

# Recognition and context memory for faces from own and other ethnic groups: A remember–know investigation

RUTH HORRY

University of Sussex, Brighton, England  
and Royal Holloway, University of London, London, England

DANIEL B. WRIGHT

Florida International University, Miami, Florida

AND

COLIN G. TREDoux

University of Cape Town, Cape Town, South Africa

People are more accurate at recognizing faces from their own ethnic group than at recognizing faces from other ethnic groups. This other-ethnicity effect (OEE) in recognition may be produced by a deficit in recollective memory for other-ethnicity faces. In a single study, White and Black participants saw White and Black faces presented within several different visual contexts. The participants were then given an old/new recognition task. Old responses were followed by remember–know–guess judgments and context judgments. Own-ethnicity faces were recognized more accurately, were given more remember responses, and produced more accurate context judgments than did other-ethnicity faces. These results are discussed in a dual-process framework, and implications for eyewitness memory are considered.

Humans have a remarkable capacity for recognizing faces; however, people are not able to remember all types of faces equally well. Several decades of research have shown that participants recognize faces from their own ethnic group more accurately than they do faces from other ethnic groups (Malpass & Kravitz, 1969; Meissner & Brigham, 2001). We will refer to this effect here as the *own-ethnicity effect* (OEE). Responses to other-ethnicity faces are often characterized by lower hit rates and higher false alarm rates in comparison with responses to own-ethnicity faces (see, e.g., Meissner, Brigham, & Butz, 2005; Teitelbaum & Geiselman, 1997). Such recognition effects are known as *mirror effects* (Glanzer & Adams, 1990).

Dual-process theorists argue that there are two separable routes to recognition: recollection and familiarity. Recollection is the explicit episodic memory of having encountered a stimulus before, including a memory of some information about the encoding event. Familiarity is a more fluid sense of knowing that a stimulus has been encountered before absent a recall of any specific details about that encounter (for a review of dual-process theory, see Yonelinas, 2002). Meissner and colleagues (Evans, Marcon, & Meissner, 2009; Marcon, Susa, & Meissner, 2009; Meissner et al., 2005) recently suggested that the

OEE could be produced by a specific impairment in recollection for other-ethnicity faces. This could account for mirror effects in hit rates and false alarm rates for own- and other-ethnicity faces.

Tulving's (1985) remember–know (RK) procedure is a widely used paradigm for investigating participants' subjective experiences of their recognition decisions. When an item is recognized as old, the participant makes a second judgment, which concerns his or her reasons for making a positive identification. If the participant is able to recall any aspects of the encoding event in which a stimulus was encountered, then he or she makes a remember (R) response. If a participant feels that an item is familiar, but is unable to recall any information about the encoding event, then he or she makes a know (K) response. Many studies also include a guess (G) option, for when recognition decisions are made with no memorial basis (e.g., Konstantinou & Gardiner, 2005). The advantages and disadvantages of the RK procedure have been discussed by Yonelinas (2002) and Gardiner, Ramponi, and Richardson-Klavehn (2002). Although some authors have equated R and K responses to recollection and familiarity, respectively, other authors (e.g., Donaldson, 1996; Dunn, 2004) have cautioned against these interpretations of the

---

R. Horry, ruth.horry@rhul.ac.uk

---

RK procedure. This issue remains contentious, and it will be returned to in the Discussion.

Recollection and familiarity can both produce accurate item recognition. Dual-process theorists argue that memory for context, however, relies much more heavily on recollection than on familiarity. R responses are often associated with higher context-memory accuracy than are K responses, although context accuracy may still be above chance for K responses (Meiser & Sattler, 2007). Meiser, Sattler, and Weißer (2008) showed that the binding of multiple elements of an encoding event into a coherent memory representation is evident in R, but not in K, responses. Although familiarity may allow the retrieval of some residual context information, therefore, recollection is required for retrieving more complex and specific context information.

In a recent study, Horry and Wright (2008) showed that White participants were less accurate at identifying the contexts in which correctly recognized Black faces were seen than at identifying the contexts in which correctly recognized White faces were seen. This study included participants from only one ethnic group, limiting the generalizability of the results. Wells and Olson (2001) argued that a priori differences in the discriminability of the two stimulus groups can create apparent cross-ethnic effects, which are actually artifacts of the stimulus selection. Recruiting participants from both ethnic groups greatly reduces the likelihood of this if a crossover interaction can be shown.

In the present article, we present a single study that improves our understanding of the OEE in three important ways. First, using Tulving's (1985) RK procedure, we add to the emerging evidence that the OEE is characterized by a marked decrease in reports of remembering for other-ethnicity faces in comparison with own-ethnicity faces. Second, we show that memory for context is impaired in cross-ethnic identifications, and that this generalizes beyond White participants. Third, whereas most studies of the OEE have been conducted in the United States or in Europe, we present data from White and Black participants living and studying in South Africa. A small number of OEE investigations have been conducted in Africa (Chiroro, Tredoux, Radaelli, & Meissner, 2008; Chiroro & Valentine, 1995; Wright, Boyd, & Tredoux, 2001, 2003), and crossover interactions were found in some studies but not in others.

