THESIS

METACOGNITION IN THE OWN-RACE BIAS:

INFLUENCES ON RESTUDY SELECTION

Submitted by

Sarah DeLozier

Department of Psychology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2015

Master's Committee:

Advisor: Matthew G. Rhodes

Benjamin Clegg
Lucy Troup
Rosa Martey

Copyright by Sarah DeLozier 2015

All Rights Reserved

ABSTRACT

METACOGNITION IN THE OWN-RACE BIAS: INFLUENCES ON RESTUDY SELECTION

The own-race bias (ORB), refers to the finding that learners demonstrate poorer memory for faces of other races or ethnicities, compared to those of their own. This memory bias has been examined from an encoding perspective, yet little has been done to examine the possible contribution of metacognition. Under the assumption that monitoring affects control, recognition memory was examined for own-race (White) and other-race (Black) faces. Pilot data suggested that participants might be aware of the ORB, as evinced by their monitoring judgments. The experiment permitted participants to select faces for restudy, make delayed JOLs and restudy selections, then restudy selected faces prior to test. Results demonstrated similar monitoring accuracy for own-race and other-race faces, suggesting that the ORB may not be due to monitoring deficits. Measures of control provided some evidence that learners make similar control judgments for own- and other-race faces, and follow-up experiments are discussed.

ACKNOWLEDGEMENTS

Thanks first to my advisor, Matt Rhodes, for all the mentorship and extensive feedback along the way, – I wouldn't be half as enthusiastic about grad school without it! Thanks also to my parents, who have supported me time and time again.

Finally, thanks to Chris, who moved to the other side of the U.S. so I could do this, and who reminds me there's a world outside of school.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

During a span of over forty years, research on memory for own-race and other-race faces has sought to explain why individuals have difficulty recognizing faces from other races and/or ethnicities, as compared to their own, a finding termed the *own-race bias* (Hourihan, Fraundorf, Benjamin, 2012; Malpass & Kravitz, 1969; Meissner, Brigham, & Butz, 2005). In the seminal article on this phenomenon, Malpass and Kravitz (1969) had Black and White learners study photographs of both Black and White faces. Learners were given a recognition test, in which all studied faces and an equal number of unstudied faces were individually presented, and were asked to indicate whether they recognized that face as having been studied previously. The results revealed that learners were better at recognizing previously-studied faces of their own race compared with other-race faces. Subsequent research on the own-race bias has focused on processes at encoding (for recent reviews, see Meissner & Brigham, 2001; Young, Hugenberg, Bernstein, & Sacco, 2012), largely neglecting awareness and control of learning, termed *metacognition* (Nelson & Leonesio, 1988), as a potential factor influencing memory for faces. In this thesis, I examined 1) whether learners were aware of the own-race bias, and 2) whether learners could effectively control their learning to improve memory for other-race faces.

Theories of the Own-Race Bias

Learners demonstrate better memory for own-race than other-race faces; a finding termed the own-race bias[1] (ORB; Hourihan et al., 2012; Malpass & Kravitz, 1969). The ORB is robust

[1] Chiroro, Tredoux, Radaelli, and Meissner (2008) utilize the more general term, *in-group face recognition advantage*. This label more accurately encompasses their finding of differences in recognition for same-race individuals of different cultures. However, although I agree in principle with this label, the term *own-race bias* is used in the current paper to link these experiments with the wider literature examining recognition for own- and other-race faces.

across a variety of populations, having been found in children (Pezdek, Blandon-Gitlin, & Moore, 2003; Sangrigoli, Pallier, Argenti, Ventureyra, & de Schonen, 2005), college students (Bernstein, Young, & Hugenberg, 2007; Hehman, Mania, & Gaertner, 2010), and older adults (Brigham & Williamson, 1979), as well as across racial and cultural groups (Kassin, Ellisworth, & Smith, 1989). Although the ORB has been repeatedly demonstrated, there is no consensus on the mechanisms driving the effect (Meissner & Brigham, 2001).

One of the primary explanations of the ORB, the *perceptual-expertise* account, suggests that the ORB is due to more frequent exposure to and differentiation of own-race than other-race faces (Malpass & Kravitz, 1969; G. Rhodes, Brake, Taylor, & Tan, 1989; Turk, Handy, & Gazzinga, 2005; but see G. Rhodes, Hayward, & Winkler, 2006). For example, G. Rhodes et al. (1989) had learners study a series of own-race and other-race faces. When learners were later tested in a recognition task, the faces were presented either upright or upside-down. As learners could not fully use their previous experience (i.e., expertise) with inverted own-race faces, recognition of inverted faces was more impaired for own-race than other-race faces, evinced in both recognition accuracy and response times. This greater impairment for own-race faces suggests that expertise plays a role in recognition for own-race faces.

In an alternative method of examining expertise, Chiroro, Tredoux, Radaelli, and Meissner (2008) tested Black and White South African learners on Black and White American and South African faces. Learners demonstrated the ORB, but only for South African faces, showing that recognition was attuned to local exposure to faces. That is, South African individuals exhibited an ORB for only South African faces. This regional bias for facial recognition is consistent with the perceptual-expertise account, which would hold that South African individuals would be primarily exposed (and attuned) to South African faces of their own race (i.e., culture). However, although expertise appears to be a reliable moderator of the ORB (Hugenberg et al., 2010), it has been difficult to account for expertise exclusively via measures of contact. For example, in their meta-analysis of the ORB, Meissner and Brigham (2001) found that self-reported measures of interracial contact (a means of examining expertise) accounted for less than 2% of the variability in the ORB between subjects. It is important to note that these measures included questions asking people to answer whether they have friends of another race, however, and that responses may have been biased by desire to respond in a socially acceptable manner. Regardless, it appears likely that a substantial portion of the variability in the ORB is not captured in measures of contact with other-race individuals, limiting the power of a perceptual-expertise explanation of the ORB.

An alternative explanation of the ORB, the *social-categorization* account, suggests that learners rapidly categorize faces as either in-group or out-group at encoding and are less motivated to differentiate between out-group compared with in-group faces (Bernstein, Young, & Hugenberg, 2007; Hourihan, Fraundorf, & Benjamin, 2013; Hugenberg, Miller & Claypool, 2007). Evidence for this account demonstrates that learners can be motivated to differentiate between out-group members when these individuals are important or the outcome is self-relevant (DeLozier & Rhodes, 2015; Hugenberg et al., 2007; Shriver & Hugenberg, 2010; Young & Hugenberg, 2011). For example, Shriver and Hugenberg (2010) presented occupational labels with studied faces, identifying each face as belonging to either a low-power (i.e., janitor, mechanic) or high-power (doctor, CEO) individual. For high-power faces, recognition was equal for Black and White faces, eliminating the ORB. For low-power faces, however, the ORB remained. In this study, high-powered jobs consisted of positions of prestige or authority over others, making recognition of these high-powered individuals both important and self-relevant.

Overall, it appears that learners can differentiate between other-race faces when motivated (but see Hugenberg, Young, Bernstein, & Sacco, 2010, for an account in which perceptual-expertise and social-categorization work conjointly).

Awareness of the Own-Race Bias

A primary question driving this thesis is whether learners are aware of the ORB (i.e., whether they believe memory differs for own-race and other-race faces) and how this might affect their memory performance. For example, if learners are unaware that their memory will be poorer for other-race faces, they may not be aware of the need to differentiate between own-race and other-race faces. Few studies have examined learners' beliefs about other-race faces (Hourihan et al., 2012; Rhodes, Sitzman, & Rowland, 2013; Schmechel, O'Toole, Easterly, & Loftus, 2006, Smith, Stinson, & Prossor, 2004), and those that have examined the issue suggest that people may be largely unaware of the ORB. For example, Schmechel et al. (2006) surveyed approximately 1,000 jurors regarding several factors influencing eyewitness testimony. They observed that nearly half (48%) believed identification of own-race and other-race faces was equally reliable. An additional 11% were either unsure or believed identification of other-race faces would be more reliable. Further supporting this idea, in a series of three experiments, Rhodes et al. (2013) had participants study own- and other-race faces and asked them to indicate, for each face, the likelihood that it would be recognized on a future test. Although participants were better at recognizing own-race than other-race faces, their memory predictions were similar for faces of both races, failing to reflect the ORB.

Research suggests that when learners are made aware of the ORB it can be reduced (Hugenberg et al., 2007; Young, Bernstein, & Hugenberg, 2010; Young & Hugenberg, 2012). For example, Hugenberg et al. (2007) reduced the ORB by informing learners of the ORB and warning them of the necessity of individuating other-race faces at study. These instructions focused on affecting learners' knowledge and beliefs (or monitoring) of their own learning and towards a specific means of directing (or controlling) their learning efforts. Since then, other studies (Young et al., 2010; Young & Hugenberg, 2012) have replicated and extended these findings. Young et al. (2010) found that instructions to individuate other-race faces eliminated the ORB only when learners were instructed prior to encoding, and not prior to a recognition phase. Accordingly, the ORB may reflect processes occurring during encoding and under some degree of learner control. What processes might these be, and how might they be examined? The answers to these questions may lie in metacognition.

