has often been reported that people follow a triangular scanning pattern when looking at faces (e.g., Yarbus, 1967, but see Blais, Jack, Scheepers, Fiset, & Caldara, 2008, on cultural differences). Specifically, most fixations are directed to the eyes (Janik, Wellens, Goldberg, & Dell'Osso, 1978; Laidlaw, Risko, & Kingstone, 2012), second most to the mouth, and only few fixations to other features, such as to the nose, ears, or hair (Walkersmith, Gale, & Findlay, 1977). But do people use the same pattern of eye movements for ingroup and outgroup faces? And is overt visual attention generally related to how well ingroup and outgroup faces are recognized? There are several suggestions as to how eye movements may be linked to the ORE in recognition memory.

Based on perceptual-expertise theories of the ORE, one idea is that when looking at outgroup faces, observers fail to attend to those physiognomic features of a face that are most helpful in distinguishing between individuals of a racial or ethnic outgroup (*diagnostic features account*; e.g., Hills, Cooper, & Pake, 2013b; Hills & Pake, 2013; Wang et al., 2015). For example, it has been suggested that attention to the eyes may be more beneficial for recognizing White faces; conversely, attending to the lower part of a face rather than the eyes may be more beneficial for recognizing Black faces (e.g., Ellis, Deregowski, & Shepherd, 1975; Hills et al., 2013b; Hills & Pake, 2013) and Asian faces (e.g., Wang et al., 2015). In line with this idea, directing participants' eye movements toward diagnostic features (e.g., instructing White participants to attend to the lower part of Black faces) has been found to improve outgroup face recognition (Hills & Pake, 2013; but see Wittwer, Tredoux, Py, & Paubel, 2019). The diagnostic features hypothesis, thus, suggests that the ORE occurs because observers use the *same* viewing strategy for ingroup and outgroup faces, leading them to miss diagnostic information of outgroup faces.

In contrast, the social-cognitive account of the ORE—that ingroup faces are individuated and outgroup faces are processed on a category level—proposes that observers use *different* viewing strategies when looking at ingroup versus outgroup faces (e.g., Hugenberg et al., 2013; Kawakami et al., 2014; Levin, 1996, 2000). Specifically, it has been suggested that individuated processing leads to more attention to the eyes; therefore, ingroup faces are assumed to receive more fixations to the eyes as compared to outgroup faces (*individuation vs. categorization account*; Kawakami et al., 2014). This aligns with previous research showing that the eye region is of general importance for face recognition (Caldara, Zhou, & Miellet, 2010; Malcolm, Lanyon, Fugard, & Barton, 2008; McKelvie, 1976; Peterson & Eckstein, 2012). It is therefore feasible that reduced attention to the eyes of outgroup faces contributes to impaired recognition of outgroup faces. Confirming this idea, some studies have found more fixations to the eyes of ingroup faces compared to outgroup faces, and this effect has been found for faces of the own racial or ethnic group as well as own-age faces, own-sex faces, and even own minimal-group faces (e.g., Brielmann, Bülthoff, & Armann, 2014; Goldinger, He, & Papesh, 2009; Hansen, Rakhshan, Ho, & Pannasch, 2015; Kawakami et al., 2014; Man & Hills, 2016; Proietti et al., 2015; Wu, Laeng, & Magnussen, 2012). However, several other studies did not find increased attention to eyes of racial ingroup faces (Blais et al., 2008; Caldara et al., 2010; Hills, Cooper & Pake, 2013a; McDonnell et al., 2014; Nakabayashi, Lloyd-Jones, Butcher, & Liu, 2012). Further, only one study reported that differences in attention to the eyes were related to the ORE in recognition memory (Kawakami et al., 2014; but see Correll & Hudson, 2020). This inconsistency in findings underlines a need for further research.

In addition, yet another theoretical approach predicts that eye movements differ for ingroup and outgroup faces because ingroup faces are processed more holistically than outgroup faces (*holistic processing account*; e.g., Blais et al., 2008; Kelly et al., 2011; Wang et al., 2015). Differences in holistic processing of ingroup versus outgroup faces have indeed been discussed to contribute to the ORE (e.g., Degutis, Mercado, Wilmer, & Rosenblatt, 2013; Michel, Caldara, & Rossion, 2006; Tanaka & Gordon, 2011; but see Horry, Cheong, & Brewer, 2015; Mondloch et al., 2010; Zhao, Hayward, & Bülthoff, 2014). The holistic-processing account predicts that observers attend more to the nose of ingroup faces compared to outgroup faces, based on the idea that attention to the nose—being the center of a face—may promote holistic processing (e.g., Blais et al., 2008; Kelly et al., 2011; Wang et al., 2015). It is thus feasible that reduced attention to the nose of outgroup faces may reduce holistic processing of outgroup faces, which in turn impairs outgroup face recognition.

In summary, several theoretical accounts make inconsistent predictions, suggesting that eye movements may be directed to the same or to different facial features for ingroup versus outgroup faces. In addition to these accounts, the cultural background of the observer may additionally influence preference for eye movements to specific facial features (*culture-specific processing account;* e.g., Blais et al., 2008; Fu, Hu, Wang, Quinn, & Lee, 2012; Kelly et al., 2011; Miellet, Vizioli, He, Zhou, & Caldara, 2013; Tan, Stephen, Whitehead, & Sheppard, 2012). It has been found that White observers fixate more on the eye region, while Asian observers fixate more on the nose region (e.g., Blais et al., 2008); however, other studies did not observe any cultural differences (Caldara et al., 2010; Or, Peterson, & Eckstein, 2015).

Finally, it has been suggested that eye movements to ingroup and outgroup faces may also differ in terms of overall eye-movement activity level. Specifically, it has been observed that ingroup faces are viewed with a more active scanning behavior (i.e., overall more eye movements) than outgroup faces (*more active scanning account*, Goldinger et al., 2009; Klama & Milton, 2012; Vargas, 2017; Wu et al., 2012; but see Hu, Wang, Fu, Quinn, & Lee, 2014; Nakabayashi et al., 2012). Observers displayed shorter fixations, as well as more saccades (i.e., rapid motions from one fixation to the other), and longer saccadic amplitudes (i.e., distance travelled by the eye between two fixations) when studying ingroup faces compared to outgroup faces (Goldinger et al., 2009; Klama & Milton, 2012; Wu et al., 2012; but see Nakabayashi et al., 2012). This has been interpreted as more rapid and efficient sampling of information from ingroup faces compared to outgroup faces (Wu et al., 2012), but more active scanning may also reflect higher motivation to encode ingroup faces (cf. Kaspar & König, 2011). The more-active-scanning approach is in line with research showing that eye movement activity during encoding is beneficial for face learning (Heisz, Pottruff, & Falk, 2013, Henderson, Williams, & Falk, 2005). Thus, differences in eye-movement activity during encoding of ingroup and outgroup faces may also contribute to the ORE in recognition memory.

Against the background of these highly diverse and sometimes contradictory empirical findings and explanations, interpretation and generalizability of results from the wealth of previous ORE eye-tracking studies is challenging. For example, while some studies found more attention to the eyes of ingroup faces compared to outgroup faces (e.g., Brielmann et al., 2014; Goldinger et al., 2009; Hansen et al., 2015; Kawakami et al., 2014; Wu et al., 2012), others have found more attention to the nose of ingroup faces (e.g., Fu et al., 2012; Hu et al., 2014; Kelly et al., 2011), and yet other studies have found no such differences (e.g., Caldara et al., 2010; Hills & Pake, 2013; McDonnell, Bornsteing, Laub, Mills, & Dodd, 2014; Nakabayashi et al., 2012). In addition, previous research is also inconsistent in regard to the relationship between eye-movement patterns to ingroup versus outgroup faces and the ORE in recognition memory. For example, while one study with White participants found that increased attention to the eyes of White ingroup faces was correlated with increased ingroup recognition performance (Kawakami et al., 2014; but see Correll & Hudson, 2020), another study with White participants found that attention to the hair, but not the eyes, was related to increased ingroup recognition performance (McDonnell et al., 2014), and a third study with Asian participants found that attention to the nose of Asian ingroup faces was related to faster ingroup face recognition (Wang et al., 2015). A fourth study found no relation between eye movements and recognition performance (e.g., Fu et al., 2012), and many other studies do not report the relation between eye movements and recognition of ingroup versus outgroup faces (e.g., Arizpe, Kravitz, Walsh, Yovel, & Baker, 2016; Blais et al., 2008; Hu et al., 2014; Klama & Milton, 2012; Nakabayashi et al., 2012). This high level of inconsistency may—at least partly—depend on differences in experimental procedures and data analysis technique (see Arizpe et al., 2016). To conclude, despite a large amount of research, inconsistent use of methodology and inconsistent results limit the interpretability and generalizability of previous research. In the sense of accumulative science, further research is needed to disentangle effects of the different underlying mechanisms.

THE PRESENT RESEARCH

The present study aimed at weighing in on this research domain by comparing eye movements and recognition for faces of four different ethnic groups. Specifically, we tested White German participants and contrasted effects for White ingroup faces with faces from a relative *proximal* outgroup (Middle Eastern faces[2]) and two relative *distal* outgroups (Black and Asian faces). We defined the level of outgroup proximity based on the level of social and societal status and relevance, contact, and exposure during daily life activities.[3] We suggest that comparing eye movements and recognition across several different outgroups may contribute to the previous literature in several important ways.

2. We would like to acknowledge that people from the geographic region of the Middle East belong to various different ethnic groups. By using the broad category "Middle Eastern," we follow the suggestion of the Arab-American Institute, who use the term MENA (Middle Eastern and North African) as a combined category.

3. Whereas individuals with Middle Eastern descent constitute a relatively large minority group (about 15% of the population in the area where participants were recruited and 11% in the general German population), individuals of Asian or African descent constitute comparably small minorities (each about 3% of the local population and each less than 1.5% of the general German population; Statistisches Bundesamt. (2017). Bevölkerung und Erwerbstätigkeit. Bevölkerung mit Migrationshintergrund: Ergebnisse des Mikrozensus 2016. https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/MigrationIntegration/Migrationshintergrund2010220167004.pdf?__blob=publicationFile). Therefore, most White German participants have more exposure and a higher probability of interaction during everyday life with people of Middle Eastern descent. This is also reflected in results of the contact measure used in this study.