South Africa provides a unique societal context for exploring the OEE, because of its recent history of enforced segregation and because of its ethnic multiplicity. The data were collected at a large university with a diverse student population. In this particular setting, there is considerable opportunity for students to interact with people from different ethnic backgrounds, so one might expect smaller cross-ethnic effects. Even within multicultural societies or organizations, however, people are much more likely to have meaningful and intimate relationships with people from their ethnic in-group than with people from other ethnic groups (Dixon, Tredoux, & Clack, 2005; Tredoux & Dixon, 2009). Participants from such areas could provide

some new insights into the subtleties and complexities of cross-ethnic effects, which would strengthen the generalizability of the results from the United States and Europe.

## METHOD

### Participants and Design

Forty-one White and 44 Black undergraduates at the University of Cape Town took part in the study for credit in an introductory psychology course or for payment (approximately \$6). Nine of the participants were male (2 White, 7 Black). The mean age was 21 years; ages ranged from 17 to 43.

The experiment had a 2 (ethnicity of participant: White or Black)  $\times$  2 (ethnicity of face: White or Black)  $\times$  4 (context: basketball court, classroom, factory, jail cell) mixed design. The first factor was between groups, and the latter two factors were repeated measures. The dependent variables were face-recognition measures (hits, false alarms,  $d'$ , and  $C$ ), decision strategy (R, K, and G judgments), and context-memory accuracy.

### Materials and Apparatus

Eighty White faces and 80 Black faces of young males were used as stimuli. All of the Black faces and approximately half of the White faces were taken from a South African database, and the remaining White faces were taken from an Australian database. Although it would have been ideal to source all of the stimulus faces locally (see Chiroro et al., 2008), the South African database did not include a sufficient number of young white males. Both databases had been created specifically for research, and they were of comparable quality.

Two photographs of each face were used in the experiment—one with a full frontal pose and one with a three-quarter-profile pose. A different view of each face was presented at encoding and testing to ensure that participants were recognizing the faces rather than any low-level features of the particular photograph (Bruce, 1982). Each photograph was cropped so that the necks and shoulders were removed. The cropped faces were then pasted onto the four backgrounds using Adobe Photoshop. Which faces appeared on which backgrounds was counterbalanced between participants. Which faces were targets and which were foils was also counterbalanced. The experiment was designed and run using the E-Prime software system.

### Procedure

Participants were tested in groups of up to 5. Each participant sat at an individual PC and worked independently. Participants were told that they would see a series of faces presented on different backgrounds, and that they would later be asked to remember the faces and their associated backgrounds.

The experiment was broken down into 10 short blocks, each with a learning phase that was followed by a testing phase. During each learning phase, 8 targets appeared (one target for each combination of ethnicity and context). Each target and its associated background was shown for 8 sec, with a 1-sec interstimulus interval. In each testing phase, the 8 targets were seen again, along with eight new faces (four White and four Black). The recognition test was self-paced, and each face remained on screen until the participant had responded to all of the questions. In total, 10 targets of each ethnicity appeared on each background throughout the experiment, giving a total of 40 White and 40 Black targets. Forty distractors from each group were also seen. Participants were not told the proportion of old to new faces. The order of the blocks was randomized for each participant, and the order in which the faces were presented within each block was also randomized.

Participants made an old/new recognition decision for each face. Following an old judgment, participants made a remember-know-guess (RKG) judgment, followed by a context-memory decision. Instructions for the RKG judgments were provided on the screen

and were also printed on a sheet of paper so that the participants could refer to them at any time. The instructions were adapted from Conway and Dewhurst (1995) to make them suitable for use in a face-recognition experiment. Care was taken not to mention the encoding contexts in the RKG instructions in order to avoid confounds between reports of remembering and memory for context (Meiser & Sattler, 2007; Meiser et al., 2008). The instructions were as follows:

Sometimes when we recognize a face we have seen before, we can consciously REMEMBER specific details about the previous occurrence of the face. At other times, we simply KNOW that we have seen a face before, even though we cannot recall specific details about the event.

For each face, you will have 3 options:

1. REMEMBER: When you see a face in this test that you recognize, you may be able to remember specific details about seeing that face before. You may for example recollect the thoughts or feelings that the face evoked when you saw it, or an association that you formed with another face, or some aspect of the face’s physical appearance. In short, any additional detail which supports your belief that the face appeared before.

2. KNOW: For other faces, you may know that they appeared before, even though you cannot remember any specific details of their occurrence.

3. GUESS: For some faces, you may have simply guessed, without any reason to believe that the face appeared before.

For the context-memory judgments, participants were required to press 1 on the computer keyboard for the basketball court, 2 for the classroom, 3 for the factory, or 4 for the jail. This was a forced choice task, with no option to respond “don’t know.”

