Research Article

The Other-Race Effect Develops During Infancy

Evidence of Perceptual Narrowing

David J. Kelly,1 Paul C. Quinn,2 Alan M. Slater,3 Kang Lee,4 Liezhong Ge,5 and Olivier Pascalis1

1University of Sheffield, Sheffield, United Kingdom; 2University of Delaware; 3University of Exeter, Exeter, United Kingdom; 4University of Toronto, Toronto, Ontario, Canada; and 5Zhejiang Sci-Tech University, Hangzhou, People’s Republic of China

ABSTRACT—Experience plays a crucial role in the development of face processing. In the study reported here, we investigated how faces observed within the visual environment affect the development of the face-processing system during the 1st year of life. We assessed 3-, 6-, and 9-month-old Caucasian infants’ ability to discriminate faces within their own racial group and within three other-race groups (African, Middle Eastern, and Chinese). The 3-month-old infants demonstrated recognition in all conditions, the 6-month-old infants were able to recognize Caucasian and Chinese faces only, and the 9-month-old infants’ recognition was restricted to own-race faces. The pattern of preferences indicates that the other-race effect is emerging by 6 months of age and is present at 9 months of age. The findings suggest that facial input from the infant’s visual environment is crucial for shaping the face-processing system early in infancy, resulting in differential recognition accuracy for faces of different races in adulthood.

Human adults are experts at recognizing faces of conspecifics and appear to perform this task effortlessly. Despite this impressive ability, however, adults are more susceptible to recognition errors when a target face is from an unfamiliar racial group, rather than their own racial group. This phenomenon is known as the other-race effect (ORE; see Meissner & Brigham, 2001, for a review). Although the ORE has been widely reported, the exact mechanisms that underlie reduced recognition accuracy for other-race faces, and precisely when this effect emerges during development, remain unclear.

The ORE can be explained in terms of a modifiable face representation. The concept of a multidimensional face-space architecture, first proposed by Valentine (1991), has received much empirical support. According to the norm-based coding model, individual face exemplars are represented as vectors within face-space according to their deviation from a prototypical average. The prototype held by each person represents the average of all faces that person has ever encoded and is therefore unique. Although it is unclear which dimensions are most salient and used for recognition, it is likely that dimensions vary between individuals and possibly within each person over time. The prototype (and therefore the entire face-space) continually adapts and is updated as more faces are observed within the environment. Consequently, individuating face-space dimensions of a person living in China are expected to be optimal for recognition of other Chinese persons, but not, for example, for recognition of African individuals.

Other authors have hypothesized that the dimensions of the face prototype present at birth are broad and develop according to the type of facial input received (Nelson, 2001). According to this account, predominant exposure to faces from a single racial category tunes face-space dimensions toward that category. Such tuning might be manifested at a behavioral level in differential responding to own- versus other-race faces, for example, in spontaneous visual preference and a recognition advantage for own-race faces.

Recent findings regarding spontaneous preference have confirmed the impact of differential face input on the tuning of the face prototype during early infancy. It has been demonstrated that selectivity based on ethnic facial differences emerges very early in life, with 3-month-old infants preferring to look at faces from their own group, as opposed to faces from other
The purpose of the study reported here was to clarify the developmental origins of the ORE during the first months of life. Using the VPC task, we assessed the ability of 3-, 6-, and 9-month-old Caucasian infants to discriminate within own-race (Caucasian) faces and within three categories of other-race faces (African, Middle Eastern, and Chinese). This task measures relative interest in the members of pairs of stimuli, each consisting of a novel stimulus and a familiar stimulus observed during a prior habituation period. Recognition of the familiar stimulus is inferred from the participant’s tendency to fixate on the novel stimulus. Previous studies have found that 3-month-old infants can perform this task even when they are exposed to different views of faces (e.g., full view vs. 3/4 profile) during the habituation period and the recognition test (Pascalis, de Haan, Nelson, & de Schonen, 1998). We also varied face views between familiarization and testing, a procedure that is preferable to using identical pictures in the habituation and testing phases because it ensures that face recognition—as opposed to picture recognition (i.e., image matching)—is tested. Our selection of which age groups to test was based on previous research demonstrating that the ORE is found in infancy (3-month-olds; Sangrigoli & de Schonen, 2004b) and that the face-processing system appears to undergo a period of tuning between 6 and 9 months of age (Pascalis et al., 2002).

**METHOD**

Participants
In total, 192 Caucasian infants were included in the final analysis. There were 64 subjects in each of three age groups: 3-month-olds (age range = 86–102 days; 33 females, 31 males), 6-month-olds (age range = 178–196 days; 31 females, 33 males), and 9-month-olds (age range = 268–289 days; 30 females, 34 males). All participants were healthy, full-term infants. Within each age group, the infants were assigned in equal numbers (n = 16) to the four testing conditions (Caucasian, African, Middle Eastern, and Chinese). The infants were recruited from the maternity wing of the Royal Hallamshire Hospital, Sheffield, United Kingdom. In each age group, we tested additional infants who were excluded from the final analysis. Twenty-two 3-month-old infants were excluded because of failure to habituate (n = 4), side bias during testing (>95% looking time to one side; n = 15), or fussiness (n = 3); sixteen 6-month-old infants were excluded because of failure to habituate (n = 7), side bias during testing (n = 3), parental interference (n = 2), or fussiness (n = 4); and eleven 9-month-old infants were excluded because of a failure to habituate (n = 3) or fussiness (n = 8).