First, recognition accuracy may vary depending on the type of outgroup: The relative recognition disadvantage may be smaller for the proximal outgroups than for distal outgroups (see Sporer, 2001). Such a gradual effect in recognition memory may provide a valuable basis for interpretation of effects in eye movements. Second, including several outgroups enables us to separate general ingroup/outgroup effects from outgroup-specific effects in eye movements. For example, if differences in eye movements result from (mere) ingroup/outgroup categorization, we should find the same pattern of eye movements for all outgroups. Third, the comparison of relative proximal and distal outgroups is informative for the role of visual exposure or relative expertise with processing outgroup faces, an account which is frequently discussed as underlying the ORE (e.g., Hancock & Rhodes, 2008). Particularly, following the visual expertise account, ingroup/outgroup effects on eye movements should be more pronounced for distal outgroups compared to proximal outgroups. Fourth, ethnicity-specific physiognomic characteristics of each outgroup may additionally influence eye movements via bottom-up processes (i.e., sizes, shapes, or perceptual contrasts of facial features that systematically differ between groups; see Wang et al., 2015). These perceptual differences may lead to additional differences in eye movements to each outgroup.

Similar to previous research, we measured recognition memory via an old/new recognition paradigm. In this paradigm, participants first view a sequence of faces and are instructed to memorize them (i.e., study phase). Subsequently, they are presented with the same faces intermixed with novel faces (i.e., recognition phase), in which they classify each face as "old" or "new". We tracked participants eye movements during both study and recognition phase, allowing us to compare effects for initial encoding and memory retrieval. Our analyses of eye movements focus on eye movements to specific facial features (i.e., number of fixations and visit durations for eyes, nose, and mouth), as well as on overall level of eye-movement activity (i.e., number and durations of fixations, number of saccades, and saccadic amplitudes). As a first step, we compared recognition accuracy for faces of the different ethnicities. As a second step, we compared measures of eye movements for faces of the different ethnicities. As a final step, we investigated if eye movements were related to recognition accuracy and may thus have contributed to the ORE in face recognition.

The variety of previous findings and theoretical accounts allows several (opposing) predictions regarding ingroup/outgroup effects on eye movements and their relation to recognition memory: Following the individuation-categorization account, reduced attention to the eyes of outgroup faces should be related to *reduced* recognition of outgroup faces; alternatively, following the diagnostic-features approach, reduced attention to the eyes (and increased attention to nose and mouth) of outgroup faces should be related to *increased* recognition of (non-White) outgroup faces. Additionally, based on the more-active-scanning account, reduced eye-movement activity for outgroup faces should be related to reduced recognition of outgroup faces as compared to ingroup faces (e.g., Wu et al., 2012). However, given the inconsistent use of methods as well as contradictory findings in the literature, these predictions may not directly apply to the specific context and methods of our study. For example, it is unclear which specific eye-tracking measures should be used to test these assumptions, as there are numerous ways to operationalize viewing behavior (e.g., previous research is highly inconsistent in defining what constitutes the eye region). Also, some previous work used a different procedure for the recognition memory test (e.g., presenting more than one face at a time; Kawakami et al., 2014), making it difficult to generalize previous evidence. Based on these considerations, we decided against specifying explicit a priori hypotheses, emphasizing the exploratory nature of this research.

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. All data files and analysis scripts are publicly accessible via the Open Science Framework (https://osf.io/27tpc/).

METHOD

PARTICIPANTS

The final sample included 62 students of various faculties of Hamburg University (14 male, 47 female, 1 not reported; Age M = 25.18, SD = 6.76, range 18 to 52). All participants reported either normal or corrected to normal vision (13 wore glasses, 8 wore contact lenses). All participants self-identified as White German. We excluded data of seven further participants from analyses: One self-identified as Asian, two reported having lived in East Asian countries, two had insufficient quality of eye-tracking data (49% and 76% data loss), and two had performances far below chance level in the recognition task (average performances of 20% and 27% correct responses).[4] Participants were reimbursed with course credit or received compensation of 6 Euro.

MATERIALS AND APPARATUS

We used 64 headshot photographs of young male adults of four different ethnicities: 16 White faces (two from Langner et al., 2010; six from Paulus et al., 2012; eight from a Google Image search), 16 Middle Eastern faces (seven from Paulus et al., 2012; nine from a Google Image search), 16 Black faces (five from Minear & Park, 2004; 11 from a Google Image search), and 16 East Asian (all from a Google Image search). All individuals displayed neutral facial expressions, had no distinctive features (e.g., glasses, jewelry), and were shown in full frontal view (see Figure 2). Images were standardized in size (375 x 450 pixels, 12.4° x 15.7° visual angle), background color (light grey), vertical position of the eyes, and black-colored clothing. For each participant, one of four fixed sets of 32 faces (eight of each ethnicity) was presented during the study phase, and all 64 faces were presented during the recognition test. The order of stimulus presentation was fixed within each of the four stimulus sets.

4. The pattern of results remains the same when all participants are included.

We presented stimuli on the monitor of a Tobii T120 Eye Tracker with a resolution of 1280 x 1024 pixels. This eye-tracking system has a maximum accuracy of 0.5°, a maximum precision of 0.2°, and records eye movements at 120 Hz. Gaze data was collected with binocular bright and dark pupil tracking. Stimulus presentation of the recognition task was controlled by Tobii Studio eye tracking software. Subsequent stimulus ratings and questionnaires were collected using the experimental software Inquisit (2014).

DESIGN AND PROCEDURE

Participants were tested in individual sessions by a White male experimenter. After providing initial consent, participants were informed that the study consisted of three parts: a face learning task, a recognition task, and subsequent stimulus ratings and questionnaires. Also, participants were informed that their eye movements would be recorded during all tasks. Participants were seated on a chair with adjustable height in order to ensure optimal positioning for recording of eye movements. Average viewing distance to the monitor was approximately 57 cm ($SD = 3.7$). Before the experiment, participants completed a nine-point calibration procedure, allowing the eye tracker to map participants' actual eye gaze positions onto landmark positions of the computer monitor. After calibration, participants completed the recognition task followed by additional measures. At the end, participants gave informed consent to allow inclusion of their data.

Recognition Task. The recognition task employed a repeated-measures design with face ethnicity (White, Middle Eastern, Black, East Asian) as the within-participants factor. During the study phase, participants were asked to memorize 32 faces presented successively in a fixed pseudo-randomized sequence. Each face was presented for 5 seconds on the center of the screen, followed by a 700 ms blank screen inter-stimulus interval. In the immediately following recognition phase, we presented all 64 faces in a fixed pseudo-random sequence in the center of the screen. Participants' task was to decide whether a face had been presented during the learning task or not. They responded by pressing the F1 key if they recognized the face as "old" and the F12 key if the face was "new."

Additional Measures. After the recognition task, participants completed stimulus ratings of attractiveness, likability, distinctiveness, and prototypicality in individually randomized order (see Supplement, Table S1). Subsequently, they completed German adaptions of the Intergroup Anxiety Scale (Britt, Bonieci, Vescio, Biernat, & Brown, 1996) and the Motivation to Control Prejudiced Reactions Scale (Plant & Devine, 1998)—see Supplement for detailed descriptions and results. Afterwards, participants completed a contact measure, in which they estimated the percentage of members of the three outgroups (termed "people of Middle Eastern descent," "people of African descent," and "people of East Asian descent") among their close friends and family as well as among their broad circle of acquaintances. We calculated contact scores by computing the mean of both items. Finally, participants completed a demographics questionnaire.

RESULTS

CONTACT MEASURE

Results showed that participants reported significantly more contact with people of Middle Eastern descent ($M = 10.10\%$, $SD = 10.13$) compared to people of African descent ($M = 6.78\%$, $SD = 10.17$), $t(59) = 2.556$, $p = .013$, $d_z = 0.33$, 95% CI = [0.07; 0.59] and people of East Asian descent ($M = 5.94\%$, $SD = 9.12$), t(59) = 3.129, $p = .003$, $d_z = 0.40$, 95% CI = [0.14; 0.67]. This differential pattern of contact confirms our proximal versus distal outgroup distinction.[5]

RECOGNITION

To analyze recognition data, we computed signal detection discrimination parameter d' and response bias c (see Green & Swets, 1966; Snodgrass & Corwin, 1988) for each group of faces. Both measures are based on hits (i.e., correctly identifying an old face as being "old") and false alarms (i.e., incorrectly classifying a new face as being "old").[6] Additionally, we conducted separate analyses of hits and false alarms to test for a so-called mirror effect (i.e., more hits and less false alarms for ingroup compared to outgroup faces; see Meissner & Brigham, 2001). Descriptive statistics of recognition measures are illustrated in Figure 1.

Discrimination parameter d' indicates the relative difficulty of separating old from new targets from each group, and it is standardized to range from 0 to approximately 4, with higher values representing better recognition performance. All d' values were significantly larger than zero, all ts(61) ≥ 17.703, ps < .001, ds ≥ 2.25, indicating that participants recognized faces of all groups above chance level. Response bias c indicates participants' general tendency to responding either "old" or "new," and it centers around zero: An exact value of zero indicates that there is no response bias; a positive value of c indicates a tendency to responding "new," thus taking the risk of missing old items; a negative value of c indicates a tendency to responding "old," thus, taking the risk of committing false alarms. For White and Middle Eastern faces, c values were larger than zero, $t(61) = 4.989$, $p < .001$, $d = 0.63$, 95% CI [0.36; 0.90] and $t(61) = 2.938$, $p = .005$, $d = 0.37$, 95% CI [0.11; 0.63], respectively, indicating a tendency for responding "new" for both groups and, thus, trying to prevent false alarms but risking misses. Conversely, c

5. Due to a programming error, only 60 participants completed the contact items. Participants' contact scores were not related to recognition performance for the respective groups, all rs(58) ≤ .10, all ps ≥ 0.435. Note that the sample size of $n = 60$ only provided sufficient test power for detecting relatively big effect sizes ($r > .35$).