**RESULTS**

**Face-Recognition Accuracy and Response Bias**

Hit rates and false alarm rates were analyzed separately, since previous research has shown that the OEE is often characterized by a marked increase in false alarm rates for other ethnicity-faces compared to those for own-ethnicity faces (Meissner & Brigham, 2001). The hit rates and false alarm rates were then used to calculate the signal-detection measures *d'* and *C*<sup>1</sup> (see Wright, Horry, & Skagerberg, 2009, for formulae). Separate 2 (ethnicity of participant) × 2 (ethnicity of face) mixed ANOVAs were conducted for each outcome measure.<sup>2</sup>

Table 1 shows the mean hit rates, false alarm rates, *d'*, and *C* for White and Black participants with White and Black stimuli. Across all four measures, the interaction term between ethnicity of participant and ethnicity of face was significant and large [hit rates,  $F(1,83) = 60.83, p < .001, \eta_p^2 = .42$ ; false alarm rates,  $F(1,83) = 45.72, p < .001, \eta_p^2 = .36$ ; for *d'*,  $F(1,83) = 55.69, p < .001, \eta_p^2 = .40$ ; *C*,  $F(1,83) = 55.69, p < .001, \eta_p^2 = .40$ ]. White participants had higher hit rates for White faces than for Black faces [ $t(40) = 8.00, p < .001, d = .98$ ] and higher false alarm rates for Black faces than for White faces [ $t(40) = 6.43, p < .001, d = .63$ ]. Black participants showed the opposite pattern, with higher hit rates for Black faces than for White faces [ $t(43) = 3.24, p = .002, d = .50$ ] and higher false alarm rates for White faces than for Black faces [ $t(43) = 3.04, p = .004, d = .33$ ].

The signal-detection measure *d'* confirmed that White participants were more accurate with White faces than with Black faces [ $t(40) = 7.25, p < .001, d = .74$ ], whereas Black participants were more accurate with Black faces than with White faces [ $t(43) = 3.12, p = .003, d = .27$ ]. The criterion measure *C* showed that White participants were more conservative with White faces than with Black faces [ $t(40) = 7.25, p < .001, d = .53$ ], whereas Black participants were more conservative with Black faces than with White faces [ $t(43) = 3.12, p = .003, d = .24$ ].

Across all of the face-recognition measures, the effect sizes appear to have been larger for White participants than for Black participants. The data were recoded so that, for each measure, a difference score was created (own-ethnicity minus other-ethnicity). Between-groups comparisons were then carried out to compare the difference scores for White and Black participants. The mean differences for White participants were significantly larger than the mean differences for Black participants across all four measures [hit rates,  $t(83) = 2.91, p = .005, d = .64$ ; false alarm rates,  $t(83) = 2.57, p = .01, d = .57$ ; for *d'*,  $t(83) = 3.19, p = .002, d = .70$ ; *C*,  $t(83) = 2.97, p = .004, d = .65$ ].

**RKG Responses**

The way in which RKG responses are analyzed is a contentious issue, and one’s position will depend on assumptions about how recollection and familiarity operate at the trial level. A discussion of these issues was presented by Gardiner, Java, and Richardson-Klavehn (1996). Some authors argue that the independence remember-know (IRK) procedure, developed by Yonelinas and Jacoby (1995), should be applied to know responses. In the IRK procedure, familiarity is estimated as follows:  $F = K/(1 - R)$ . Other authors (e.g., Gardiner et al., 1996) argue that the analysis should simply explore the proportion of K responses given. In the present article, we present data both for the proportion of K responses and for familiarity estimates from the IRK procedure. Table 2 shows the mean proportions of hit trials producing R, K, and G responses, and the IRK familiarity estimates, broken down by ethnicity of participant and ethnicity of face. Since G responses

**Table 1**  
Mean Hit Rates, False Alarm (FA) Rates, *d'*, and *C* for White Participants and for Black Participants With White and Black Faces

Stimuli	Measure							
	Hit Rates		FA Rates		<i>d'</i>		<i>C</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
White Participants								
White faces	.76	.11	.12	.11	1.77	0.59	-.07	.37
Black faces	.65	.12	.20	.14	1.36	0.52	-.27	.37
Black Participants								
White faces	.63	.15	.15	.12	1.71	0.59	-.46	.36
Black faces	.70	.13	.12	.10	1.87	0.59	-.38	.33

**Table 2**  
**Mean Proportions of Remember, Know, and Guess Responses in Hit Trials, Along With Transformed Independence Remember–Know (IRK) Responses, Broken Down by Ethnicity of Participant and Ethnicity of Face**

