Original Article

QJEP

Quarterly Journal of Experimental Psychology
1–13

DOI: 10.1177/17470218251346749
qjep.sagepub.com
Sage

Own- and other-race face learning in high and low variability

Siew Kei Kho[1], David R T Keeble[2], Hoo Keat Wong[2] and Alejandro J Estudillo[1]

Abstract

Research suggests that faces learned in high variability conditions (pictures taken on different days, with different viewpoints and lighting) enhanced the learning of own-race identities compared to low variability conditions (pictures taken on the same day, with similar lighting). However, it remains unclear how this variability affects the learning of other-race faces, as they are recognized differently compared to own-race faces. Thus, this study aims to examine the effect of high and low variability exposure on both own-race and other-race face learning. Chinese Malaysian and White participants were exposed to own- and other-race identities under high and low variability conditions. Identity recognition was assessed using a name verification task (Experiment 1) and an old-new recognition paradigm (Experiment 2). Results revealed enhanced learning of own-race faces under high variability conditions compared to low variability across both experiments. However, improved learning of other-race faces was evident only in the old-new recognition paradigm, not in the name verification task. These findings suggest that high variability exposure benefits other-race face recognition but not the face-name association for other-race identities.

Keywords

Face recognition; other-race effect; within-person variability

Received: 11 September 2024; revised: 17 February 2025; accepted: 25 March 2025

Introduction

Familiar faces are more easily recognized compared to unfamiliar faces (Bruce et al., 2001). Most people can recognize familiar faces in different viewing conditions (e.g., differences in lighting), but this is seemingly difficult for unfamiliar faces (Sinha et al., 2006). For example, minor differences such as viewing angle (Favelle et al., 2011), changes in lighting, viewpoint, or expression (Bruce, 1982; Estudillo, 2012; Estudillo & Bindemann, 2014; Longmore et al., 2008) could impair unfamiliar face recognition. Familiar faces are thought to have a robust representation in memory, which is built up from multiple exposures of a face in different contexts (Burton et al., 2005; Jenkins & Burton, 2011; Johnston & Edmonds, 2009). Extensive research has since investigated if faces presented in multiple exposures and variations could enhance the learning of new identities (e.g., Dowsett et al., 2016; Ritchie et al., 2021; White et al., 2014).

Research from different paradigms has shown that exposure to different face instances of an identity enhances

[1]Department of Psychology, Bournemouth University, Poole, UK
[2]School of Psychology, University of Nottingham Malaysia, Semenyih, Malaysia

Corresponding authors:
Siew Kei Kho, Department of Psychology, Bournemouth University, Poole House Talbot Campus, Poole BH12 5BB, UK.
Email: kkho@bournemouth.ac.uk

Alejandro J. Estudillo, Department of Psychology, Bournemouth University, Poole House Talbot Campus, Poole BH12 5BB, UK.
Email: aestudillo@bournemouth.ac.uk

its identification (Andrews et al., 2015; Matthews & Mondloch, 2018; Menon et al., 2015). For example, when participants are asked to match the identity of two simultaneously (Bindemann & Sandford, 2011) or sequentially (Menon et al., 2015) presented faces, accuracy increases if an additional face image of the identity is included. Similar improvements in face-sorting task performance have been found as additional photos of the target were presented for comparison (Dowsett et al., 2016; Matthews & Mondloch, 2018). Presenting multiple images of a target identity has also been shown to increase accuracy in identifying the target in surveillance video footage compared to when only a single image of the target is presented (Mileva & Burton, 2019). These studies demonstrate that multiple exposures to a face can enhance learning and recognition of a new identity, even when the faces are presented from different viewpoints (Hunnisett & Favelle, 2021).

Apart from multiple exposures, past research has also examined if different levels of variation during multiple exposures of an identity could affect identity learning (e.g., Susilo et al., 2018). For instance, a study by Ritchie and Burton (2017) examined the effect of within-person variability on face learning. In Experiment 1, participants were exposed to high variability (i.e., pictures taken several days apart, with different viewpoints and lighting conditions) and low variability (i.e., pictures taken the same day, with similar lighting conditions) images of unfamiliar identities and were tested on a name verification task. Experiment 2 was similar, but it involved a simultaneous face-matching task. The results of both experiments showed that high variability learning led to better identification. In line with this, it has been found that identification accuracy was higher when the two-image comparison presented during a sequential face-matching task was in a high variability condition compared to a low variability condition (Menon et al., 2015). Another study compared identity learning when viewing a 10-min video footage in low variability (i.e., video filmed on the same day with the same appearance and lighting) and high variability (i.e., video filmed on different days with different appearance and lighting) in children and adults (Baker et al., 2017). Children were more accurate in an identity-sorting task after viewing the video footage in the high variability condition compared to the low variability condition; however, this effect was weaker in adult participants. More recently, it has been shown that multiple exposures to identity in low variability conditions are not more effective than a single-image presentation for face learning (Matthews et al., 2024, but see Zhou et al, 2018). Altogether, the results of these studies highlight the importance of within-person variability in face learning and identification.

Although it is clear that multiple exposures to a face, especially in high variability conditions, improve the learning of a new identity for own-race faces, less research has explored whether this benefit extends to other-race faces. Recognition of other-race faces is usually more difficult compared to own-race faces, the so-called other-race effect (ORE; Meissner & Brigham, 2001). Consistent with this, research has shown that other-race faces are indeed more difficult to learn compared to own-race faces (Tüttenberg & Wiese, 2019; Wong et al., 2020). However, past research has shown that recognition accuracy for other-race face recognition increases when four different images of an identity are learned as compared to when four repetitions of the same image of an identity are learned (Cavazos et al., 2019). A similar pattern of results was also found when a line-up task was implemented (Matthews & Mondloch, 2018). These findings suggest that, similar to own-race face identification, exposure to different face instances of an identity could enhance other-race face recognition.

Previous research has shown that other-race face recognition usually relies more on external features (e.g., hairstyle) compared to internal features (e.g., shape of eyes) (Havard, 2021; Sporer & Horry, 2011; Wong et al., 2020). For instance, Wong et al. (2020) found that the ORE was observed only when internal features were presented independently, and this effect was eliminated when faces were shown with both internal and external features as a unified whole, demonstrating the importance of external features for other-race face recognition. Aligned with this, other research has indicated that the omission of external features leads to a decline in accuracy for recognizing other-race faces, as observed in old-new recognition paradigms (Sporer & Horry, 2011) and simultaneous face-matching tasks (Havard, 2021). Since other-race face recognition relies more on external features, it is possible that high within-person variability may not be beneficial for other-race face learning because high variability images involve consistent changes in external features that may hinder other-race face learning.

Recent research, however, seems to rule out this possibility. Specifically, a study by Zhou et al. (2018) examined whether different levels of variation during multiple exposures to other-race identities could affect identity learning. In the experiment, participants learned two own-race and two other-race identities either in low variability (i.e., a 1-min video filmed on the same day with the same appearance and lighting), high variability (i.e., a 1-min video filmed on three different days with different appearance and lighting), or a single image. The results revealed that both the low variability and the high variability conditions were beneficial for own-race identity learning when compared to the single-image presentation. However, only the high variability condition benefited other-race identity learning. These findings suggest that high variability learning improved identification for both own- and other-race identities. Despite these findings, only White participants were recruited for the study, which limits the generalizability of the results.

Thus, we aim to further examine the effect of high and low within-person variability exposure for both own- and other-race face learning using a cross-race design. This design is advantageous as it allows for an examination of whether these findings could be generalized across different racial groups, providing more comprehensive insights into identity learning. In addition, we implemented identity learning using photos instead of videos to assess whether the effect remains consistent across different mediums. Own- and other-race identities will be learned in high and low variability conditions, and identity recognition will first be tested by using a name verification task in Experiment 1 and then by an old-new recognition paradigm in Experiment 2. A name verification task, as in Ritchie and Burton (2017), was selected for Experiment 1 because it provides more ecologically valid results by simulating real-world scenarios where people need to recognize and associate names with faces. However, this task requires participants to learn names for both own- and other-race faces, introducing an additional cognitive load. To eliminate this extra cognitive burden, an old-new recognition paradigm was used for Experiment 2. This task was chosen because the benefits of multiple exposures with high variation may apply only to tasks with high memory demands (e.g., sequential face matching) but not to tasks without memory demands (e.g., simultaneous face matching) (Ritchie et al., 2021; Sandford & Ritchie, 2021).

