Journal of Applied Research in Memory and Cognition 1 (2012) 158–162

Contents lists available at SciVerse ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

A cross-race effect in metamemory: Predictions of face recognition are more accurate for members of our own race☆

Kathleen L. Hourihan[a,*], Aaron S. Benjamin[b], Xiping Liu[c]

[a] Department of Psychology, Memorial University of Newfoundland, Canada
[b] Department of Psychology, University of Illinois at Urbana-Champaign, United States
[c] Department of Psychology, Tianjin Normal University, China

ARTICLE INFO

Article history:
Received 15 December 2011
Received in revised form 25 June 2012
Accepted 25 June 2012
Available online 2 July 2012

Keywords:
Metamemory
Face recognition
Eyewitness memory
Cross-race effect

ABSTRACT

The Cross-Race Effect (CRE) is the well-replicated finding that people are better at recognizing faces from their own race, relative to other races. The CRE reveals systematic limitations on eyewitness identification accuracy, suggesting that some caution is warranted in evaluating cross-race identification. The CRE is problematic because jurors value eyewitness identification highly in verdict decisions. We explore how accurate people are in *predicting* their ability to recognize own-race and other-race faces. Caucasian and Asian participants viewed photographs of Caucasian and Asian faces, and made immediate judgments of learning during study. An old/new recognition test replicated the CRE: both groups displayed superior discriminability of own-race faces. Importantly, relative metamnemonic accuracy was also greater for own-race faces, indicating that the accuracy of predictions about face recognition is influenced by race. This result indicates another source of concern when eliciting or evaluating eyewitness identification: people are less accurate in judging whether they will or will not recognize a face when that face is of a different race than they are. This new result suggests that a witness's claim of being likely to recognize a suspect from a lineup should be interpreted with caution when the suspect is of a different race than the witness.

The Cross-Race Effect (CRE; also known as the Other-Race Effect or Own-Race Bias) in face recognition is one of the most replicated findings in cognitive and social psychology (see Meissner & Brigham, 2001 for a review). Across a variety of contexts, experimental methods, and ethnic groups, humans have been shown to be better at remembering faces from their own race than faces from other races. This finding is particularly important for legal and psychological scholars who study eyewitness memory, as it indicates that we are more likely to falsely identify an innocent suspect if he or she is from a different race (Brigham, Bennett, Meissner, & Mitchell, 2007; Meissner & Brigham, 2001).

Understanding the legal implications of the CRE will ultimately require a broader consideration of the ecological contexts in which eyewitness identification takes place. The literature on recognition memory takes great care to control for extraneous variables and individual differences, and the CRE has been principally demonstrated in paradigms that derive from this tradition. The agenda for the researcher interested in the *metacognition* of such judgments is to extend those well controlled recognition experiments into situations in which individuals' abilities to *monitor* their learning and memory and *control* aspects of their processing are additionally assessed. The present research takes a first step in that direction by examining how effectively learners *predict* future memory performance for own- and other-race faces. Such judgments are critically important to assess in an eyewitness setting, as an individual's assessment of how well he or she will remember a face likely plays a major role in whether he or she volunteers to attempt to pick the perpetrator from a lineup. In addition, because metamnemonic judgments often reflect accessibility of the queried materials (e.g., Benjamin, 2005; Nelson & Dunlosky, 1991), the correspondence between judgments and actual recognition might be lower for other- than own-race faces. Such a result would have legal implications, because it would imply that eyewitnesses' self-assessments of their ability to recognize a perpetrator would be less accurate for perpetrators of a different race.

Cases exist that support this idea. In 1984, Jennifer Thompson, a Caucasian woman, was sexually assaulted by a man who broke into her apartment (read more on this case at

☆ This research was supported by grant R01AG026263 to Aaron Benjamin from the National Institutes of Health and by grant 09YJAXLX021 to Xiping Liu from the Chinese Ministry of Education.

* Corresponding author at: Department of Psychology, Memorial University of Newfoundland, St. John's, NL, A1B 3X9 Canada. Tel.: +1 709 864 8771.
E-mail address: khourihan@mun.ca (K.L. Hourihan).