Stimuli	Measure							
	Remember		Know		IRK		Guess	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
White Participants								
White faces	.70	.20	.21	.17	.71	.39	.09	.10
Black faces	.54	.23	.27	.16	.63	.29	.18	.19
Black Participants								
White faces	.57	.25	.28	.23	.58	.29	.14	.11
Black faces	.61	.24	.28	.24	.66	.25	.11	.09

are dependent on R and K responses, G responses were not analyzed. All of the other measures were analyzed in separate 2 (ethnicity of participant) × 2 (ethnicity of face) mixed ANOVAs.<sup>3</sup>

Table 2 shows that both groups of participants made more R responses for own-ethnicity faces than for other-ethnicity faces, although the mean difference appears to have been larger for White participants than for Black participants. This observation was confirmed in the ANOVA, since the interaction term was significant [ $F(1,83) = 34.40, p < .001, \eta_p^2 = .29$ ]. White participants made significantly more R responses to White faces than to Black faces [ $t(40) = 6.25, p < .001, d = .74$ ]. In contrast, Black participants made slightly more R responses to Black faces than to White faces, although this difference was not statistically significant [ $t(43) = 1.72, p = .09, d = .15$ ].

The table shows that White participants made more K responses for Black faces than for White faces, whereas Black participants were equally likely to make K responses for White and Black faces. The ANOVA confirmed these observations. The interaction term was significant [ $F(1,83) = 4.43, p = .04, \eta_p^2 = .05$ ], but the difference was significant for White participants [ $t(40) = 3.00, p = .005, d = .40$ ] and not for Black participants [ $t(43) = 0.34, p = .74, d = .01$ ].

A different pattern seems to have emerged when the IRK transformation was applied to the RKG data, with higher means for own-ethnicity faces than for other-

ethnicity faces within both groups of participants. Again, the interaction between ethnicity of participant and ethnicity of face was significant [ $F(1,83) = 4.73, p = .03, \eta_p^2 = .05$ ]; however, the differences between White and Black faces were not statistically significant for either the White participants [ $t(40) = 1.36, p = .18, d = .22$ ] or the Black participants [ $t(43) = 1.74, p = .09, d = .29$ ].

**Context Memory**

Context memory was analyzed in a multilevel logistic regression. Context judgments were contingent upon participants giving an old response to a stimulus. This means that context judgments were produced for hit and false alarm trials. Since all context judgments in false alarm trials will be erroneous, only hit trials were included in the analysis. Because the number of hits varies by person, a multilevel model was used so that inappropriate weight was not given to people with different numbers of hits (Wright & London, 2009). The outcome measure, whether the context judgment was accurate or inaccurate, was binary, so a multilevel logistic regression was used (Wright, 1997). Participant variation and stimulus variation were included as random effects.

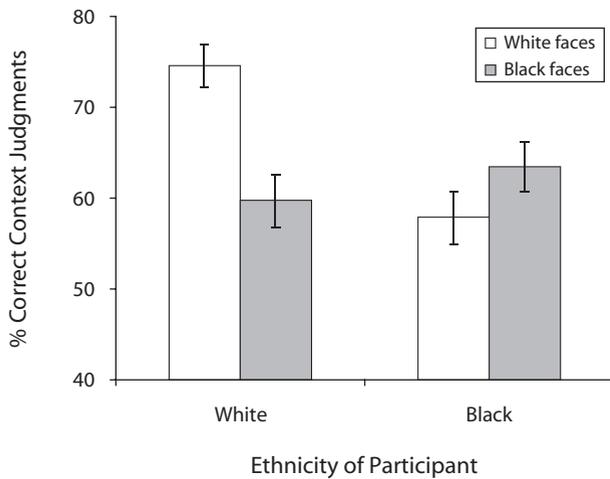
The regression model was built in two steps, shown in Table 3. In the first step (shown in the left-hand columns), ethnicity of participant and ethnicity of face were entered as predictors of context accuracy. White participants were more likely to be accurate (67.84%) than were Black participants (60.80%) ( $\beta_1 = .38, p = .01$ ). Overall, context judgments were more likely to be accurate when responding to White faces (66.79%) than when responding to Black faces (61.72%) ( $\beta_2 = .22, p = .02$ ).

In the second step of the model, the interaction between ethnicity of participant and ethnicity of face was entered. This interaction was a strong predictor of accuracy ( $\beta_3 = .86, p < .001$ ). The percentages of correct context judgments, broken down by ethnicity of participant and ethnicity of face, are shown in Figure 1. The figure shows that White participants were more accurate with White faces than with Black faces. For Black participants, there seems to have been a smaller difference in the opposite direction, with higher accuracy for Black faces than for White

**Table 3**  
**Comparison of Regression Models Predicting Context Accuracy From Ethnicity of Participant, Ethnicity of Face, and Remember–Know–Guess Response**

Parameter		Model 1			Model 2		
		Estimate	<i>SE</i>	<i>p</i>	Estimate	<i>SE</i>	<i>p</i>
Fixed Effects							
Intercept	$\beta_0$	.32	.12	.007	.99	.12	<.001
Ethnicity of participant	$\beta_1$	.38	.15	.01	.05	.17	.75
Ethnicity of face	$\beta_2$	.22	.10	.02	.25	.11	.02
OEE interaction	$\beta_3$	–	–	–	.86	.13	<.001
Remember–know	$\beta_4^{(a)}$	–	–	–	.86	.08	<.001
Remember–guess	$\beta_4^{(b)}$	–	–	–	1.45	.11	<.001
Guess–know	$\beta_4^{(c)}$	–	–	–	.59	.11	<.001
Random Effects							
Participant level	$\sigma^2_u$	.39			.39		
Stimulus level	$\sigma^2_e$	.20			.13		

Note—Difference in fit between the two models:  $\chi^2(4) = 298.66, p < .001$ .



