Japanese Psychological Research
THE JAPANESE PSYCHOLOGICAL ASSOCIATION

Japanese Psychological Research
2021
doi: 10.1111/jpr.12327

Own-Race Advantage in Visual Working Memory for Faces Reflects Enhanced Storage Capacity and Quick Encoding[1]

YUKA NISHIMURA* *Kwansei Gakuin University*

HIROYUKI TSUDA *Keio University*

HIROKAZU OGAWA *Kwansei Gakuin University*

Abstract: Previous research has shown that working memory (WM) performance for own-race faces is better than for other-race faces. We focused on the storage capacity and encoding rate to identify WM characteristics that facilitate own-race face recognition. We investigated WM's temporal dynamics for own- and other-race faces to separately identify the contribution of storage capacity and encoding rate on the own-race advantage in WM. We presented Asian participants with Asian faces as own-race faces and Black faces as other-race faces in two experiments. Experiments 1 and 2 indicated a higher storage capacity for own-race faces, and Experiment 2 also indicated an increased encoding rate for own-race faces when backward masking was used. Moreover, there was no association between storage capacity and encoding rate. These findings suggest that both storage capacity and encoding rate independently contribute to the cross-race effect in WM.

Key words: cross-race effect, visual working memory, face recognition.

The improved recognition of own-race faces compared to other-race faces is known as the cross-race effect (Hancock & Rhodes, 2008; Hayward, Crookes, & Rhodes, 2013). Reduced recognition of other-race faces is problematic in many real-world situations such as social interactions, eyewitness testimony, and person identification at airports and border checkpoints (Behrman & Davey, 2001; Meissner, Susa, & Ross, 2013). Most studies investigating the cross-race effect have used long-term recognition memory paradigms (Meissner & Brigham, 2001). However, the cross-race effect is related to a variety of cognitive processes other than long-term memory, including perception (Megreya, White, & Burton, 2011; G. Zhou et al., 2018) and attention (Hugenberg, Young, Bernstein, & Sacco, 2010).

Recent research has suggested that mechanisms involved in the early stages of face

---

*Correspondence concerning this article should be sent to: Yuka Nishimura, Department of Psychological Sciences, Graduate School of Humanities, Kwansei Gakuin University, Uegahara-1bancho, Hyogo 662-8501, Japan. (E-mail: yuka.nishimura@outlook.jp)

[1] This research was supported by JSPS KAKENHI Grant JP18K03191.

Commercial relationships: None.

© 2021 Japanese Psychological Association. Published by John Wiley & Sons Australia, Ltd.

processing, including encoding and working memory, contribute to the cross-race effect in long-term memory (Hugenberg et al., 2010; Levin, 2000; Stelter & Degner, 2018). For example, some studies have indicated a positive correlation between differences in face encoding and the cross-race effect in long-term memory (DeGutis, Mercado, Wilmer, & Rosenblatt, 2013; Susa, Meissner, & de Heer, 2010; Wiese, Kaufmann, & Schweinberger, 2014). Stelter and Degner (2018) suggested that WM's and long-term memory's cross-race effects are correlated. Working memory serves as an intermediate system that links early processing stages (perception and attention) to the memory system, and is involved in developing novel representations in long-term memory (Fukuda & Vogel, 2019; Ranganath, Cohen, & Brozinsky, 2005). Therefore, understanding the WM mechanism leading to the cross-race effect is expected to provide critical insights into the cross-race effect in online perception, attention, and long-term memory.

WM performance is better for own-race than for other-race faces (Valentine & Bruce, 1986). Several studies have highlighted the relevance of memory capacity in the own-race advantage in WM tasks. For example, Stelter and Degner (2018) examined differences in WM performance between own- and other-race faces using a N-back task, a self-ordered pointing task, and a change-detection task (note that the N-back task and the self-ordered pointing task measure executive functions such as selective attention and interference control, as well as WM's storage capacity, whereas the change-detection task is a purer measure of WM's storage capacity). They found that own-race faces were consistently remembered better across all three tasks, providing converging evidence for higher WM capacity for own-race, compared to other-race faces. Neuroimaging studies also support the higher storage capacity of WM for own-race faces. Sessa and Dalmaso (2016) measured event-related potentials during a change-detection task. They reported that the amplitude of the SPCN (Sustained Posterior Contralateral Negativity), an indicator of WM maintenance for tracking visual information retained in WM (Klaver, Talsma, Wijers, Heinze, & Mulder, 1999), is more significant for own-race faces than for other-race faces. This finding suggests that WM stores more information about own-race than for other-race faces.

Other studies have suggested that increased encoding rate explains the own-race advantage in WM tasks. Marcon and colleagues (Marcon, Meissner, Frueh, Susa, & MacLin, 2010) investigated how a sample item's presentation time affects delayed face matching. They reported significant cross-race effects in 100- and 500-ms conditions, but not when the presentation time was 1,000 or 1,500 ms, suggesting that face encoding is more accurate and/or rapid for own-race faces than for other-race faces. X. Zhou, Mondloch, and Emrich (2018) examined the nature of WM's cross-race effect by measuring memory recall precision. They used probabilistic mixture modeling of visual WM (van den Berg, Shin, Chou, George, & Ma, 2012; Zhang & Luck, 2008) and reported that recall precision was reduced for other-race faces, but not for own-race faces under a short encoding time of 200 ms. In contrast, a cross-race effect was indicated by increased guess rate for other-race faces, but not by reduced recall precision under a long encoding time of 1,500 ms. These studies demonstrate the influence of the initial memory formation stage on face recognition performance and suggest that the cross-race effect in WM tasks is caused, at least in part, by a failure to encode and consolidate other-race faces rapidly into WM. Long-term face memory research has also documented the crucial role of the encoding process on the cross-race effect (Herzmann, Minor, & Adkins, 2017; Herzmann, Willenbockel, Tanaka, & Curran, 2011).