2211-3681/$ – see front matter
http://dx.doi.org/10.1016/j.jarmac.2012.06.004

http://www.theinnocenceproject.org). Ms. Thompson reported having made considerable effort to memorize the face of her attacker (Cotton's Wrongful Conviction, n.d.) an African-American male, and was confident that she would be able to recognize him later. In 1985, and again in 1987, Ronald Cotton was convicted of assault, and served more than 10 years in prison for the crime before being exonerated by DNA evidence in 1995.

Some studies have examined metacognitive aspects of the CRE. Most prominently, researchers have examined how recognition confidence relates to accuracy. Confidence is a metacognitive assessment of accuracy that takes place at the time of, or following, the memory judgment, and is also an important factor in the courtroom. Jurors value eyewitness testimony highly in reaching a verdict (e.g., Benton, Ross, Bradshaw, Thomas, & Bradshaw, 2006), and indeed are instructed to weight the confidence of an eyewitness as an important factor in considering the value of the testimony (Neil v. Biggers, 1972). Unfortunately, research has shown that post-recognition confidence is an inconsistent predictor of accuracy in face recognition (e.g., Leippe & Eisenstadt, 2007; but see Lindsay, Read, & Sharma, 1998).

Much less has been done in understanding the metacognition of the CRE prior to the time of recognition. Smith, Stinson, and Prossor (2004) appear to be the only researchers to have collected predictive judgments from subjects in the context of a cross-race eyewitness experiment. After White subjects viewed a video of a staged theft (depicting either a White or Black perpetrator), but before being presented with a lineup, subjects were asked to rate both the clarity of their memory of the perpetrator and their confidence that they would be able to select the correct individual from the lineup. Smith et al. replicated the standard CRE in recognition accuracy: White subjects were more accurate at identifying the White perpetrator than the Black perpetrator. Importantly, they also found that pre-identification ratings of memory clarity were significantly higher in the own-race condition than in the other-race condition. Thus, there is some indication that judgments about memory differ between own- and other-race faces. Though these data indicate that subjects are more confident in their ability to recognize own-race faces, they have little to say about our ability to discriminate between faces that we will or will not remember within each group. The present experiment extends this literature by examining the correspondence between judgments and recognition for individual faces.

There has been considerable focus on the cognitive and social underpinnings of the CRE. Some theories focus on experience-based encoding differences (e.g., Valentine, 1991; Valentine & Endo, 1992). For example, it has been suggested that we learn to encode faces by focusing on features that are useful for differentiating individuals within our own race, but are suboptimal for differentiating other-race faces. However, this view suggests that individuals with extensive exposure to other races should be immune to the CRE, and in fact the amount of contact with other races typically plays only a very weak role in predicting the CRE (accounting for only 2% of variability in a meta-analysis; Meissner & Brigham, 2001; cf. He, Ebner, & Johnson, 2011).

Other theories concentrate on the social influences on face recognition. These social-cognitive theories suggest that faces are rapidly classified as in-group or out-group members (e.g., Sporer, 2001; see also Levin, 2000). In-group faces are further processed in an individuating manner, supporting subsequent recognition, whereas only category-defining features of out-group faces tend to be encoded. Other social-cognitive theories focus on the manner in which social motivation can affect the encoding and classification of faces (e.g., Hugenberg, Young, Bernstein, & Sacco, 2010).

Further emphasizing the contribution of superior encoding of own-race faces, Meissner, Brigham, and Butz (2005) proposed a dual-process account of the CRE. In their second experiment, participants provided Remember-Know-Guess responses following recognition of recently studied own- and other-race faces. Their results showed the typical CRE in overall recognition accuracy, but no difference between own- and other-race face recognition when considering familiarity-based responses (neither for hits nor false alarms). When considering recollection-based responses, however, own-race faces produced both a higher hit rate and lower false alarm rate than other-race faces. Meissner et al. concluded that own-race faces are encoded qualitatively better than other-race faces, which supports more accurate recollection-based recognition responses. Their findings have subsequently been replicated and extended in a process dissociation procedure (Marcon, Susa, & Meissner, 2009).

In the present paper, our goal is not to determine which of these accounts provides the best explanation of the mechanisms underlying the CRE, but rather to examine how the accuracy of metamemory judgments are influenced by the CRE. All of the explanations of the CRE discussed above rely at least partially on some form of differential encoding for own-race and other-race faces (regardless of whether the encoding differences are under volitional control). If the processing of own-race faces involves encoding attributes that enable more precise differentiation, then judgments of future memorability assessed at the time of encoding should also support superior differentiation.