6. In order to avoid infinite z-scores, false alarm rates of zero were replaced with 0.5/n and hit rates = 1 were replaced with (n-0.5)/n (Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. Psychological Bulletin, 98(1), 185–199. https://doi.org/10.1037/0033-2909.98.1.185).

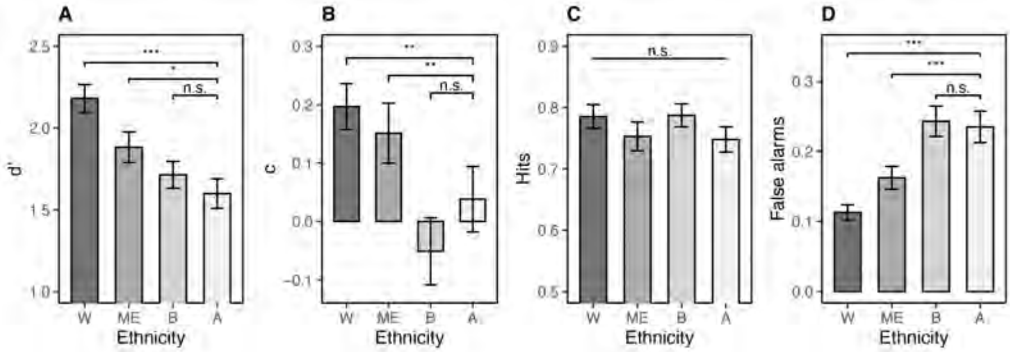

FIGURE 1. Recognition performance *d′* (A), response bias c (B), hit rates (C), and false alarm rates (D) for the different ethnicities. W = White; ME = Middle Eastern; B = Black; A = Asian.

values for Black and Asian faces did not differ from zero, $t(61) = -0.892$, $p = .376$, $d = -0.11$, 95% CI [–0.35; 0.00], and $t(61) = 0.679$, $p = .500$, $d = 0.09$, 95% CI [–0.34; 0.16], indicating that there was no response bias for these groups.

To analyze ethnicity effects on *d′*, *c*, hits, and false alarms, we conducted four univariate repeated-measures ANOVAs, with Face Ethnicity (White, Middle Eastern, Black and Asian) as within-participants factor. Given the sample size of $n = 62$ participants, we obtained sufficient test power ($1–\beta = 0.80$) to test for a small-to-medium effect ($f = 0.18$) on a significance level of $\alpha = .05$, with an observed average correlation of $r = 0.27$ between measures (Faul, Erdfelder, Lang, & Buchner, 2007). Significant main effects were followed up by a fixed set of Helmert contrasts (see Table 1): In a first contrast, we tested White ingroup faces against all outgroup faces (Middle Eastern, Black, and Asian). This contrast tests for a general ingroup versus outgroup difference. In a second contrast, we tested Middle Eastern faces against Black and Asian faces combined. This contrast tests for an effect of proximal versus distal outgroups. In a third contrast, we tested Black versus Asian faces, exploring differences between the two distal outgroups. Results of the ANOVAs and contrast effects are reported in Table 2.

In summary, we observed significant ethnicity effects on *d′*, *c*, and false alarms. Specifically, the first contrast indicated general ingroup-vs.-outgroup effects: Better recognition, a response tendency to responding “new,” and fewer false alarms for ingroup faces compared to all outgroup faces (see Table 2). Further, the second contrast revealed similar pattern of proximal-vs.-distal outgroup effects: better recognition, a response tendency to responding “new,” and fewer false alarms for proximal outgroup faces compared to distal outgroup faces. The third contrast indicated that there were no differences between the two distal outgroups on any of the measures. Finally, we did not observe any ethnicity effects on hit rates. Thus, unlike previous research, we did not observe a mirror effect (see Meissner & Brigham, 2001). This finding implies that the ethnicity effect in *d′* was mainly driven by responses to new faces (i.e., erroneously recognizing new outgroup

TABLE 1. Contrasts Coding for a General Ingroup vs. Outgroup Effect (1. Contrast), for a Proximal vs. Distal Outgroup Effect (2. Contrast), and for an Effect Between Distal Outgroups (3. Contrast)

Ethnicity	1. Contrast	2. Contrast	3. Contrast
White	3	0	0
Middle Eastern	–1	2	0
Black	–1	–1	1
Asian	–1	–1	–1

faces as old more often than new ingroup faces), and not by responses to old faces (i.e., old faces were equally well recognized across all ethnicities).

EYE MOVEMENTS

Data Handling and Analytic Approach. Eye movements were classified into fixations and saccades based on default settings of the Tobii I-VT fixation filter.[7] Fixations were defined based on eye movements slower than 3° visual angle per 100 ms. If subsequent fixations were less than 0.5° visual angle and less than 75 ms apart, they were merged into a single fixation. We excluded the first fixation of each trial from analyses, given that all images were presented in the center of the screen (following the recommendation from Caldara & Miellet, 2011). Periods of more than 75 ms data loss (i.e., during which the pupil and corneal reflections were not reliably detected, for example, due to eye blinks) were considered to be missing data. Trials with more than 30% missing data were excluded from analyses; this applied to 0.9% of all trials.

Analyses of eye movements focused on the following measures: First, we explored whether ethnicity influenced eye movements to facial features. Specifically, we analyzed visit durations (i.e., overall time spend fixating a feature) and number of fixations for each facial feature. Second, we explored whether ethnicity influenced overall eye-movement activity level. For this, we analyzed four different eye-tracking measures: number of fixations, average fixation duration, number of saccades, and saccadic amplitudes (i.e., saccade length as measured in degrees of visual angle).

For analysis of eye movements to facial features, we pre-defined six non-overlapping areas of interest (AOIs, see Wu et al., 2012), including eyes, nose, mouth, chin, forehead, and cheeks (left and right cheeks were combined to a single AOI). We manually adjusted the size of each AOI to fit the individual features of each face—see Figure 2 for examples. As a consequence, AOIs sizes vary systematically across faces of different ethnicities (see Supplement, Table S2). For this reason, we

7. See also https://www.tobiipro.com/siteassets/tobii-pro/learn-and-support/analyze/how-do-we-classify-eye-movements/tobii-pro-i-vt-fixation-filter.pdf.

TABLE 2. Results of the Univariate Repeated-Measures ANOVAs for Recognition Performance *d′*, Response Bias *c*, Hit Rates, False Alarm Rates, and Reaction Times

Comparison	*F*	df_1	df_2	*p*	η^2	90% *CI*
d′						
Univariate ANOVA	12.309	3	183	< .001	.17	[.08; .24]
1. Contrast (W – ME, B, A)	30.507	1	61	< .001	.33	[.18; .46]
2. Contrast (ME – B, A)	5.201	1	61	.026	.08	[.01; .20]
3. Contrast (B – A)	1.599	1	61	.211	.03	[.00; .12]
c						
Univariate ANOVA	5.879	3	183	.001	.09	[.02; .15]
1. Contrast (W – ME, B, A)	11.546	1	61	.001	.16	[.04; .29]
2. Contrast (ME – B, A)	8.032	1	61	.006	.12	[.02; .24]
3. Contrast (B – A)	1.395	1	61	.242	.02	[.00; .11]
Hit rates						
Univariate ANOVA	1.269	3	183	.286	.02	[.00; .05]
False alarm rates						
Univariate ANOVA	16.172	3	183	< .001	.21	[.12; .28]
1. Contrast (W – ME, B, A)	49.497	1	61	< .001	.45	[.29; .56]
2. Contrast (ME – B, A)	16.061	1	61	< .001	.21	[.07; .34]
3. Contrast (B – A)	0.103	1	61	.749	.00	[.00; .05]

Note. W = White (ingroup); ME = Middle Eastern (proximal outgroup); B = Black (distant outgroup); A = Asian (distant outgroup).

normalized values of visit duration and number of fixations, accounting for differences in AOI size (see also Hills & Pake, 2013; Laidlaw et al., 2012).[8] Normalized values were computed by dividing proportions of visit durations and proportions of numbers of fixations by the proportional size of the respective AOI in relation to the entire image. For example, if an AOI subtends 5% of the image and it receives 5% of all fixations, the normalized score for numbers of fixations will be 1, implying that the AOI received as many fixations as expected by chance (i.e., random viewing); conversely, normalized fixation scores larger than 1 imply that the AOI received more fixations than expected by chance. We observed normalized numbers of fixations larger than 1 only for eyes, nose, and mouth (see Supplement, Table S4), implying that only these regions were attended more than expected by chance. We thus focus our analyses on these three AOIs.

It should be noted that the different eye-tracking measures are not independent: For example, given that only one facial feature can be fixated at a time, more fixations to the eyes automatically imply less fixations to all other features. Also, a larger number of fixations implies shorter fixations and more saccades, given that eye movements during the study phase are measured within a fixed time frame of

8. We report descriptive statistics of non-normalized values in the Supplement (see Table S3).

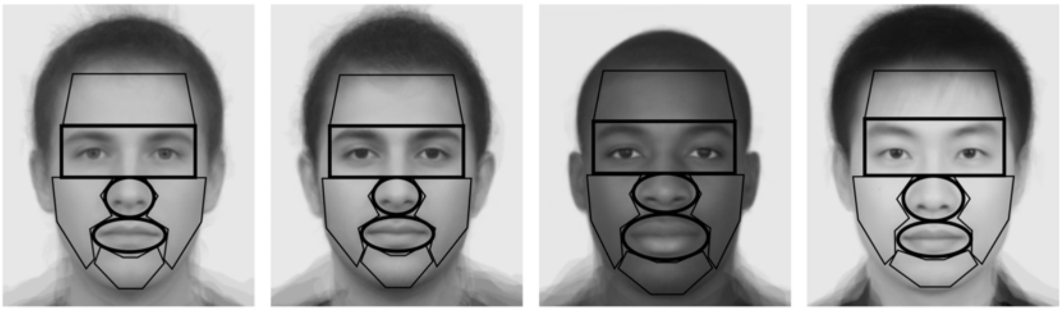

FIGURE 2. Illustration of manually-adjusted areas of interests for each ethnicity (White, Middle Eastern, Black, and Asian; from left to right). Faces are displayed as averages morphed across individual stimulus faces used for each ethnic group in this study.