Metacognition and Judgments of Learning

Metacognition, or "knowing about knowing" (Flavell, 1978), consists of processes that aim to assess one's knowledge (i.e., *monitoring*) and to self-regulate behavior based on the knowledge inferred from monitoring (i.e., *control*) (Nelson & Narens, 1990; for a review, see Koriat, 2007). For example, consider a student preparing for a chapter exam with the goal of receiving an A. In order to do this, she first *monitors* her current overall knowledge, reaching an assessment that the chapter has not been well learned. Because of this assessment, she chooses to *control* her learning by restudying the entire chapter. Her overall assessment that the chapter has not been well learned thus determines her choice of whether she should engage in further study of the chapter.

A common means of investigating how people monitor learning is to collect judgments of learning (JOLs), either immediately or after a delay (Dunlosky & Nelson, 1992; Nelson & Dunlosky, 1991; see Rhodes, in press, for a review). JOLs are often solicited by first presenting an item individually for study. The learner is then asked how likely (e.g., on a scale from 0-

100%) it is that he or she will remember the studied item on a later test. JOLs are formally measured in terms of *absolute accuracy* and *relative accuracy*. Absolute accuracy reflects whether learners' average JOLs are overconfident or underconfident (Dunlosky & Lipko, 2007). For example, if a learner's average recognition performance was 70%, absolute accuracy would be excellent if the average JOL was 70%. Likewise, average JOLs below 70% would indicate underconfidence whereas average JOLs exceeding 70% would signify overconfidence. In other words, participants' absolute accuracy reflects the difference between their average JOLs and their overall test performance. Relative accuracy refers to the extent to which participants' JOLs distinguish between individual items that are (or are not) correctly remembered on a later test. It is usually measured as an item-by-item association between JOLs and memory performance and is commonly reported via nonparametric gamma correlations as a value ranging from -1 to +1 (Nelson, 1984). Relative accuracy values approaching +1.0 would indicate that recognized items were consistently given higher JOLs than unrecognized items. Conversely, relative accuracy values approaching -1.0 would indicate that unrecognized items were consistently given higher JOLs than recognized items. Finally, relative accuracy values near 0 would indicate little to no relationship between the JOL given an item and whether that item was later recognized.

According to the predominant monitoring-affects-control hypothesis, monitoring accuracy affects how learning is controlled (Benjamin, Bjork, & Schwartz, 1998; Dunlosky & Hertzog, 1998; Metcalfe, 2002; Metcalfe & Finn, 2008; Nelson & Leonesio, 1988; Nelson & Narens, 1990, 1994; Rhodes & Castel, 2009; Thiede, 1999, but see Koriat, Ma'ayan, & Nussinson, 2006). For example, upon making a new acquaintance, an individual might monitor whether he or she will remember the acquaintance's name later on, and decide the answer is "yes". This monitoring judgment is then presumed to affect how the individual controls memory for that name, perhaps choosing to cease trying to remember it, since he or she believes it is already learned (but see Koriat, 2007; Koriat et al., 2006; Koriat et al., 2013, for a proposal that control can drive monitoring). In particular, people often seek to restudy more poorly learned information (Dunlosky & Thiede, 1998; Metcalfe, 2002; Thiede & Dunlosky, 1999). For example, if two faces were given JOLs of 50 and 100 (respectively), an individual would be more likely to select the lower-JOL face for restudy.

Importantly, monitoring can affect control over learning independently of actual memory performance (Nelson & Narens, 1990; Metcalfe & Finn, 2008; Rhodes & Castel, 2009). In an example of inaccurate monitoring and control, Rhodes and Castel (2009) had participants listen to loud or quiet words, predict their memory for those words at test, and make restudy selections. Participants gave higher JOLs to loud than to quiet words, and were more likely to select quiet words for restudy[2]. However, actual memory did not differ between loud and quiet words. Another possibility is that a learner might accurately monitor her memory and choose to restudy items she has poor memory for; yet, for a multitude of reasons, she still might not enhance her memory for the restudied items. For example, the to-be-learned items might be too difficult (Nelson & Leonesio, 1988), or she might choose an ineffective method of restudying this information (for a detailed review of effective study methods, see Dunlosky et al., 2013).

What role might monitoring and control processes play in the ORB? If monitoring of other-race faces affects how learners control their learning of those faces, then accurate (or inaccurate) monitoring judgments would be reflected in subsequent restudy decisions. That is, participants should choose to restudy faces not yet learned and choose *not* to restudy faces already learned (Metcalfe, 2002; Thiede & Dunlosky, 1999). However, a perceptual-expertise

[2] Learners selected words for restudy, but no restudy phase occurred.

account would hold that learners could not improve recognition of other-race faces without an accompanying improvement in expertise. Thus, according to a perceptual-expertise account, people should be poorer at monitoring and regulating learning of other-race than own-race faces. However, even if participants make accurate monitoring judgments and control decisions, as mentioned above, recognition for other-race faces might not be affected (Nelson & Leonesio, 1988). If processes other than acquired expertise influence learning, then accurate monitoring of the ORB should influence restudy selection and recognition for other-race faces.

Overview of the Current Study

In this thesis, I examined monitoring and control for own-race and other-race faces. First, I present data from a pilot experiment, discuss the results, and use these findings to refine the methodology for the experiment reported. Monitoring accuracy was assessed by examining the relationship between learners' monitoring judgments (JOLs) and recognition through measures of absolute and relative accuracy. Control was examined through learners' restudy choices and their relationship to performance at test. A question of particular interest was whether learners' monitoring accuracy differed for own- and other-race faces, and whether their study choices could compensate for poorer recognition of other-race faces. If monitoring accuracy affects how learners control subsequent learning, ultimately affecting memory performance (Nelson & Leonesio, 1988; Nelson & Narens, 1990), accurate monitoring should enhance face recognition. Based on prior research (Rhodes et al., 2013), which suggested a trend for higher JOLs for own-race than other-race faces, this experiment might expect to find either a small effect, or no differences at all, between JOLs for own-race and other-race faces. In contrast, it is possible that enhanced monitoring accuracy would not affect recognition for other-race faces, consistent with a perceptual-expertise account. That is, a perceptual-expertise account would hold that regardless of monitoring accuracy, recognition for other-race faces could not be enhanced without a concurrent improvement in expertise.

In the pilot experiment, participants studied individually presented own-race and other-race faces. After each face was presented, participants made a JOL and then indicated whether that face should later be restudied (although faces were selected for restudy, no restudy phase occurred). After a series of faces were presented, these faces were re-presented one-at-a-time alongside a new face of the same race and sex, in a 2-alternative, forced-choice (2AFC) recognition test. Participants were asked to indicate which of the two faces in each pair had been previously studied. If learners' JOLs are greater for own- than other-race faces, restudy selection of other-race faces should likewise reflect these differences, with learners selecting more lower-JOL faces for restudy. If JOLs do not indicate awareness of the ORB, however, restudy selection should not differ for other-race compared to own-race faces.

Pilot

Participants

Participants consisted of 35 Caucasian individuals from Colorado State University. The use of only Caucasian participants is due to the ethnic diversity of Colorado State University. According to the 2013-2014 Colorado State University Fact Book (http://www.ir.colostate.edu/factbook-fb.aspx), out of 22,565 undergraduates enrolled at the university, 16,799 self-identify as White (74.36%) and 471 self-identify as Black (2.09%).

Materials

Stimuli consisted of 132 faces of Black and White individuals, taken from a set used by Meissner, Brigham, and Butz (2005). All faces were presented on a white background, and edited to show only the face, hair, and neck. Faces were randomly divided into four blocks of 32 faces (each with 16 Black and 16 White faces). Within each set of faces, each race was evenly represented by sex. Each block of faces was counterbalanced to be equally likely to be presented on list 1 or list 2, for study or as a foil on the test. At test, each face was equally likely to be presented on the left or right side of the screen. For the study phase, all faces showed neutral expressions whereas, at test, all faces were shown smiling. The remaining four faces were used as examples during the initial experiment instructions and were not presented as stimuli during the experiment. All stimuli were presented on Dell computer monitors using E-Prime software.

Procedure

Participants were instructed that they would complete two study-test blocks. A practice test was given using 2 pairs of faces, one Black and one White, to instruct participants on the nature of the task. Data were not collected for these two pairs of faces. Each study block consisted of 32 faces presented in a random order (16 Black and 16 White), individually for 4 sec each. Immediately after studying each face, participants were given 5 sec to make a JOL, indicating how likely they were to correctly recognize that face later. Judgments were made on a scale of 50%-100%. A judgment of 50% indicated a likelihood no greater than chance of recognizing a face, whereas a judgment of 100% indicated an absolute likelihood that face would correctly be recognized later. Intermediate values represented intermediate levels of confidence. After each JOL, participants were given 5 sec to select whether or not they would like to restudy that particular face before the recognition test. However, no faces were presented for restudy throughout this experiment (see Finn, 2008; Rhodes & Castel, 2009; for a similar procedure with verbal stimuli). Faces remained visible during each JOL and restudy selection. This procedure continued until all 32 faces from that block had been presented.

After all study faces in the first block were presented, participants were given a two-alternative forced choice (2AFC) recognition test. In this recognition test, two faces were presented simultaneously, matched by race and sex. Within each pair of faces, one was previously studied, or “Old”, whereas one was taken from a set of new faces and had never-before been seen, or “New”. Learners were asked to indicate which of the two faces they had previously studied until all 32 pairs of old-new faces had been presented. This entire procedure was repeated with the 2 remaining block of faces, for a total of two study-test blocks.