**Figure 1.** The percentage of correct context judgments on hit trials, broken down by ethnicity of face and ethnicity of participant. Error bars represent 95% confidence intervals.

faces. To confirm these observations, separate regressions were run on the data from White and Black participants. White participants were more accurate with White faces (75.45%) than with Black faces (59.70%) ( $\beta = .80, p < .001$ ), whereas Black participants were more accurate with Black faces (63.45%) than with White faces (57.87%) ( $\beta = .29, p = .01$ ).

RKG responses were also added to the second step of the regression model. Context judgments were more likely to be accurate in R trials (73.67%) than in K trials (54.11%) ( $\beta_4^{(a)} = .86, p < .001$ ) or in G trials (39.83%) ( $\beta_4^{(b)} = 1.45, p < .001$ ). K trials were also more accurate than G trials ( $\beta_4^{(c)} = .59, p < .001$ ).

Context responses always followed old decisions and RKG judgments. A possible confound could have occurred if the RKG responses influenced participants' later ability to remember contextual information. For example, providing an R judgment may have generated additional retrieval cues, allowing easier access to the representation of the encoding context. A further regression was run, using only data from R trials ( $N = 2,909$  trials). The interaction between ethnicity of participant and ethnicity of face remained a strong predictor of accuracy ( $\beta_3 = .88, p < .001$ ). White participants were more accurate with White faces (82.13%) than with Black faces (70.02%) ( $\beta = .71, p < .001$ ). Although Black participants were slightly more accurate with Black faces (70.92%) than with White faces (67.68%), this difference was nonsignificant ( $\beta = -.18, p = .19$ ).

To confirm the validity of these analyses, the context data were used to calculate a measure of source identification that is independent of item recognition, the *average conditional source identification measure* (ACSIM). This measure can vary between 0 and 1, with higher scores indicating more accurate memory for context. The calculation of this measure was described in some detail by Murnane and Bayen (1996). The ACSIMs were then analyzed in a

**Table 4**  
Mean Average Conditional Source Identification Measure Values for White and Black Faces, Broken Down by Ethnicity of Participant

Ethnicity of Participant	White Faces		Black Faces	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
White	.74	.15	.59	.19
Black	.57	.14	.63	.18

2 (ethnicity of face)  $\times$  2 (ethnicity of participant) mixed ANOVA. The mean ACSIMs are shown in Table 4, broken down by ethnicity of face and ethnicity of participant. The table shows that for White participants, the mean ACSIM was higher for White faces than for Black faces. For Black participants, the means were in the opposite direction.

This observation was confirmed by a significant interaction in the ANOVA [ $F(1,82) = 36.57, p < .001, \eta_p^2 = .31$ ]. White participants were significantly more accurate with White faces than with Black faces [ $t(40) = 6.13, p < .001, d = .92$ ], whereas Black participants were more accurate with Black faces than with White faces [ $t(43) = 2.27, p = .03, d = .34$ ]. The main effects in the analysis were also significant. Overall, accuracy was higher for White faces ( $M = .66, SD = .17$ ) than for Black faces ( $M = .61, SD = .18$ ) [ $F(1,82) = 8.71, p = .004, d = .28$ ]. White participants were more accurate ( $M = .66, SD = .15$ ) than Black participants ( $M = .60, SD = .14$ ) [ $F(1,82) = 3.79, p = .06, d = .43$ ]. The ACSIM analysis confirms the findings from the regression analysis that memory for context is impaired in other-ethnicity identifications compared with own-ethnicity identifications.

## DISCUSSION

White and Black participants were shown White and Black faces presented within several different visual contexts. In a recognition test, participants attempted to discriminate old faces from new faces, and made RKG judgments and context judgments for faces identified as old. The study had three main aims: (1) to provide evidence that own-ethnicity faces are accompanied by more reports of remembering than are other-ethnicity faces (see, e.g., Evans et al., 2009; Marcon et al., 2009; Meissner et al., 2005); (2) to show that people have difficulty in retrieving contextual information about correctly recognized other-ethnicity faces, and that this effect generalizes past White participants; and (3) to show that the OEE can be found outside of North America and Europe, even in an area of considerable ethnic diversity.