5000ms (although this does not apply to the same degree to the recognition phase, during which stimulus presentation was participant-paced). Bearing this restriction in mind, we nonetheless decided to perform separate analyses for each individual eye-tracking measure, in order to facilitate interpretation of results. Further, each eye-tracking measure was analyzed separately for study and recognition phase. Given that the ORE in recognition was driven by responses to new faces (i.e., increased false alarms for outgroup faces) rather than to old faces (i.e., no difference in hits for ingroup versus outgroup faces), we analyzed eye movements during recognition separately for old versus new faces (i.e., exploring if ethnicity effects on eye movements are more pronounced for new faces than for old faces).

Eye-tracking measures were analyzed on a trial-by-trial basis with linear mixed-effects models, treating both participants and items as random factors (Judd, Westfall, & Kenny, 2012). Mixed-effects models were fitted with the R package Ime4 (Bates, Maechler, Bolker, & Walker, 2015; R Core Team, 2017); p-values were calculated using Satterthwaite approximate degrees of freedom with the R package lmerTest (Kuznetsova, Brockhoff, & Christensen, 2017). Our models include intercepts for items and participants as random effects but no random slopes.

Our first aim was to test for ethnicity effects on the different eye-tracking measures. For each eye-tracking measure, we entered group contrasts as fixed effects into a linear mixed effects model, coding for a general ingroup versus outgroup effect, a proximal versus distal outgroup effect, and an effect between distal outgroups (see Table 1). Our next aim was to explore if eye-tracking measures were related to recognition accuracy, regardless of face ethnicity. We fitted mixed-effects logistic regression models, predicting recognition accuracy for each item (coded as 1 = correct or 0 = incorrect) by the respective eye-tracking measures. Our final aim was to link ethnicity effects in eye-tracking measures to ethnicity effects in recognition by exploring if eye movements mediate ethnicity effects on recognition. Given that we observed ethnicity effects on recognition only for *new* faces (i.e., false alarm rates), only eye movements to new faces could be potential mediators for recognition effects. Further, we decided to run mediation analyses only for

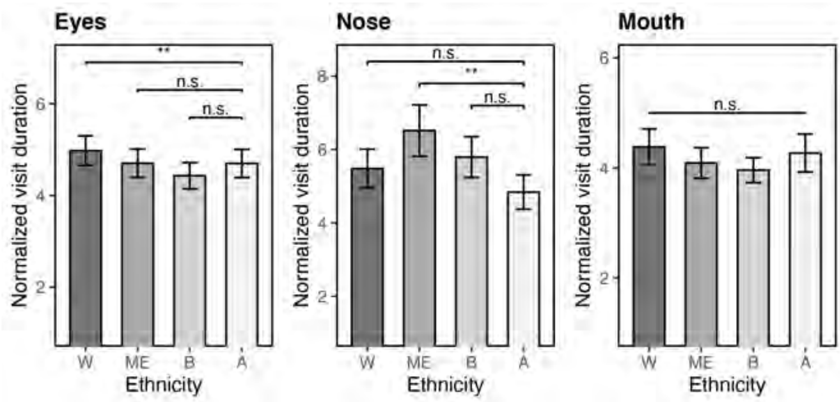

FIGURE 3. Illustration of normalized visit durations for eyes, nose, and mouth during study phase as a function of face ethnicity and trial type. W = White (ingroup); ME = Middle Eastern (proximal outgroup); B = Black (distant outgroup); A = Asian (distant outgroup). Error bars represent standard errors of the mean.

eye-tracking measures that were related to recognition accuracy overall. Foreshadowing the results, none of the eye-tracking measures were related to recognition of new faces (i.e., there was no correlation between the potential mediator eye movements and the outcome variable recognition). We therefore did not test any mediation models.

For ease of comprehension, only significant predictors are reported in this article; complete results are reported in the Supplement.

Study Phase.

Eye movements to facial features. Descriptive statistics of normalized visit durations and normalized number of fixations to facial features of White, Middle Eastern, Black, and Asian faces during the study phase are depicted in Figure 3 and Table 3, respectively. Both measures were analyzed separately for eyes, nose, and mouth via linear mixed-effects models. Subsequently, we analyzed if eye movements to eyes, nose, and mouth were related to recognition performance via logistic mixed-effects models.

Eyes. Participants fixated longer, $\beta = 0.093$, 95% CI [0.026; 0.161], $t(58.84) = 2.684$, $p = .009$, and more often, $\beta = 0.085$, 95% CI [0.019, 0.150], $t(58.96) = 2.519$, $p = .014$, to the eyes of White faces compared to all outgroup faces during the study phase. There was no difference in fixation durations and number of fixations to the eyes of Middle Eastern, Black, and Asian faces (see Supplement, Table S5). Further, eye movements to the eyes during the study phase did not predict subsequent recognition performance (see Supplement, Table S10).

Nose. There was no difference in visit duration and number of fixations to the nose of White faces compared to all outgroup faces (see Supplement, Table S5). However, participants fixated longer, $\beta = 0.407$, 95% CI [0.119, 0.694], $t(59.52) = 2.748$,

TABLE 3. Descriptive Statistics of Normalized Number of Fixations for Eyes, Nose, and Mouth, as a Function of Face Ethnicity During the Study and Recognition Phase (Standard Deviations in Parentheses)

	White	Middle Eastern	Black	Asian
Study phase				
Eyes	4.967 (2.415)	4.776 (2.300)	4.456 (2.195)	4.657 (2.317)
Nose	6.265 (4.528)	6.929 (5.010)	6.180 (4.131)	3.916 (2.294)
Mouth	4.066 (2.119)	3.713 (1.841)	3.786 (1.671)	5.295 (3.667)
Recognition phase				
Old faces				
Eyes	5.583 (3.214)	4.975 (3.251)	4.415 (2.939)	4.240 (3.041)
Nose	9.305 (8.415)	10.670 (10.024)	8.699 (7.637)	7.982 (6.715)
Mouth	4.019 (4.740)	4.549 (4.700)	4.271 (3.503)	4.792 (4.707)
New faces				
Eyes	5.481 (3.554)	5.008 (3.200)	4.270 (2.774)	4.041 (2.684)
Nose	9.533 (9.358)	10.209 (8.764)	8.310 (6.738)	8.148 (6.575)
Mouth	4.402 (4.707)	4.632 (4.320)	4.794 (3.271)	5.181 (5.020)

p = .008, and more often, β = 0.403, 95% CI [0.106, 0.700], t(59.42) = 2.636, p = . 011, to the nose of Middle Eastern faces compared to Black and Asian faces. There was no difference in eye movements to noses of Black and Asian faces. Also, eye movements to the nose during the study phase did not predict subsequent recognition performance (see Supplement, Table S10).

Mouth. Number of fixations and visit durations to the mouth during the study phase did not differ for faces of different ethnicities (see Supplement, Table S5). Also, eye movements to the mouth during the study phase did not predict subsequent recognition performance (see Supplement, Table S10).

Overall eye-movement activity level. Descriptive statistics of number of fixations, fixation durations, number of saccades, and saccadic amplitudes for White, Middle Eastern, Black, and Asian faces during the study phase are illustrated in Figure 4. Results of linear mixed-effects models demonstrated no ethnicity effects on any of the measures (see Supplement, Table S6). However, regardless of face ethnicity, more fixations, β = 0.296, 95% CI [0.043, 0.548], *Wald* z = 2.327, p = .020, shorter fixations, β = –0.151, 95% CI [–0.266, –0.037], *Wald* z = –2.607, p = .009, as well as wider-ranging saccades, β = 0.275, 95% CI [0.035, 0.528], *Wald* z = 2.218, p = .027, during the study phase were related to increased recognition accuracy. Conversely, number of saccades during the study phase were not related to recognition performance (see Supplement, Table S11).

Summary. During the study phase, participants fixated longer and more often to the eyes of White faces compared to Middle Eastern, Black, and Asian faces.

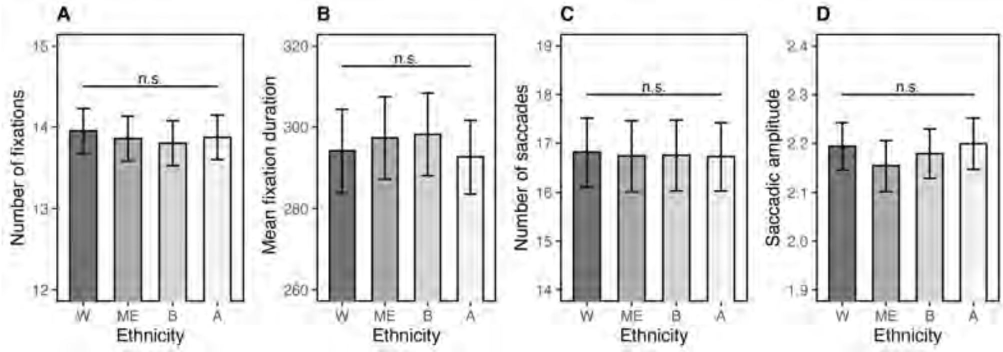

FIGURE 4. Number of fixations, mean fixation durations (in ms), number of saccades and length of saccades (in degrees visual angle) during the study phase as a function of face ethnicity and recognition accuracy. W = White; ME = Middle Eastern; B = Black; A = Asian. Error bars represent standard errors of the mean.

Further, participants fixated longer and more often to the nose of Middle Eastern faces compared to Black and Asian faces. Finally, there were no ethnicity effects on overall eye-movement activity level. We conclude that eye movements during the study phase were unrelated to the ORE in recognition memory: First, participants recognized old faces of all ethnicities at equal rates, suggesting that differential processing of ingroup and outgroup faces during the study phase did not affect recognition memory. Further, even though overall eye-movement activity during study was related to recognition performance, this effect was unrelated to the ORE, given that there were no ethnicity effects on eye-movement activity.

Recognition Phase.

Eye movements to facial features. Descriptive statistics of normalized visit durations and normalized number of fixations to facial features of White, Middle Eastern, Black, and Asian faces during the recognition phase are illustrated in Figure 5 and Table 3, respectively. Both measures were analyzed separately for eyes, nose, and mouth in linear mixed-effects models. Subsequently, we analyzed if eye movements to eyes, nose, and mouth were related to recognition performance using logistic mixed-effects models. For all analyses, we tested separate models for old and new faces, given that ethnicity effects on recognition differed for old and new faces.

