Are individuals' familiarity judgments diagnostic of prior contact?

Kathy Pezdek* and Stacia Stolzenberg

Department of Psychology, Claremont Graduate University, Claremont, CA, USA

(Received 6 July 2012; final version received 25 December 2012)

The extensive eyewitness memory research literature has been restricted to memory for strangers. Although it is often assumed that eyewitnesses are more accurate identifying familiar than unfamiliar individuals, little is known about whether individuals' familiarity judgments are diagnostic of prior contact. Caucasian and Asian sophomores (N = 139) in two small private high schools viewed yearbook pictures of (a) graduated students from their school who were seniors (fourth year) when participants were freshmen (first year) (familiar) and (b) unfamiliar individuals, and responded whether each was 'familiar'. The design was completely crossed; familiar faces at each school served as unfamiliar faces at the other school. Based on a' data, the cross-race effect resulted for familiarity judgments. Also, although individuals' familiarity judgments were diagnostic of prior contact, accuracy was low (mean hit rate = 0.42; mean false alarm rate = 0.23), rendering an eyewitness's report of having seen a perpetrator casually in the past of limited forensic value.

Keywords: eyewitness memory; cross-race effect; face recognition memory

Introduction

Eyewitness evidence is critical for solving crimes, and in the absence of incriminating physical evidence, it is often the sole source of evidence for determining the perpetrator's identity. However, studies consistently report that eyewitness misidentifications are the leading cause of erroneous convictions (Huff, 1987; Huff, Rattner, & Sagarin, 1996; Penrod & Cutler, 1999). The Innocence Project (2012) reported that 75% of the first 292 DNA exonerated cases were the result of misidentifications. There is a wealth of scientific research on the psychological factors that affect the accuracy of eyewitness memory, and several reviews of this research are available (Pezdek, 2007, 2009; Wells, Memon, & Penrod, 2006; Wells & Olson, 2003).

One factor likely to affect the accuracy of eyewitness memory and identification is whether the eyewitness has previously seen the perpetrator. It would seem a matter of common sense that eyewitnesses would be more likely to identify correctly someone they know well than someone who they do not know at all. For example, if a store clerk is robbed and tells the police that he can identify the shooter because he is a regular customer, the clerk is probably correct. In fact, at least one court has held the exclusion of an eyewitness expert harmless when the witness claimed to have been familiar with
the defendant and seen him on a daily basis for ‘well over a year’, *Hagar v. United States*, 856 A.2d 1143 (D.C. Cir. 2004). And, in *People v. Rodriguez*, 79 N.Y.2d 445, 451, 593 N.E. 2d 268, 583 N.Y.S.2d 814 (1992), the New York Court of Appeals included ‘the number of times [the witness] viewed defendant prior to the crime’, in a list of factors one might consider in evaluating the reliability of an eyewitness identification. Despite these legal rulings, in fact, we do not know how accurately individuals can look at a person and determine whether they have casually encountered that person in the past.

Consider the following situation relevant to our study. An eyewitness observes a drive-by-shooting at dusk from across the street and sees the driver of the suspect’s vehicle only briefly. However, the eyewitness tells the police that he thinks he can pick the suspect out of a lineup because he had seen him in a park last summer. Later, reading the eyewitness’s report that he had seen the suspect previously, defense counsel decides that the evidence against his client is relatively strong and encourages his client to take a plea rather than take the case to a jury. The question of interest in this study is whether eyewitnesses’ familiarity judgments are actually diagnostic of prior contact; is the above decision by defense counsel well informed?

This study specifically examines familiarity judgments for which the context is nebulous. If an eyewitness reports that he had seen the suspect previously, that ‘he is a regular customer’, or ‘we played on the same soccer team’, the police can actually use this contextual information to move the investigation forward and help determine the suspect’s identity. However, eyewitnesses often simply report that, ‘I’ve seen him around the neighborhood’, or ‘I saw him at the park last summer’. In these cases – the situations of interest in this study – although the context of the familiarity judgment is nebulous, the familiarity judgment may still be considered diagnostic of prior contact and, as indicated in the example above, influential in defense counsel’s decision regarding the strength of the evidence and therefore the wisdom of plea bargaining the case.