The first aim of this study was met, since the analyses of the RK judgments clearly showed that the OEE had dissociable effects on R and K responses. White participants made significantly more R judgments for White faces than for Black faces, whereas Black participants showed a nonsignificant trend in the opposite direction. The effects on knowing were smaller and more difficult to interpret, since the relationship between K responses and familiarity was not straightforward. When the proportions

of K responses were analyzed, White participants were somewhat more likely to make K responses to Black faces than to White faces. Black participants, however, showed very little difference in their use of K responses for White and Black faces. When the IRK procedure was used, a small yet significant interaction emerged between ethnicity of participant and ethnicity of face. For both groups of participants, the mean IRK estimate was slightly higher for own-ethnicity faces than for other-ethnicity faces, although these differences were nonsignificant.

These results support those of Meissner et al. (2005), who found an increase in R responses for own-ethnicity faces but similar IRK estimates for own- and other-ethnicity faces in an old/new recognition task. The RK procedure has many advantages—for example, it is easy to incorporate into many experimental designs and is easy for participants to understand. It also provides a very inclusive definition of recollection, since participants are able to make R responses if they can recall any aspect of the encoding event. There are also disadvantages associated with the RK procedure, however. Some authors have suggested that participants may use R and K responses to reflect high and low confidence, respectively (e.g., Donaldson, 1996; Dunn, 2004). Participants may not always have direct access to their memory processes, so the procedure will inevitably be associated with some measurement error. The analysis of RK responses is a contentious issue, since there are still debates concerning the relationship between recollection and familiarity. The decision to use the IRK procedure (Yonelinas & Jacoby, 1995) or to analyze the proportion of K responses (see, e.g., Gardiner et al., 1996) can influence the pattern of results. One should always be very careful when discussing RK responses as if they map directly onto underlying memory processes.

The second aim of the study was also met, since participants from both groups were less accurate at identifying the contexts associated with other-ethnicity faces than at identifying the contexts associated with own-ethnicity faces. Analyzing memory for context can be difficult, since context recognition can be confounded with item recognition. To confirm the validity of the OEE for context memory, the analysis was conducted in two ways. First, a multilevel regression was run on hit-trial data only; second, ACSIM values were analyzed in a traditional ANOVA (Murnane & Bayen, 1996). Using both of these methods, the results extended those of Horry and Wright (2008) to Black participants.

Context memory was much more accurate on R trials than on K trials. This is not surprising, since reports of remembering are more often associated with increased memory for source information than are K responses (Meiser & Sattler, 2007); however, K and G responses also produced correct context judgments at a higher rate than chance. It may be that the K and G responses were not “process pure,” and that they included some identifications in which there was some degree of recollection present. It is also possible that some context memory can be supported by familiarity (Meiser & Sattler, 2007). This

position is supported in the implicit-learning literature (Scott & Dienes, 2008).

Failures of contextual memory can have very serious consequences in some situations, such as when a witness identifies a suspect in a criminal case (Lindsay, 2007). For example, a substantial body of research has shown that searching mugshots after witnessing a crime is very harmful to later lineup identifications (Deffenbacher, Bornstein, & Penrod, 2006). In such a case, the witness may find a particular individual familiar yet be unable to correctly identify the source of that familiarity. Thus, a face that was seen in the mugshot search may be misidentified as that of the perpetrator. Our results indicate that failures of context memory may occur more frequently in other-ethnicity identifications than in own-ethnicity identifications. This could have serious implications for the legal system, in which eyewitness errors are a leading cause of wrongful conviction (Scheck, Neufeld, & Dwyer, 2003).

The third main aim of this paper was to examine the OEE in a population outside of North America or Europe. A handful of studies carried out in Africa have produced inconsistent results, with some but not others finding crossover effects (Chiroro et al., 2008; Chiroro & Valentine, 1995; Wright et al., 2001, 2003). South Africa provides a particularly interesting societal context for testing the OEE because of its social and political history and its ethnic diversity. In the present study, across all of the face-recognition measures, significant crossover effects were found. Both White and Black participants identified own-ethnicity faces more accurately than they identified other-ethnicity faces. The effect sizes in all cases were larger for White participants than for Black participants, however. Such asymmetry is not uncommon in OEE studies, and it is difficult to interpret. The two groups may have differed in the amount of meaningful contact that they had had with members of other groups. There could also have been a priori differences in the memorability of the two stimulus sets.