In sum, studies on the cross-race effect in WM have suggested two putative mechanisms underlying the own-race advantage: (a) higher WM *capacity* for own-race faces; and (b) an increased *encoding rate* for own-race faces into WM. Although these are not necessarily mutually exclusive, it is essential to clarify the relationship between them because the two mechanisms might not function independently. The encoding rate could be a fundamental factor in the own-race advantage, and differences in memory

© Japanese Psychological Association 2021.

capacity for own- and other-race faces might be an artifact created by insufficient encoding time. The reduced memory precision for other-race faces caused by insufficient encoding time (X. Zhou et al., 2018) supports the view that reduced recognition of other-race faces under a short encoding-time (Marcon et al., 2010) was caused by reduced fidelity of WM representations for other-race faces, rather than a failure in the storage of other-race faces. Stelter and Degner (2018) presented faces for 1,500 ms (Set Size 3), 2,000 ms (Set Size 4), or 2,500 ms (Set Size 5) using a change-detection task and reported that WM capacity was higher for own-race faces. Because the presentation time of 500 ms per face might be insufficient to fully encode a face (Curby, Glazek, & Gauthier, 2009; Eng, Chen, & Jiang, 2005; Gao & Bentin, 2011), the difference in the change-detection performance in Stelter and Degner (2018) could have been caused by encoding limits rather than capacity limits. If the limits at the initial encoding and consolidation stages play a significant role in the cross-race effect, then we would expect little or no difference in WM capacity for own- and other-race faces if participants were given ample encoding time. X. Zhou et al. (2018) used guess rate estimates and suggested that the number of stored faces in WM was larger for own-race faces when the encoding time was as long as 1,500 ms for a face. However, whether the guess rate analysis provides a reliable estimate of WM capacity is debatable (W. J. Ma, 2018). Therefore, whether WM capacity depends on the race of a face remains unclear, and it is as yet unknown whether the bottleneck in the encoding stage thoroughly explains the cross-race effect in WM.

This study evaluated WM capacity and encoding rate for face images to identify the underlying causes of WM's cross-race effect. If differences in encoding rate are a significant bottleneck for creating WM representations of other-race faces, then the own-race advantage should disappear given a sufficiently long sample array presentation time. On the other hand, if differences in WM's storage capacity caused the cross-race effect, then superior performance for own-race faces should be evident under a long presentation condition, in which encoding limits are minimal. We conducted two experiments in which WM capacity was measured at several time points during the sample array presentation to address this issue. The duration of face stimuli presentation was either 250, 500, 750, 1,000, or 1,500 ms. We calculated Cowan's K (Cowan, 2001) for each condition, and curve fitting was applied to describe the increase in K as a function of the presentation time. We separately estimated the encoding rate, defined as the curve's slope at time zero, and WM capacity, defined as the curve's height. We expected to test how the capacity and/or the encoding rate contributed to WM advantage for own-race faces.

In this study, Asians consisting of Japanese and several Chinese/Korean participants were presented with Asian faces as own-race faces and Black faces as other-race faces. In Experiment 1, face stimuli were presented without backward masking following the sample array, and in Experiment 2, backward masking was added to control the time available for encoding face stimuli strictly. We found higher WM capacity for the own-race than other-race faces in both experiments. In addition, an increase in the encoding rate for own-race faces was also observed in Experiment 2.

Experiment 1

We examined how the race of a face and the presentation duration of a sample array affected WM performance to identify the time course of WM's cross-race effect. We used a variant of the change-detection paradigm, in which participants were presented with two face images of the same race, followed by a test face that was presented after a 1-s delay. Participants were asked to respond as to whether or not the test face matched either of the two sample faces. A concurrent digit memory task was included to suppress verbal encoding of faces.

Method

Participants. Students (n = 30, eight men and 22 women, mean age 19.8 years, age range 18–23 years), including 26 Japanese students,

© Japanese Psychological Association 2021.

and four students from China and Korea, who were enrolled in Kwansei Gakuin University participated in the experiment. All the participants had normal or corrected-to-normal visual acuity. The experimental procedure was approved by the Kwansei Gakuin University Institutional Review Board for Behavioral Research with Human Participants. All participants gave their written informed consent before taking part in the study. Participants received course credits for participating in the study.