An example of how superior differentiation can support both enhanced memory and metamemory accuracy can be seen in the metacognition literature for the case of word frequency. Recognition of uncommon words is superior to recognition of common words; this result derives at least in part from the fact that uncommon words are more distinctive from one another than common words. Though it was not the focus of that study, Benjamin (2003) reported a number of conditions in which metacognitive accuracy was assessed separately for uncommon and common words on a recognition task. In the relevant conditions from those experiments[1], metacognitive accuracy was higher for uncommon words in four out of four cases. The case of word frequency effects in recognition provides a concrete example of how differentiation affects the accuracy of memory and metamemory similarly.

In the context of word recognition, predictions may relate more strongly to recollection-based responses than to familiarity-based responses (e.g., Daniels, Toth, & Hertzog, 2009). As discussed above, encoding of own-race faces results in more accurate recollection-based recognition than does encoding of other-race faces, contributing to the overall CRE in recognition. Moreover, other-race faces are generally perceived as less distinctive from one another than are own-race faces (e.g., Meissner et al., 2005; Valentine & Endo, 1992). Given the greater perceived distinctiveness and higher recollection of own-race faces, it seems likely that predictions for own-race faces should be more strongly related to subsequent recognition accuracy than predictions for other-race faces.

In the present experiment, we used a standard recognition paradigm combined with a *judgment of learning* procedure to assess metamemory and memory judgments for own-race and other-race faces. This recognition procedure has been widely used in the face recognition literature (e.g., Hugenberg et al., 2010; Meissner et al., 2005; Valentine, 1991) and allows for the collection of multiple judgments across a wide range of own- and other-race faces from each subject. We tested two groups of subjects, one residing in the United States (Caucasian) and one residing in China (Asian). Each group of subjects studied an equal number of photographs of Caucasian and Asian faces. For each face, they were asked to

[1] These conditions include Experiment 1, the two Test 1 conditions from Experiment 2, and the Test 1 condition from Experiment 3.

predict the probability that they would recognize that face on a later recognition test: a *judgment of learning* (JOL)[2]. We then examined both memory accuracy and relative metamnemonic accuracy (the degree to which the JOLs provided at study actually predicted the later recognition outcome for a given face) for own-race and other-race faces. We predicted that recognition accuracy would be superior for own-race faces than for other-race faces, replicating the standard CRE. More importantly, we predicted that relative metamnemonic accuracy (i.e., the item-by-item accuracy of JOLs) would also be better for own-race faces than for other-race faces.

1. Method

1.1. Subjects

One hundred and two individuals participated in this experiment. The Caucasian subjects ($n = 50$) were students at the University of Illinois at Urbana-Champaign who participated for course credit or payment. The Asian subjects ($n = 52$) were students at Tianjin Normal University who received a token gift for their participation.

1.2. Materials

The stimulus pool consisted of 100 black and white pictures of faces on a plain background. The pool contained an equal number of Caucasian and Asian faces. The Caucasian faces were obtained from the Center for Vital Longevity Face Database (Minear & Park, 2004). An equal number of male and female faces were selected from the subset of the database that had been normed on various measures (Kennedy, Hope, & Raz, 2009). The Asian faces were photos taken of volunteers at the Institute of Psychology at the Chinese Academy of Sciences for use in research, and also included an equal number of males and females. All pictures showed the full face from the neck up with neutral expression, and were not modified to obscure hairstyle or any other distinctive features.

A random subset of 25 Caucasian and 25 Asian faces were selected for study for each subject; the remaining 25 Caucasian and 25 Asian faces were reserved for distractor items at test. Randomizations were carried out on an individual basis such that each subject saw a different random subset of faces at study and test, but always an equal number of Caucasian and Asian faces. E-Prime version 1.1 (Psychology Software Tools, Pittsburgh, PA) was used to randomize and present stimuli, and to record subject responses.