Based on the findings discussed earlier, which showed that multiple exposures with high within-person variability are more advantageous for own-race identity learning compared to low within-person variability (Baker et al., 2017; Menon et al., 2015; Ritchie & Burton, 2017), we expect enhanced face learning for own-race identities learned in high variability conditions compared to low variability conditions. In addition, if other-race faces are predominantly identified using external features (Havard, 2021; Sporer & Horry, 2011; Wong et al., 2020), we would expect that recognition performance would be similar for identities learned in low- and high variability conditions. Conversely, if participants could successfully make use of the internal features of other-race faces when exposed to the high variability images with consistent changes in external features, we would expect that enhanced learning of other-race faces would occur for identities learned in high variability conditions than low variability conditions in line with Zhou et al. (2018), as high variation in face appearance leads to detailed encoding of internal features (Devue et al., 2021; Reedy & Devue, 2019).

Experiment 1

The current experiment aimed to investigate own- and other-race face learning in high and low within-person variability exposure. Experiment 1 partially replicates Ritchie and Burton's (2017) experiment, where participants learned identities in high- and low variability conditions and were tested with a name verification task. However, the current experiment included both own- and other-race identities in the task (i.e., White and Chinese), and we recruited both Chinese Malaysian and White participants for this experiment. In addition, this study builds on Zhou et al.'s (2018) findings by incorporating a cross-race design to explore whether this effect is consistent across different racial groups. We also used photos instead of videos in the identity learning phase to assess whether the benefits of high variability exposure remain consistent across different mediums and included a name verification task to examine whether high variability exposure could aid other-race face-name association.

Methods

Design. A mixed design was implemented. The within-subject factors were variability (high and low) and stimuli (Chinese and White), and the between-subject factor was participants' race (Chinese Malaysian and White). The dependent variable was the rate-correct score (RCS), a measure of efficiency in solving the task that combines accuracy and reaction times (Woltz & Was, 2006). We include this efficiency measure to avoid any potential speed-accuracy trade-offs (Gueugneau et al., 2017; Heitz, 2014; Liesefeld et al., 2015; Wickelgren, 1977). We use RCS as it has been shown to be more efficient in detecting an effect and accounting for a larger proportion of the variance compared to other integrative measures of speed and accuracy (Vandierendonck, 2017). RCS is calculated by the number of correct trials divided by the sum of reaction times for correct and incorrect trials. The value of RCS indicates the number of correct trials per second, whereas a higher value of RCS denotes higher efficiency.

Participants. Chinese Malaysian participants were recruited by advertising the study on the University of Nottingham Malaysia's social media page, while White participants were recruited through the Psychology Department's SONA system at Bournemouth University. In total, 125 Chinese Malaysian and 156 White participants took part in this experiment, but the final sample included 103 Chinese Malaysian (79 females, 2 others) and 91 White (76 females, 2 others) aged between 18 to 67 years ($M = 22.32$ years, $SD = 5.42$ years). Data from 87 participants were excluded from further analysis due to: (a) racial background being neither Chinese Malaysian or White (12); (b) median reaction time below 500 ms or accuracy rates below chance level (50%) (32); (c) inaccurate responses in the learning stage indicating inattentive behavior during the experiment (41); (d) accidental repetition of the experiment (1); or (e) familiar with more than half (i.e., more than five) of the identities presented in a single block (1).

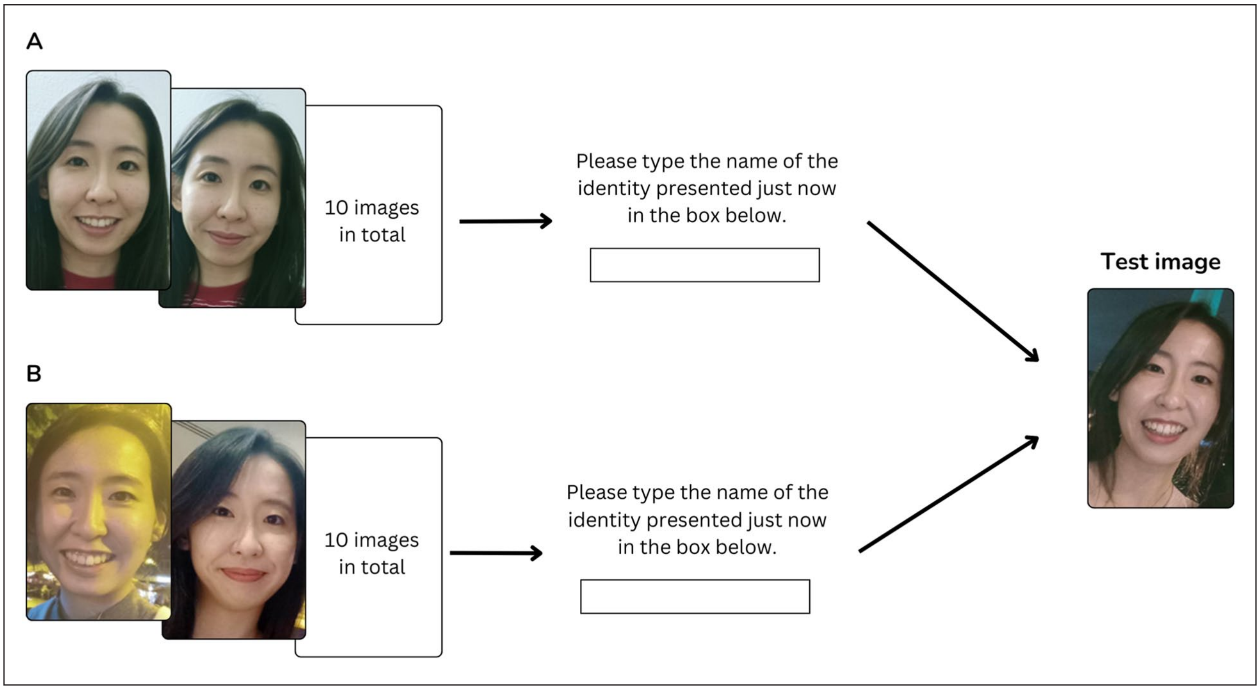

Figure 1. Sample stimuli for low variability (A) and high variability (B) conditions, featuring an identity that did not appear in the experiment. The test images were high variability images that were not used during the study phase. Actual stimuli used in experiment are not presented due to copyright restriction on the images.

An a priori power analysis was conducted using G*Power 3.1 (Faul et al., 2009) for a mixed ANOVA comparing the own- and other-race identities learned in high and low variability conditions of Chinese Malaysian and White participants. The effect size for variability was based on Ritchie and Burton (2017), where $\eta_p^2=.28$ and Ritchie et al. (2021), where $\eta_p^2=.52$ and .20, large effect size. A large effect size estimate ($\eta_p^2=.14$) was entered into the power analysis with the following parameters: alpha=.05, power=0.95. The power analysis suggested that $N=40$ is required to detect a difference between the variability conditions with 95% probability.

All participants provided informed consent to participate in the study. Participants were compensated with either course credits or RM5 for participation. The study was reviewed and approved by the Science and Engineering Research Ethics Committee (SEREC) at the University of Nottingham Malaysia (approval code: KSK270521).