In a typical eyewitness memory study – of which there are hundreds (see Lindsay, Ross, Read, & Toglia, 2007; Pezdek, 2011) – there is a presentation phase in which participants view never-before-seen faces, followed by a test phase to assess memory for presented faces. A different methodology is necessary to assess how well individuals can determine whether a person is familiar from a prior time in their life, that is, non-stranger identification. To assess the accuracy of this type of familiarity judgment there is no presentation phase, only a test phase. A test is constructed that includes faces of (a) individuals who were seen at some point in the participant’s past (i.e., our operational definition of non-strangers) and (b) individuals never seen before (i.e., our operational definition of strangers), and as each is presented, participants decide whether they have ever seen this person before. It is important to note that in this study, familiarity was an independent variable and was defined as a dichotomous variable.

The cognitive processes (Burton, Bruce, & Hancock, 1999; Klitzky & Forrest, 1984) and neural systems (Gobbini & Haxby, 2007) involved in recognizing and scanning (Althoff & Cohen, 1999) familiar as compared with unfamiliar faces are fundamentally different. In reviewing the literature on this topic, Hancock, Bruce, and Burton (2000) concluded:

We do generally identify familiar faces with little effort, despite possibly large variations of lighting, viewpoint and expressions, and ‘disguises’ such as beards, spectacles and hats. (p. 330)
However, these studies that have addressed recognition memory for familiar faces have primarily assessed recognition of people who are famous or very well known, usually known by name. On the familiarity continuum, there is another category of faces that has not been examined, that is, memory for casually familiar faces – faces of people seen before but not necessarily well known.

Few previous studies have examined how well individuals can recognize casually familiar individuals. In a classic study, Bahrick, Bahrick, and Wittlinger (1975) presented participants with photographic lineups composed of high school yearbook pictures between 3 months to 47 years after high school graduation. Face recognition accuracy declined from 90% correct at 3 months to 71% correct after 47 years. Although this level of face recognition memory after 47 years is impressive, participants in this study likely encountered target individuals during the years subsequent to graduation, and this would have served to reestablish memory for each.

Vokey and Read (1988), on the other hand, assessed the accuracy of face familiarity judgments after a minimal time delay. They had participants view 209 photographs of yearbook pictures from an unknown yearbook and rate whether each had appeared previously in this test sequence; none of the individuals presented had ever been seen by the participants and none were repeated in the sequence. The overall false-positive rate was 16%. Thus, even with a minimal delay interval, individuals were likely to recognize as ‘old’, 16% of the new faces. Together, these results suggest that it is challenging for people to view faces and determine which ones they have seen before. Specifically, when people report that they have seen a person before, the false alarm rate \( p(\text{familiar}/\text{unfamiliar}) \) and hit rate \( p(\text{familiar}/\text{familiar}) \) are unimpressive relative to what might be considered diagnostically useful in a forensic setting. Thus, if an eyewitness reports to the police that he thinks he can pick a suspect out of a lineup because, for example, he had seen him in a park last summer, his ability to assess whether he has in fact seen the suspect previously may not be reliable. Clearly, additional research is necessary to address this question, especially within the context of people’s memory for individuals casually encountered in the course of their normal life and for which their prior contact is nebulous or not verifiable. Unlike the majority of the research on familiarity, this study focuses on casual familiarity rather than memory for well known or famous people. It is important to note that the ability to classify a previously seen face as familiar is not the same as the ability to recognize a person from a crime scene because he was seen previously. This study addresses the first of these issues and not the second.

This study examines whether individuals’ familiarity judgments are diagnostic of prior contact, and further, if the accuracy of familiarity judgments is affected by whether the individual and the face being viewed are of the same or different race. Generally, individuals display higher recognition accuracy for faces of their own race than faces of another race, a phenomenon known as the cross-race effect (CRE) or own-race bias. Over the past three decades, the CRE has been verified as a robust construct \( d = 0.30 \) (Meissner & Brigham, 2001) and has been observed across a wide range of ages (Pezdek, Blandon-Gitlin, & Moore, 2003) and cultural and racial groups (Platz & Hosch, 1988; Wright, Boyd, & Tredoux, 2001) including the racial groups used in this study, Asians and Caucasians (Goldinger, He, & Papesh, 2009; Ng & Lindsay, 1994). However, no previous studies have assessed the extent to which the CRE affects judgments of familiarity.
Method