1.3. Procedure

Each subject completed the experiment individually in a small room with one computer. The Caucasian subjects completed the task as part of a 1-h session containing other, unrelated tasks, none of which used face stimuli or a similar procedure; the Asian subjects did not complete any other tasks in the session. The entire experiment took approximately 15 min. Prior to study, subjects were informed that they would be studying a series of faces, one at a time, for a later recognition memory test. They were instructed that they would be making a recognition prediction for each face after they had studied it.

The experimental program was identical for both groups of subjects, with the exception that instructions were printed in English for the Caucasian subjects, but in Chinese for the Asian subjects.

Table 1
Recognition memory performance for Caucasian and Asian subjects.

	Caucasian subjects		Asian subjects	
Hits				
Caucasian faces	.79	(.02)	.70	(.02)
Asian faces	.74	(.02)	.73	(.02)
False alarms				
Caucasian faces	.13	(.02)	.20	(.02)
Asian faces	.23	(.02)	.12	(.01)
d′				
Caucasian faces	2.19	(0.11)	1.56	(0.08)
Asian faces	1.51	(0.08)	1.98	(0.08)

Note: standard errors are shown in parentheses below the respective means.

Each study trial began with presentation of a face on a white background at the center of the screen for 3000 ms. Face stimuli were fit to fill the screen without distorting the aspect ratio of the original picture. Following a 1000 ms blank screen, the JOL screen appeared. Subjects were instructed to "Please indicate how likely you think it is that you will later recognize the face that you just studied." The numbers 1 through 9 were displayed at the bottom of the screen, with "I am sure that I will NOT remember this face" displayed below the number 1 and "I am sure that I WILL remember this face" displayed below the number 9 (the program for the Asian subjects displayed these messages in Chinese). This prediction screen stayed visible until the subject pressed a key from 1 to 9 to respond. A 1000 ms blank screen preceded the next trial.

Following completion of the study phase, the instructions for the recognition test appeared on the screen. Subjects were told that they would be shown one face at a time, some of which would be the faces they had just studied and some of which would be new faces. They were instructed to press the "m" key if they recognized the presented face as one they had studied and to press the "c" key if they believed the face was new. Each test trial began with the presentation of the face at the center of the screen, and the face remained visible until the subject responded with a keypress. The keypress labels for "studied" and "new" remained visible at the bottom of the screen while the face was visible. A 1000 ms blank screen preceded the next trial. All 100 faces in the stimulus pool were presented in random order during the test phase.

2. Results

2.1. Recognition performance

Mean hits, false alarms, and computed d' values[3] are displayed in Table 1. We discuss only the d' analysis, although we note that a similar pattern of results is observed in separate analyses of hits and false alarms. The d' scores were analyzed in a 2 (face: Caucasian vs. Asian) × 2 (group: Caucasian vs. Asian) mixed ANOVA, with face as a repeated-measures factor and group as a between-subjects factor. The main effect of face was marginally significant ($F(1,100) = 3.52$, $\eta_p^2 = 0.034$, $p = .06$) while the main effect of group was not significant ($F(1,100) = 0.60$, $\eta_p^2 = 0.006$, $p > 1$). Importantly, these effects were qualified by a significant face × group interaction ($F(1,100) = 62.76$, $\eta_p^2 = 0.386$, $p < .001$). As expected, planned comparisons revealed that both groups showed a significant CRE in recognition accuracy: Caucasian subjects showed superior recognition of Caucasian faces ($M = 2.19$), relative to Asian faces ($M = 1.51$, $t(49) = 6.52$, $\eta_p^2 = 0.645$, $p < .001$) and Asian subjects

[2] Though we refer to these predictions as JOLs, we note that the instructions to the participants were phrased in terms of their confidence in their ability to later recognize a face, rather than as an assessment of how well they had learned a face.

[3] Because hit and false alarm rates of 1 or 0 cannot be used to compute d', one half of an item hit was subtracted from any perfect hit rates and one half of an item false alarm was added to any perfect false alarm rates to allow computation of d' (e.g., Snodgrass & Corwin, 1988).