Apparatus and materials. The White stimuli used in the face-learning task were identical to Ritchie and Burton (2017), which were kindly provided by the authors. Ten identities (five males and five females) were included for each stimulus race, totaling up to 20 identities. Identities used in the White stimuli consist of Australian celebrities (radio hosts, comedians, etc.), and identities used in the Chinese stimuli consist of Chinese celebrities (athletes, Esports players, etc.). Participants in this experiment should not be familiar with any of the identities shown in the task. In total, there were 20 high variability images and 10 low variability images for each identity (i.e., 10 high variability and 10 low variability images for the study phase, and 10 high variability images for the test phase). The high variability images differed in terms of the person (hairstyle, age, clothing, facial expression, etc.) and conditions (background, lighting, quality of image, etc.), whereas the low variability images differed in terms of facial expression and head angle but not in terms of hairstyle, age, clothing, and conditions (background, lighting, quality of image, etc.). For the Chinese stimuli, the high variability images were obtained by searching for the name of the celebrity on Google Images, whereas the low variability images were screenshots of interview videos found on YouTube by searching for the name of the celebrity. The images (260×390 pixels) were presented on a gray background. A sample of low and high variability images can be found in Figure 1.

Procedure

Testable (https://www.testable.org/) was used to run the online experiment (Rezlescu et al., 2020). The task was presented in two blocks, Chinese stimuli and White stimuli, and the presentation order of these blocks was randomized. Each block consists of two phases: learning and test. In the learning phase, 10 identities (10 images for each identity) were presented with their actual names above the image to be memorized. Faces were presented with their actual names (i.e., White faces were presented with White names and Chinese faces were presented with Chinese names), rather than common English names, as previous studies have found that associating faces with atypical names may influence the ORE (Stelter & Degner, 2018). Each image was presented for 5,000 ms with an intertrial interval of 500 ms. Five identities were presented in the low variability condition, and five were presented in the high variability condition. The identities used for high and low variability conditions were counterbalanced. The name of the identity remained on screen throughout the presentation of images of each identity. At the end of the presentation of each identity, participants were asked to type the name of the identity that they had viewed. This was done to ensure that participants were attentive during the learning phase. Participants were allowed to take breaks in between the presentation of identities if required. The learning phase took approximately 10 min.

The test phase of the task consists of a name verification task, which consists of 100 trials (10 identities × 10 trials). Images presented in the test phase were novel high variability images. In each trial, the name was presented for 1,500 ms, followed by the test images, which were presented until a response. There was an intertrial interval of 500 ms. Half of the trials were matched with the correct name, and the other half were mismatched (five matched trials and five mismatched trials for each identity). The names presented were only from the 10 identities' names, no novel name was introduced. Female names were only used for female identities, and male names were only used for male identities in the mismatched trials. Participants were asked to indicate if the name matched the image presented as quickly and as accurately as possible. The keys used for response were "z" and "m." The keys used for the "match" or "does not match" response, either right hand ("m") or left hand ("z"), were counterbalanced. The test phase took about 5 min to complete. At the end of each block, participants were asked if they were familiar with any of the identities shown in the task. The whole experiment lasted approximately 30 min.

Results

All data analysis was conducted using JASP (JASP Team, 2022). Data from participants who had typed the name of the identity with one incorrect letter during the learning phase were included in the analysis (e.g., typing Fiffi when the actual name is Fifi). For participants (two Chinese Malaysians) who reported familiarity with less than half of the identities shown in the task in a single block (i.e., one identity), test trials involving the familiar identity were removed prior to the analysis. Mixed ANOVAs were conducted to explore potential differences between own- and other-race identities learned in high and low variability conditions. The datasets generated are available in the Open Science Framework (OSF) repository (https://osf.io/s43tb/).

Efficiency. A 2 (variability: high vs. low) × 2 (face race: Chinese vs. White) × 2 (participant race: Chinese Malaysian vs. White) mixed ANOVA was conducted on the efficiency scores calculated by RCS (Figure 2). The analysis revealed a significant main effect of variability on efficiency, $F(1, 192)=25.010$, $p<.001$, $\eta_p^2=.115$. Efficiency for the high variability condition ($M=0.529$, $SD=0.241$) was greater compared to the low variability condition ($M=0.495$, $SD=0.221$). The analysis also revealed a significant main effect of face race, $F(1, 192)=78.159$, $p<.001$, $\eta_p^2=.289$, where White stimuli ($M=0.562$, $SD=0.245$) had higher efficiency compared to Chinese stimuli ($M=0.462$, $SD=0.206$). No main effect of participant race was found, $F(1, 192)=2.609$, $p=.108$, $\eta_p^2=.013$.

A significant interaction effect of face race and participant race was found, $F(1, 192)=70.254$, $p<.001$, $\eta_p^2=.268$. Chinese Malaysian participants showed no difference in efficiency between Chinese and White stimuli, $F(1, 102)=.120$, $p=.730$, $\eta^2=.001$, while White participants showed higher efficiency for White stimuli ($M=0.639$, $SD=0.233$) than Chinese stimuli ($M=0.431$, $SD=0.169$), $F(1, 90)=130.637$, $p<.001$, $\eta^2=.592$. In addition, Chinese Malaysian participants ($M=0.489$, $SD=0.208$) showed higher efficiency compared to White participants ($M=0.431$, $SD=0.169$) for Chinese stimuli, $F(1, 192)=4.40$, $p=.037$, $\eta^2=.022$. In contrast, for White stimuli, Chinese Malaysian participants showed lower efficiency ($M=0.494$, $SD=0.208$) compared to White participants ($M=0.639$, $SD=0.233$), $F(1, 192)=20.966$, $p<.001$, $\eta^2=.098$.

A significant interaction effect of variability and participant race was found, $F(1, 192)=6.181$, $p=.014$, $\eta_p^2=.031$. Chinese Malaysian participants showed higher efficiency for high variability condition ($M=0.516$, $SD=0.202$) than low variability condition ($M=0.467$, $SD=0.190$), $F(1, 102)=33.440$, $p<.001$, $\eta^2=.247$, while White participants showed no difference in efficiency for high and low variability conditions, $F(1, 90)=2.657$, $p=.107$, $\eta^2=.029$. In terms of the low variability condition, Chinese Malaysian participants ($M=0.467$, $SD=0.190$) showed lower efficiency compared to White participants ($M=0.527$, $SD=0.189$), $F(1, 192)=4.827$, $p=.029$, $\eta^2=.025$. No difference between Chinese Malaysian participants and White participants was found in the high variability condition, $F(1, 192)=.937$, $p=.334$, $\eta^2=.005$. Analysis showed no interaction effect of variability and face race, $F(1, 192)=3.376$, $p=.068$, $\eta_p^2=.017$, and variability, face race, and participant race, $F(1, 192)=1.923$, $p=.167$, $\eta_p^2=.010$.

Discussion

Overall, our results showed that participants performed better in terms of efficiency for identities learned in a high variability condition compared to the identities learned in a low variability condition. In addition, White participants