Selection of research site

For this study, it was necessary to test a sufficiently large group of people (participants) who we could be sure had casual exposure to another group of people (target individuals) in their past. We sought an ecologically valid situation in which participants had casually encountered the target individuals in the course of their normal life. With this in mind, we selected students from two small private high schools in the Los Angeles Metropolitan area. At each school, there are approximately 750 students total in four grades; both of these are Catholic all-girls schools. Both schools are on small campuses where all students come in contact with each other on a daily basis; they eat lunch in the same courtyard, take classes in the same two buildings, participate together in after-school sports and clubs, attend assemblies and rallies together, etc. Thus, in each school, the likelihood of multiple sightings of every student on a regular basis is high even for individuals not acquainted or in the same grade. It is very unlikely that students at each school had any contact with students at the other school; the two schools were 31 miles apart, in different cities within the densely populated metropolitan Los Angeles area, and did not share any athletic or after-school programs. Also, students at both schools commuted to campus from a relatively large geographic area that included many different cities. Thus, it is unlikely that the participants encountered the already graduated students after they had graduated and gone away to college as most did.

Participants from school A (N = 75; White = 27, Asian = 48; M age = 15.32, SD = 0.43) and school B (N = 64; White = 36, Asian = 28; M age = 15.37, SD = 0.57) were sophomores (second-year students) each tested on their ability to recognize as familiar, yearbook pictures of students who were seniors (fourth-year students) when the participants were freshman (first-year students). In other words, the target individuals had graduated at the end of the year prior to when this study was conducted. Students were excluded if they (a) had not attended this high school since they were freshman or (b) had a sibling in the target class. Thus, the target individuals were schoolmates whom participants would have seen on a regular basis for an entire school year, only 1 year prior. At both schools, the percentage of non-Hispanic Caucasians was about 22% and the percentage of Asian/Pacific Islanders was about 14%. These were the groups from which participants were recruited for this study.

Participants and design

A total of 139 high school sophomores participated. The design was a 2 (race of participant: White or Asian) × 2 (same- vs. cross-race target face) × 2 (familiar or unfamiliar target face) mixed factorial design, with race of participant as the only between-subjects factor. The design was completely crossed so that familiar faces at each school served as unfamiliar faces at the other school. Thus, each face equally often served as a familiar and an unfamiliar face; consequently, the faces themselves were not confounded with familiarity conditions.

Procedure and materials

Individuals participated together in regularly scheduled classes. At each school, six or seven different classes participated with approximately 12 participants in each
class. Participants were presented 40-yearbook pictures – half pictures of now graduated students who were seniors (fourth year) at their school when participants were freshman (first year) (familiar) and half pictures of individuals from the other high school (unfamiliar). In selecting the yearbook pictures from each school, we specifically excluded a small set of student council officers, sports stars, and public individuals; excluded were five or six students at each school.

The test sequence included a randomly arranged sequence of 20-yearbook pictures of White students (10 familiar and 10 unfamiliar), and 20-yearbook pictures of Asian students (10 familiar and 10 unfamiliar). The test sequence was constructed as a PowerPoint presentation, presented by an LCD projector, larger than life-size, to a screen.

Participants were instructed to look at each face and decide whether they had seen this person before. Each face was presented for 10 s and was immediately preceded by a warning tone. At the offset of the face, participants had 10 s to respond (a) whether the individual presented was familiar or not and (b) their confidence in each of the previous responses on a scale from 1 (low) to 5 (high). The two experimenters monitored each classroom to prevent students from looking at the response sheet of others.

Results

The accuracy of familiarity judgments

Familiarity judgments were coded in terms of the Signal Detection measure of $d'$. As suggested by Macmillan and Creelman (2005), hit rates and false alarm rates equal to 0 were converted using $1/(2N)$, and proportions equal to 1 were converted using $1-1/(2N)$. The Signal Detection measure of criterion, $c$, was examined as well to assess how cautious participants were in calling a face familiar. The use of $c$ as a measure of response bias followed suggestions by MacMillan and Creelman (1990). However, because of the forensic relevance of examining the hit rate data (i.e., the probability of correctly recognizing the real perpetrator) separately from the false alarm rate data (i.e., the probability of misidentifying an individual who was not the real perpetrator), the results were analyzed in terms of these additional measures as well.