Eyes. For both old and new faces, number of fixations and visit durations to the eyes differed between faces of different ethnicities: Participants fixated longer (old faces: $\beta = 0.271$, 95% CI [0.160, 0.383], $t(58.08) = 4.754$, $p < .001$; new faces: $\beta = 0.249$, 95% CI [0.157, 0.342], $t(59.71) = 5.224$, $p < .001$), and more often (old faces: $\beta = 0.261$, 95% CI [0.152, 0.370], $t(57.88) = 4.647$, $p < .001$; new faces: $\beta = 0.263$, 95% CI [0.168, 0.359], $t(59.98) = 5.361$, $p < .001$), to the eyes of White faces compared to

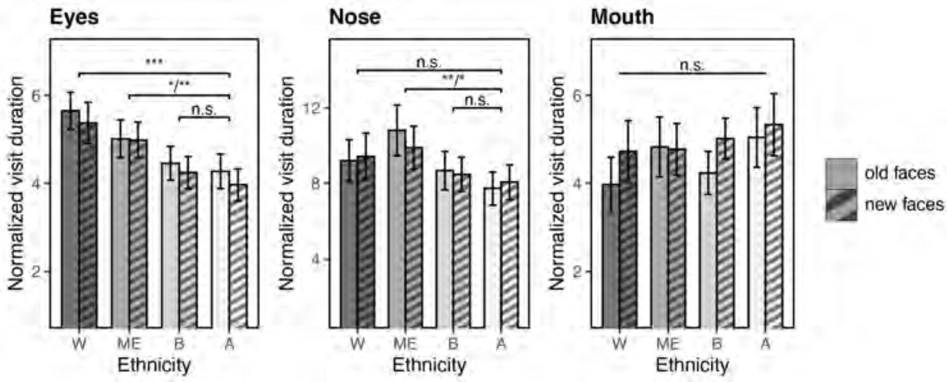

FIGURE 5. Illustration of normalized visit durations for eyes, mouth, and nose for old faces during the recognition phase as a function of face ethnicity. Dark shades represent old faces; light shades represent new faces. * = $p < .05$; ** = $p < .01$; *** = $p < .001$; values before the dash represent significance level for old faces; values after the dash represent significance level for new faces. W = White (ingroup); ME = Middle Eastern (proximal outgroup); B = Black (distant outgroup); A = Asian (distant outgroup). Error bars represent standard errors of the mean.

Middle Eastern, Black, and Asian faces. Further, participants fixated longer (old faces: $\beta = 0.209$, 95% CI [0.053, 0.366], $t(57.51) = 2.598$, $p = .012$; new faces: $\beta = 0.293$, 95% CI [0.162, 0.423], $t(59.07) = 4.351$, $p < .001$), and more often (old faces: $\beta = 0.210$, 95 % CI [0.056, 0.364], $t(57.34) = 2.650$, $p = .010$; new faces: $\beta = 0.285$, 95% CI [0.151, 0.420], $t(58.08) = 4.125$, $p < .001$), to the eyes of Middle Eastern faces compared to Black and Asian faces. There was no difference in fixations to the eyes of Black and Asian faces (see Supplement, Tables S7 and S8).

In addition, for old faces, more fixations to the eyes, $\beta = 0.042$, 95% CI [0.006, 0.076], *Wald* $z = 2.276$, $p = .023$, and longer fixations to the eyes, $\beta = 0.041$, 95% CI [0.006, 0.079], *Wald* $z = 2.306$, $p = .021$, were related to higher recognition accuracy. For new faces, fixations to the eyes were not related to recognition accuracy (see Supplement, Table S10).

Nose. For both old and new faces, participants fixated equally long and equally often to the nose of White faces compared to Middle Eastern, Black, and Asian faces combined. However, participants fixated longer (old faces: $\beta = 0.893$, 95% CI [–0.371, 0.426], $t(57.69) = 3.082$, $p = .003$; new faces: $\beta = 0.536$, 95% CI [0.065, 1.005], $t(59.53) = 2.211$, $p = .031$) and more often (old faces: $\beta = 0.789$, 95% CI [–0.367, 0.467], $t(58.09) = 2.603$, $p = .012$; new faces: $\beta = 0.657$, 95% CI [0.175, 1.140], $t(59.9) = 2.644$, $p = .011$) to the nose of Middle Eastern faces compared to Black and Asian faces. There was no difference in fixations to noses of Black and Asian faces (see Supplement, Tables S7 and S8). For both old and new faces, fixations to the nose during recognition were unrelated to recognition performance (see Supplement, Table S10).

Mouth. Number of fixations and visit durations to the mouth did not differ between faces of different ethnicities—neither for old nor for new faces (see

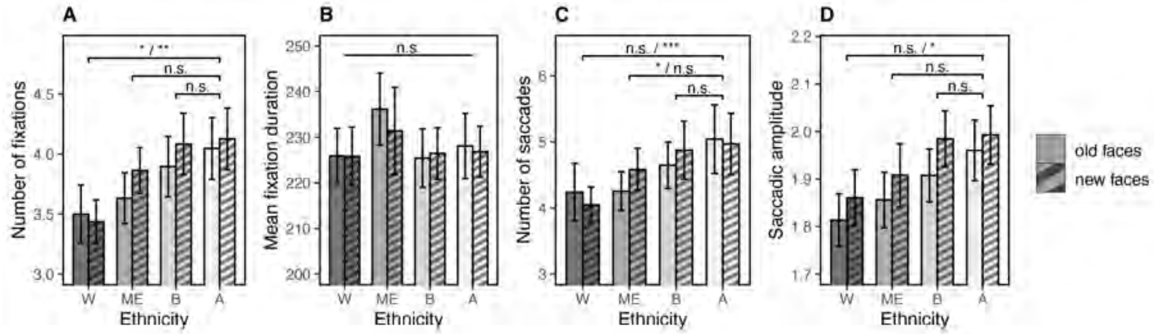

FIGURE 6. Number of fixations, mean fixation durations (in ms), number of saccades and length of saccades (in degrees visual angle) during the recognition phase as a function of face ethnicity and recognition accuracy. W = White; ME = Middle Eastern; B = Black; A = Asian. Error bars represent standard errors of the mean.

Supplement, Tables S7 and S8). Subsequent analyses demonstrated that for old faces, a longer visit duration, $\beta = -0.033$, 95% CI [–0.050, –0.016], Wald $z = -3.898$, $p < .001$, and more fixations, $\beta = 0.036$, 95% CI [–0.054, –0.017], Wald $z = -3.876$, $p < .001$, to the mouth decreased the likelihood of recognizing a face. For new faces, fixations to the mouth were unrelated to recognition accuracy (see Supplement, Table S10).

Overall eye-movement activity level. Descriptive statistics of number of fixations, fixation durations, number of saccades, and saccadic amplitudes for White, Middle Eastern, Black, and Asian faces during the recognition phase are illustrated in Figure 6. We first tested for ethnicity effects using linear mixed effects models. Next, we analyzed if overall eye-movement activity was related to recognition performance using logistic mixed-effects models. All models were tested separately for old and new faces.

Number of fixations. For both old and new faces, participants made less fixations to White faces compared to Middle Eastern, Black, and Asian faces, and this effect was more pronounced for new faces, $\beta = -0.149$, 95% CI [–0.228, –0.070], $t(59.5) = -3.666$, $p < .001$, as compared to old faces, $\beta = -0.092$, 95% CI [–0.179, –0.004], $t(58.55) = -2.031$, $p = .047$. There was no difference in number of fixations to Middle Eastern, Black, and Asian faces (see Supplement, Table S9). For old faces, less fixations were related to higher recognition accuracy, $\beta = -0.149$, 95% CI [–0.192, –0.108], Wald $z = -7.052$, $p < .001$. For new faces, number of fixations was unrelated to recognition accuracy (see Supplement, Table S11).

Fixation duration. There was no ethnicity effect on fixation durations during recognition, neither for old nor for new faces (see Supplement, Table S9). Further,

TABLE 4. Reaction Times in ms for Old and New Faces as a Function of Face Ethnicity (Standard Deviations in Parentheses)

	White	Middle Eastern	Black	Asian
Old faces	1151 (879)	1230 (892)	1283 (994)	1368 (1244)
New faces	1138 (732)	1284 (782)	1339 (932)	1381 (980)

fixation durations were unrelated to recognition performance (see Supplement, Table S11).

Number of saccades. For old faces, participants made less saccades to Middle Eastern faces compared to Black and Asian faces, β = –0.200, 95% CI [–0.362, –0.039], $t(58.87) = -2.404$, $p = .019$; there were no other ethnicity effects (see Supplement, Table S9). Further, lower numbers of saccades were related to higher recognition accuracy for old faces, β = –0.099, 95% CI [–0.132, –0.068], Wald $z = -6.081$, $p < .001$. For new faces, participants made fewer saccades to White faces compared to all outgroup faces, β = –0.192, 95% CI [–0.290, –0.094], $t(60.03) = -3.814$, $p < .001$; there were no other ethnicity effects (see Supplement, Table S9). In addition, for new faces, number of saccades was unrelated to recognition accuracy (see Supplement, Table S11).

Saccadic amplitudes. For old faces, there was no ethnicity effect on saccadic amplitude (see Supplement, Table S9). Shorter-ranging saccades were related to higher recognition accuracy for old faces, β = –0.356, 95% CI [–0.539, –0.178], Wald $z = -3.910$, $p < .001$. For new faces, saccades for White faces were shorter ranging than for Middle Eastern, Black and, Asian faces, β = –0.026, 95% CI [–0.047, –0.004], $t(59.89) = -2.275$, $p = .027$, but number of saccades was unrelated to recognition accuracy (see Supplement, Table S11).

Auxiliary analysis of reaction times. We additionally analyzed reaction times to facilitate interpretation of ethnicity effects in eye-movement activity level during recognition. We again tested separate mixed-effects model for responses to old and new faces, including previous contrasts as predictors (see Table 4 and Supplement Table S12 for full results). For both old and new faces, reaction times were faster for White faces compared to Middle Eastern, Black, and Asian faces (old faces: β = –35.740, 95% CI [–63.84, –7.64], $t(57.31) = -2.467$, $p = .017$; new faces: β = –48.970, 95% CI [–73.67, –24.27], $t(58.85) = -3.846$, $p < .001$).