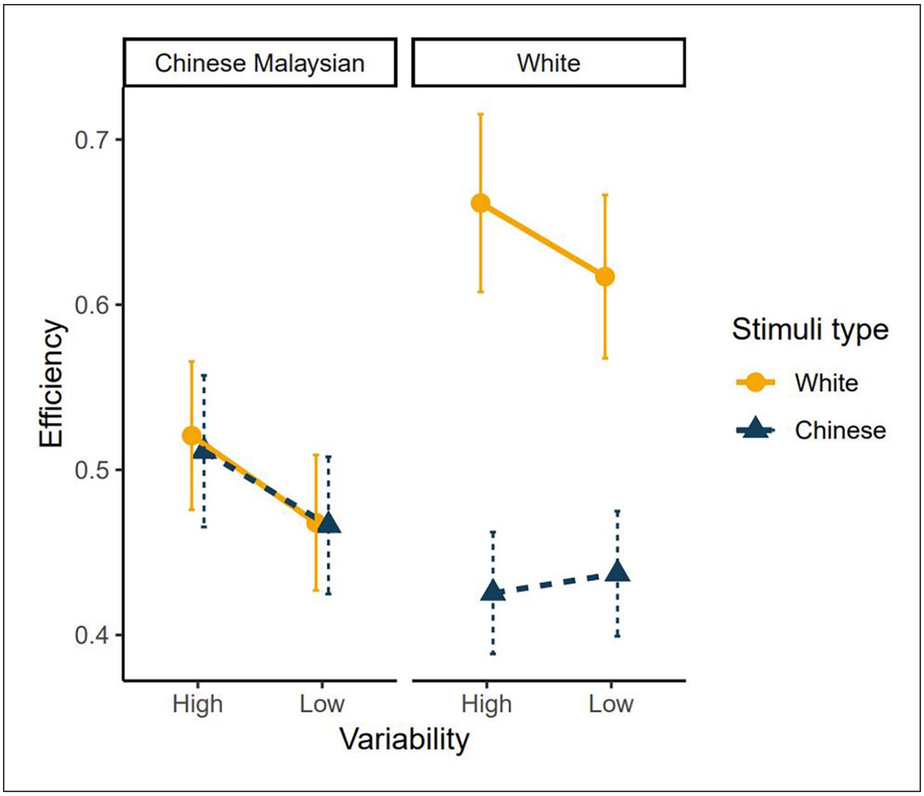

Figure 2. Efficiency plotted separately for White and Chinese Malaysian participants in Experiment 1.
Note. Error bars represent 95% confidence intervals.

performed better in terms of efficiency for White stimuli compared to Chinese stimuli. However, Chinese Malaysian participants performed equally well for White stimuli and Chinese stimuli. Although White participants presented the expected ORE for Chinese faces, which is consistent with past work (Meissner & Brigham, 2001; Wong et al., 2020), Chinese Malaysians did not present an ORE for White faces. This finding is in line with Tan et al. (2012), who found that Chinese Malaysian participants recognized East Asian and Western faces equally well (but see Estudillo et al., 2020). The absence of ORE for White faces may be due to high exposure to Western culture in Malaysia, as evident from the preference for Hollywood films over local films in Malaysia (Kit & Chuan, 2012; Sriganeshvarun & Abdul Aziz, 2019). This is aligned with the contact hypothesis, where we tend to develop a higher level of perceptual expertise for faces that are more often seen in our everyday lives (Rossion & Michel, 2011). Viewing Hollywood films may have increased Chinese Malaysian participants' perceptual expertise for White faces, which in turn reduced the ORE for White faces. However, it is important to note that while some studies propose that interracial exposure and the magnitude of the ORE are not significantly correlated (Wong et al., 2020), others have suggested a reduction in the ORE with increased exposure (Estudillo et al., 2020; Singh et al., 2022).

Since Chinese Malaysians did not show an ORE for White faces, only White participants could be used to examine if other-race identities (i.e., Chinese) learned in a high variability condition had better recognition compared to the other-race identities learned in a low variability condition. Based on the results and graphs in Figure 2 for White participants, there were no differences in efficiency for other-race identities learned in high variability conditions and low variability conditions.

One limitation of this experiment is that the name verification task requires participants to memorize the name and the face to perform accurately during the testing phase. However, White participants may be unfamiliar with Chinese names, which could deter face learning and name-matching accuracy for the Chinese identities. This is demonstrated by the high percentage of White participants who entered names inaccurately during the learning stage: 33 of the 41 participants who did so were White participants, whereas the remaining 8 were Chinese Malaysian participants. Therefore, in Experiment 2, we employed an old-new recognition paradigm that does not require precise name memorization during the testing phase.

Experiment 2

In Experiment 2, participants learned own- and other-race identities in high and low variability conditions as in

Experiment 1, but they were tested with an old-new face recognition paradigm as opposed to a name verification task. Similar to Experiment 1, we recruited Chinese Malaysian and *White* participants for this experiment.

Methods

Design. A mixed design was implemented. The within-subject factors were variability (high and low) and stimuli (Chinese and White), and the between-subject factor was participant race (Chinese Malaysian and White). Similar to Experiment 1, the dependent variable was efficiency, but in this experiment, we also included *d*-prime (d'), which evaluates participants' ability to distinguish between signal (stimuli) and noise (absence of stimuli) (Macmillan & Creelman, 2005; Stanislaw & Todorov, 1999). We included d' for Experiment 2 because, in contrast to Experiment 1, the old-new recognition paradigm enables the calculation of false alarm rates.

Participants. Chinese Malaysian participants were recruited by advertising the study on the University of Nottingham Malaysia's social media page, while White participants were recruited through the Psychology Department's SONA system at Bournemouth University. In total, 129 Chinese Malaysian and 135 White participants took part in this experiment, but the final sample included 95 Chinese Malaysian (63 females, 1 other) and 96 White (84 females, 2 others) aged between 18 to 67 years (M=21.59 years, SD=4.78 years). Data from 73 participants were excluded from further analysis due to: (a) racial background being neither Chinese Malaysian nor White (1); (b) median reaction time below 500 ms or accuracy rates below chance level (50%) (31); (c) inaccurate responses in the learning stage indicating inattentive behavior during the experiment (40); or (d) accidental repetition of the experiment (1).

The results of an a priori power analysis conducted using G*Power 3.1 (Faul et al., 2009) were as in Experiment 1. A large effect size (η_p^2=.14) was estimated and the power analysis suggested that N=40 is required to detect a difference between the variability condition with 95% probability. All participants have provided informed consent to participate in the study. Participants were compensated with either course credits or RM5 for participation. The study has been reviewed and approved by the SEREC at the University of Nottingham Malaysia (approval code: KSK270521).

Apparatus and materials. For both White and Chinese stimuli, the high variability images were obtained by searching for the name of the celebrity on Google Images, whereas the low variability images were screenshots of interview videos found on YouTube by searching for the name of the celebrity. We employed a new set of White and Chinese stimuli in this study to prevent familiarity with the stimuli used in Experiment 1. Twenty identities (10 males and 10 females) were included for each race, totaling up to 20 identities. Identities used in the White stimuli consist of American and Australian celebrities (athletes, models, television presenters, etc.), and identities used in the Chinese stimuli consist of Chinese celebrities (athletes, etc.). Participants recruited in this experiment should not be familiar with any of the identities shown in the task. In total, there were 15 high variability images and 10 low variability (i.e., 10 high variability and 10 low variability images for the study phase, and 5 high variability images for the test phase) images for each identity. The images (260×390 pixels) were presented on a gray background.

Procedure

Testable was used to run the online experiment (Rezlescu et al., 2020). The task was presented in two blocks, Chinese stimuli and White stimuli, in randomized order. Each block consists of two phases: learning and test. The learning phase was as in Experiment 1. In this study, six Chinese names were modified to facilitate name memorization in the learning phase (e.g., MoSheung to MoShen). The names were included in the experiment to aid participants in differentiating the faces and to ensure that participants were attentive during the learning phase. However, it was not required for participants to recognize the names during the test phase.

The test phase of the task consisted of a recognition memory task, which consisted of 100 trials (10 identities $\times$ 10 trials). Images presented in the test phase were novel high variability images. In each trial, the test images without names were presented until a response. There was an intertrial interval of 500 ms. Half of the trials were images of identities which have been presented in the learning stage, and the other half were novel identities. Participants were asked to indicate if the identity shown had been presented in the learning stage or not as quickly and as accurately as possible. The keys used for response were "z" and "m." The keys used for "have seen the identity in the learning stage" or "have not seen the identity before" responses, either the right hand ("m") or left hand ("z"), were counterbalanced. The test phase took about 5 min to complete. Identities used in the learning phase and novel identities in the test phase were counterbalanced. At the end of each block, participants were asked if they were familiar with any of the identities shown in the task. The whole experiment lasted approximately 30 min.