Overall, levels of performance were examined first with descriptive statistics. The mean memory level for hit rate, false alarm rate, $d'$, and $c$ rate data are presented in Table 1. How commonly did participants misidentify as familiar, individuals whom they had never seen before? The overall false alarm rate was $0.23$ (SD = $0.16$); 23% of the unfamiliar faces were misidentified as familiar individuals. How accurately can individuals recognize schoolmates who they had probably seen daily for 1 year, only 1 year ago? The overall hit rate was $0.42$ (SD = $0.20$). In all conditions, the miss rate was equal to or higher than the hit rate. And, the hit rate data reported were actually

<table>
<thead>
<tr>
<th>$p$ (Hit)</th>
<th>$p$ (False alarm)</th>
<th>$d'$</th>
<th>Criterion ($c$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.42 (0.20)</td>
<td>0.23 (0.16)</td>
<td>0.78 (0.82)</td>
<td>0.63 (0.66)</td>
</tr>
</tbody>
</table>
inflated somewhat by the fact that a few of the participants indicated that they were friends with one or two target individuals.

We next assessed whether individuals’ familiarity judgments were diagnostic of prior contact. Here, we focus on the response of participants (familiar or unfamiliar) because in the real-world forensic context, we know what the eyewitness’s response is; what we want to know is whether the eyewitness is accurate given that response. In this analysis, we considered only faces that were responded to as ‘familiar’ and compared the conditional probability that a face was familiar given that it was responded to as ‘familiar’ ($p = 0.66$) to the conditional probability that a face was unfamiliar given that it was responded to as ‘familiar’ ($p = 0.34$). This difference was statistically significant, $t(137) = 10.54$, $p < 0.001$, $d = 1.80$. These findings suggest that individuals’ familiarity judgments were diagnostic of prior contact. However, given the high diagnostic error rate – 34% of faces classified as ‘familiar’ had not been seen before – familiarity judgments are of limited forensic utility.

Separate 2 (race of participant) $\times$ 2 (same- vs. cross-race target face) ANOVAs were performed on the $d^\prime$, $c$, hit rate, and false alarm rate data. These results are presented in Figure 1. Regarding the analysis of the $d^\prime$ data, as predicted, same-race faces ($d^\prime = 0.97$, $SD = 1.18$) were more accurately recognized than cross-race faces ($d^\prime = 0.56$, $SD = 1.06$), $F(1, 137) = 7.79$, $p < 0.01$, $MS_e = 0.09$, $\eta^2_p = 0.05$. In addition, there was a significant interaction of race of participant $\times$ same- vs. cross-race target face, $F(1, 137) = 5.55$, $p < 0.05$, $MS_e = 0.10$, $\eta^2_p = 0.04$. Whereas Asian participants were equally accurate recognizing same-race ($d^\prime = 0.86$, $SD = 1.01$) and cross-race faces ($d^\prime = 0.80$, $SD = 1.36$), $t(59) = 0.27$, $p = 0.78$, two-tailed, $d = 0.05$, White participants were more accurate recognizing same-race ($d^\prime = 1.05$, $SD = 1.29$) than cross-race faces ($d^\prime = 0.40$, $SD = 0.77$), $t(87) = 4.09$, $p < 0.001$, $d = 0.61$. Thus, overall, there was a CRE on $d^\prime$ data for White, but not Asian participants. The main effect of race of participant was not significant.

There were similar findings in the analyses of the hit rate data. As predicted, the mean hit rate was higher for same-race faces ($M = 0.46$, $SD = 0.22$) than for cross-race faces ($M = 0.37$, $SD = 0.20$), $F(1, 137) = 19.14$, $p < 0.001$, $MS_e = 0.02$, $\eta^2_p = 0.11$. In addition, there was a significant interaction of race of participant $\times$ same- vs. cross-race face, $F(1, 137) = 10.31$, $p < 0.01$, $MS_e = 0.04$, $\eta^2_p = 0.07$.

---

**Figure 1.** Mean hit rate, false alarm rate, $d^\prime$ rate, and criterion ($c$) rate data for White and Asian participants in the same-race and cross-race conditions with standard error bars indicated.
cross-race target face, \( F(1, 137) = 5.55, p < 0.05, \) MS\(_{c} = 0.02, \) \( \eta^2_p = 0.04. \) Whereas Asian participants were equally accurate recognizing same-race (\( M = 0.42, \) SD = 0.22) and cross-race faces (\( M = 0.38, \) SD = 0.29), \( t(60) = 1.13, p = 0.26, \) two-tailed, \( d = 0.15, \) White participants were more accurate recognizing same-race (\( M = 0.50, \) SD = 0.23) than cross-race faces (\( M = 0.37, \) SD = 0.21), \( t(87) = 5.94, p < 0.001, \) \( d = 0.59. \) Thus, consistent with results reported elsewhere (Meissner & Brigham, 2001), overall, there was a CRE on \( d \) data only for White participants. The main effect of race of participant was not significant.