Throughout this article, the OEE in item and context memory has been interpreted within a dual-process framework. Supporters of dual-process models argue that R and K responses reflect recollection and familiarity, respectively (e.g., Gardiner et al., 1996; Gardiner et al., 2002), and that memory for contextual information relies heavily on recollection (Meiser & Sattler, 2007; Meiser et al., 2008). The present study, when related to dual-process theory, suggests that recollection is specifically impaired for other-ethnicity faces; however, the results of this study are also consistent with single-process models of memory. Dunn (2004) and Donaldson (1996) argued that R responses are made to items that are recognized with high confidence, and that apparent dissociations between R and K responses can be explained by a simple memory-strength model. DeCarlo (2003) also argued that source memory can be explained by an extended multivariate signal-detection model, which does not assume the existence of two separate memory processes. Glanzer, Hilford, and Kim (2004) showed that receiver operating characteristics (ROCs) for item recognition and source

recognition share similar properties (e.g., the ROCs are convex, the  $z$ -ROCs are linear), and that items that influence item recognition have very similar effects on source recognition. A memory-strength interpretation of the present results would suggest that own-ethnicity faces produce stronger activation than do other-ethnicity faces. This leads to decreased item recognition, fewer reports of remembering, and impaired memory for contextual details of other-ethnicity faces.

The results could also potentially be explained by biases in decision making rather than in memory. As is commonly found in studies of the OEE, participants were more conservative when identifying own-ethnicity faces than when identifying other-ethnicity faces (see Meissner & Brigham, 2001). The participants in the present study were aware that they would be required to make RKG and context judgments for faces that they identified as old. For own-ethnicity faces, participants may therefore have become increasingly conservative, identifying faces only when they felt that they could make an R judgment and/or when they felt reasonably confident of making a correct context judgment. Differences in response bias for item recognition could therefore have exaggerated or even created the differences in RK judgments and context accuracy for own- and other-ethnicity faces. Future research should attempt to disentangle the effects of decision making at the item-recognition stage from later judgments. For example, participants could make old/new judgments, RK judgments, and context judgments in entirely separate phases of the testing procedure.

To conclude, both White and Black participants showed item-recognition and context-recognition advantages for own-ethnicity faces in comparison with other-ethnicity faces. Own-ethnicity faces were more likely than other-ethnicity faces to be accompanied by reports of remembering. Failures to remember contextual details about familiar faces could have harmful consequences within the legal system, since innocent individuals could be mistakenly identified by witnesses. Further research should explore whether such errors are more likely to occur in other-ethnicity identifications than in own-ethnicity identifications.

#### AUTHOR NOTE

This research was funded by the U.K. Economic and Social Research Council. Correspondence related to this article may be sent to R. Horry, Royal Holloway Eyewitness Group, Department of Psychology, Royal Holloway University of London, Egham, Surrey TW20, England (e-mail: ruth.horry@rhul.ac.uk).

#### REFERENCES

- BRUCE, V. (1982). Changing faces: Visual and nonvisual coding processes in face recognition. *British Journal of Psychology*, *73*, 105-116.
- 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. doi:10.3758/PBR.15.6.1089
- CHIRORO, P. [M.], & VALENTINE, T. (1995). An investigation of the contact hypothesis of the own-race bias in face recognition. *Quarterly Journal of Experimental Psychology*, *48A*, 879-894.
- CONWAY, M. A., & DEWHURST, S. A. (1995). The self and recollective experience. *Applied Cognitive Psychology*, *9*, 1-19. doi:10.1002/acp.2350090102
- DECARLO, L. T. (2003). Source monitoring and multivariate signal detection theory, with a model for selection. *Journal of Mathematical Psychology*, *47*, 292-303. doi:10.1016/S0022-2496(03)00005-1
- DEFFENBACHER, K. A., BORNSTEIN, B. H., & PENROD, S. D. (2006). Mugshot exposure effects: Retroactive interference, mugshot commitment, source confusion, and unconscious transference. *Law & Human Behavior*, *30*, 287-307. doi:10.1007/s10979-006-9008-1
- DIXON, J. [A.], TREDOUX, C. [G.], & CLACK, B. (2005). On the microecology of racial division: A neglected dimension of segregation. *South African Journal of Psychology*, *35*, 395-411.
- DONALDSON, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition*, *24*, 523-533.
- DUNN, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review*, *111*, 524-542. doi:10.1037/0033-295X.111.2.524
- EVANS, J. R., MARCON, J. L., & MEISSNER, C. A. (2009). Cross-racial lineup identification: Assessing the potential benefits of context reinstatement. *Psychology, Crime, & Law*, *15*, 19-28. doi:10.1080/10683160802047030
- GARDINER, J. M., JAVA, R. I., & RICHARDSON-KLAVEHN, A. (1996). How level of processing really influences awareness in recognition memory. *Canadian Journal of Experimental Psychology*, *50*, 114-122. doi:10.1037/1196-1961.50.1.114
- GARDINER, J. M., RAMPONI, C., & RICHARDSON-KLAVEHN, A. (2002). Recognition memory and decision processes: A meta-analysis of remember, know, and guess responses. *Memory*, *10*, 83-98. doi:10.1080/09658210143000281
- GLANZER, M., & ADAMS, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *16*, 5-16. doi:10.1037/0278-7393.16.1.5
- GLANZER, M., HILFORD, A., & KIM, K. (2004). Six regularities of source recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*, 1176-1195. doi:10.1037/0278-7393.30.6.1176
- HORRY, R., & WRIGHT, D. B. (2008). I know your face but not where I saw you: Context memory is impaired for other-race faces. *Psychonomic Bulletin & Review*, *15*, 610-614. doi:10.3758/PBR.15.3.610
- KONSTANTINOVA, I., & GARDINER, J. M. (2005). Conscious control and memory awareness when recognising famous faces. *Memory*, *13*, 449-457. doi:10.1080/09658210444000016
- LINDSAY, D. S. (2007). Autobiographical memory, eyewitness reports, and public policy. *Canadian Psychology*, *48*, 57-66.
- MALPASS, R. S., & KRAVITZ, J. (1969). Recognition for faces of own and other race. *Journal of Personality & Social Psychology*, *13*, 330-334. doi:10.1037/h0028434
- MARCON, J. L., SUSAN, 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. doi:10.3758/PBR.16.1.99
- MEISER, T., & SATTLER, C. (2007). Boundaries of the relation between conscious recollection and source memory for perceptual details. *Consciousness & Cognition*, *16*, 189-210. doi:10.1016/j.concog.2006.04.003
- MEISER, T., SATTLER, C., & WEIßER, K. (2008). Binding of multidimensional context information as a distinctive characteristic of remember judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *34*, 32-49. doi:10.1037/0278-7393.34.1.32
- 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, & Law*, *7*, 3-35. doi: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. doi:10.1002/acp.1097
- MURNANE, K., & BAYEN, U. J. (1996). An evaluation of empirical measures of source identification. *Memory & Cognition*, *24*, 417-428.
- SCHECK, B., NEUFELD, P., & DWYER, J. (2003). *Actual innocence: When justice goes wrong and how to make it right*. New York: New American Library.
- SCOTT, R. B., & DIENES, Z. (2008). The conscious, the unconscious, and familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *34*, 1264-1288. doi:10.1037/a0012943