Results[1]

Data from participants who had typed the name of the identity with just one incorrect letter during the learning phase were included in the analysis. For participants who

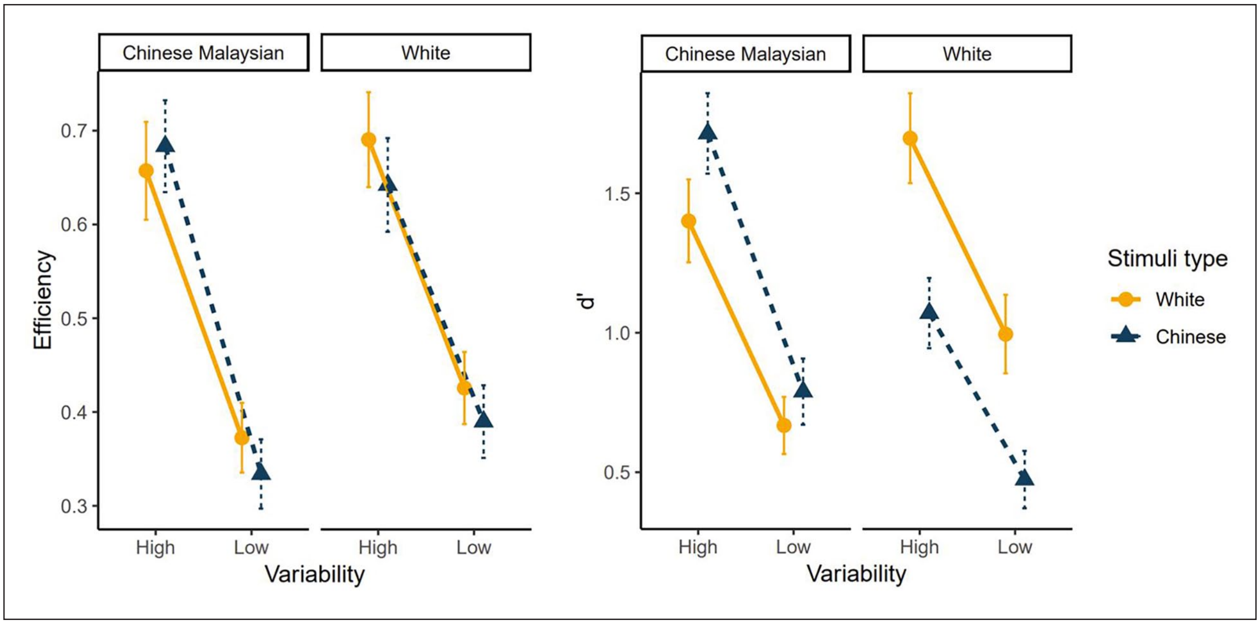

Figure 3. Efficiency and *d*-prime plotted separately for White and Chinese Malaysian participants in Experiment 2.
Note. Error bars represent 95% confidence intervals.

were familiar with fewer than half of the identities shown in the task (i.e., fewer than 5 out of 10 identities in each block) (four Chinese Malaysians), test trials involving the familiar identity were removed prior to the analysis. For the efficiency analysis, we conducted our analysis using only the trials featuring identities that were presented during the learning stage (i.e., hit trials) and excluded trials with novel identities (for a similar procedure, see Longmore et al., 2008). The reason was that the identities presented during the learning stage varied on two factors (high and low variability), while the distractors only varied on one factor (novel identities). Since the variability manipulation only applies to old but not new trials, *d*′ was calculated based on the same false alarm rate for high variability and low variability conditions. The psycho package in R was used to calculate *d*′ (Makowski, 2018). Mixed ANOVAs were conducted to explore potential differences between own- and other-race identities learned in high and low variability conditions. The datasets generated are available in the OSF repository (https://osf.io/s43tb/).

Efficiency. A 2 (variability: high vs. low) × 2 (face race: Chinese vs. White) × 2 (participant race: Chinese Malaysian vs. White) mixed ANOVA was conducted on efficiency calculated by RCS (Figure 3). The analysis revealed a significant main effect of variability, $F(1, 189) = 524.991$, $p < .001$, $\eta_p^2 = .735$. Efficiency for high variability condition ($M = 0.668$, $SD = 0.248$) was higher compared to the low variability condition ($M = 0.381$, $SD = 0.188$). No main effect of face race, $F(1, 189) = 3.862$, $p = .051$, $\eta_p^2 = .020$, or participant race, $F(1, 189) = 1.080$, $p = .300$, $\eta_p^2 = .006$, was found.

Results showed a significant interaction effect of variability and participant race, $F(1, 189) = 5.420$, $p = .021$, $\eta_p^2 = .028$. High variability condition ($M = 0.666$, $SD = 0.209$) had higher efficiency compared to low variability condition ($M = 0.408$, $SD = 0.163$) for White participants, $F(1, 95) = 248.565$, $p < .001$, $\eta^2 = .723$, and Chinese Malaysian participants (high variability condition: $M = 0.670$, $SD = 0.221$; low variability condition: $M = 0.353$, $SD = 0.152$), $F(1, 94) = 276.833$, $p < .001$, $\eta^2 = .747$. Chinese Malaysian participants and White participants showed no difference in efficiency in the high variability condition, $F(1, 189) = .017$, $p = .897$, $\eta^2 = 8.923\text{e-}5$. In the low variability condition, White participants ($M = 0.408$, $SD = 0.163$) showed higher efficiency compared to Chinese Malaysian participants ($M = 0.353$, $SD = 0.152$), $F(1, 189) = 5.692$, $p = .018$, $\eta^2 = .029$.

Analysis revealed no interaction effect of face race and participant race, $F(1, 189) = 2.128$, $p = .146$, $\eta_p^2 = .011$. No interaction effect was found between variability and face race, $F(1, 189) = 1.547$, $p = .215$, $\eta_p^2 = .008$, and between variability, face race, and participant race, $F(1, 189) = 3.359$, $p = .068$, $\eta_p^2 = .017$.

d*-Prime.* A 2 (variability: high vs. low) × 2 (face race: Chinese vs. White) × 2 (participant race: Chinese Malaysian vs. White) mixed ANOVA was conducted on *d*′ (Figure 3). The analysis revealed a significant main effect of variability on *d*′, $F(1, 189) = 540.223$, $p < .001$, $\eta_p^2 = .741$. *d*′ for the high variability condition ($M = 1.470$, $SD = 0.761$) was higher compared to the low variability condition ($M = 0.731$, $SD = 0.605$). The analysis also revealed a significant main effect of face race, $F(1, 189) = 17.028$,

$p \leq .001$, $\eta_p^2 = .083$, where White stimuli ($M = 1.191$, $SD = 0.791$) had higher *d′* compared to Chinese stimuli ($M = 1.011$, $SD = 0.760$). No effect of participant race was found, $F(1, 189) = 1.353$, $p = .246$, $\eta_p^2 = .007$. Results showed a significant interaction effect between variability and participant race, $F(1, 189) = 7.976$, $p = .005$, $\eta_p^2 = .040$, and between face race and participant race, $F(1, 189) = 83.901$, $p < .001$, $\eta_p^2 = .307$. No interaction effect of variability and face race was found, $F(1, 189) = .608$, $p = .436$, $\eta_p^2 = .003$.

A significant interaction was found between variability, face race, and participant race, $F(1, 189) = 7.108$, $p = .008$, $\eta_p^2 = .036$. To further explore this three-way interaction, we ran a 2 (variability: high vs. low) × 2 (face race: Chinese vs. White) ANOVA for Chinese Malaysian participants and White participants separately. For White participants, we found a significant main effect of variability, $F(1, 95) = 253.052$, $p < .001$, $\eta_p^2 = .727$, where the high variability condition ($M = 1.384$, $SD = 0.607$) had a higher *d′* compared to the low variability condition ($M = 0.734$, $SD = 0.502$). The analysis also revealed a significant main effect of face race, $F(1, 95) = 84.264$, $p < .001$, $\eta_p^2 = .470$, whereby White participants showed higher *d′* for White stimuli ($M = 1.346$, $SD = 0.695$) compared to Chinese stimuli ($M = 0.772$, $SD = 0.494$). No interaction effect of variability and face race was found, $F(1, 95) = 1.870$, $p = .175$, $\eta_p^2 = .019$.