Results

Correct identification of a previously studied face is termed a *hit,* whereas incorrectly identifying a new face as “studied” is termed a *false alarm*. False alarm rates are the inverse of hit rates in a 2AFC recognition task; therefore, recognition was examined using only the proportion of hits out of all possible responses (see Rhodes et al., 2013). To more directly assess the focus of this paper – monitoring and control processes in the own-race bias – results are first presented for measures of monitoring and then for measures of control. The alpha level for all statistical tests was set to .05.

Monitoring. As mentioned previously, monitoring processes refers to one's assessment of their own cognitive processes (Nelson & Narens, 1990). Thus, the following results reflect whether participants correctly assess (i.e., monitor) their memory for own-race and other-race faces.

Absolute Accuracy. The absolute accuracy of learners' predicted recognition performance, or the extent to which JOLs are underconfident or overconfident, was calculated by comparing mean JOLs with recognition performance (expressed as a percentage) through creation of a 'Measure' factor (cf. Rhodes et al., 2013; Zimmerman & Kelley, 2010). A 2 (Measure: Predicted, Recognition) x 2 (Race of Face: Black, White) repeated-measures analysis of variance (ANOVA) was conducted (Figure 1). Overall, actual recognition was higher than predicted by learners' JOLs, $F(1, 34) = 24.80$, $\eta^2 = .42$. A main effect was found for Race of Face, $F(1, 34) = 46.60$, $\eta^2 = .58$, qualified by a Measure x Race of Face interaction, $F(1, 34) = 11.04$, $\eta^2 = .25$. Pairwise comparisons demonstrated that recognition was better for White faces than for Black faces, $t(34) = 5.45$, $d = 1.08$, and that JOLs were higher for White faces than for Black faces, $t(34) = 4.58$, $d = .45$. Thus, the interaction indicates that the effect of race of face was larger for recognition than JOLs.

Relative Accuracy. Relative accuracy was examined through nonparametric gamma correlations (Nelson, 1984). Gamma correlations range from -1.0 to +1.0, with a value of +1 indicating that recognized faces were given higher JOLs than unrecognized faces. A value of -1 indicates that unrecognized faces were given higher JOLs than recognized faces, and a value of 0 indicates no relationship between recognition and JOLs. Mean gamma correlations indicated that participants' JOLs differentiated between faces that were recognized and that they failed to recognize, both for Black faces (*M* = .19, *SE* = .05), $t(34) = 3.79$, and White faces (*M* = .25, *SE* = .08), $t(34) = t(32) = 3.19$. Relative accuracy did not differ significantly between Black and White faces, $t(32) = .56$, $p = .581$. Thus, the relationship between JOLs and recognition was similar for White and Black faces.

Control. Metacognitive control, or the self-regulation of learning (Nelson & Narens, 1990), was examined through a multitude of measures. My primary interest was in whether participants controlled their learning differently for own-race and other-race faces, as well as whether monitoring affects these control decisions. Measures of absolute accuracy demonstrated that learners correctly assessed Black faces as being less likely to be recognized than White faces. However, does this assessment also manifest itself in self-regulation of memory? Participants selected a greater proportion of Black faces (*M* = .30, *SE* = .05) than White faces (*M* = .27, *SE* = .05) for restudy, $t(30) = 3.43$, $d = .15$.

Recognition was examined for faces participants selected or rejected for restudy in a 2 (Restudy Selection: selected, rejected) x 2 (Race of Face: Black, White) repeated-measures ANOVA (Figure 2). A main effect for restudy selection indicated that, overall, recognition was greater for faces participants rejected for restudy (*M* = .86, *SE* = .02) than for the faces they selected for restudy (*M* = .76, *SE* = .03), $F\,(1, 22) = 11.40$, $\eta^2 = .34$. Recognition was also greater for White than Black faces, $F\,(1, 22) = 8.75$, $\eta^2 = .29$. The restudy selection by race of face interaction was not significant, $F < 1$. Thus, although faces selected for restudy were less likely to be recognized, the effect of race of face on recognition did not interact with restudy selection.

The correlation between JOLs and restudy selection differed significantly from zero for both Black faces (G = -.56, SE = .10), $t(24) = 5.91$, and White faces (G = -.46, SE = .10), $t(23) = 4.81$, and was negative for both. Thus, faces associated with lower JOLs were more likely to be selected for restudy. This correlation did not differ significantly between Black and White faces, $t(22) = .67$, $p = .509$, $d = .14$.

Discussion

Pilot results demonstrated a robust ORB as, on average, participants correctly recognized a greater proportion of own-race than other-race faces. Measures of participants' monitoring likewise reflected this memory bias, as JOLs were lower for other-race than own-race faces. In support of the idea that monitoring is related to control (Nelson & Narens, 1990), gamma correlations indicated that participants' JOLs differentiated between own-race and other-race faces they later recognized (or failed to recognize) at test. However, this relationship was modest (G = .19 for Black faces), and inconsistent with the findings of Rhodes et al. (2013; Experiments 1 and 2), in which G = .03, .05 (respectively) for Black faces. Monitoring also appeared to influence restudy selection, as participants chose to restudy faces of both races that were more poorly learned, and were more likely to select Black faces than White faces for restudy. With these pilot results as a foundation, the following experiment pursued the concept of monitoring and control over memory in regard to the ORB in greater depth.

Experiment

Pilot results provided a valuable springboard for the experiment reported. Most importantly, the pilot results suggested that participants' monitoring judgments were not only reflective of the ORB, but that participants were also willing to control their learning (via restudy choices) in an effort to enhance their memory for other-race faces. However, as participants were not permitted to restudy the faces they selected, it was not apparent whether the ORB could be reduced for restudied faces. The following experiment remedied this by permitting participants to restudy the faces they selected for restudy, prior to a final recognition test. Specifically, participants were permitted to choose whether to restudy a face prior to the final test and, unlike the pilot experiment, restudied those selected faces.

A foundational hypothesis of this research is that monitoring accuracy affects control of learning. For example, if participants were inaccurate at determining which faces they have or have not yet learned, then their restudy selections should likewise be inaccurate. Conversely, if participants were accurate at assessing their learning for faces, then their restudy selections should be informed by this accuracy. Accordingly, an important question is whether participants can make accurate JOLs for other-race faces when conditions are optimal (i.e., when they are most likely to enhance monitoring accuracy). One means of enhancing monitoring accuracy has been to delay JOLs for a brief period, rather than soliciting JOLs immediately after study, thereby greatly increasing relative accuracy for these items (Connor, Dunlosky, & Hertzog, 1997; Koriat & Bjork, 2006; Nelson & Dunlosky, 1991; for a review, see Rhodes & Tauber, 2011). For example, Koriat and Bjork (2006) had participants study a series of word-pairs; after the pairs had been studied, participants were shown the first word of each previously-studied word-pair (e.g., *kittens* - ?), and were asked to judge the likelihood they would recall the second word of that word-pair on a later test (delayed JOLs). Compared to JOLs made immediately after studying each word-pair (immediate JOLs), delaying JOLs led to better relative accuracy (i.e., JOLs were more likely to discriminate between word-pairs participants later recalled or failed to recall). Moreover, when JOLs are delayed, memory performance itself is enhanced relative to immediate JOLs (Rhodes & Tauber, 2011), although the benefit appears quite small (g = .08). The mechanisms causing benefits of delayed-JOLs are not fully understood (Nelson, Narens, & Dunlosky, 2004), with evidence suggesting delayed-JOLs may utilize information from long-term memory, or factors such as a match between cues available at study and at test (for a review, see Rhodes & Tauber, 2011). The current article assumes, as in immediate JOLs, that delayed JOLs reflect monitoring of memory performance.

In order to enhance relative accuracy in this experiment, JOLs were delayed (Nelson & Dulosky, 1991). If participants can accurately indicate which faces they will or will not recognize at test, then this monitoring accuracy should influence restudy selection, potentially enhancing recognition for faces. However, if recognition for other-race faces is contingent upon expertise, then regardless of monitoring accuracy or restudy selection, the ORB should not be affected. This would be comparable with the "labor-in-vain" findings of Nelson and Leonesio (1988), who reported that large increases in study time resulted in minimal-to-no enhancements for memory.

Accordingly, in the experiment reported, participants studied own- and other-race faces and made delayed JOLs. Specifically, JOLs were delayed by individually presenting each of eight faces for study, before asking for JOLs. After 8 faces were presented, the first of those 8 faces was re-presented or "flashed" for 500 ms, and participants were asked to make a JOL and restudy selection for that face. This procedure continued for each of the subsequent 8 faces, in sets of 8, for a total of 32 faces. Selected faces were then restudied prior to test. After test, this study-restudy-test procedure repeated for a total of two study-restudy-test blocks.

I also examined two other factors in the current experiment that might bear on the own-race bias. First, participants in the experiment were given a measure of interracial contact. Although Meissner and Brigham's (2001) meta-analysis found that expertise accounted for approximately 2% of the variability in the ORB, they theorized this might be artificially deflated due to a lack of variance in expertise. In support of this possibility, a significant effect of other-race contact has been found when using participants with a wide range of other-race contact experience (Hancock & Rhodes, 2008; G. Rhodes et al. (2009). For example, Hancock and Rhodes (2008) used an "extreme difference" approach of examining the ORB amongst recent immigrants as well as amongst more traditional participants, and found that other-race contact did affect the ORB. Thus, to more fully examine the perceptual-expertise account of the own-race bias, a survey of other-race contact (modified from Hancock & Rhodes, 2008) was administered at the end of the experiment. I predict that as self-reported other-race contact increases, the ORB will be reduced.