- TEITELBAUM, S., & GEISELMAN, R. E. (1997). Observer mood and cross-racial recognition of faces. *Journal of Cross-Cultural Psychology*, **28**, 93-106. doi:10.1177/0022022197281006
- TREDoux, C. G., & DIXON, J. A. (2009). Mapping the multiple contexts of racial isolation: The case of Long Street, Cape Town. *Urban Studies*, **46**, 761-777. doi:10.1177/0042098009102128
- TULVING, E. (1985). Memory and consciousness. *Canadian Psychology*, **26**, 1-12.
- WELLS, G. L., & OLSON, E. A. (2001). The other-race effect in eyewitness identification: What do we do about it? *Psychology, Public Policy, & Law*, **7**, 230-246. doi:10.1037/1076-8971.7.1.230
- WRIGHT, D. B. (1997). Extra-binomial variation in multilevel logistic models with sparse structures. *British Journal of Mathematical & Statistical Psychology*, **50**, 21-29.
- WRIGHT, D. B., BOYD, C. E., & TREDoux, C. G. (2001). A field study of own-race bias in South Africa and England. *Psychology, Public Policy, & Law*, **7**, 119-133. doi:10.1037/1076-8971.7.1.119
- WRIGHT, D. B., BOYD, C. E., & TREDoux, C. G. (2003). Inter-racial contact and the own-race bias for face recognition in South Africa and England. *Applied Cognitive Psychology*, **17**, 365-373. doi:10.1002/acp.898
- WRIGHT, D. B., HORRY, R., & SKAGERBERG, E. M. (2009). Functions for traditional and multilevel approaches to signal detection theory. *Behavior Research Methods*, **41**, 257-267. doi:10.3758/BRM.41.2.257
- WRIGHT, D. B., & LONDON, K. (2009). Multilevel modeling: Beyond the basic applications. *British Journal of Mathematical & Statistical Psychology*, **62**, 439-456. doi:10.1348/000711008X327632
- YONELINAS, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language*, **46**, 441-517. doi:10.1006/jmla.2002.2864
- YONELINAS, A. P., & JACOBY, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory & Language*, **34**, 622-643. doi:10.1006/jmla.1995.1028

#### NOTES

1. To calculate  $d'$  and  $C$  across the different levels of context, hit rates were combined with overall false alarm rates for White and Black faces. The foils had not appeared on any encoding background and so could not be broken down by context. A flattening constant of .5 was added to each cell, since  $d'$  and  $C$  cannot be calculated if hit rates or false alarm rates are equal to 1 or 0.

2. Analyses including encoding context revealed no significant main effects or interactions on any outcome measure (minimum  $p = .16$ ); therefore, context will not be discussed further.

3. The same analyses were conducted on R and K responses for false alarm trials. No significant effects were found [main effects, maximum  $F(1,83) = 2.40, p = .13, d = .34$ ; interaction terms, maximum  $F(1,83) = 1.78, p = .18, \eta_p^2 = .02$ ].

(Manuscript received April 14, 2009;  
revision accepted for publication September 12, 2009.)