Stimuli and apparatus. We used 40 images of Asian faces (20 men and 20 women) as own-race face stimuli, and 40 images of Black faces (20 men and 20 women) as other-race face stimuli. The face images were taken from the Chicago Face Database (D. S. Ma, Correll, & Wittenbrink, 2015). The front-facing faces were presented with neutral expressions. None of the participants had seen these faces before the experiment. The images were normalized to equate for mean luminance and cropped with an oval-shaped mask to remove the hair. They were displayed at a size of 3.11° (width) × 4.11° (height) of a visual angle from a viewing distance of 57 cm. Stimuli were presented on a 24.5-in. ASUS monitor (ROG SWIFT PG258Q, with a screen resolution of 1,920 × 1,080 pixels). The experiment was controlled by PsychoPy v1.85.3 (Peirce, 2007).

Design and procedure. Figure 1a shows the procedure of a single trial. First, a fixation cross was presented at the center of the screen for 1,000 ms, followed by a three-digit number presented at the center of the screen for 1,000 ms, which the participants memorized. Then, the central fixation cross was presented again for 1,000 ms, followed by two face images presented on each side of the fixation cross for a variable duration. The sample array presentation duration was randomized in each trial such that it was either 250, 500, 750, 1,000, or 1,500 ms. A three-digit number was presented again, after a blank period of 1,000 ms with the central fixation cross. The participants reported whether the digit matched the digit that was presented at the beginning of the trial. A test face was presented at the center of the screen after the response. The participants reported whether the test face matched any of the two preceding faces in that trial.

Own- and other-race faces were tested in separate trial blocks. The order of the blocks was randomized. The sample array's face pairs were randomly determined across trials and participants, and they were always of the same sex. Half the trials were "old" trials in which the test face matched one of the two sample faces, and

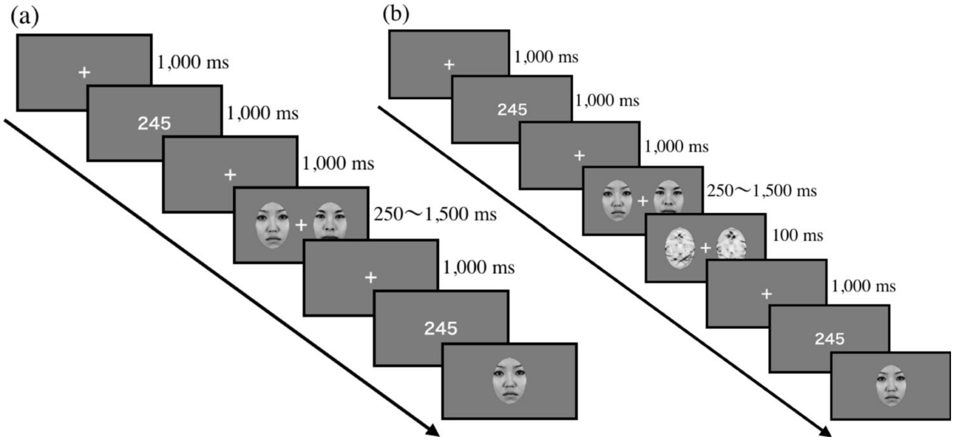

Figure 1 The procedure of a single trial in Experiment 1 (a) and Experiment 2 (b).

© Japanese Psychological Association 2021.

the other half were “new” trials. The test face used in a “new” trial could appear as a to-be-remembered face in other trials. Each participant completed a 1-hr session consisting of 480 trials (48 trials per combined duration × race conditions).

After completing the task, each participant’s experience of interacting with other-race people was assessed by using a questionnaire. Participants responded with “yes” or “no” to questions that inquired (a) if they have many other-race acquaintances; (b) if other-race people are living in their neighborhood; (c) if they meet other-race people in their own time; (d) if there were other-race students in their high school; and (e) if they had lived in a place where most of the residents were other-race people. It took about 3 min to complete the questionnaire.

Analysis. The Cowan’s K was calculated for each experimental condition by subtracting the false alarm rate from the hit rate and multiplying it by the set size, which was 2. We fitted the equation below to model the temporal dynamics of K as a function of presentation duration:

$$K = a\left(1 - e^{-bt}\right) \tag{1}$$

where K is Cowan’s K, t is the presentation duration of the sample array, a is the storage capacity parameter that determines the maximum range of the function (height of the curve), and b is the rate parameter that determines the time to reach the maximum value of the function. This formula is based on an ordinary differential equation,

$$\frac{\mathrm{d}K}{\mathrm{d}t} = b(a - K) \tag{2}$$

which assumes that the rate of increase in K is proportional to the difference between the current and maximum values of K. This formulation has been successfully used in modeling the temporal dynamics of WM performance for orientation stimuli (Bays, Gorgoraptis, Wee, Marshall, & Husain, 2011). The tangent of the curve at $t = 0$, which corresponds to $a \times b$, was used to measure the encoding rate.

Curve fitting was conducted for each participant by the least-squares method. A generalized linear mixed model (GLMM) was used with participants as random intercepts to examine the effects of experimental manipulation (encoding time and race) on the estimates of storage capacity and encoding rate. We also examined the effects of experimental manipulations on the false alarm rate and the hit rate, using a GLMM with participants as random intercepts. In the models, the error structure was Gamma distribution, and the link function was the inverse function. The analyses were conducted using the free statistical software R (Version 4.0.0; R Core Team, 2013) and the glmer function of the lme4 package (Bates, Mächler, Bolker, & Walker, 2015).