We then repeated linear mixed-effects models for number of fixations, number of saccades, and saccadic amplitudes during recognition while controlling for reaction times (see Supplement, Table S13). Results revealed that ethnicity effects were no longer significant, with reaction times being the only predictor for number of fixations (old faces: β = 2.727, 95% CI [2.649, 2.802], $t(1749) = 70.384$, $p < .001$; new faces: β = 2.472, 95% CI [2.407, 2.538], $t(1908) = 73.850$, $p < .001$), number of saccades (old faces: β = 3.456, 95% CI [3–320, 3.588], $t(1850) = 50.742$, $p < .001$; new faces: β = 2.981, 95% CI [2.868, 3.093], $t(1924) = 51.985$, $p <.001$), and saccadic amplitudes (old faces: β = 0.244, 95% CI [0.216, 0.273], $t(1920) = 16.618$, $p < .001$; new faces: β = 0.218, 95% CI [0.188; 0.249], $t(1932) = 14.038$, $p <.001$) during recognition (with the only exception that the proximal vs. distal outgroups contrast for number of saccades to old faces remained significant, β = –0.123, $t(60) = –2.443$, $p = .018$).

Summary. During recognition of both old and new faces, participants fixated more on the eyes of White faces compared to Middle Eastern, Black, and Asian faces. Also, they fixated more on the eyes of Middle Eastern faces compared to Black and Asian faces. While fixations to the eyes were related to better recognition performance for old faces, this was not the case for new faces. Given that the ORE occurred only for new faces and not for old faces, we conclude that fixations to the eyes during recognition were not related to the ORE in recognition.

Regarding overall eye-movement activity during recognition, ethnicity effects were more pronounced for new faces compared to old faces: New White faces received overall less fixations, less saccades and shorter-ranging saccades than new Middle Eastern, Black, and Asian faces. However, these effects disappeared when controlling for reaction times. Further, for old faces, higher eye-movement activity (i.e., less fixations, less saccades, and shorter-ranging saccades) was related to higher recognition accuracy; however, this was not the case for new faces. Given that the ORE occurred only for new faces but not for old faces, we suggest that eye-movement activity level during recognition was not related to the ORE in recognition.

DISCUSSION

The present study compared recognition memory and eye movements for ingroup faces with proximal and distal outgroup faces and explored the relation between eye movements and the ORE. Results of the recognition task replicated a typical ORE: White participants were substantially better at recognizing White (i.e., ingroup) faces compared to all outgroup faces, explaining 33 percent of the variance in recognition accuracy. Further, recognition accuracy differed between faces of proximal and distal outgroups: Participants were better at recognizing Middle Eastern faces (i.e., proximal outgroup) compared to Black and Asian faces (i.e., distal outgroups), explaining another eight percent of variance. Notably, the ORE in recognition memory was entirely driven by responses to new faces—with higher false alarm rates for outgroup faces than ingroup faces—and not by impaired recognition of outgroup faces compared to ingroup faces presented during the study phase.

In addition, we observed a similar pattern of ORE effects in response bias: Participants showed a greater bias to respond “new” to ingroup compared to all outgroup faces; similarly, participants showed a greater bias to respond “new” to proximal compared to distal outgroup faces. In conclusion, the present data replicate the ORE, and the magnitude of the ORE varied gradually for proximal and distal outgroups, providing a valuable basis for interpretation of eye-tracking data.

Our analysis of eye movements followed two aims: First, we investigated if eye movements differed for ingroup faces compared to proximal and distal outgroup faces. Second, we explored if eye movements were related to recognition accuracy. Eye movements were analyzed separately for study and recognition phases, allowing us to disentangle effects occurring during encoding versus retrieval. Given that the ORE in recognition memory was entirely driven by responses to new faces, we additionally explored if differences in eye movements to ingroup, proximal outgroup, and distal outgroup faces were more pronounced for new faces than for old faces. In the following paragraphs, we summarize and discuss eye-tracking results against the background of different theoretical approaches derived from the literature.

According to the categorization-individuation account, people are expected to pay more attention to the eyes of ingroup compared to outgroup faces. Confirming this idea, present results showed that participants fixated longer and more often to the eyes of ingroup faces as compared to all outgroup faces. This effect emerged consistently during study and during recognition, regardless whether a face was old or new. These findings replicate evidence from previous studies (e.g., Goldinger et al., 2009; Kawakami et al., 2014; Wu et al., 2012). However, contrary to some previous studies (e.g., Caldara et al., 2010; Kawakami et al., 2014; Malcolm et al., 2008; McKelvie, 1976; Peterson & Eckstein, 2012), and conflicting with the categorization-individuation account, fixations to the eyes were not related to recognition accuracy. This finding implies that differences in eye movements to eyes of ingroup and outgroup faces did not contribute to the ORE in recognition memory in our study.

Based on the holistic-processing account, people are expected to pay more attention to the nose of ingroup faces compared to outgroup faces, given that attention to the center of the face may foster holistic processing. However, present results demonstrate that ingroup faces did not receive more fixations to the nose than outgroup faces. Unexpectedly, Middle Eastern faces received more fixations to the nose as compared to Black and Asian faces. This effect may have been caused by specific bottom-up processes due to perceptual characteristics of Middle Eastern faces. Given the lack of previous eye-tracking research with Middle Eastern faces, we need further research to replicate this finding and understand its underlying mechanism. Eye movements to the nose were not related to recognition accuracy, suggesting that increased attention to the nose of Middle Eastern faces—if interpreted as an indicator of increased holistic processing—does not account for the recognition advantage of Middle Eastern compared to Black and Asian faces.

According to the diagnostic-feature account, attention to features that help discriminate between individuals of a certain ethnic group is expected to be related to higher recognition accuracy. For example, people who attend more to the mouth and nose (as compared to the eyes) of Black and Asian faces are expected to have better recognition for Black and Asian faces (e.g., Hills et al., 2013b; Wang et al., 2015). Unlike previous research, the present study showed that participants fixated equally long and equally often to the nose and mouth of all faces, regardless of ethnicity (c.f., Goldinger et al., 2009; McDonnell et al., 2014). Further, across all ethnicities, fewer fixations to the mouth during recognition was related to higher recognition accuracy of old faces. Overall, these findings do not provide any evidence in favor of the diagnostic-feature account.

Finally, based on the more-active-scanning account, people are expected to display more active eye movement patterns for ingroup than outgroup faces. The present findings regarding eye movement activity differ between study and recognition phase: During the study phase, there was the same level of eye-movement activity for ingroup and outgroup faces. Thus, participants did not engage in more active encoding of ingroup faces, contradicting with the more-active-scanning account as well as with previous findings (e.g., Goldinger et al., 2009; Klama & Milton, 2012; Wu et al., 2012). Importantly, across all ethnicities, faces viewed with a higher level of eye-movement activity during the study phase were better recognized. This finding is in line with previous research, showing that eye-movement activity during encoding is beneficial for face learning (Heisz et al., 2013, Henderson et al., 2005). Conversely, during recognition, participants made fewer fixations, fewer saccades, and shorter-ranging saccades for ingroup faces than for outgroup faces, but these effects were no longer significant when controlling for response times. Given that stimulus presentation in the recognition phase was self-paced, faster responses automatically led to shorter presentations times, which may automatically imply fewer fixations and fewer saccades. This should be considered when interpreting effects in eye-movement activity during recognition, given that we do not know why both measures are related. Particularly, we do not know if (a) longer response times are caused by more active eye movements, (b) increased eye-movement activity is a mere byproduct of prolonged response times, or (c) both effects are caused by a common underlying mechanism. Finally, eye-movement activity during recognition was related to recognition accuracy for old faces, but not for new faces. Specifically, correctly identified old faces received fewer fixations, fewer saccades, and shorter-ranging saccades. This effect is consistent with previous research and can be interpreted as an eye-movement-based memory effect (i.e., reprocessing effect; see Althoff & Cohen, 1999). Conversely, eye-movement activity for new faces did not vary with recognition accuracy. Given that the ORE in recognition was only observed for new faces, we thus conclude that eye-movement activity during recognition is unrelated to the ORE in recognition memory.

In summary, the present research does not provide any evidence that eye movements during encoding or retrieval contributed to the ORE in recognition memory. However, it should be noted that contrary to previous ORE eye-tracking studies, the ORE was only found for new faces (i.e., in false alarm rates). This means that since participants were equally good at recognizing old faces from all ethnic groups, we would not expect any relation between eye movements during the study phase and the ORE. We cannot rule out that this particular pattern of finding an ORE only for new faces (i.e., in false alarms) may have been caused by the specific set-up of our recognition task. Different from most ORE eye-tracking experiments, our recognition task included three different outgroups; as a consequence, participants saw three times as many outgroup faces than ingroup faces. This may have prompted feelings of being threatened (e.g., in the sense of being outnumbered) or motivated attending to racial outgroups, perhaps due to their (perceived unusual) majority status in the experimental context. These two mechanisms could have caused shifts in visual attention and memory for ingroup and outgroup faces (see Wilson & Hugenberg, 2010; Wilson, See, Bernstein, Hugenberg, & Chartier, 2014). In addition, the difference in base rates may have shifted the participants' response criterion to produce more false alarms. Finding an ORE only in false alarm rates may also suggest that participants applied different retrieval processes for ingroup and outgroup faces. Future research may further investigate ingroup-outgroup differences in retrieval strategies (e.g., by using processing tree models dissociating between different retrieval processes; see Meiser & Broder, 2002) and relate these effects to differences in eye movements.