For Chinese Malaysian participants, we found a significant interaction effect of variability and face race, $F(1, 94) = 5.656$, $p = .019$, $\eta_p^2 = .057$. Simple main effects analysis revealed that Chinese Malaysian participants showed higher *d′* for Chinese stimuli ($M = 1.715$, $SD = 0.708$) compared to White stimuli ($M = 1.401$, $SD = 0.730$) in the high variability condition, $F(1, 94) = 16.601$, $p < .001$, $\eta^2 = .150$, while no difference was found in the low variability condition, $F(1, 94) = 3.301$, $p = .072$, $\eta^2 = .034$. In addition, Chinese Malaysian participants showed higher *d′* for the high variability condition ($M = 1.715$, $SD = 0.708$) compared to the low variability condition ($M = 0.789$, $SD = 0.582$) for Chinese stimuli, $F(1, 94) = 250.078$, $p < .001$, $\eta^2 = .727$, and White stimuli, $F(1, 94) = 116.099$, $p < .001$, $\eta^2 = .553$ (high variability: $M = 1.401$, $SD = 0.730$; low variability: $M = 0.668$, $SD = 0.502$).

We also ran complementary analyses to explore the three-way interaction for each race and variability level. For Chinese stimuli, a 2 (variability: high vs. low) × 2 (participant race: Chinese vs. White) ANOVA revealed a significant interaction between variability and participant race, $F(1, 189) = 16.212$, $p < .001$, $\eta_p^2 = .079$. Simple main effects analysis revealed that for Chinese stimuli, Chinese Malaysian participants showed higher *d′* ($M = 1.715$, $SD = 0.708$) compared to White participants ($M = 1.071$, $SD = 0.621$) in the high variability condition, $F(1, 189) = 44.676$, $p < .001$, $\eta^2 = .191$, and the low variability condition, $F(1, 189) = 15.930$, $p < .001$, $\eta^2 = .078$ (Chinese Malaysian participants: $M = 0.789$, $SD = 0.582$; White participant: $M = 0.474$, $SD = 0.508$). Additionally, higher *d′* was found for the high variability condition ($M = 1.715$, $SD = 0.708$) compared to the low variability condition ($M = 0.789$, $SD = 0.582$) for Chinese Malaysian participants, $F(1, 94) = 250.078$, $p < .001$, $\eta^2 = .727$, and White participants, $F(1, 95) = 109.992$, $p < .001$, $\eta^2 = .537$ (high variability: $M = 1.071$, $SD = 0.621$; low variability: $M = 0.474$, $SD = 0.508$) in terms of Chinese stimuli.

For White stimuli, we found a significant main effect of variability, $F(1, 189) = 268.310$, $p < .001$, $\eta_p^2 = .587$, where the high variability condition ($M = 1.550$, $SD = 0.776$) had a higher *d′* compared to the low variability condition ($M = 0.832$, $SD = 0.627$). The analysis also revealed a significant main effect of participant race, $F(1, 189) = 12.121$, $p < .001$, $\eta_p^2 = .060$, where White participants ($M = 1.346$, $SD = 0.695$) had a higher *d′* compared to Chinese Malaysian participants ($M = 1.034$, $SD = 0.531$). No significant interaction effect of variability and participant race was found, $F(1, 189) = .123$, $p = .726$, $\eta_p^2 = 6.502\text{e-}4$.

Finally, we also ran a 2 (face race: Chinese vs. White) × 2 (participant race: Chinese vs. White) ANOVA for the high variability and the low variability conditions separately. In terms of the high variability condition, we found a significant interaction effect of face race and participant race, $F(1, 189) = 74.979$, $p < .001$, $\eta_p^2 = .284$. Simple main effects analysis revealed that in the high variability condition, Chinese Malaysian participants showed a higher *d′* ($M = 1.715$, $SD = 0.708$) compared to White participants ($M = 1.071$, $SD = 0.621$) for Chinese stimuli, $F(1, 189) = 44.676$, $p < .001$, $\eta^2 = .191$. In terms of White stimuli, a lower *d′* was found for Chinese Malaysian participants ($M = 1.401$, $SD = 0.730$) compared to White participants ($M = 1.697$, $SD = 0.795$), $F(1, 189) = 7.211$, $p = .008$, $\eta^2 = .037$. In addition, higher *d′* was found for Chinese stimuli ($M = 1.715$, $SD = 0.708$) compared to White stimuli ($M = 1.401$, $SD = 0.730$) for Chinese Malaysian participants, $F(1, 94) = 16.601$, $p < .001$, $\eta^2 = .150$. For White participants, it was found that Chinese stimuli ($M = 1.071$, $SD = 0.621$) had lower *d′* compared to White stimuli ($M = 1.697$, $SD = 0.795$), $F(1, 95) = 66.944$, $p < .001$, $\eta^2 = .413$, in the high variability condition.

In terms of the low variability condition, a significant interaction effect of face race and participant race was found, $F(1, 189) = 44.023$, $p < .001$, $\eta_p^2 = .189$. Simple main effects analysis revealed that in the low variability condition, Chinese Malaysian participants showed higher *d′* ($M = 0.789$, $SD = 0.582$) compared to White participants ($M = 0.474$, $SD = 0.508$) for Chinese stimuli, $F(1, 189) = 15.930$, $p < .001$, $\eta^2 = .078$. In terms of White stimuli, lower *d′* was found for Chinese Malaysian participants ($M = 0.668$, $SD = 0.502$) compared to White participants ($M = 0.995$, $SD = 0.694$), $F(1, 189) = 13.923$, $p < .001$, $\eta^2 = .069$. In addition, no difference in *d′* was found between Chinese stimuli and White stimuli for Chinese

Malaysian participants, $F(1, 94)=3.301$, $p=.072$, $\eta^2=.034$. For White participants, it was found that Chinese stimuli ($M=0.474$, $SD=0.508$) had lower d' compared to White stimuli ($M=0.995$, $SD=0.694$), $F(1, 95)=55.219$, $p<.001$, $\eta^2=.368$, in the low variability condition.

Discussion

Similar to Experiment 1, our results showed that participants performed better in terms of efficiency and d' for identities learned in the high variability condition compared to the identities learned in the low variability condition. Although our results showed no ORE for White and Chinese Malaysian participants in the efficiency measure, both White and Chinese Malaysian participants exhibited an ORE for other-race faces in the d' measure. Interestingly, the ORE exhibited by Chinese Malaysian participants was only observed in the high variability condition and was not evident in the low variability condition. The contrasting results between efficiency and d' measure could potentially be attributed to the fact that efficiency solely encompasses old trials (faces that were previously learned), whereas d' encompasses both old and new trials (faces that were novel). This suggests that the ORE may be more pronounced when making decisions to reject faces (i.e., indicating that a face was not seen before) as opposed to confirming familiarity (i.e., acknowledging that a face was previously learned). In addition, the efficiency score combines both accuracy and reaction time, whereas d' includes only accuracy. This suggests that the ORE may be less pronounced when reaction time is also considered.

Despite these findings, the precise reasons behind the presence of the ORE in Chinese Malaysian participants exclusively in the high variability condition and its absence in the low variability condition remain unclear. Based on our results, own- and other-race identities learned in the high variability condition had higher efficiency and d' compared to the low variability condition. This demonstrates that identities learned in high variability conditions benefited both own- and other-race face learning.

General discussion

We aimed to examine the effect of high and low within-person variability exposure for own- and other-race face learning. Own- and other-race identities were learned in high and low variability conditions and identity recognition was tested using a name verification task in Experiment 1 and an old-new recognition paradigm in Experiment 2.