Results

Four of 30 participants had interactions with other-race people. We did not consider these interaction effects on the following analysis because the number of participants interacting with other-race people was small, and the interactions had no statistical effect on dependent variables.

Figure 2a shows Cowan’s K as a function of the sample array’s presentation duration. Curves are model fits described above. The solid (●) and dotted (▲) lines correspond to own-race and other-race conditions, respectively. Figure 2b shows the estimate of storage capacity as measured by the parameter in each race condition. Figure 2c shows the encoding rate estimate as measured by $a \times b$ in each race condition. The GLMM analysis showed a statistical effect of race on storage capacity, $b = -0.19$, $SE = 0.02$, $p < .01$, which demonstrated higher storage capacity for own-race compared to the other-race condition. The effect of the race of a face on the encoding rate was not significant, $b = -0.01$, $SE = 0.01$, $p = .21$. We also examined the correlation between storage capacity and encoding rate to examine if these two measures were interrelated (Jannati, McDonald, & Di Lollo, 2015). However, results indicated no significant association between the two in the own-race, $r = -.32$, $p = .09$, and the other-race conditions, $r = -.11$, $p = .54$.

The results of the hit and false alarm rates are summarized in Table 1. There was a statistical effect of the presentation duration on the hit rate, $b = -0.31$, $SE = 0.05$, $p < .01$, showing that longer presentation durations led to higher hit rates. However, there was no significant effect of race on hit rate, $b = 0.00$, $SE = 0.02$, $p = .98$. There was no significant effect of the presentation duration on the false alarm rate, $b = 0.06$, $SE = 0.73$, $p = .94$, whereas there was a significant effect of race on the false alarm rate, $b = -1.05$, $SE = 0.31$, $p < .01$, showing that the false alarm rate was increased in the other-race condition.

Discussion

Experiment 1 suggested that WM's own-race advantage is reflected in higher storage capacity, but not in the increased encoding rate for own-race faces. Notably, the results also indicated that the false alarm rate, but not the hit rate, was affected by the race of a face to be maintained in WM. This finding suggests that errors in the other-race condition were caused by failures in distinguishing test faces from sample faces, rather than by missing faces (e.g., a failure in the storage of faces). This result may have important implications for interpreting the cross-race effect in WM and will be discussed in the General Discussion.

Experiment 2

No backward masking was used in Experiment 1 after the presentation of the sample array, similar to previous studies examining the cross-race effect in face memory (Marcon et al., 2010; Sessa & Dalmaso, 2016; X. Zhou et al., 2018). However, not using a mask could be problematic for examining the temporal aspect of WM performance. Encoding of sensory information into WM is considered to persist until attention is distracted from sensory traces (Woodman & Vogel, 2008). The encoding process would continue even after the sample stimuli offset if backward masking was not used (Ricker, 2015). Therefore, the lack of masking in Experiment 1 could have obscured the possible effects of presentation duration on task performance.

In Experiment 2, we added backward masking following the sample array to control the time available for encoding faces into WM strictly. The experimental procedure was identical to that of Experiment 1, except for the addition of masking.

Method

Participants. Students ($n = 30$, 10 men and 20 women, mean age, 20.2 years, age range 18–26 years), including 29 Japanese students and an international student from China

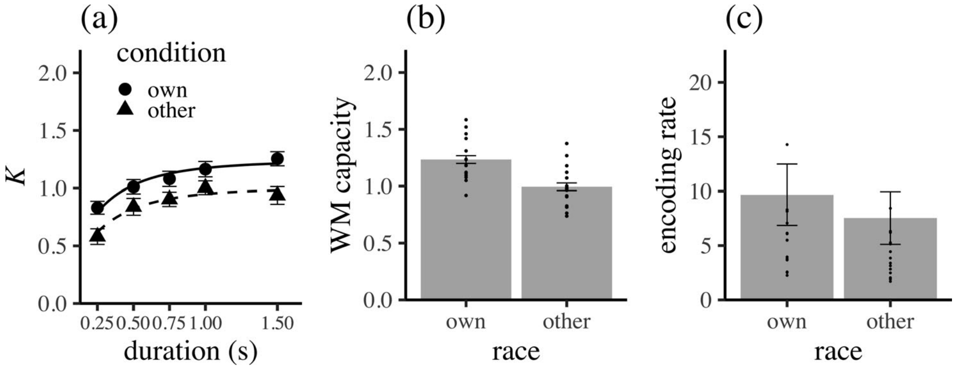

Figure 2 Results of Experiment 1. (a) Cowan's *K* as a function of presentation duration of the sample array. Curves are model fits. (b) Estimates of working memory (WM) capacity (the *a* parameter of the model). (c) Estimates of encoding rate. Error bars indicate the SEM.

© Japanese Psychological Association 2021.

Table 1 Hit rates and false alarm rates in Experiment 1

	Own race					Other race				
Duration	0.25	0.50	0.75	1.00	1.50	0.25	0.50	0.75	1.00	1.50
Hit (*SD*)	0.57 (0.13)	0.71 (0.13)	0.74 (0.13)	0.77 (0.14)	0.80 (0.13)	0.59 (0.18)	0.73 (0.16)	0.77 (0.15)	0.77 (0.15)	0.77 (0.16)
FA (*SD*)	0.16 (0.12)	0.20 (0.11)	0.20 (0.14)	0.19 (0.11)	0.17 (0.11)	0.30 (0.14)	0.31 (0.15)	0.31 (0.13)	0.27 (0.12)	0.30 (0.14)

enrolled in Kwansei Gakuin University, participated in the experiment. All the participants had normal or corrected-to-normal visual acuity. The experimental procedure was approved by the Kwansei Gakuin University Institutional Review Board for Behavioral Research with Human Participants. All participants gave their written informed consent before taking part in the study. Participants received course credits for participating in the study.