Second, the design used permits me to examine the potential effect of task experience on recognition memory and metacognition for own- and other-race faces. Some evidence suggests that absolute accuracy, or participants' over- or under-confidence, may become more accurate as participants gain task experience. For example, Tauber and Rhodes (2010) had participants study and make JOLs for a series of words. After completing a test for these words, participants again studied and made JOLs for a new series of words. When compared to the first study-test trial, overall JOLs on the second study-test trial more accurately reflected participants' actual performance. In the current experiment, I examined experience as a function of performance changes across blocks and report sub-analyses of block when block impacted performance. If participants' monitoring and recognition do not change with task experience, then the absolute accuracy of their JOLs and restudy selections for Black and White faces should not differ from block 1 to block 2. However, if participants' monitoring judgments for other-race faces become more reflective of recognition performance with task experience, then JOLs and restudy selections for Black faces should be more accurate on Block 2 than on Block 1.

Participants

Participants consisted of 124 Caucasian individuals from Colorado State University. As in the pilot experiment, testing only Caucasian students reflects the ethnic diversity of Colorado State University.

Materials

All materials were identical to those from the pilot experiment.

Procedure

As in the pilot experiment, participants were instructed that they would complete two study-test blocks. A practice test was given using 2 pairs of faces, one Black and one White, to instruct participants on the nature of the task. Data were not analyzed for these two pairs of faces. Each study block consisted of 32 faces (16 Black and 16 White). Within each study block, faces were randomly assigned to sets of 8 faces with the constraint that each set contained an equal number of Black and White faces. For each randomized set of 8 faces, faces were presented individually for 4 s each, with a 500-ms blank screen in between each face presentation. After all 8 faces were presented, the first face of that set was individually re-presented, or "flashed" for 500 ms (for a similar procedure, see Kimball & Metcalfe, 2003).

Immediately after this flash, participants were given 5 s to make a delayed JOL, followed by 5 s to indicate whether they would like to restudy that face. This delayed-JOL procedure then repeated for each of the 8 faces in the set, and was repeated with each subsequent set of 8 faces. Therefore, each JOL and restudy selection was delayed, occurring approximately 1.5 minutes after the face had been initially presented for study. After all 32 faces had been studied, and all JOLs and restudy selections had been made, all faces selected for restudy (if any) were individually re-presented in random order for 4 s each. Immediately after all faces had been restudied, participants completed the 2AFC recognition test. As in the pilot, this entire study-test procedure repeated for both blocks of faces.

Results and Discussion

All results focus only on reliable interactions, with higher-order interactions favored over lower-order interactions. Results are presented first for measures of monitoring, then for measures of control. For both monitoring and control, results are presented first across blocks, then separately for blocks 1 and 2 whenever Block interacts with a variable. As in the pilot experiment, the alpha level was set to .05 for all statistical tests.

Monitoring.

Absolute Accuracy. Absolute accuracy was examined in a 2 (Measure: Predicted, Recognition) x 2 (Race of Face: Black, White) x 2 (Block: 1, 2) repeated-measures ANOVA (Figure 3). Main effects were found for measure, $F(1, 122) = 161.87$, $\eta^2 = .57$, and race of face, $F(1, 122) = 233.22$, $\eta^2 = .66$, but the main effect for block was not significant, $F(1, 122) = 1.24$, $p = .269$. The race of face by measure interaction was significant, $F(1, 122) = 18.40$, $\eta^2 = .13$, and the race of face by block by measure interaction was also significant, $F(1, 122) = 8.77$, $\eta^2 = .07$. In order to unpack these interactions, I examined results separately by block.

Results of a 2 (Measure: Predicted, Recognition) x 2 (Race of Face: Black, White) repeated-measures ANOVA (Table 1) revealed that, on Block 1, actual recognition was greater than predicted by participants' JOLs, $F(1, 123) = 153.36$, $\eta^2 = .56$. A main effect was found for race of face, $F(1, 123) = 156.78$, $\eta^2 = .56$, qualified by a reliable measure by race of face interaction, $F(1, 123) = 26.79$, $\eta^2 = .18$. A priori pairwise comparisons revealed that JOLs were higher for White faces than for Black faces, $t(123) = 7.56$, $d = .50$, and that recognition was also higher for White faces than for Black faces, $t(123) = 9.70$, $d = .99$. Thus, the effect of race of face was greater for recognition than for JOLs on Block 1.

Regarding Block 2, results of a 2 (Measure: Predicted, Recognition) x 2 (Race of Face: Black, White) repeated-measures ANOVA (Table 1) further indicated that, as in Block 1, actual recognition exceeded predicted recognition, $F(1, 123) = 122.04$, $\eta^2 = .50$. A main effect was found for race of face, $F(1, 123) = 120.43$, $\eta^2 = .50$, but the measure by race of face interaction was not significant, $F(1, 123) = 3.05$, $p = .083$. Thus, both recognition and JOLs were greater for White faces than for Black faces, and the effect of race of face did not differ reliably across measure on Block 2.

Relative Accuracy. As in the pilot experiment, relative accuracy was examined via gamma correlations between JOLs and recognition for each face. However, correlations were examined only for those faces participants rejected for restudy, as JOLs could not be meaningfully examined for restudied faces. That is, JOLs predict future memory based upon memory at the time of judgment. However, given that restudying may alter memory, the relationship between JOLs and memory should be weaker for restudied faces than for faces participants chose not to restudy. Learners were able to use JOLs to correctly distinguish between individual faces they later recognized (or failed to recognize) at test, ($G = .27$, SE =

.04), *t*(105) = 6.56, both for Black faces (*M* = .20, *SE* = .05), *t*(98) = 4.22, and for White faces (*M* = .25, *SE* = .07), *t*(86) = 3.71. Moreover, there was no reliable difference in relative accuracy for individual Black and White faces, *t*(81) = .71, *p* = .482. Thus, relative accuracy was largely similar for Black and White faces.

Mean gamma correlations for Blocks 1 and 2 may be found in Table 2. On Block 1, participants' JOLs reliably distinguished between individual faces they later did or did not recognize, *t*(94) = 2.81, but these results were not reliable when examined separately by race of face, either for Black faces, *t*(84) = 1.52, *p* = .133, or for White faces, *t*(59) = 1.28, *p* = .206. In addition, JOL accuracy did not differ reliably between Black and White faces, *t*(50) = 1.21, *p* = .231. As in Block 1, JOLs in Block 2 distinguished between individual faces they later did or did not recognize at test, *t*(92) = 6.34. However, unlike Block 1, JOLs reliably distinguished between recognized faces and faces participants failed to recognize, both for Black faces, *t*(76) = 4.62, and for White faces, *t*(69) = 5.06. As in Block 1, JOL accuracy did not reliably differ between Black and White faces, *t*(56) = .46, *p* = .649.

Control. In order to examine whether restudy selections differed by race of face, a 2 (Race of Face: Black, White) x 2 (Block: 1, 2) repeated-measures ANOVA was conducted only on those faces selected for restudy[3]. Overall, a greater proportion of Black faces (*M* = .45, *SE* = .03) than White faces (M = .35, *SE* = .03) were selected for restudy, *F*(1, 123) = 61.55, η^2 = .33. No significant differences in restudy selection were found by block, *F*(1, 123) = 1.07, *p* = .304,

[3] If participants did not choose whether to restudy or not restudy a face (or make a JOL) within the time limit, the program moved on to the next face. Restudy choices were missing for 1.1% of White faces and 1.0% of Black faces. JOLs were missing for 2.6% of White faces and 2.4% of Black faces. Whenever JOLs or restudy selection were relevant factors, those analyses were conducted only for the proportion of faces for which participants provided a response.

and the race of face by block interaction was not significant, *p* = .888. Thus, a greater proportion of Black than White faces were selected for restudy, regardless of block.

If participants choose to restudy faces assessed as more difficult to learn, then faces selected for restudy should be given lower JOLs than faces not selected for restudy. This was confirmed in a 2 (Restudy Selection: Selected, Not Selected) x 2 (Race of Face: Black, White) x 2 (Block: 1, 2) repeated-measures ANOVA (see Figure 4). JOLs were reliably lower for faces selected for restudy than for faces that were rejected, $F(1, 75) = 276.41$, $\eta^2 = .79$. JOLs were also lower for Black faces than for White faces, $F(1, 75) = 13.14$, $\eta^2 = .15$, and JOLs were lower on Block 1 ($M = 73.29$, $SE = .69$) than on Block 2 ($M = 74.42$), $F(1, 75) = 5.29$, $\eta^2 = .07$. No other findings reliably differed when examined separately by block.

How did participants use their monitoring judgments to select or reject individual faces for restudy? (See Table 3). Overall, participants were more likely to select low-JOL faces for restudy, ($G = -.75$, $SE = .03$), $t(100) = 23.00$. This pattern of results held for both Black faces, ($G = -.76$, $SE = .03$), $t(97) = 23.09$, and for White faces, ($G = -.76$, $SE = .04$), $t(97) = 21.22$, and did not differ significantly by race of face, $t(94) < 1$. Similar results were found when examined separately by block, except that for Black faces, participants chose to restudy faces with higher JOLs on Block 2 than on Block 1, $t(82) = 2.79$. In summary, for both Black and White faces, participants used their monitoring judgments similarly when making control decisions, supporting the monitoring-affects-control hypothesis.