We found enhanced own-race face learning for identities learned in the high variability condition compared to the low variability condition across Experiment 1 and Experiment 2. This finding is in line with previous work, which found that multiple exposures to own-race faces in high within-person variability stimuli sets is more advantageous for identity learning as compared to low within-person variability, as demonstrated in a face-matching task (Menon et al., 2015; Ritchie & Burton, 2017), name verification task (Ritchie & Burton, 2017), and identity-sorting task (Baker et al., 2017).

To examine if identities learned in a high variability condition enhanced other-race face learning compared to the identities learned in a low variability condition, we mainly examined White participants in Experiment 1, as Chinese Malaysian participants did not exhibit an ORE for White faces. In Experiment 1, we found no difference in the performance of other-race face learning for identities learned in the high variability condition and the low variability condition. However, we implemented a name verification task in Experiment 1, which required participants to precisely memorize the name and the face to perform accurately during the testing phase. While Chinese participants may be familiar with White names (Kit & Chuan, 2012; Sriganeshvarun & Abdul Aziz, 2019), White participants may be unfamiliar with Chinese names, which could deter face and name-matching accuracy for the Chinese identities. This is demonstrated by the high percentage of White participants who entered the names inaccurately during the learning stage: 33 of the 41 participants who did so were White participants, whereas the remaining eight were Chinese Malaysian participants.

Instead of name verification, in Experiment 2, an old-new recognition paradigm was implemented. Experiment 2 revealed that identities learned in the high variability condition benefited other-race face learning in comparison to the low variability condition. While it has been demonstrated that exposure to identities with within-person variability can improve other-race face recognition compared to a single image of identities (Cavazos et al., 2019; Matthews & Mondloch, 2018), our findings indicate that identities learned in high variability condition could further improve other-race face recognition compared to identities learned in low variability condition, which is in line with the findings of Zhou et al. (2018). This suggests that different levels of variation during multiple exposures to other-race faces could affect identity learning, where higher variation of faces would lead to improved other-race face learning.

Our findings revealed that even though there were consistent changes in external features in the high variability condition, and previous research suggests that external features are typically prioritized when processing other-race faces (Havard, 2021; Sporer & Horry, 2011; Wong et al., 2020), participants in our study were able to rely on the internal features of other-race faces in the high variability condition when the external features were consistently changing. According to the cost-effective mechanism for face learning (Devue et al., 2021; Reedy & Devue, 2019), variability in face appearance could lead to detailed encoding of internal features and enhanced learning of faces. Altogether, our results across Experiment 1 and Experiment 2 suggest that the identities learned in high variability conditions may only benefit other-race face recognition but not the association of other-race faces and names when the names are unfamiliar. In addition, high variability benefits own-race face recognition and the association of own-race faces and names.

Our results also show inconsistencies in the presence of the ORE across Experiment 1 and Experiment 2 and in different measures (i.e., efficiency and *d′*). In Experiment 1, White participants exhibited the ORE for Chinese faces, while Chinese Malaysian participants did not exhibit the ORE for White faces. As discussed earlier, this may be due to high exposure to Western culture in Malaysia (Kit & Chuan, 2012; Sriganeshvarun & Abdul Aziz, 2019). In Experiment 2, we found no ORE for either White or Chinese Malaysian participants in the efficiency measure, while all participants exhibited the ORE in the *d′* measure. The contrasting results between the efficiency and *d′* measures could be because efficiency solely encompasses old trials (faces that were previously learned), while *d′* encompasses both old and new trials (faces that were novel). This suggests that the ORE may be more pronounced when making decisions to reject faces (i.e., indicating that a face was not seen before) as opposed to confirming familiarity (i.e., acknowledging that a face was previously learned). In addition, the efficiency score combines both accuracy and reaction time, whereas *d′* includes only accuracy, suggesting that the ORE may be less pronounced when reaction time is also considered. Thus, based on results from both experiments, the presence of ORE may depend on the task characteristics and the measures used.

Interestingly, we also found that the ORE exhibited by Chinese Malaysian participants was only observed in the high variability condition and was not evident in the low variability condition in Experiment 2. However, the precise reasons behind the presence of the ORE in Chinese Malaysian participants exclusively in the high variability condition remain unclear. It is possible that a combination of factors, including exposure to White faces in Malaysia, differences in tasks and samples across Experiments 1 and 2, and stimuli selection, might explain this finding.

Despite our findings, our study is subject to several limitations. First, we did not incorporate an eye-tracking task to validate participants' attention to specific facial features, whether internal or external. The absence of eye-tracking data limits our ability to draw strong conclusions about the participants' reliance on internal features in the high variability images. Including an eye-tracking measure would provide fixation data on the external and internal features of the face when low and high variability images are presented. Second, the stimuli comprising White faces obtained from Ritchie and Burton (2017) and additional stimuli generated for Experiments 1 and 2 were not controlled for levels of variability across faces. Thus, it is possible that the identities employed in our experiment exhibit varying degrees of variability in the high variability condition. A pilot study could be conducted in which participants rate the level of variability for each set of high variability images for each identity before the actual study. These limitations underscore the need for future research to address these factors and enhance the robustness of our findings.

In sum, we found enhanced own-race face learning for identities learned in a high variability condition compared to a low variability condition. Our results revealed that identities learned in high variability conditions benefit only other-race face recognition, but not face-name association of other-race faces, as compared to identities learned in low variability conditions. This suggests that high within-person variation during multiple exposures to faces could lead to detailed encoding of internal features, which refines the resolution of the representation not only for own-race faces but also for other-race faces.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Siew Kei Kho https://orcid.org/0000-0001-5675-9655
Hoo Keat Wong https://orcid.org/0000-0002-0387-3590
Alejandro J. Estudillo https://orcid.org/0000-0002-7760-318X

Data accessibility statement

The data from the present experiment are publicly available at the Open Science Framework website: https://osf.io/s43tb/.

Supplemental material

The Supplementary Material is available at: qjep.sagepub.com.

Note

1. Analyses conducted on median reaction for correct trials and criterion for Experiment 2 and on efficiency for Experiment 1 and 2 could be found in Supplemental Material.

References

Andrews, S., Jenkins, R., Cursiter, H., & Burton, A. M. (2015). Telling faces together: Learning new faces through exposure to multiple instances. *Quarterly Journal of Experimental Psychology*, *68*(10), 2041–2050. https://doi.org/10.1080/17470218.2014.1003949

Baker, K. A., Laurence, S., & Mondloch, C. J. (2017). How does a newly encountered face become familiar? The effect of within-person variability on adults' and children's perception of identity. *Cognition*, *161*, 19–30. https://doi.org/10.1016/j.cognition.2016.12.012

Bindemann, M., & Sandford, A. (2011). Me, myself, and I: Different recognition rates for three photo-IDs of the same person. *Perception*, *40*(5), 625–627. https://doi.org/10.1068/p7008

Bruce, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology*, *73*(1), 105–116. https://doi.org/10.1111/j.2044-8295.1982.tb01795.x

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied*, *7*(3), 207–218. https://doi.org/10.1037/1076-898X.7.3.207

Burton, A. M., Jenkins, R., Hancock, P. J. B., & White, D. (2005). Robust representations for face recognition: The power of averages. *Cognitive Psychology*, *51*(3), 256–284. https://doi.org/10.1016/j.cogpsych.2005.06.003

Cavazos, J. G., Noyes, E., & O'Toole, A. J. (2019). Learning context and the other-race effect: Strategies for improving face recognition. *Vision Research*, *157*, 169–183. https://doi.org/10.1016/j.visres.2018.03.003

Devue, C., de Sena, S., & Wright, J. (2021). *A cost-efficient face learning mechanism: The impact of stability in appearance on the resolution of facial representations*. https://doi.org/10.31234/osf.io/fqd3v

Dowsett, A. J., Sandford, A., & Burton, A. M. (2016). Face learning with multiple images leads to fast acquisition of familiarity for specific individuals. *Quarterly Journal of Experimental Psychology*, *69*(1), 1–10. https://doi.org/10.1080/17470218.2015.1017513