Stimuli and apparatus. The identical face stimuli and apparatus as in Experiment 1 were used. The masking stimulus was created by superimposing five face images: one upright image and four images with the orientation rotated either −135°, −45°, 45°, or 135°. The face images were randomly picked from the Chicago Face Database and an in-house database. These face images were not presented in the memory task. The size and shape of the masking stimulus were identical to those of face images.

Design, procedure, and analysis. The task's design and procedure were identical to those of Experiment 1 except for masking (Figure 1b). Masking images were presented at the same locations as the sample images for 100 ms. Data were analyzed identically to those of Experiment 1.

Results

Four of 30 participants had interactions with other-race people. We did not consider these interactions in the analysis because the number of participants having interactions with other-race people was small, and the interactions had no statistical effect on dependent variables.

Figure 3 shows the results of Experiment 2. The GLMM analysis indicated a statistically significant effect of race on storage capacity, $b = -0.10$, $SE = 0.02$, $p < .01$, which indicated

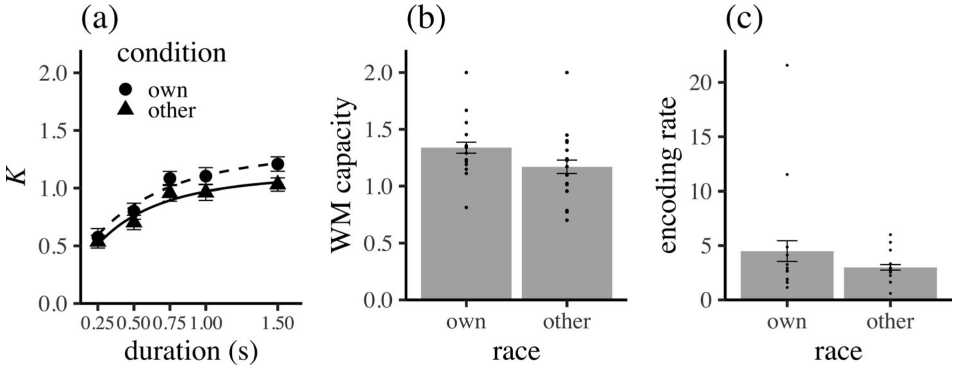

Figure 3 Results of Experiment 2. (a) Cowan's *K* as a function of presentation duration of the sample array. Curves are model fits. (b) Estimates of working memory (WM) capacity (the *a* parameter of the model). (c) Estimates of encoding rate. Error bars indicate the SEM.

© Japanese Psychological Association 2021.

Table 2 Hit rates and false alarm rates in Experiment 2

	Own race					Other race				
Duration	0.25	0.50	0.75	1.00	1.50	0.25	0.50	0.75	1.00	1.50
Hit (*SD*)	0.49 (0.17)	0.62 (0.15)	0.74 (0.15)	0.75 (0.15)	0.79 (0.12)	0.52 (0.18)	0.64 (0.16)	0.75 (0.18)	0.76 (0.14)	0.77 (0.13)
FA (*SD*)	0.21 (0.16)	0.22 (0.14)	0.20 (0.14)	0.20 (0.14)	0.19 (0.13)	0.26 (0.15)	0.29 (0.14)	0.28 (0.12)	0.28 (0.15)	0.25 (0.13)

better storage capacity for the own-race than the other-race condition (Figure 3b). Moreover, there was a significant effect of race on the encoding rate, $b = -0.07$, $SE = 0.02$, $p = < .01$ (Figure 3c). There was no significant correlation between storage capacity and encoding rate in the own-race, $r = -.11$, $p = .56$, and the other-race conditions, $r = -.33$, $p = .07$.

The results of the hit and false alarm rates are summarized in Table 2. The effect of presentation duration on the hit rate was significant, $b = -0.55$, $SE = 0.08$, $p < .01$, indicating that a longer presentation duration led to a higher hit rate. However, the effect of race on the hit rate was not significant, $b = -0.02$, $SE = 0.02$, $p = .50$. Moreover, the effect of presentation duration on the false alarm rate was not significant, $b = 0.25$, $SE = 0.52$, $p = .63$, whereas there was a significant effect of race on false alarm rate, $b = -0.58$, $SE = 0.22$, $p < .01$.

Discussion

Similar to Experiment 1, the storage capacity of WM was higher for own-race faces. Experiment 2 also replicated the effects of race and encoding time on hit and false alarm rates. Importantly, however, Experiment 2 indicated an increased encoding rate in the own-race condition, which was not the case in Experiment 1. Cowan's *K* differences between own- and other-race conditions were small and obscure in Experiment 2 in shorter presentation time conditions (Figure 3a). This finding suggests that the task was equally challenging in each race condition when the presentation time was short. Together with the result of Experiment 1, these findings suggest the crucial role of the encoding process in the WM cross-race effect.