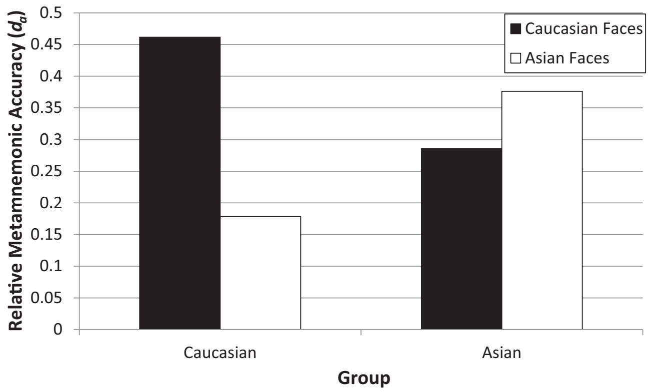

Fig. 1. Relative metamnemonic accuracy (d_a) for Caucasian and Asian subjects.

showed superior recognition of Asian faces ($M = 1.98$), relative to Caucasian faces ($M = 1.56$, $t(51) = 4.56$, $\eta_p^2 = 0.290$, $p < .001$).

2.2. Metamemory

The accuracy of predictions was assessed using the detection-theoretic measure d_a, which treats the prediction ratings as criteria that partition the range of evidence (see Benjamin & Diaz, 2008; Green & Swets, 1966, Masson & Rotello, 2009). The values are shown in Fig. 1. The d_a scores were analyzed in a 2 (face: Caucasian vs. Asian) × 2 (group: Caucasian vs. Asian) mixed ANOVA, with face as a repeated-measures factor and group as a between-subjects factor. Neither the main effect of face ($F(1,100) = 1.84$, $\eta_p^2 = 0.018$, $p > .10$), nor the main effect of group ($F(1,100) = 0.02$, $\eta_p^2 = 0.001$, $p > .10$) were significant. Critically, the face × group interaction was significant, ($F(1,100) = 6.89$, $\eta_p^2 = 0.064$, $p < .05$). As can be seen in Fig. 1, a CRE in relative metamnemonic accuracy for faces was observed. Planned comparisons revealed that the Caucasian subjects had superior relative metamnemonic accuracy for Caucasian faces ($M = 0.46$, $SD = 0.37$), relative to Asian faces ($M = 0.18$, $SD = 0.57$, $t(49) = 3.13$, $\eta_p^2 = 0.166$, $p < .01$). The Asian subjects showed a smaller CRE in relative metamnemonic accuracy, with the accuracy for Asian faces ($M = 0.38$, $SD = 0.61$) being numerically larger, but not statistically different from the accuracy for Caucasian faces ($M = 0.29$, $SD = 0.46$, $t(51) = 0.83$, $\eta_p^2 = 0.013$, $p = .41$). A substantial majority of Asian subjects did, however, exhibit numerically higher metamemory accuracy for Asian faces than for Caucasian faces; a non-parametric (sign) test comparing the metamnemonic d_a scores in the Asian subjects produced a result much closer to conventional levels of statistical significance ($Z = -.180$, $p = .071$).

3. Discussion

The aim of the current study was to examine the metacognitive accuracy of predictions of recognition for own- and other-race faces. Caucasian and Asian subjects viewed Caucasian and Asian faces, and made JOLs concerning future recognition of those faces. The results indicate that metamemory is more accurate for classes of faces that are most like our own: Relative metamemory accuracy was higher for own-race faces than other-race faces.

As expected, we also observed a CRE in recognition accuracy. In both groups of subjects, own-race faces were more accurately recognized than other-race faces. This well documented effect has been thought to reveal more effective encoding of own-race faces (e.g., Meissner et al., 2005; Sporer, 2001; Valentine, 1991), and the parallel effects on metacognitive judgments support this claim: More effective encoding leads to more accurate recognition predictions. One caveat regarding our findings is that the bias in metamemory was smaller in the Chinese subjects than in the American subjects, not reaching conventional levels of statistical significance, though numerically in the correct direction.

Why might the two populations of subjects have differences in their relative metamnemonic accuracy for own- and other-race faces? It may be that individuals in the Chinese group of subjects simply vary more in their relative metamemory skill for own- vs. other-race faces, compared to the American subjects. This may be caused by differential experience or attitudes towards other-race individuals as a result of exposure in popular media. For example, many current American TV shows are aired in China, but few Chinese TV shows are aired in the United States. Thus, some individuals in our group of Chinese subjects may actually have better metamemory for Caucasian faces than other individuals in that group, due to years of American TV viewing, while the same is not true of individuals in our group of American subjects. The relative size of the own-race bias effect in recognition accuracy speaks to this point as well: The own-race bias effect in memory observed in the Chinese subjects ($\eta_p^2 = 0.290$) was smaller than the own-race bias effect observed in the American subjects ($\eta_p^2 = 0.645$). Regardless of the reason, the own-race bias appears to be present in both memory and metamemory.