Does restudying enhance recognition for the restudied faces? In particular, can restudying enhance memory for other-race faces? A 2 (Race of Face: Black, White) x 2 (Restudy Selection: Selected, Not Selected) x 2 (Block: 1, 2) repeated-measures ANOVA (Figure 5) found that recognition was better for White faces than for Black faces, $F(1, 76) = 40.19$, $\eta^2 = .35$.

Recognition was also greater for faces participants restudied than for faces not chosen for restudy, $F(1, 76) = 14.90$, $\eta^2 = .16$. No main effect was found for block, $F(1, 76) = 1.76$, $p = .189$, but this was qualified by a significant race of face by block interaction, $F(1, 76) = 5.24$, $\eta^2 = .06$. A priori pairwise comparisons revealed that, amongst faces participants selected to restudy or not to restudy, the difference in recognition for Black and White faces was smaller on Block 2, $t(123) = 6.59$, $d = .67$, than on Block 1, $t(123) = 9.43$, $d = .98$. Recognition for restudied White faces ($M = .94$, $SE = .01$) was greater than recognition for White faces participants chose not to restudy ($M = .90$, $SE = .01$), $t(98) = 3.51$, $d = .40$. Most important, recognition for restudied Black faces ($M = .85$, $SE = .02$) was greater than recognition for Black faces participants chose not to restudy ($M = .79$, $SE = .02$), $t(98) = 2.66$, $d = .34$, suggesting that restudying can improve recognition of other-race faces.

Self-Reported Contact. As in Hancock and Rhodes (2008), regression analyses[4] were used to examine whether recognition for other-race faces was affected by contact with other-race faces (see Appendix for materials). Using a scale ranging from 1-6, with increasing numbers indicating greater experience with individuals of that race, participants reported more experience with own-race (White) individuals ($M = 5.24$, $SE = .06$) than with other-race (Black) individuals ($M = 3.24$, $SE = .08$; Table 4)[5]. A simple linear regression indicated that self-reported contact with Black individuals did not reliably predict recognition for Black faces, $R^2 < .01$.

Notably, it is possible that the failure to find an effect of self-reported contact with Black individuals on the ORB may be due to the low levels of contact participants reported having with

---

[4] Hancock and Rhodes (2008) used multiple linear regression in order to assess the effects of participant race (White or Chinese) and self-reported contact on recognition for other-race faces. As all of the participants in the current study were White, however, only a simple linear regression needed to be run.

[5] Self-reported contact results were missing for 17.7% of participants, due to either experimenter error or failure to comply with the survey instructions.

Black individuals. It is possible that, among individuals with greater levels of expertise with other-race faces, the ORB may be reduced. Therefore, the data were divided into quartiles as a function of average self-reported contact with Black individuals, and the primary analyses of interest (absolute accuracy, relative accuracy, and recognition by restudy selection) were examined only on those participants reporting the highest level of contact with Black individuals (4^{th} quartile). As the power of these analyses is greatly reduced, they are reported simply as indicators of whether similar results might be found in a population with higher levels of contact with other-race individuals.

Amongst participants with high levels of self-reported contact with Black individuals, absolute accuracy was examined using a 2 (Measure: Predicted, Recognition) x 2 (Race of Face: Black, White) repeated-measures analysis (Figure 6). As in the overall analysis, actual recognition was greater than predicted by participants' JOLs, $F(1, 24) = 36.07$, $\eta^2 = .60$. A main effect was found for race of face, $F(1, 24) = 42.53$, $\eta^2 = .64$, but the measure by race of face interaction was not significant, $F(1, 24) = 1.03$, $p = .320$, indicating that both actual and predicted recognition were higher for White faces than for Black faces.

Relative accuracy was likewise examined amongst participants with high levels of self-reported contact with Black individuals, through nonparametric gamma correlations (Nelson, 1984). Mean gamma correlations indicated that, for faces that were not restudied, participants' JOLs did not differentiate between faces that were recognized and those that they failed to recognize, ($M = .14$, $SE = .10$), $t(17) = 1.31$, $p = .209$, either for Black faces ($M = -.01$, $SE = .15$), $t(16) = .06$, $p = .957$, or for White faces ($M = .22$, $SE = .20$), $t(14) = 1.08$, $p = .297$. Relative accuracy also did not differ significantly between Black faces and White faces, $t(14) = 1.08$, $p =$

.297. Thus, the relationship between JOLs and recognition was similar for both White and Black faces.

Finally, recognition was examined for faces which high-contact participants selected or rejected for restudy in a 2 (Restudy Selection: selected, rejected) x 2 (Race of Face: Black, White) repeated-measures ANOVA (Figure 7). A main effect for restudy selection indicated that, as in the overall analyses, recognition was greater for faces participants restudied (M = .92, SE = .01), than for faces rejected for restudy (M = .86, SE = .02), $F(1, 16) = 16.26$, $\eta^2 = .50$. Recognition was also greater for White than Black faces, $F(1, 16) = 15.07$, $\eta^2 = .49$. However, unlike in the overall analyses, the restudy selection by race of face interaction was significant, $F(1, 16) = 4.52$, $\eta^2 = .22$. Results of a priori follow-up comparisons revealed that restudying enhanced recognition for Black faces, $t(17) = 2.72$, $d = .88$, but that restudying did not significantly enhance recognition for White faces, $t(18) = 1.30$, $p = .209$.

General Discussion

Recognition is poorer for faces of other races or ethnicities than one's own – a finding termed the own-race bias (ORB; Meissner & Brigham, 2001). Although different theories have been proposed for why the ORB might occur, little work has examined the influence of metacognitive processes on this memory bias. A widely held view in metamemory posits that monitoring accuracy affects how learners control their learning (Nelson & Narens, 1990; 1994). Given that monitoring affects control, one factor contributing to the ORB may be that learners are unaware that the bias exists and do not exert appropriate control over learning. If learners are unaware of the ORB, it should be reflected in both learners' monitoring and control of study.

Accordingly, monitoring and control of learning other-race faces was examined in a pilot experiment and subsequent study. Participants studied a series of own-race (White) and other-race (Black) faces and made both monitoring and control decisions. Recognition was greater for White faces than Black faces, consistent with the ORB. Additionally, JOLs on average were lower for Black than for White faces, reflective of the ORB in memory performance. However, the ORB was not apparent in measures of relative accuracy. That is, JOLs distinguished between individual faces that later were or were not recognized, regardless of race of face. Restudy selections revealed that participants selected a greater proportion of Black than White faces for restudy, and that lower-JOL faces were more likely to be selected for restudy. Most important, recognition was greater for Black faces that were restudied than for Black faces that were not restudied, demonstrating that recognition for other-race faces can be enhanced via a manipulation that should have no impact on expertise. However, despite this recognition enhancement, restudying did not benefit recognition for Black faces to a greater extent than for White faces. Thus, although restudying *improved* memory for other-race faces, this improvement was no greater for Black faces than for White faces, failing to result in a reduction of the ORB.

How do these results contribute to theory? First, consider the perceptual-expertise account of the ORB. According to this account, the ORB is driven by a lack of expertise with other-race faces (G. Rhodes et al., 1989). Enhanced recognition for restudied other-race faces fails to support the perceptual-expertise account, as this enhancement was not accompanied by greater expertise for Black faces. In fact, the perceptual-expertise account fails to make predictions regarding awareness of cognition (i.e., metacognition), and thereby cannot account for either the monitoring or control differences for own- and other-race faces. Further difficulties with this account as an explanation for the ORB were found in participants' self-reported expertise with other-race faces. As demonstrated by the regression for measures of self-reported contact, expertise with other-race faces accounted for less than 1% of the variance in recognition between own- and other-race faces, comparable to the 2% variance reported by Meissner and Brigham (2001) in their meta-analysis. Although Hancock and Rhodes (2008) found a greater impact of expertise on recognition, their design purposefully utilized participants with extreme differences in other-race expertise.

The current study results suggest that, amongst the participants sampled for this study, perceptual-expertise does not contribute significantly to the ORB. Analyses confined only to participants reporting high levels (within the 4th quartile) of contact with Black individuals revealed that absolute accuracy and relative accuracy were similar amongst this subgroup of participants, further suggesting that expertise does not affect recognition accuracy. However, unlike results for participants with a range of experience with Black faces, for participants with high levels of contact with Black faces, restudying enhanced recognition for Black, but not White faces. These findings suggest that expertise may play a larger role in the ORB (amongst participants with high levels of other-race contact) than found in the overall results of the current study (perhaps even reducing the ORB). However, it is also possible that this difference in findings may be due to the reduced power of analyzing a subset of this study's population. These possibilities may be examined using a population of individuals with high levels of other-race contact, after ensuring sufficient power to find significant effect sizes if they do exist.