General Discussion

The previous literature has suggested two distinct mechanisms underlying the cross-race effect in WM: higher storage capacity (Sessa & Dalmaso, 2016; Stelter & Degner, 2018; X. Zhou et al., 2018) and increased encoding rate for own-race faces (Marcon et al., 2010). However, there have been no systematic investigations of the relationship between these two mechanisms. Moreover, previous studies did not quantify the encoding rate (Marcon et al., 2010; X. Zhou et al., 2018), making it challenging to identify the actual differences in the encoding rate between own- and other-race conditions. In contrast, the current study explored the temporal dynamics of WM performance in more detail and separately estimated the storage capacity and the encoding rate. Experiments 1 and 2 indicated that WM's storage capacity was higher for own-race faces. Moreover, Experiment 2, in which the sample array was followed by backward masking to control for the time available for face encoding strictly, indicated an increased encoding rate as well as a higher storage capacity for own-race faces. These results provide evidence that both enhanced storage capacity and quick encoding of faces underlie WM's own-race face advantage.

We also found that the correlation between storage capacity and the encoding rate was weak and non-significant for own- and other-race conditions, suggesting that these two processes independently contributed to WM's cross-race effect. Specific studies using stimuli such as letters (Finke et al., 2005; Habekost & Starrfelt, 2009) and color hue (Jannati et al., 2015) have reported positive correlations between the encoding rate and WM's storage capacity. Habekost and Starrfelt (2009)

© Japanese Psychological Association 2021.

proposed putative mechanisms underlying the correlation between storage capacity and encoding rate. These mechanisms include shared or partially overlapping brain networks between the two processes, suggested by lesion studies, and the contribution of general cognitive efficiency factor, suggested by findings that storage capacity and encoding rate are correlated with IQ. However, there is evidence that individual differences in WM capacity for objects of expertise such as faces and cars (Curby & Gauthier, 2007; Curby et al., 2009) are not associated with individual differences in encoding speeds for those objects (Curby & Gauthier, 2009). Moreover, storage capacity and encoding rate for letter stimuli are only moderately correlated (about $r = .4$, Habekost & Starrfelt, 2009; Finke et al., 2005). Based on previous findings and the present results, we suggest that the own-race advantage for face stimuli is mediated by relatively distinct storage capacity and processing speed mechanisms.

Another exciting implication of our findings is related to the results of the hit rate and the false alarm rate. In both experiments, the cross-race effect was observed for the false alarm rate (i.e., a higher false alarm rate for other-race faces), but not for the hit rate. This finding suggests that the difficulty in recognizing other-race faces was caused in part by spurious feelings of familiarity, presumably due to the higher perceptual similarity between other-race faces (see Valentine, 1991, for an account of the cross-race effect based on the face space framework). Several studies have suggested that perceptual expertise for faces leads to higher WM precision, but not higher WM capacity for faces (Lorenc, Pratte, Angeloni, & Tong, 2014; Scolari, Vogel, & Awh, 2008, but see X. Zhou et al., 2018). The present result indicating a higher false alarm rate for other-race faces might reflect a decrease in memory precision for other-race faces. Alternatively, according to the category-individuation model (Hugenberg et al., 2010), which proposes that own-race faces are easier to discriminate because observers are trained to selectively attend to characteristic features that are identity-diagnostic for their own race, other-race faces might be represented with identical fidelity as own-race faces, but the facial encoding strategy might be adapted to own-race faces. Future studies exploring the nature of facial representation, including memory precision and encoding strategy, are expected to further our understanding of the cross-race effect in recognition memory.

Experiment 2 indicated an increase in encoding rate for own-race faces. This seems to be associated with a decline in Cowan's *K* for the own-race condition at shorter presentation durations: There were only a few differences in Cowan's *K* between own- and other-race conditions under shorter encoding conditions (250 and 500 ms) in Experiment 2, compared to Experiment 1 (comparing Figures 2a and 3a), which resulted in a steeper curve fitting for the own-race condition that led to an increased encoding rate. Why was the detrimental effect of backward masking more pronounced for own-race faces? The floor effect may explain the result because the Cowan's *K* values were already low in the shorter encoding conditions for other-race, leaving less room for performance decline. In fact, the hit rate in the 250-ms condition of Experiment 2 was at chance level (0.49 for own-race and 0.52 for other-race). However, the floor effect might not fully explain the result because the false alarm rate was still far better than chance (0.21 for own-race and 0.26 for other-race).

Alternatively, backward masking might be more detrimental to encoding own-race faces. We know that encoding can continue for several hundred milliseconds after the offset of the sample array (Brockmole, Wang, & Irwin, 2002; Di Lollo & Dixon, 1988; Irwin & Thomas, 2008; Vogel & Machizawa, 2004) if the sensory memory trace is not overwritten by backward masking (Ricker, 2015). If the continued encoding of faces after stimulus offset is more effective for own-race faces, it would explain the pronounced detrimental effect of masking on own-race faces. Although we have no direct evidence of this possibility, studies by Xie and Zhang (2017b, 2018) have shown that encoding in WM is faster for familiar than unfamiliar stimuli, which raises the interesting possibility that faster encoding of familiar stimuli is caused by stimulus processing that occurs after stimulus offset as well as during stimulus presentation.