On the other hand, there were no significant effects in the analysis of the false alarm rate data.

In the analysis of the criterion data, only the main effect of same- vs. cross-race target face was significant, \( F(1, 137) = 4.72, p = 0.03, \) MS\(_{c} = 0.07, \) \( \eta^2_p = 0.03. \) The mean response criterion was significantly lower for same-race faces (\( c = 0.56, \) SD = 0.73) than for cross-race faces (\( c = 0.68, \) SD = 0.81). This finding suggests that participants were less cautious calling same-race faces familiar than cross-race faces.

**Confidence results**

If eyewitnesses indicated that they thought a face was familiar, was their confidence in this response related to whether the face was, in fact, familiar? The next analysis assessed this question by comparing the mean confidence ratings for familiar and unfamiliar faces to which a response of ‘familiar’ had been made. For faces that were responded to as ‘familiar’, the mean confidence rating (scale = 1–5) for faces that were actually familiar (\( M = 3.41, \) SD = 0.72) was significantly higher than that for faces that were unfamiliar (\( M = 2.90, \) SD = 0.95), \( t(126) = 6.13, p < 0.001, d = 0.61. \)^3 This suggests that although individuals are not very accurate determining whether a face is familiar or not, when a face is responded to as ‘familiar’, individuals’ subjective experience of the familiarity of the face is higher for faces that actually are familiar.

The accuracy–confidence correlation was also examined. In a meta-analysis of 30 studies on this topic, Sporer, Penrod, Read, and Cutler (1995) found that the accuracy–confidence relationship was \( r = 0.29, \) accounting for only 8% of the variance in eyewitness accuracy. Consistent with this finding, the overall accuracy–confidence relationship in this study was low and not significant (\( r = +0.16, p = 0.06. \)). In addition, the accuracy–confidence correlation was assessed separately for (a) familiar and unfamiliar faces and (b) same- and cross-race faces. The accuracy–confidence correlation was significant for unfamiliar (\( r = +0.25, p < 0.002 \)) but not familiar faces (\( r = +0.12, p = 0.14 \)) and for same- (\( r = +0.33, p < 0.003 \)) but not cross-race faces (\( r = +0.09, p = 0.46 \)). Mean confidence ratings were significantly higher for familiar faces (\( M = 3.54, \) SD = 0.59) than unfamiliar faces (\( M = 3.31, \) SD = 0.57, \( t(126) = 5.31, p < 0.001, d = 0.40, \) and significantly higher for same-race (\( M = 3.60, \) SD = 0.67) than cross-race faces (\( M = 3.09, \) SD = 0.55), \( t(126) = 36.65, p < 0.001, d = 0.83. \) Together, these later findings suggest that if an eyewitness observes a crime and says that he can identify the perpetrator because he has casually seen him before, the eyewitness’s statement should not necessarily be considered reliable. However, the eyewitness’s recognition confidence is likely to be more indicative of accuracy for (a) unfamiliar than familiar faces and (b) same- than cross-race faces.
Discussion

This study addresses two issues. First, are individuals’ familiarity judgments diagnostic of prior contact? Specifically, if an eyewitness says that she had seen a suspect previously, is it likely that she has actually seen that suspect previously? Yes. In this study, if a face was responded to as ‘familiar’, it was significantly more likely to have actually been familiar ($p = 0.66$) than unfamiliar ($p = 0.34$). Also, if a face was responded to as ‘familiar’, it received a higher confidence rating if it was familiar than if it was not.