This new finding adds to the important implications that the CRE has for eyewitness memory. Our subjects predicted how likely (or unlikely) they would be to successfully recognize each studied face, and these judgments predicted actual recognition more accurately for own-race faces than for other-race faces. This implies that when people witness a crime, their prediction of their future ability to pick the suspect out of a lineup is likely to be less accurate for other-race suspects than for own-race suspects. In the context of a line-up identification procedure, the relation between pre-identification confidence and subsequent recognition accuracy has already been shown to be weak (e.g., Cutler & Penrod, 1989). Combined with the fact that recognition accuracy is also worse for other-race faces, our results further add to the evidence that eyewitnesses are less accurate when observing other-race suspects than own-race suspects.

3.1. Practical applications

Our results showed that recognition predictions for other-race faces are less accurate than predictions for own-race faces. These results fit well with the existing literature showing a weak relationship between eyewitness recognition and post-decision confidence (e.g., Leippe & Eisenstadt, 2007), and an even weaker relationship between line-up identification accuracy and pre-identification confidence (e.g., Cutler & Penrod, 1989). The practical ramifications of the CRE in metacognitive judgments can be devastating. As described above, Ronald Cotton spent over a decade in prison after being falsely convicted of assault on the basis of a cross-race eyewitness: the victim of a crime committed by a different individual (Cotton's Wrongful Conviction, n.d.). Based on Ms. Thompson's high confidence in her ability to recognize her attacker, Detective Mike Gauldin, the officer in charge of the case, never doubted the victim's incorrect identification of Mr. Cotton. Despite a second victim's failure to select Mr. Cotton from the same line-up shown to Ms. Thompson (a fact that was not admissible in the subsequent court case), the police operated on the basis that a victim's self-reported confidence was enough to assure her recognition accuracy.

Despite what researchers have known for decades about the poor relationship between face recognition accuracy and post-recognition confidence (e.g., Leippe & Eisenstadt, 2007), particularly in the case with cross-race faces (e.g., Brigham et al., 2007), police officers still highly value eyewitness confidence. In a survey of over 500 U.S. law enforcement officers, Wise, Safer, and Maro

(2011) found that only 30% of officers in departments that have instituted eyewitness reform procedures correctly state that eyewitness confidence is not a good predictor of accuracy. (Only 19% of officers in departments without eyewitness reform procedures correctly agreed with this statement.) Thus, even after receiving some instruction on what experts currently know about eyewitness factors, most police officers likely still consider eyewitness confidence to be a reliable indicator of accuracy.

Our results speak to an understudied aspect of the police–witness interaction: how police proceed with their investigation prior to witness exposure to a line-up. It seems likely that police would put more effort into an investigation and locating suspects to include in a line-up when they have a witness who reports high confidence in his or her ability to recognize the suspect, as in the case of Ms. Thompson described above (cf. Cutler & Penrod, 1989). Had Ms. Thompson not been so certain of her recognition ability, perhaps the police might have spent longer investigating possible suspects prior to constructing a line-up to show the victims. In the case of suspects who are of a different race from the eyewitness, confidence at the time of encoding is even less likely to relate to subsequent recognition accuracy than for same-race faces, and police should practice more caution in their investigative search to locate potential suspects.

4. Conclusions

The present findings are consistent with several views on the origin of the CRE. If features of greater differentiating value are sampled from own- than other-race faces, the variability in those features will affect later recognition and current metamnemonic judgments similarly. This could occur either by examining features that are only helpful for differentiating own-race individuals (i.e., encoding all faces in the same manner, but less successfully for other-race faces; e.g., Valentine, 1991; Valentine & Endo, 1992), or by focusing on category-defining features at a cost to individuating features (e.g., Sporer, 2001). Alternatively, subjects simply may spend less effort encoding other-race faces (e.g., Goldinger, He, & Papesh, 2009); this reduction in effort then plays out in poorer memory (and likely less recollection; Meissner et al., 2005) and metamemory. While the precise mechanisms are still unclear, it is clear that both memory and metamemory are superior for own-race faces, underscoring the importance for caution in relying on eyewitness confidence in motivating police investigations, particularly in cross-race circumstances.