The observation that encoding of faces into WM proceeds in a coarse-to-fine manner (Gao & Bentin, 2011) might provide another explanation of the detrimental effects of backward masking in the shorter encoding condition. Low-spatial frequency information is encoded and stored in WM faster, within 500 ms than high-spatial frequency information, which requires more than 800 ms to complete. Therefore, it would be difficult to encode high-spatial frequency information under the short encoding conditions (250 and 500 ms) of Experiment 2. High-spatial frequency information is crucial for discriminating face identities (Fiorentini, Maffei, & Sandini, 1983). Therefore, backward masking that disrupted high-spatial frequency information encoding might have impaired face recognition under shorter encoding conditions.

What is the mechanism that underlies the own-race advantage of WM capacity and the encoding rate? Holistic processing of visual information is generally considered the hallmark of face processing (Richler & Gauthier, 2014; Richler, Palmeri, & Gauthier, 2012), and might cause the own-race advantage in WM's storage capacity. For example, Curby and Gauthier (2010) suggested that objects of expertise such as faces are efficiently encoded and represented in WM due to the tighter binding of features, or the creation of larger featural units, thereby maximizing the use of an inherently limited WM system. We note that although some studies have suggested the role of holistic processing in the cross-race effect (Bukach, Cottle, Ubiwa, & Miller, 2012; McKone, Brewer, MacPherson, Rhodes, & Hayward, 2007; Michel, Rossion, Han, Chung, & Caldara, 2006; G. Zhou et al., 2018), others have not (Harrison, Gauthier, Hayward, & Richler, 2014; Horry, Cheong, & Brewer, 2015).

Other studies provide insights into the role of long-term memory in WM's performance advantage. For example, familiarity with a specific class of objects increases WM's storage capacity for those objects (Xie & Zhang, 2017a), and face processing speed is faster for own-race faces (Sporer, Trinkl, & Guberova, 2007). Therefore, experience with a class of objects, such as own-race faces, alters how it is perceptually processed and encoded/represented in memory, which may underlie WM's own-race advantage. In addition, as discussed above, the experience might also affect the precision of WM representations. It would be interesting for future studies to explore how holistic processing and long-term memory are associated with different aspects of WM's own-race advantage, including storage capacity, memory precision, and encoding speed.

In conclusion, this study showed that the own-race advantage in WM for faces is driven by both higher storage capacity and rapid encoding for own-race faces. Previous research had provided only partial evidence for the involvement of storage capacity or encoding rate in WM's cross-race effect. This study provides evidence that both factors independently contribute to WM's cross-race effect and raise the prospect that WM's own-race advantage is mediated by distinct mechanisms, including memory capacity and processing speed.

Conflict of Interest

The authors have no conflicts of interest directly relevant to the content of this article.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Bays, P. M., Gorgoraptis, N., Wee, N., Marshall, L., & Husain, M. (2011). Temporal dynamics of encoding, storage, and reallocation of visual working memory. *Journal of Vision*, *11*, 6.

Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior*, *25*, 475–491.

Brockmole, J. R., Wang, R. F., & Irwin, D. E. (2002). Temporal integration between visual images and visual percepts. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 315–334.

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

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–114.

© Japanese Psychological Association 2021.

Curby, K. M., & Gauthier, I. (2007). A visual short-term memory advantage for faces. *Psychonomic Bulletin & Review*, *14*, 620–628.

Curby, K. M., & Gauthier, I. (2009). The temporal advantage for individuating objects of expertise: Perceptual expertise is an early riser. *Journal of Vision*, *9*(6), 7.

Curby, K. M., & Gauthier, I. (2010). To the trained eye: Perceptual expertise alters visual processing. *Topics in Cognitive Science*, *2*, 189–201.

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 94–107.

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.

Di Lollo, V., & Dixon, P. (1988). Two forms of persistence in visual information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 671–681.

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin and Review*, *12*, 1127–1133.

Finke, K., Bublak, P., Krummenacher, J., Kyllingsbæk, S., Müller, H. J., & Schneider, W. X. (2005). Usability of a theory of visual attention (TVA) for parameter-based measurement of attention I: Evidence from normal subjects. *Journal of the International Neuropsychological Society*, *11*, 832–842.

Fiorentini, A., Maffei, L., & Sandini, G. (1983). The role of high spatial frequencies in face perception. *Perception*, *12*, 195–201.

Fukuda, K., & Vogel, E. K. (2019). Visual short-term memory capacity predicts the "bandwidth" of visual long-term memory encoding. *Memory & Cognition*, *47*, 1481–1497.

Gao, Z., & Bentin, S. (2011). Coarse-to-fine encoding of spatial frequency information into visual short-term memory for faces but impartial decay. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1051–1064.

Habekost, T., & Starrfelt, R. (2009). Visual attention capacity: A review of TVA-based patient studies. *Scandinavian Journal of Psychology*, *50*, 23–32.

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

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

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

Herzmann, G., Minor, G., & Adkins, M. (2017). Neural correlates of memory encoding and recognition for own-race and other-race faces in an associative-memory task. *Brain Research*, *1655*, 194–203.