Next, consider the social-categorization account of the ORB. According to this account, the ORB is driven by a lack of motivation to differentiate between out-group (or other-race) faces (Bernstein et al., 2007). This categorization of in-group versus out-group impacts the processing of faces, such that out-group faces are processed in terms of similarities, whereas in-group faces are processed in terms of differences. Likewise, this categorical processing then affects recognition for these faces. This account, therefore, provides room to acknowledge potential cognitive or metacognitive contributions to the ORB. Results of the current study indicated that participants were aware of a potential recognition deficit for other-race faces (evinced by lower JOLs for other-race faces), and that learners were further motivated to learn faces of both races which they perceived as poorly learned (evinced by lower JOLs for faces selected than those rejected for restudy, as well as a greater proportion of Black than White faces selected for restudy). That is, participants appeared to both identify deficits in face recognition and to demonstrate motivation to work towards reducing these deficits. The fact that recognition was enhanced for restudied other-race faces further suggests that participants were *successfully* motivated to differentiate between other-race faces, once they were able to control their learning by restudying selected faces. I note that full confirmation of this possibility requires future work to manipulate whether participants' restudy selections were honored or dishonored (cf., Kornell & Metcalfe, 2006). Under this paradigm, after selecting or rejecting faces for restudy, some participants would be permitted to restudy (or reject for restudy) these faces, as they had previously indicated (honor condition). However, other participants' restudy selections would not be followed – they would instead restudy faces they had rejected for restudy, and would not be permitted to restudy the faces they had selected (dishonor condition). Through this means, the accuracy of restudy selections can be examined. If participants are accurate in their restudy selections, then recognition performance should be greater in the honor than the dishonor condition. In all, results of the current study indicate that theoretical accounts of the ORB should account for monitoring and control of own- and other-race faces.

The current study was not alone in finding differences in monitoring accuracy for own- and other-race faces (Hourihan et al., 2012, Rhodes et al., 2013). Hourihan et al. (2012) reported greater relative accuracy for individual own-race than other-race faces, indicating that participant

JOLs were more predictive of recognition accuracy for White than for Black faces. Further, Rhodes et al. (2013) reported lower aggregate JOLs for Black faces than for White faces (*d* = .13), indicating, as in the current study, that participants were aware of a greater difficulty in learning other-race (versus own-race) faces. Similar to our results, Rhodes et al. did not find significant differences in relative accuracy for Black and White faces; however, in Experiment 1, relative accuracy differed significantly from 0 for White, but not Black faces, whereas in Experiment 2, relative accuracy did not differ from 0 for either Black or White faces. Results from the current experiment, however, found that participants could use their JOLs to differentiate between faces they did or did not recognize at test, both for Black and White faces. For this difference in results, one possible contributing factor is that the current study utilized delayed JOLs in an effort to do this very thing – enhance relative accuracy (Nelson & Dunlosky, 1991). As a second contributing factor the current study utilized a greater number of subjects, achieving significance with greater power.

An important limitation of the current study is that although measures of monitoring and control indicate sensitivity towards the ORB, it remains unclear whether participants made effective control decisions to enhance their memory for own-race and other-race faces. In order to accurately examine this, it would be necessary to utilize a honor/dishonor paradigm (as in Rhodes et al., 2013) to determine whether participant study choices were effective. For example, although participants gave lower JOLs to faces they selected for restudy (and demonstrated enhanced memory for faces they restudied), it is possible that they would have recognized these faces even without restudy opportunities. Under an honor/dishonor paradigm, the impacts of monitoring and control accuracy could be examined separately from each other. If control accuracy were poor, then face recognition would be no better (or perhaps worse) if their study decisions are honored than if their study decisions were dishonored. Conversely, if monitoring accuracy were poor, then recognition for faces rejected for restudy should be enhanced if these restudy rejections were dishonored, rather than honored. Lastly, due to the nature of the population sampled, both reported experiments used only White or Caucasian individuals. Although previous work has indicated that both Black and White participants demonstrate the ORB (Bennett-Day, 2007, Meissner et al., 2005), future work should examine whether these results generalize in a crossover design.

As a unique contribution to the literature, this study examined whether monitoring and control (as well as recognition) for faces changed with task experience. To examine this idea, results were further examined by block (i.e., the first and second halves of the experiment). Regardless of race of face, relative accuracy was poor during the first half of the experiment, but improved with task experience, as in the second half of the experiment JOLs were positively and significantly related to recognition for both Black and White faces (i.e., faces recognized at test were more likely to have been given high JOLs). Prior research examining monitoring in the ORB (Hourihan et al., 2012; Rhodes et al., 2013; Schmechel et al., 2006; Smith et al., 2004) has not examined whether monitoring changes with task experience, nor has it examined whether accuracy itself can change with task experience. As accuracy was enhanced for other-race as well as own-race faces, these findings suggest that participants may be able to modify their metamemory to reduce the ORB, even without an accompanying enhancement in expertise, inconsistent with a perceptual-expertise account of the ORB. Further research should be conducted on this area.

In summary, measures of monitoring and control revealed awareness of the ORB and an accompanying effort to reduce this bias. Relative accuracy improved with task experience for faces of both races, demonstrating an ability to update and modify monitoring and control judgments to better conform to actual recognition performance. Future research will be conducted to examine whether honoring or dishonoring control decisions reveals further information about the accuracy of monitoring and control in the ORB. These results provide evidence that current theories of the ORB should account for monitoring, and its influence on any opportunities for further learning. In everyday situations, if individuals are aware of difficulties in recognizing other-race faces, they may often have the opportunity to continue learning those faces. The current research indicates that these additional learning opportunities may enhance recognition for own- and other-race faces.

Table 1. Mean recognition and predicted recognition for Black and White faces in Experiment, examined separately by block.

	Block 1		Block 2	
	Recognized	Predicted	Recognized	Predicted
Black	0.81 (.01)	0.72 (.01)	0.83 (.01)	0.72 (.01)
White	0.92 (.01)	0.76 (.01)	0.90 (.01)	0.77 (.01)

Table 2. Mean gamma correlations between JOLs and recognition for faces rejected for restudy in Experiment as a function of block.

	Block 1	Block 2
All Races	0.16 (.06)	0.34 (.05)
Black	0.10 (.07)	0.30 (.06)
White	0.12 (.09)	0.38 (.08)

Table 3. Mean gamma correlations between JOLs and restudy selection for Black and White faces in Experiment. Standard errors of the mean are in parentheses.

Block	All Blocks	Block 1	Block 2
Race of Face			
All Races	-.75 (.03)	-.71 (.04)	-.81 (.04)
Black	-.76 (.03)	-.71 (.04)	-.82 (.04)
White	-.76 (.04)	-.75 (.04)	-.81 (.04)

Table 4. Mean self-reported contact with own-race and other-race individuals. Standard errors of the mean are in parentheses.

		Race of Face	
	Questions	Own-Race	Other-Race
(1)	I know lots of Black people	-	3.46 (.10)
(2)	I interact with White people during recreational periods	5.03 (.07)	-
(3)	I live, or have lived in an area where I interact with White people	5.49 (.07)	-
(4)	I live, or have lived in an area where I interact with Black people	-	3.99 (.12)
(5)	I interact with Black people during recreational periods	-	3.90 (.11)
(6)	I interact with White people on a daily basis	5.56 (.07)	-
(7)	I socialize a lot with White people	5.44 (.07)	-
(8)	I went to a high school where I interacted with Black students	-	4.17 (.12)
(9)	I socialize a lot with Black people	-	3.58 (.10)
(10)	I know lots of White people	5.59 (.06)	-
(11)	I generally only interact with Black people	-	1.90 (.08)
(12)	I interact with Black people on a daily basis	-	3.71 (.11)
(13)	I went to a high school where I interacted with White students	5.48 (.08)	-
(14)	I generally only interact with White people	3.54 (.12)	-
(15)	I have lived in a country where the predominant race is Black	-	1.96 (.09)

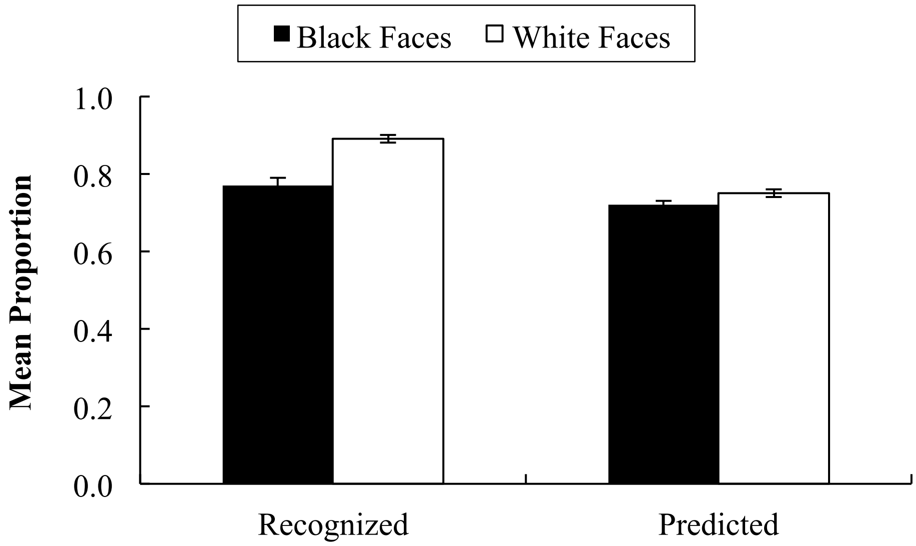

Figure 1. Mean proportion of faces recognized and predicted recognized by race of face in Pilot, collapsed across Blocks 1 and 2. Error bars represent standard error of the mean.