Why were the rates of correctly classifying faces as familiar so low in our study? Perhaps they were not so low. Surprisingly, the overall hit rate, false alarm rate, and $d'$ data presented in Table 1, when compared to the meta-analytic results of research on identification of strangers’ faces reported by Shapiro and Penrod (1986) and Steblay, Dysart, Fulero, & Lindsay (2003), suggest that the accuracy rate at which faces of non-strangers are correctly classified as familiar is comparable to the level of recognition accuracy for strangers. These findings suggest that looking at the face of a perpetrator and recognizing that the face has been seen before (the task parallel to the familiarity judgments in this study) are likely to involve cognitive processes that are similar to those involved in looking at a face in a lineup and recognizing that the face has been seen before (the typical eyewitness identification task that involves memory for strangers). This process is described by numerous general models of recognition memory including the encoding specificity model first reported by Tulving and Thomson (1973). According to this model, the initially viewed face is encoded into memory, the second face is encoded into memory, and the probability that the second face will match the first is determined by the extent to which the encoding of the second face arouses the same perceptual or affective information as was encoded with the first face. What is important here is the prediction that follows from this. If the processes involved in recognizing that a non-stranger is familiar are similar to those involved in recognizing that a stranger has been seen before, then the estimator variables (cf. Wells, 1978) known to produce high rates of misidentifications in eyewitness recognition memory generally (Pezdek, 2011) would also account for inaccuracies making the type of familiarity judgments required in this study, with both prone to error under the same predictable conditions.

The rates of correctly classifying faces as familiar in this study were not likely depressed because the participants had not previously seen the target students. The participant and target students at each school shared the same restricted physical space for significant periods of time most days. Both schools are on small campuses. Students in all four grades eat lunch in the same courtyard, take classes in the same two buildings, participate together in after-school sports and clubs, attend assemblies and rallies together, etc. Thus, in each school, the likelihood of multiple sightings of every student on a daily basis is high even for individuals not acquainted or in the same grade. A second interpretation of the low rate of correctly classifying faces as familiar relates to the fact that the test pictures in this study were yearbook pictures. It might be argued that individuals in yearbook pictures look dissimilar to their appearance in everyday life, thus making this task more difficult. However, the same argument can be made for mug-shot photographs as representations of real perpetrators.
Burton et al. (1999) proposed a cognitive processing model to account for familiarity judgments of faces. However, this model, which uses a connectionist architecture, was developed to account for recognition of well known or famous individuals, for whom the corresponding mental representation would include not only face information but also name, social information, etc. Many of the processes in this model are not relevant to processing individuals about whom there is limited non-visual social information.

Steblay, Dietrich, Ryan, Raczyński, and James (2011) tested identification accuracy from lineups for an individual who was a recent graduate of the college where the study was conducted. After the lineup, participants were asked whether they were familiar with any of the actors prior to the study. Those who reported having previously seen the perpetrator were more likely to identify him than were those for whom the perpetrator was reported to be a stranger. Nonetheless, even in the best condition, identification accuracy for familiar individuals was only 62.9%, well below what might be required to eliminate ‘reasonable doubt’ in the courts. However, Steblay et al. (2011) did not assess whether individual’s judgments of familiarity were actually diagnostic of prior contact, which is the focus of our study.

The second issue addressed in this study is whether the well-established CRE previously reported with recognition memory for unfamiliar faces also applies to the accuracy of familiarity judgments for non-strangers. The answer is yes based on the $d’$ results. However, contrary to the findings in typical cross-race studies, in this study, significant effects with hit rate but not false alarm rate data primarily accounted for this effect. Several cognitive models have been proposed to account for the CRE, including the multidimensional space (MDS) framework proposed by Valentine (1991) and Valentine and Endo (1992). According to this framework, faces are represented in a hypothetical space that retains each face on the basis of various dimensions including both configural and featural aspects of faces developed from an individual’s prior experience. It is proposed that the representational space for own-race faces includes a better differentiation of the variant and invariant dimensions of faces than does the representational space for cross-race faces. Consequently, own-race faces are spread more evenly throughout the multidimensional ‘face space’ and are more distinctively represented and more accurately recognized. On the other hand, less distinction is made between the variant and the invariant dimensions of cross-race faces, and as a consequence cross-race faces are more tightly clustered in the face space and thus are less distinctively represented and less accurately recognized.

According to this MDS framework, it would also be predicted that familiar faces would be spread more evenly throughout the multidimensional ‘face space’ than unfamiliar faces, and consequently, familiar faces would be more distinctively represented. It is not then clear whether face familiarity would trump the same-race face dimension and thus eliminate the CRE in making familiarity judgments for familiar faces. The results of this study suggest that this does not occur. Even with casually familiar faces, sensitivity distinguishing between familiar and unfamiliar faces was greater for same- than cross-race faces, suggesting again that estimator variables known to affect the accuracy of eyewitness recognition memory generally would also be predicted to affect familiarity judgments.