Another important aim of the study was to compare recognition and eye movements between proximal and distal outgroups. We observed several differences between proximal versus distal outgroup faces: Participants were better at recognizing proximal outgroup faces than distal outgroup faces. Also, response bias was more conservative for proximal outgroup faces compared to distal outgroup faces. Moreover, during recognition, proximal outgroup faces received more fixations to the eyes than distal outgroup faces. Finally, participants displayed (numerically) fewer fixations and fewer saccades during recognition of proximal outgroup faces compared to distal outgroup faces. This pattern of results suggests that recognition and eye movements for the different outgroup faces were not only characterized by mere ingroup/outgroup categorization. Instead, recognition and eye movements regarding proximal outgroup faces mimic effects for ingroup faces. There are several potential explanations for these observations: First, participants may have had more experience with proximal outgroup faces compared to distal outgroup faces, due to more exposure and more frequent interactions in daily life—as attested by their responses in the contact measures. Second, participants may also have had greater motivation to individuate proximal outgroup faces given the more important role within society. Particularly, relatively more individuation of proximal than distal outgroup faces may have led to more attention to the eyes of proximal than distal outgroup faces. Future research may systematically disentangle effects of exposure and individuation on eye movements by experimentally manipulating each factor individually. Finally, effects may have also been caused by specific perceptual characteristics of the selected groups. Specifically, the physiognomy of Middle Eastern faces may have been perceived to be more similar to White faces compared to Black and Asian faces. Comparing effects of White participants with Middle Eastern participants may provide further information regarding the influence of physiognomic characteristics on eye movements.

LIMITATIONS

The present study has several limitations that should be acknowledged when interpreting our findings. First, our research included only White German participants.

It is usually beneficial to include participants from all studied ethnic groups, in order to rule out that effects are driven by biases in stimulus selection (see Sporer, 2001). The design of our study may have particularly benefitted from participants from the other three ethnic groups, because differences between participant groups may have provided more information regarding the use of diagnostic features and cultural differences in eye movements. Further, the main finding of our study—namely, differences in attention to the eyes of ingroup versus outgroup faces—has, to our knowledge, mainly been observed for White participants (Brielmann et al., 2014, Exp. 1; Goldinger et al., 2009; Kawakami , Amodio, & Hugenberg, 2017; Wu et al., 2012; but see Goldinger et al., 2009, Exp. 2). It would therefore be highly informative to further investigate if, and under what circumstances, this effect generalizes to non-White participants.

Second, we limited analyses of eye movements to two groups of eye-tracking measures: fixations to AOIs and overall eye-movement activity. There are, however, further eye-tracking measures and analysis techniques available. For example, while we have used pre-defined AOIs, other studies have used data-driven analysis techniques, such as heat maps, to analyze attention to facial features (e.g., Arizpe et al., 2016; Caldara et al., 2010; Fu et al., 2012). Further, other studies have additionally measured pupil sizes, demonstrating that pupils were larger during encoding of outgroup compared to ingroup faces (e.g., Goldinger et al., 2009; Shang & Fu, 2016), which may imply higher mental effort during encoding of outgroup faces (cf. Kahneman, 1973). Comparing pupil sizes for ingroup, proximal outgroup, and distal outgroup faces may have also been informative in our study. However, our study design did not allow reliable interpretation of pupil sizes because we did not match stimuli in luminance, which is a crucial perceptual confound affecting pupil dilation (e.g., Porter, Troscianko, & Gilchrist, 2007). We note that the large number of potential dependent variables measurable with eye-tracking methods leave researchers with many "researcher degrees of freedom" (Wicherts et al., 2016). For example, there are different ways of classifying eye movements as fixations and saccades (Hessels et al., 2018). Similarly, there are many ways of defining what constitutes the eye region, either including the eyebrows and the bridge of the nose (e.g., Goldinger et al., 2009; Wu et al., 2012; see Sadr, Jarudi, & Sinha, 2003, on the importance of eye brows for face recognition) or including only the eyes (e.g., Fu et al., 2012; Hu et al., 2014). These decisions may partly account for inconsistent findings across studies. We therefore argue that researchers should be more transparent about their analysis decisions (see also Arizpe et al., 2016) and that they should openly share their raw data whenever possible for re- and meta-analyzes. This is a crucial precondition for accumulation of scientific knowledge, allowing higher-order integration of results. Future research may address this issue by conducting a systematic review of previous evidence, considering differences in analysis techniques.

Finally, another important limitation of our study, as well as of previous eye-tracking research, is the use of univariate analyses, focusing on single eye-tracking measures at a time. This approach neglects the fact that selected measures may potentially be dependent, given that they are all sampled simultaneously.

Therefore, future research may also make use of more advanced analysis techniques, such as machine learning as a data-driven approach, in order to consider the wealth of eye-movement information.

In summary, one important conclusion of this research is that recognition and eye movements differ for ingroups as well as for proximal and distal outgroups. Recognition was better for ingroup faces than for outgroup faces; further, recognition was better for proximal outgroup faces than distal outgroup faces. Ingroup faces received more attention to the eyes than outgroup faces. Also, participants made less fixations and saccades as well as shorter saccades during recognition of ingroup faces, which may imply less effort during retrieval of ingroup faces. However, eye movements were not related to differential recognition of ingroup, proximal outgroup, and distal outgroup faces.

REFERENCES

Althoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(4), 997–1010. https://doi.org/10.1037/0278-7393.25 .4.997

Arizpe, J., Kravitz, D. J., Walsh, V., Yovel, G., & Baker, C. I. (2016). Differences in looking at own- and other-race faces are subtle and analysis-dependent: An account of discrepant reports. *PloS One, 11*(2), e0148253. https://doi.org/10.1371/journal.pone.0148253

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PloS One, 3*(8), e3022. https://doi.org/10.1371/journal.pone.0003022

Brielmann, A. A., Bülthoff, I., & Armann, R. (2014). Looking at faces from different angles: Europeans fixate different features in Asian and Caucasian faces. *Vision Research, 100*, 105–112. https://doi.org/10.1016/j.visres.2014.04.011

Britt, T. W., Bonieci, K. A., Vescio, T. K., Biernat, M., & Brown, L. M. (1996). Intergroup anxiety: A person × situation approach. *Personality and Social Psychology Bulletin, 22*(11), 1177–1188. https://doi.org/10.1177/01461672962211008

Caldara, R., & Miellet, S. (2011). iMap: A novel method for statistical fixation mapping of eye movement data. *Behavior Research Methods, 43*(3), 864–878. https://doi.org/10.3758/s13428-011-0092-x

Caldara, R., Zhou, X., & Miellet, S. (2010). Putting culture under the spotlight reveals universal information use for face recognition. *PloS One, 5*(3), e9708. https://doi.org/10.1371/journal.pone.0009708

Correll, J., & Hudson, S. M. (2020). An error in the analysis of "An eye for the I". *Journal of Personality and Social Psychology, 119*(5), 1030–1036. https://doi.org/10.1037/pspa0000200

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

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*(12), 1827–1837. https://doi.org/10.1016/0042-6989(95)00294-4

Ellis, H. D., Deregowski, J. B., & Shepherd, J. W. (1975). Descriptions of white and black faces by white and black subjects. *International Journal of Psychology, 10*(2), 119–123. https://doi.org/10.1080/00207597508247325

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/bf03193146

Fu, G., Hu, C. S., Wang, Q., Quinn, P. C., & Lee, K. (2012). Adults scan own- and other-race faces differently. *PloS One, 7*(6), e37688. https://doi.org/10.1371/journa .pone.0037688

Goldinger, S. D., He, Y., & Papesh, M. H. (2009). Deficits in cross-race face learning: insights from eye movements and pupillometry. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(5), 1105–1122. https://doi.org/10.1037/a0016548

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

Hancock, K. J., & Rhodes, G. (2008). Contact, configural coding and the other-race effect in face recognition. *British Journal of Psychology, 99*, 45–56. https://doi.org/10.1348/000712607x199981

Hansen, B. C., Rakhshan, P. J., Ho, A. K., & Pannasch, S. (2015). Looking at others through implicitly or explicitly prejudiced eyes. *Visual Cognition, 23*(5), 612–642. https://doi.org/10.1080/13506285.2015.1063554

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

Heisz, J. J., Pottruff, M. M., & Shore, D. I. (2013). Females scan more than males: A potential mechanism for sex differences in recognition memory. *Psychological Science, 24*(7), 1157–1163. https://doi.org/10.1177/0956797612468281

Henderson, J., Williams, C., & Falk, R. (2005). Eye movements are functional during face learning. *Memory and Cognition, 33*(1), 98–106. https://doi.org/10.3758/BF03195300

Hessels, R. S., Niehorster, D. C., Nyström, M., Andersson, R., & Hooge, I. T. C. (2018). Is the eye-movement field confused about fixations and saccades? A survey among 124 researchers. *Royal Society Open Science, 5*(8), 180502. https://doi.org/doi:10.1098/rsos.180502

Hills, P. J., Cooper, R. E., & Pake, J. M. (2013a). First fixations in face processing: The more diagnostic they are the smaller the face-inversion effect. *Acta Psychologica, 142*(2), 211–219. https://doi.org/10.1016/j.actpsy.2012.11.013

Hills, P. J., Cooper, R. E., & Pake, J. M. (2013b). Removing the own-race bias in face recognition by attentional shift using fixation crosses to diagnostic features: An eye-tracking study. *Visual Cognition, 21*(7), 876–898. https://doi.org/10.1080/13506285.2013.834016

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. https://doi.org/10.1016/j.cognition.2013.08.012

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics, 57*(6), 787–795. https://doi.org/10.3758/BF03206794

Horry, R., Cheong, W., & Brewer, N. (2015). The other-race effect in perception and recognition: Insights from the complete composite task. *Journal of Experimental Psychology: Human Perception and Performance, 41*(2), 508. https://doi.org/10.1037/xhp0000042

Hu, C., Wang, Q., Fu, G., Quinn, P. C., & Lee, K. (2014). Both children and adults scan faces of own and other races differently. *Vision Research, 102*, 1–10. https://doi.org/10.1016/j.visres.2014.05.010

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*, 1392–1417. https://doi.org/10.1080/13506285.2013.821429

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*, 1168–1187. https://doi.org/10.1037/a0020463

Inquisit 4.0.2.0 [Computer software]. (2014). Millisecond Software LLC. http://www.millisecond.com

Janik, S. W., Wellens, A. R., Goldberg, M. L., & Dell'Osso, L. F. (1978). Eyes as the center of focus in the visual examination of human faces. *Perceptual and Motor Skills, 47*(3), 857–858. https://doi.org/10.2466/pms.1978.47.3.857

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103*(1), 54. https://doi.org/10.1037/a0028347

Kahneman, D. (1973). *Attention and effort*. Upper Saddle River, NJ: Prentice-Hall.