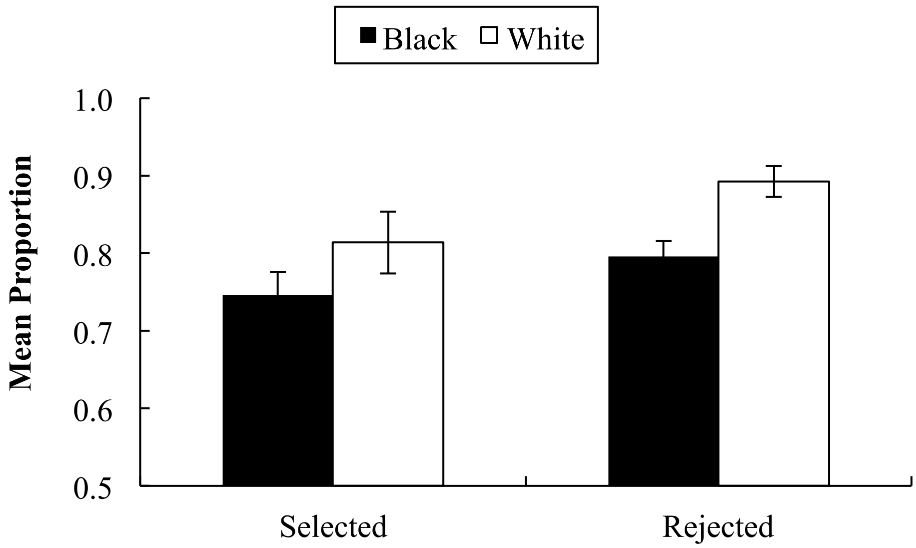

Figure 2. Mean proportion of faces recognized by restudy selection and race of face in Pilot, collapsed across Blocks 1 and 2. Error bars represent standard error of the mean.

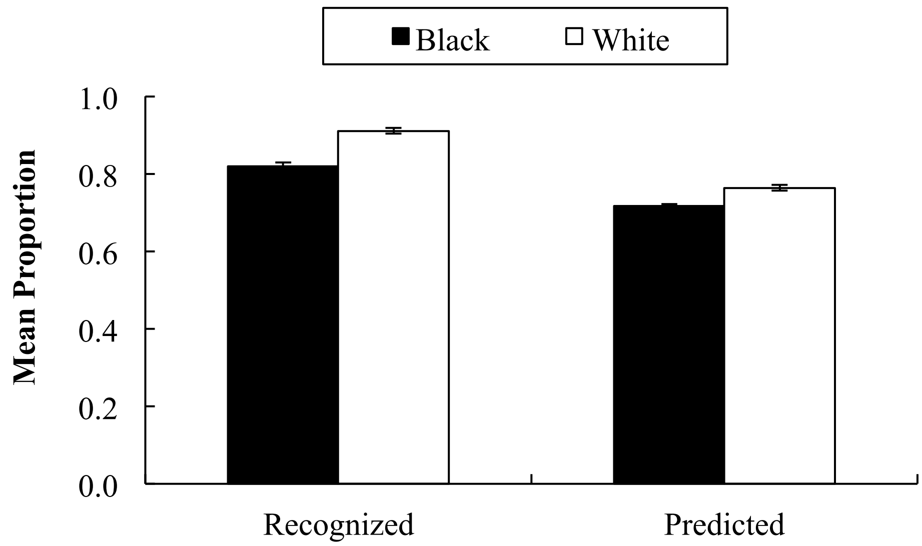

Figure 3. Mean proportion of faces recognized and predicted recognized by race of face in Experiment, collapsed across Blocks 1 and 2. Error bars represent standard error of the mean.

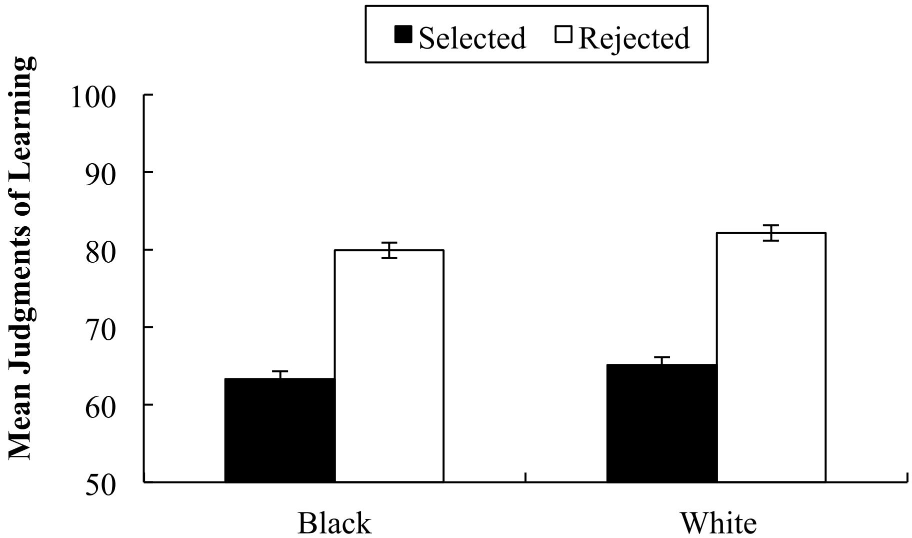

Figure 4. Mean judgments of learning (JOLs) in Experiment for Black and White faces by restudy selection. Error bars represent standard error of the mean.

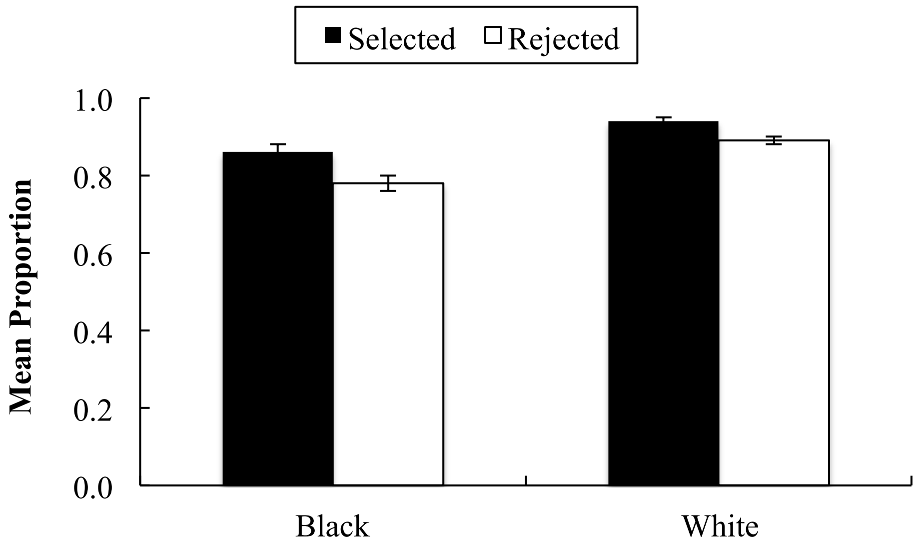

Figure 5. Mean proportion recognized by restudy selection and race of face in Experiment, collapsed across Blocks 1 and 2. Error bars represent standard error of the mean.

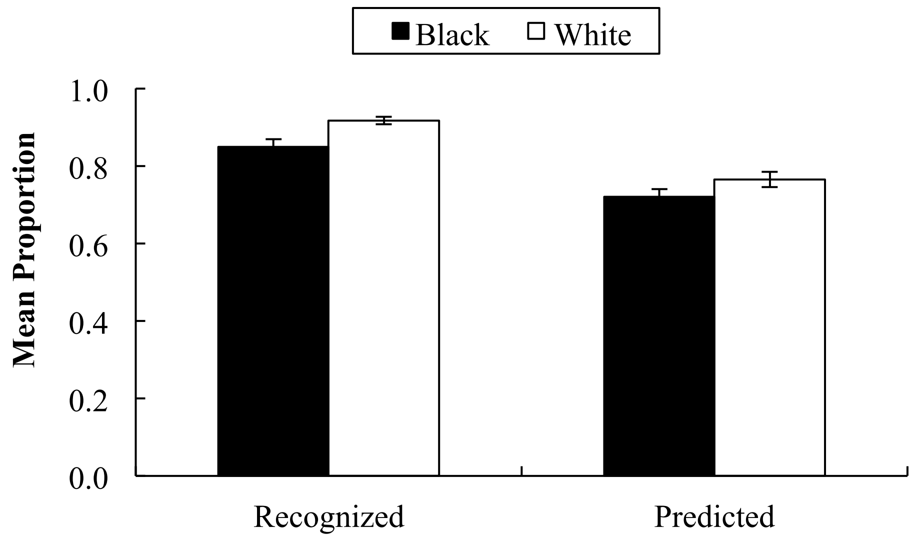

Figure 6. Mean proportion of faces recognized and predicted recognized by race of face amongst participants with high levels of self-reported contact with Black individuals in Experiment, collapsed across Blocks 1 and 2. Error bars represent standard error of the mean.