Estudillo, A. J. (2012). Facial memory: The role of the pre-existing knowledge in face processing and recognition. *Europe's Journal of Psychology*, *8*(2), 231–244. https://doi.org/10.5964/ejop.v8i2.455

Estudillo, A. J., & Bindemann, M. (2014). Generalization across view in face memory and face matching. *I-Perception*, *5*(7), 589–601. https://doi.org/10.1068/i0669

Estudillo, A. J., Lee, J. K. W., Mennie, N., & Burns, E. (2020). No evidence of other-race effect for Chinese faces in Malaysian non-Chinese population. *Applied Cognitive Psychology*, *34*(1), 270–276. https://doi.org/10.1002/acp.3609

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Favelle, S. K., Palmisano, S., & Avery, G. (2011). Face viewpoint effects about three axes: The role of configural and featural processing. *Perception*, *40*(7), 761–784. https://doi.org/10.1068/p6878

Gueugneau, N., Pozzo, T., Darlot, C., & Papaxanthis, C. (2017). Daily modulation of the speed–accuracy trade-off. *Neuroscience*, *356*, 142–150. https://doi.org/10.1016/j.neuroscience.2017.04.043

Havard, C. (2021). The importance of internal and external features in matching own and other race faces. *Perception*, *50*(10), 861–875. https://doi.org/10.1177/03010066211043464

Heitz, R. P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Frontiers in Neuroscience*, *8*, 150. https://doi.org/10.3389/fnins.2014.00150

Hunnisett, N., & Favelle, S. (2021). Within-person variability can improve the identification of unfamiliar faces across changes in viewpoint. *Quarterly Journal of Experimental Psychology*, *74*(11), 1873–1887. https://doi.org/10.1177/17470218211009771

JASP Team. (2022). *JASP* (Version 0.16.3) [Computer software]. https://jasp-stats.org/

Jenkins, R., & Burton, A. M. (2011). Stable face representations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1571), 1671–1683. https://doi.org/10.1098/rstb.2010.0379

Johnston, R. J., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory*, *17*(5), 577–596. https://doi.org/10.1080/09658210902976969

Kit, L. C., & Chuan, L. W. (2012). Hollywood films and local films: A quantitative study on malaysian young generations' preference. *Procedia – Social and Behavioral Sciences*.

Liesefeld, H. R., Fu, X., & Zimmer, H. D. (2015). Fast and careless or careful and slow? Apparent holistic processing in mental rotation is explained by speed-accuracy trade-offs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1140–1151. https://doi.org/10.1037/xlm0000081

Longmore, C. A., Liu, C. H., & Young, A. W. (2008). Learning faces from photographs. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 77–100. https://doi.org/10.1037/0096-1523.34.1.77

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Lawrence Erlbaum Associates.

Makowski, D. (2018). The psycho package: An efficient and publishing-oriented workflow for psychological science. *Journal of Open Source Software*, *3*(22), 470. https://doi.org/10.21105/joss.00470

Matthews, C. M., & Mondloch, C. J. (2018). Finding an unfamiliar face in a line-up: Viewing multiple images of the target is beneficial on target-present trials but costly on target-absent trials. *British Journal of Psychology*, *109*(4), 758–776. https://doi.org/10.1111/bjop.12301

Matthews, C. M., Ritchie, K. L., Laurence, S., & Mondloch, C. J. (2024). Multiple images captured from a single encounter do not promote face learning. *Perception*, *53*(5–6), 299–316. https://doi.org/10.1177/03010066241234034

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

Menon, N., White, D., & Kemp, R. I. (2015). Variation in photos of the same face drives improvements in identity verification. *Perception*, *44*(11), 1332–1341. https://doi.org/10.1177/0301006615599902

Mileva, M., & Burton, A. M. (2019). Face search in CCTV surveillance. *Cognitive Research: Principles and Implications*, *4*, 37. https://doi.org/10.1186/s41235-019-0193-0

Reedy, M., & Devue, C. (2019). *New perspective on face learning: Stability modulates resolution of facial representations in the optimal observer*. https://doi.org/10.31219/osf.io/6s8c7

Rezlescu, C., Danaila, I., Miron, A., & Amariei, C. (2020). More time for science: Using Testable to create and share behavioral experiments faster, recruit better participants, and engage students in hands-on research. *Progress in Brain Research*, *253*, 243–262. https://doi.org/10.1016/bs.pbr.2020.06.005

Ritchie, K. L., & Burton, A. M. (2017). Learning faces from variability. *Quarterly Journal of Experimental Psychology*, *70*(5), 897–905. https://doi.org/10.1080/17470218.2015.1136656

Ritchie, K. L., Kramer, R. S. S., Mileva, M., Sandford, A., & Burton, A. M. (2021). Multiple-image arrays in face matching tasks with and without memory. *Cognition*, *211*, 104632. https://doi.org/10.1016/j.cognition.2021.104632

Rossion, B., & Michel, C. (2011). An experience-based holistic account of the other-race face effect. In A. Calder, G. Rhodes, M. Johnson, & J. Haxby (Eds.), *Oxford handbook of face perception* (pp. 215–244). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199559053.013.0012

Sandford, A., & Ritchie, K. L. (2021). Unfamiliar face matching, within-person variability, and multiple-image arrays. *Visual Cognition*, *29*(3), 143–157. https://doi.org/10.1080/13506285.2021.1883170

Singh, B., Mellinger, C., Earls, H. A., Tran, J., Bardsley, B., & Correll, J. (2022). Does cross-race contact improve cross-race face perception? A meta-analysis of the cross-race deficit and contact. *Personality and Social Psychology Bulletin*, *48*(6), 865–887. https://doi.org/10.1177/01461672211024463

Sinha, P., Balas, B., Ostrovsky, Y., & Russell, R. (2006). Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE*, *94*(11), 1948–1961. https://doi.org/10.1109/JPROC.2006.884093

Sporer, S. L., & Horry, R. (2011). Recognizing faces from ethnic in-groups and out-groups: Importance of outer face features and effects of retention interval. *Applied Cognitive Psychology*, *25*(3), 424–431. https://doi.org/10.1002/acp.1709

Sriganeshvarun, N., & Abdul Aziz, M. M. (2019). Local films or hollywood films: Malaysian's preferences. *INTI Journal*, *2019*, 45.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Stelter, M., & Degner, J. (2018). Recognizing emily and latisha: Inconsistent effects of name stereotypicality on the other-race effect. *Frontiers in Psychology*, *9*(April), 1–11. https://doi.org/10.3389/fpsyg.2018.00486

Susilo, T., Ritchie, K. L., & Devue, C. (2018). *Within-person variability promotes learning of internal facial features and facilitates perceptual discrimination and memory*. https://doi.org/10.31219/osf.io/5scnm

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

Tüttenberg, S. C., & Wiese, H. (2019). Learning own- and other-race facial identities from natural variability. *Quarterly Journal of Experimental Psychology*, *72*(12), 2788–2800. https://doi.org/10.1177/1747021819859840

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior Research Methods*, *49*(2), 653–673. https://doi.org/10.3758/s13428-016-0721-5

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

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, *41*(1), 67–85. https://doi.org/10.1016/0001-6918(77)90012-9

Woltz, D. J., & Was, C. A. (2006). Availability of related long-term memory during and after attention focus in working memory. *Memory and Cognition*, *34*(3), 668–684. https://doi.org/10.3758/BF03193587

Wong, H. K., Stephen, I. D., & Keeble, D. R. T. (2020). The own-race bias for face recognition in a multiracial society. *Frontiers in Psychology*, *11*, 208. https://doi.org/10.3389/fpsyg.2020.00208

Zhou, X., Matthews, C. M., Baker, K. A., & Mondloch, C. J. (2018). Becoming familiar with a newly encountered face: Evidence of an own-race advantage. *Perception*, *47*(8), 807–820. https://doi.org/10.1177/0301006618783915