References

Benjamin, A. S. (2003). Predicting and postdicting the effects of word frequency on memory. *Memory & Cognition*, *31*, 297–305.

Benjamin, A. S. (2005). Response speeding mediates the contribution of cue familiarity and target retrievability to metamnemonic judgments. *Psychonomic Bulletin & Review*, *12*, 874–879.

Benjamin, A. S., & Diaz, M. (2008). Measurement of relative metamnemonic accuracy. In J. Dunlosky, & R. A. Bjork (Eds.), *Handbook of memory and metamemory* (pp. 73–94). New York, NY: Psychology Press.

Benton, T. R., Ross, D. F., Bradshaw, E., Thomas, W. N., & Bradshaw, G. S. (2006). Eyewitness memory is still not common sense: Comparing jurors, judges, and law enforcement to eyewitness experts. *Applied Cognitive Psychology*, *20*, 115–129.

Brigham, J. C., Bennett, L. B., Meissner, C. A., & Mitchell, T. L. (2007). The influence of race on eyewitness memory. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Memory for people The handbook of eyewitness psychology* (pp. 257–281). Mahwah, NJ: Lawrence Erlbaum Associates.

Cotton's Wrongful Conviction (n.d.). Available from http://www.pbs.org/wgbh/pages/frontline/shows/dna/cotton/.

Cutler, B. L., & Penrod, S. D. (1989). Forensically relevant moderators of the relation between eyewitness identification accuracy and confidence. *Journal of Applied Psychology*, *74*, 650–652.

Daniels, K. A., Toth, J. P., & Hertzog, C. (2009). Aging and recollection in the accuracy of judgments of learning. *Psychology and Aging*, *24*, 494–500. http://dx.doi.org/10.1037/a0015269

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

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

He, Y., Ebner, N. C., & Johnson, M. K. (2011). What predicts the own-age bias in face recognition memory? *Social Cognition*, *29*, 97–109.

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. http://dx.doi.org/10.1037/a0020463

Kennedy, K. M., Hope, K., & Raz, N. (2009). Lifespan adult faces: Norms for age, familiarity, memorability, mood, and picture quality. *Experimental Aging Research*, *35*, 268–275. http://dx.doi.org/10.1080/03610730902720638

Leippe, M. R., & Eisenstadt, D. (2007). Eyewitness confidence and the confidence–accuracy relationship in memory for people. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Memory for people The handbook of eyewitness psychology* (pp. 377–425). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Lindsay, S. D., Read, J. D., & Sharma, K. (1998). Accuracy and confidence in person identification: The relationship is strong when witnessing conditions vary widely. *Psychological Science*, *9*, 215–218.

Marcon, J. L., Susa, K. J., & Meissner, C. A. (2009). Assessing the influence of recollection and familiarity in memory for own- versus other-race faces. *Psychonomic Bulletin & Review*, *16*, 99–103. http://dx.doi.org/10.3758/PBR.16.1.99

Masson, M. E. J., & Rotello, C. M. (2009). Sources of bias in the Goodman–Kruskal gamma coefficient measure of association: Implications for studies of metacognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 509–527. http://dx.doi.org/10.1037/a0014876

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in face memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, *7*, 1–35. http://dx.doi.org/10.1037//1076-8971.7.1.3

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

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

Neil v. Biggers (1972). 409 U.S. 188.

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.

Smith, S. M., Stinson, V., & Prossor, M. A. (2004). Do they all look alike? An exploration of decision-making strategies in cross-race facial identification. *Canadian Journal of Behavioural Science*, *36*, 146–154.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*, 34–50.

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law*, *7*, 36–97. http://dx.doi.org/10.1037//1076-8971.7.1.36

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

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

Wise, R. A., Safer, M. A., & Maro, C. M. (2011). What U.S. law enforcement officers know and believe about eyewitness factors, eyewitness interviews, and identification procedures. *Applied Cognitive Psychology*, *25*, 488–500. http://dx.doi.org/10.1002/acp.1717