Kaspar, K., & König, P. (2011). Overt attention and context factors: The impact of repeated presentations, image type, and individual motivation. *PloS One, 6*(7), e21719. https://doi.org/10.1371/journal.pone.0021719

Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). Intergroup perception and cognition: An integrative framework for understanding the causes and consequences of social categorization. *Advances in Experimental Social Psychology, 55*, 1–80. https://doi.org/10.1016/bs.aesp.2016.10.001

Kawakami, K., Friesen, J., & Vingilis-Jaremko, L. (2018). Visual attention to members of own and other groups: Preferences, determinants, and consequences. *Social and Personality Psychology Compass, 12*(4), e12380. https://doi.org/doi:10.1111/spc3.12380

Kawakami, K., Williams, A., Sidhu, D., Choma, B. L., Rodriguez-Bailon, R., Canadas, E., Chung, D., & Hugenberg, K. (2014). An eye for the I: Preferential attention to the eyes of ingroup members. *Journal of Personality and Social Psychology, 107*(1), 1–20. https://doi.org/10.1037/a0036838

Kelly, D. J., Jack, R. E., Miellet, S., De Luca, E., Foreman, K., & Caldara, R. (2011). Social experience does not abolish cultural diversity in eye movements. *Frontiers in Psychology, 2*(95), 1–11. https://doi.org/10.3389/fpsyg.2011.00095

Klama, E., & Milton, F. (2012). Differences in eye movements between same and other race face recognition. *Proceedings of the Annual Meeting of the Cognitive Science Society, 34*(34), 1804–1809. https://escholarship.org/uc/item/9rr2s7cz

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–12. https://doi.org/10.18637/jss.v082.i13

Laidlaw, K. E. W., Risko, E. F., & Kingstone, A. (2012). A new look at social attention: Orienting to the eyes is not (entirely) under volitional control. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 1132–1143. https://doi.org/10.1037/a0027075

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion, 24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(6), 1364–1382. https://doi.org/10.1037/0278-7393.22.6.1364

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), 559–574. https://doi.org/10.1037S0096-3445.129.4.559

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin, 98*(1), 185–199. https://doi.org/10.1037/0033-2909.98.1.185

Malcolm, G. L., Lanyon, L. J., Fugard, A. J., & Barton, J. J. (2008). Scan patterns during the processing of facial expression versus identity: An exploration of task-driven and stimulus-driven effects. *Journal of Vision, 8*(8), 1–9. https://doi.org/10.1167/8.8.2

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. https://doi.org/10.1080/13506285.2017.1301614

McDonnell, G. P., Bornstein, B. H., Laub, C. E., Mills, M., & Dodd, M. D. (2014). Perceptual processes in the cross-race effect: Evidence from eyetracking. *Basic and Applied Social Psychology, 36*(6), 478–493. https://doi.org/10.1080/01973533.2014.958227

McKelvie, S. J. (1976). The role of eyes and mouth in the memory of a face. *American*

Journal of Psychology, 89(2), 311–323. https://doi.org/10.2307/1421414

Meiser, T., & Broder, A. (2002). Memory for multidimensional source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(1), 116–137. https://doi.org/https://doi.org/10.1037/0278-7393.28.1.116

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), 3–35. https://doi.org/10.1037//1076-8971.7.1.3

Meissner, C. A., Brigham, J. C., & Butz, D. A. (2005). Memory for own-and other-race faces: A dual-process approach. *Applied Cognitive Psychology, 19*(5), 545–567. https://doi.org/10.1002/acp.1097

Michel, C., Caldara, R., & Rossion, B. (2006). Same-race faces are perceived more holistically than other-race faces. *Visual Cognition, 14*(1), 55–73. https://doi.org/10.1080/13506280500158761

Miellet, S., Vizioli, L., He, L., Zhou, X., & Caldara, R. (2013). Mapping face recognition information use across cultures. *Frontiers in Psychology, 4*, 34. https://doi.org/10.3389/fpsyg.2013.00034

Minear, M., & Park, D. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers, 36*(4), 630–633. https://doi.org/10.3758/BF03206543

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

Nakabayashi, K., Lloyd-Jones, T. J., Butcher, N., & Liu, C. H. (2012). Independent influences of verbalization and race on the configural and featural processing of faces: A behavioral and eye movement study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(1), 61–77. https://doi.org/10.1037/a0024853

Or, C. C. F., Peterson, M. F., & Eckstein, M. P. (2015). Initial eye movements during face identification are optimal and similar across cultures. *Journal of Vision, 15*(13), 12. https://doi.org/10.1167/15.13.12

Paulus, A., Rohr, M., Neuschwander, M., Seewald, M., & Wentura, D. (2012). *Validation of the Saarland University Emotion Database.* [Unpublished manuscript]. Universität des Saarlandes, Saarbrücken, Germany.

Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences of the United States of America, 109*(48), e3314–3323. https://doi.org/10.1073/pnas.1214269109

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75*(3), 811–832. https://doi.org/10.1037/0022-3514.75.3.811

Porter, G., Troscianko, T., & Gilchrist, I. D. (2007). Effort during visual search and counting: insights from pupillometry. *Quarterly Journal of Experimental Psychology (2006), 60*(2), 211–229. https://doi.org/10.1080/17470210600673818

Proietti, V., Cassia, V. M., dell'Amore, F., Conte, S., & Bricolo, E. (2015). Visual scanning behavior is related to recognition performance for own and other-age faces. *Frontiers in Psychology, 6*, Article 1684. https://doi.org/10.3389/fpsyg.2015.01684

R Core Team. (2017). *R: A language and environment for statistical computing.* In R Foundation for Statistical Computing. https://www.R-project.org/

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

Rodin, M. J. (1987). Who is memorable to whom: A study of cognitive disregard. *Social Cognition, 5*, 144–165. https://doi.org/10.1521/soco.1987.5.2.144

Sadr, J., Jarudi, I., & Sinha, P. (2003). The role of eyebrows in face recognition. *Perception, 32*(3), 285–293. https://doi.org/10.1068/p5027

Shang, J., & Fu, X. (2016). Differences of affective learning with own-race and other-race faces: An eye-tracking study. In D. Harris (Ed.), *Engineering psychology and* *cognitive ergonomics: 13th International Conference, EPCE 2016, held as Part of HCI International 2016, Toronto, ON, Canada, July 17–22, 2016, Proceedings* (pp. 90–96). New York: Springer International Publishing. https://doi.org/10.1007/978-3-319-40030-3_10

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*(1), 34. https://doi.org/10.1037/0096-3445.117.1.34

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology Public Policy and Law, 7*(1), 36–97. https://doi.org/10.1038//1076-8971.7.1.36

Statistisches Bundesamt. (2017). *Bevölkerung und Erwerbstätigkeit. Bevölkerung mit Migrationshintergrund: Ergebnisse des Mikrozensus 2016* [Results from the German Microcensus 2016]. https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/ MigrationIntegration/Migrationshintergrund2010220167004.pdf?__blob=publicationFile

Tan, C. B. Y., Stephen, I. D., Whitehead, R., & Sheppard, E. (2012). You look familiar: How Malaysian Chinese recognize faces. *PloS One, 7*(1), e29714. https://doi.org/10.1371/journal.pone.0029714

Tanaka, J. W., & Gordon, I. (2011). Features, configuration and holistic face processing. In G. Calder, J. V. Rhodes, & M. J. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 177–194). Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199559053.013.0010

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology, 43*(2), 161–204. https://doi.org/10.1080/14640749108400966

Vargas, O. H. (2017). Cross-race effect reduction: Insights from eye-tracking and social background analyses [undergraduate honors thesis, University of Colorado Boulder]. https://scholar.colorado.edu/honr_theses/1523

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

Walkersmith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye-movement strategies involved in face perception. *Perception, 6*(3), 313–326. https://doi.org/10.1068/p060313

Wan, L., Crookes, K., Reynolds, K. J., Irons, J. L., & 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*, 91–115. https://doi.org/10.1016/j.cognition.2015.07.011

Wang, Q., Xiao, N. G., Quinn, P. C., Hu, C. S., Qian, M., Fu, G. Y., & Lee, K. (2015). Visual scanning and recognition of Chinese, Caucasian, and racially ambiguous faces: Contributions from bottom-up facial physiognomic information and top-down knowledge of racial categories. *Vision Research, 107*, 67–75. https://doi.org/10.1016/j.visres.2014.10.032

Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., van Aert, R. C. M., & van Assen, M. A. L. M. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid p-hacking. *Frontiers in Psychology, 7*, 1832. https://doi.org/10.3389/fpsyg.2016.01832

Wilson, J. P., & Hugenberg, K. (2010). When under threat, we all look the same: Distinctiveness threat induces ingroup homogeneity in face memory. *Journal of Experimental Social Psychology, 46*(6), 1004–1010. https://doi.org/https://doi.org/10.1016/j.jesp.2010.07.005

Wilson, J. P., See, P. E., Bernstein, M. J., Hugenberg, K., & Chartier, C. (2014). Differences in anticipated interaction drive own group biases in face memory. *PloS One, 9*(3), e90668. https://doi.org/10.1371/journal.pone.0090668

Wittwer, T., Tredoux, C. G., Py, J., & Paubel, P.-V. (2019). Training participants to focus on critical facial features does not decrease own-group bias. *Frontiers in Psychology, 10*(2081). https://doi.org/10.3389/fpsyg.2019.02081

Wu, E. X., Laeng, B., & Magnussen, S. (2012). Through the eyes of the own-race bias:

eye-tracking and pupillometry during face recognition. *Social Neuroscience, 7*(2), 202–216. https://doi.org/10.1080/17470919.2011.596946

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Springer. https://doi.org/https://doi.org/10.1007/978-1-4899-5379-7

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. https://doi.org/10.1177/1088868311418987

Zhao, M., Hayward, W. G., & Bülthoff, I. (2014). Holistic processing, contact, and the other-race effect in face recognition. *Vision Research, 105*, 61–69. https://doi.org/10.1016/j.visres.2014.09.006