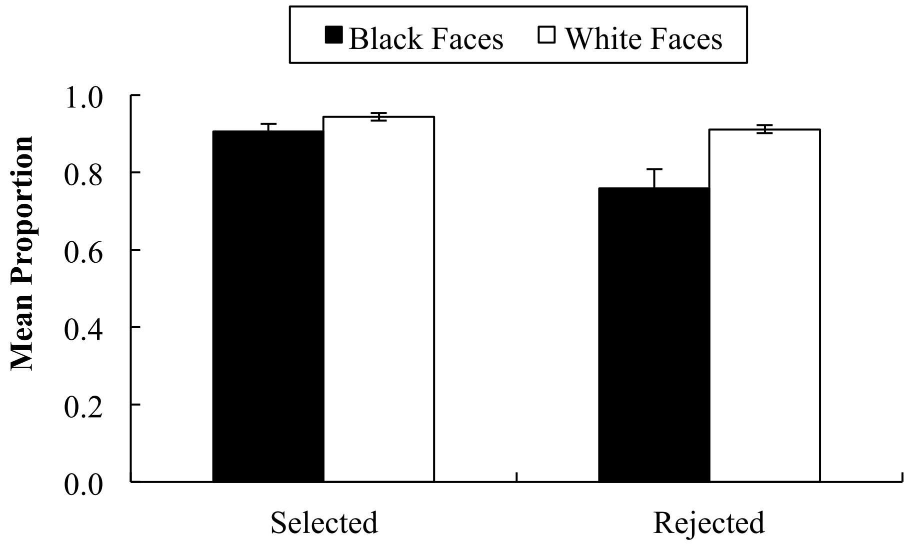

Figure 7. Mean proportion recognized by restudy selection and race of face amongst participants with high levels of self-reported contact with Black individuals in Experiment, collapsed across Blocks 1 and 2. Error bars represent standard error of the mean.

References

Benjamin, A. S., Bjork, R. A., Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General, 127*, 55–68.

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science*, *18*, 706–712.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444.

Brigham, J. C. & Williamson, N. L. (1979). Cross-racial recognition and age: When you're over 60, do they still "all look alike?" *Personality & Social Psychology Bulletin, 5*, 218-222.

Chiroro, P. M., Tredoux, C. G., Radaelli, S., & Meissner, C. A. (2008). Recognizing faces across continents: The effect of within-race variations on the own-race bias in face recognition. *Psychonomic Bulletin & Review*, *15*, 1089–1092.

Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging, 12*, 50–71.

DeLozier, S., & Rhodes, M. G. (2015). The impact of value-directed remembering on the own-race bias. *Acta Psychologica, 154*, 62-68.

Dunlosky, J., & Hertzog, C. (1998). Aging and deficits in associative memory: What is the role of strategy production? *Psychology and Aging*, *13*, 597–607.

Dunlosky, J., & Lipko, A. A. R. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science, 16*, 228–233.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition*, *20*, 374-380.

Dunlosky, J., Serra, M., Baker, J.M.C., (2007). Metamemory applied. In F.T. Durso, R.S. Nickerson, S.T. Dumais, S. Lewandowsky (Eds.), *Handbook of Applied Cognition* (pp. 137–162. New York: Wiley.

Dunlosky, J., & Thiede, K. W. (1998). What makes people study more? An evaluation of factors that affect self-paced study. *Acta Psychologica, 98*, 37–56.

Hehman, E., Mania, E. W., & Gaertner, S. L. (2010). Where the division lies: Common ingroup identity moderates the cross-race facial-recognition effect. *Journal of Experimental Social Psychology*, *46*, 445–448.

Hourihan, K. L., Benjamin, A. S., & Liu, X. (2012). A cross-race effect in metamemory: Predictions of face recognition are more accurate for members of our own race. *Journal of Applied Research in Memory and Cognition*, *1*, 158–162.

Hourihan, K. L., Fraundorf, S. H., & Benjamin, A. S. (2013). Same faces, different labels: Generating the cross-race effect in face memory with social category information. *Memory & Cognition*, *41*, 1–37.

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

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review*, *117*, 1168–1187.

Kassin, S. M., Ellsworth, P. C., & Smith, V. L. (1989). The "general acceptance" of psychological research on eyewitness testimony: A survey of the experts. *American Psychologist, 44*, 1089-1098.

Kimball, D. R., & Metcalfe, J. (2003). Delaying judgments of learning affects memory, not metamemory. *Memory & Cognition*, *31*, 918–929.

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, and E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 289–325). New York: Cambridge UP.

Koriat, A., & Bjork, R. A. (2006). Illusions of competence during study can be remedied by manipulations that enhance learners' sensitivity to retrieval conditions at test. *Memory & Cognition*, *34*, 959–972.

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General*, *133*, 643-656.

Kornell, N., & Bjork, R. A. (2009). A stability bias in human memory: Overestimating remembering and underestimating learning. *Journal of Experimental Psychology: General*, *138*, 449-468.

Koriat, A., & Bjork, R. A. (2006). Illusions of competence during study can be remedied by manipulations that enhance learners' sensitivity to retrieval conditions at test. *Memory & Cognition, 34*, 959–972.

Koriat, A., & Goldsmith, M. (2013). Memory as something that can be counted versus memory as something that can be counted on. *Basic and Applied Memory Research: Volume 1: Theory in Context; Volume 2: Practical Applications,* 3.

Kornell, N., & Metcalf, J. (2006). Study efficacy and the region of proximal learning framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 32,* 609-622.

Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S. K. (2011). The ease-of-processing heuristic and the stability bias: Dissociating memory, memory beliefs, and memory judgments. *Psychological Science*, *22*, 787–794.

Leonesio, R. J., & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16*, 464-470.

Maki, R. H., Shields, M., Wheeler, A. E., & Zacchilli, T. L. (2005). Individual differences in absolute and relative metacomprehension accuracy. *Journal of Educational Psychology*, *97*, 723.

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, *13*, 330–334.

McCabe, J. (2011). Metacognitive awareness of learning strategies in undergraduates. *Memory & Cognition*, *39*, 462–76.

McGillivray, S., & Castel, A. D. (2011). Betting on memory leads to metacognitive improvement by younger and older adults. *Psychology and Aging*, *26*, 137–42.

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., Brigham, J. C., & Butz, D. A. (2005). Memory for own- and other-race faces: A dual-process approach. *Applied Cognitive Psychology*, *19*, 545–567.

Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? *Journal of Experimental Psychology: General*, *131*, 349–363.

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, *15*, 174–179.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *84*, 93–116.

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science, 2*, 267-270.

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the "labor-in-vain effect". *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 676–686.

Nelson, T.O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G.H. Bower (Ed.) *The psychology of learning and motivation* (Vol. 26, pp. 125-141). New York: Academic Press.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.

Nelson, T. O., Narens, L., & Dunlosky, J. (2004). A revised methodology for research on metamemory: Pre-judgment recall and monitoring (PRAM). *Psychological Methods, 9*, 53-69.

Pezdek, K., Blandon-Gitlin, I., & Moore, C. (2003). Children's face recognition memory: More evidence for the cross-race effect. *Journal of Applied Psychology, 88*, 760–763.

Rhodes, G., Brake, S., Taylor, K., & Tan, S. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology*, *80*, 313–331.

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

Rhodes, G., Locke, V., Ewing, L., & Evangelista, E. (2009). Race coding and the other-race effect in face recognition. *Perception*, *38*, 232-241.

Rhodes, M. G. (In Press). Judgments of Learning: Methods, data, and theory. In J. Dunlosky & S. K. Tauber (Eds.), *The Oxford Handbook of Metamemory*. Oxford University Press.

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General, 137*, 615–625.

Rhodes, M. G., & Castel, A. D. (2009). Metacognitive illusions for auditory information: effects on monitoring and control. *Psychonomic Bulletin & Review*, *16*, 550–554.

Rhodes, M. G., Sitzman, D. M., & Rowland, C. A. (2013). Monitoring and control of learning own-race and other-race faces. *Applied Cognitive Psychology*, *27*, 553-563.

Rhodes, M. G., & Tauber, S. K. (2011). The influence of delaying judgments of learning on metacognitive accuracy: A meta-analytic review. *Psychological Bulletin*, *137*, 131–148.

Sangrigoli, S., Pallier, C., Argenti, A. M., Ventureyra, V. A. G., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science, 16*, 440–444.

Schmechel, R. S., O'Toole, T. P., Easterly, C., & Loftus, E. F. (2006). Beyond the ken? Testing jurors' understanding of eyewitness reliability evidence. *Jurimetrics*, *46*, 177–214.

Shriver, E. R., Young, S. G., Hugenberg, K., Bernstein, M. J., & Lanter, J. R. (2008). Class, race, and the face: Social context modulates the cross-race effect in face recognition. *Personality & Social Psychology Bulletin*, *34*, 260–274.

Son, L. K., & Kornell, N. (2009). Simultaneous decisions at study: time allocation, ordering, and spacing. *Metacognition and Learning*, *4*, 237–248.

Thiede, K. W. (1999). The importance of monitoring and self-regulation during multitrial learning. *Psychonomic Bulletin & Review*, *6*, 662–667.

Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1024–1037.

Turk, D. J., Handy, T. C., & Gazzaniga, M. S. (2005). Can perceptual expertise accountfor the own-race bias in face recognition? A split-brain study. *Cognitive Neuropsychology*, *22*, 877–883.

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

Young, S. G., & Hugenberg, K. (2012). Individuation motivation and face experience can operate jointly to produce the own-race bias. *Social Psychological and Personality Science, 3*, 1–8.

Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and motivation in face recognition: A critical review of theories of the Cross-Race Effect. *Personality and Social Psychology Review, 16*, 116–42.

Zimmerman, C. A., & Kelley, C. M. (2010). "I'll remember this!" Effects of emotionality on memory predictions versus memory performance. *Journal of Memory and Language*, *62*, 240-253.