Although there was a significant CRE in this study for judgments of familiarity with both $d’$ and hit rate data, this effect interacted with the race of participants on
both measures; there was a CRE for White but not Asian participants. This result cannot be an artifact of the specific Asian target faces used; the design was completely crossed so that the familiar faces at each school served as the unfamiliar faces at the other school. Thus, because each face equally often served as a familiar and an unfamiliar face, the faces themselves were not confounded with familiarity conditions. This finding, that the CRE is generally weaker for minority than majority group members, is consistent with results from the meta-analysis of Meissner and Brigham (2001). However, some other researchers have reported a CRE with both White and Asian participants, although the effect was significant on $d'$ and false alarm rate data but not hit rate data, and the effect was not related to intercultural contact (Ng & Lindsay, 1994).

There are several caveats to consider. First, in this study, judgments of familiarity were made from a 10 s look at a photograph of each target individual. In the real world, an eyewitness may have a better opportunity to observe a suspect before determining that they are familiar. Second, it is difficult to estimate the guessing rate in the task used in this study and thus what might be the probability of being correct with a ‘familiar’ judgment in the real world. Consequently, the quantitative findings of this study, as reported in Table 1, might not precisely apply to real eyewitness memory. Third, casual familiarity is clearly not one thing, and the level of familiarity among schoolmates would not be the same as the level of familiarity among all other overlapping cohorts. On this same point, although the level of familiarity was presumed to be quite high for students who spent a year together in the same small school, the 3-year test delay in this study may have influenced the low recognition accuracy rates. We know from meta-analytic results reported by Deffenbacher, Bornstein, McGorty, and Penrod (2008) that there is a reliable association ($r = 0.18$, $d = 0.37$) between longer retention intervals and positive forgetting of faces. Fourth, although the two schools used in this study were 31 miles apart in different cities within the densely populated Los Angeles Metropolitan area, we had no way of guaranteeing that in fact all unfamiliar faces had never been seen before. Finally, in this study, familiarity was a dichotomous independent variable. A topic of future research is how these results are affected by the amount and type of exposure to an individual.

There are forensically relevant applications of this research. It is clear from these findings that the ability to recognize that a non-stranger is familiar is subject to low hit rates and relatively high false alarm rates typical of recognition memory for strangers, thus rendering familiarity judgments of limited forensic utility. These findings run counter to what appears to be general forensic practice. For example, if an identifying eyewitness says that he has seen the perpetrator before, prosecutors may be less inclined to offer a plea bargain and defense attorneys more inclined to recommend a plea bargain because they assume that this eyewitness evidence is stronger than it would be without the expressed level of familiarity. The results of this study suggest that these assumptions may be ill founded. Estimator variables known to produce high rates of misidentifications in eyewitness recognition memory generally would be predicted to affect the accuracy of familiarity judgments as well. The relatively low hit rate and high false alarm rate reported here suggest that an eyewitness’s report that a perpetrator had been casually seen before is likely to be of limited forensic value in predicting actual familiarity.
Acknowledgements
This article is based on work supported by a Fletcher Jones Foundation grant to the first author. We thank Sister Monica Collins and the teachers and students at St. Lucy’s High School in Glendora, California, and Mrs. Karen Barnes and the teachers and students at Saint Joseph’s High School in Lakewood, California. Without their cooperation, this study would not have been possible. We also thank Elise Mayberry, Luke Meyer, John McCormick, and Chandrima Batacharia for their help coding these data, and Amber Gaffney and David Rast for their conceptual contributions to this project.

Notes
1. It has been estimated that in the USA, approximately 90% of cases are resolved through plea bargaining (Libuser, 2001). So significant is this trend that the effect has been referred to as ‘the disappearing trial’ (Higginbotham, 2004). Attorneys’ decisions whether and how to plea bargain are largely based on their perceptions of the strength of the evidence against the defendant (Burke, 2007).
2. To avoid verbosity, hereafter in this manuscript, non-Hispanic Caucasian participants and faces will be referred to as White.
3. Only 127 participants were included in this analysis because 12 participants of the total sample of 139 had no responses in one of the two conditions.
4. Further, when the d’ data were analyzed with a 2 (race of participant) × 2 (same- vs. cross-race target face) × 2 (high school A or B) ANOVA, there was no significant main effect of school, nor were any of the interactions involving the school factor significant.

References
Hagar v. United States, 856 A.2d 1143 (D.C. Cir. 2004).