Herzmann, G., Willenbockel, V., Tanaka, J. W., & Curran, T. (2011). The neural correlates of memory encoding and recognition for own-race and other-race faces. *Neuropsychologia*, *49*, 3103–3115.

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*, 508–524.

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.

Irwin, D. E., & Thomas, L. E. (2008). Visual sensory memory. In S. J. Luck & A. Hollingworth (Eds.), *Visual memory* (pp. 9–42). New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195305487.003.0002

Jannati, A., McDonald, J. J., & Di Lollo, V. (2015). Individual differences in rate of encoding predict estimates of visual short-term memory capacity (*K*). *Canadian Journal of Experimental Psychology*, *69*, 213–220.

Klaver, P., Talsma, D., Wijers, A. A., Heinze, H.-J., & Mulder, G. (1999). An event-related brain potential correlate of visual short-term memory. *NeuroReport*, *10*, 2001–2005.

Levin, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories. *Journal of Experimental Psychology: General*, *129*, 559–574.

Lorenc, E. S., Pratte, M. S., Angeloni, C. F., & Tong, F. (2014). Expertise for upright faces improves the precision but not the capacity of visual working memory. *Attention, Perception, & Psychophysics*, *76*, 1975–1984.

Ma, W. J. (2018). Problematic usage of the Zhang and Luck mixture model. *bioRxiv*. https://doi.org/10.1101/268961

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*, 1122–1135.

Marcon, J. L., Meissner, C. A., Frueh, M., Susa, K. J., & MacLin, O. H. (2010). Perceptual identification and the cross-race effect. *Visual Cognition*, *18*, 767–779.

McKone, E., Brewer, J. L., MacPherson, S., Rhodes, G., & Hayward, W. G. (2007). Familiar other-race faces show normal holistic processing and are robust to perceptual stress. *Perception*, *36*, 224–248.

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

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*, 3–35.

Meissner, C. A., Susa, K. J., & Ross, A. B. (2013). Can I see your passport please? Perceptual discrimination of own- and other-race faces. *Visual Cognition*, *21*, 1287–1305.

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

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13.

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ranganath, C., Cohen, M. X., & Brozinsky, C. J. (2005). Working memory maintenance contributes to long-term memory formation: Neural and behavioral evidence. *Journal of Cognitive Neuroscience*, *17*, 994–1010.

Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin*, *140*, 1281–1302.

Richler, J. J., Palmeri, T. J., & Gauthier, I. (2012). Meanings, mechanisms, and measures of holistic processing. *Frontiers in Psychology*, *3*, 553.

Ricker, T. (2015). The role of short-term consolidation in memory persistence. *AIMS Neuroscience*, *2*, 259–279.

Scolari, M., Vogel, E. K., & Awh, E. (2008). Perceptual expertise enhances the resolution but not the number of representations in working memory. *Psychonomic Bulletin & Review*, *15*, 215–222.

Sessa, P., & Dalmaso, M. (2016). Race perception and gaze direction differently impair visual working memory for faces: An event-related potential study. *Social Neuroscience*, *11*, 97–107.

Sporer, S. L., Trinkl, B., & Guberova, E. (2007). Matching faces: Differences in processing speed of out-group faces by different ethnic groups. *Journal of Cross-Cultural Psychology*, *38*, 398–412.

Stelter, M., & Degner, J. (2018). Investigating the other-race effect in working memory. *British Journal of Psychology*, *109*, 777–798.

Susa, K. J., Meissner, C. A., & de Heer, H. (2010). Modeling the role of social-cognitive processes in the recognition of own- and other-race faces. *Social Cognition*, *28*, 523–537.

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

Valentine, T., & Bruce, V. (1986). The effect of race, inversion and encoding activity upon face recognition. *Acta Psychologica*, *61*, 259–273.

van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*, *109*, 8780–8785.

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*, 748–751.

Wiese, H., Kaufmann, J. M., & Schweinberger, S. R. (2014). The neural signature of the own-race bias: Evidence from event-related potentials. *Cerebral Cortex*, *24*, 826–835.

Woodman, G. F., & Vogel, E. K. (2008). Selective storage and maintenance of an object's features in visual working memory. *Psychonomic Bulletin & Review*, *15*, 223–229.

Xie, W., & Zhang, W. (2017a). Familiarity increases the number of remembered Pokémon in visual short-term memory. *Memory & Cognition*, *45*, 677–689.

Xie, W., & Zhang, W. (2017b). Familiarity speeds up visual short-term memory consolidation. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 1207–1221.

Xie, W., & Zhang, W. (2018). Familiarity speeds up visual short-term memory consolidation: Electrophysiological evidence from contralateral delay activities. *Journal of Cognitive Neuroscience*, *30*, 1–13.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*, 233–235.

Zhou, G., Liu, J., Xiao, N. G., Wu, S. J., Li, H., & Lee, K. (2018). The fusiform face area plays a greater role in holistic processing for own-race faces than other-race faces. *Frontiers in Human Neuroscience*, *12*, 220.

Zhou, X., Mondloch, C. J., & Emrich, S. M. (2018). Encoding differences affect the number and precision of own-race versus other-race faces stored in visual working memory. *Attention, Perception, & Psychophysics*, *80*, 702–712.

(Received August 13, 2020; accepted December 16, 2020)

© Japanese Psychological Association 2021.