**Stimuli**
The stimuli were 24 color images of male and female adult faces (age range = 23–27 years) from four different ethnic groups (African, Asian, Middle Eastern, and Caucasian). All faces had dark hair and dark eyes so that the infants would be unable to demonstrate recognition on the basis of these features. The
images were photos of students. The Africans were members of the African and Caribbean Society at the University of Sheffield; the Asians were Han Chinese students from Zhejiang Sci-Tech University, Hangzhou, China; the Middle Easterners were members of the Pakistan Society at the University of Sheffield; and the Caucasians were psychology students at the University of Sheffield.

For each ethnic group, we tested male and female faces in separate conditions. The images for each combination of ethnic group and gender consisted of a habituation face and two test faces, a novel face and the familiar face in a new orientation. The two faces in the test phase were always in the same orientation, and this orientation differed from the orientation of the face seen during habituation. In one orientation condition, infants were habituated to full-view faces and saw test faces in 3/4-profile views; in the other orientation condition, the views were reversed. Equal numbers of infants were assigned to the two orientation conditions. Figure 1 displays examples of the stimuli.

All photos were taken with a Canon S50 digital camera and subsequently cropped using Adobe Photoshop to remove the neck and background details. All individual pictures were then mounted on a uniform dark-gray background, and the stimuli were resized to the same dimensions to ensure uniformity. Sixteen independent observers rated a pool of 32 faces for attractiveness and distinctiveness, using a scale from 1 to 10, and the final set of 24 faces was selected so as to match gender, attractiveness, and distinctiveness within each face pair.

Procedure

All infants were tested in a quiet room at the department of psychology at the University of Sheffield. They were seated on their mother's lap, approximately 60 cm from a screen onto which the images were projected. Each infant was randomly assigned to one of the four ethnic-group conditions (African, Asian, Middle Eastern, or Caucasian). Within each of these four conditions, infants were tested with either male or female faces; testing was counterbalanced appropriately, with half the infants assigned to the male-faces condition and half the infants assigned to the female-faces condition. Equal numbers of infants were tested in the male and female conditions. Before the session started, all mothers were instructed to fixate centrally above the screen and to remain as quiet as possible during testing.

Habituation Phase

Each infant was first presented with a single face projected onto a screen measuring 45 cm × 30 cm. The face measured 18 cm × 18 cm (14° visual angle). The experimenter observed the infant’s eye movements on a control monitor from a black-and-white closed-circuit television camera (specialized for low-light conditions) that was positioned above the screen. Time was recorded and displayed on the control monitor using a Horita (Mission Viejo, CA) II TG-50 time coder; video was recorded at 25 frames per second.

The experimenter recorded the infant’s attention to the face by holding down the “z” key on a keyboard whenever the infant fixated on the image. When the infant looked away from the image, the experimenter released the key. If the infant’s attention was averted for more than 2 s, the image disappeared from the screen. The experimenter then presented the image again and repeated the procedure. The habituation phase ended when the infant’s looking time on a presentation was equal to or less than 50% of the average looking time from the infant’s first two presentations. Thus, our measure of looking time was the sum of looking time across all presentations until the habituation criterion was reached.

Test Phase

The test phase consisted of two trials. First, two face images (novel and familiar), each measuring 18 cm × 18 cm (14° visual angle), were presented in a new orientation. The test faces were always in the same orientation, and this orientation differed from the orientation of the face seen during habituation. In one orientation condition, infants were habituated to full-view faces and saw test faces in 3/4-profile views; in the other orientation condition, the views were reversed. Equal numbers of infants were assigned to the two orientation conditions. Figure 1 displays examples of the stimuli.

Fig. 1. Sample stimuli from the Chinese male and Middle Eastern female conditions. The habituation face is shown at the top of each triad. The test faces (novel and familiar) are shown underneath.
angle), were presented on the screen. The images were separated by a 9-cm gap and appeared in the bottom left and bottom right corners of the screen. When the infant first looked at the images, the experimenter pressed a key to begin a 5-s countdown. At the end of the 5 s, the images disappeared from the screen. The faces then appeared with their left/right position on the screen reversed. As soon as the infant looked at the images, another 5-s countdown was initiated. Eye movements were recorded throughout, and the film was digitized for frame-by-frame analysis by two independent observers who used specialized computer software to code looking time to each of the two faces. The observers were blind to both gender and ethnic-group condition and to the screen positions of the faces being viewed by the infants. The average level of interobserver agreement was high (Pearson $r = .93$). Recognition was inferred from a preference for the novel face stimulus across the two 5-s test trials.

RESULTS

Habituation Trials

A preliminary analysis revealed no significant gender differences for stimuli or participants, so data were collapsed across stimulus gender and participant’s gender in subsequent analyses. Habituation time (total looking time across trials) was analyzed in a 3 (age: 3, 6, or 9 months) × 4 (face ethnicity: African, Middle Eastern, Chinese, or Caucasian) × 2 (face orientation: full face or 3/4 view) between-subjects analysis of variance (ANOVA). The ANOVA yielded only a significant effect of age, $F(2, 189) = 73.193, p < .001, \eta^2 = .35$. Post hoc Tukey’s honestly significant difference (HSD) tests revealed that 3-month-olds ($M = 53.19\%$) showed significantly greater preference for the novel face ($p < .003$) than did 9-month-olds ($M = 53.19\%$). There were no main effects of face ethnicity or face orientation.

To investigate novelty preferences within each age group, we conducted one-way-between-groups ANOVAs on the percentage of time spent looking at the novel stimuli in the four face-ethnicity conditions. A significant effect of face ethnicity was found for 9-month-old infants, $F(3, 60) = 3.105, p < .033, \eta^2 = .134$, but not for 3- or 6-month-old infants. These results suggest that novelty preferences differed between face-ethnicity conditions only within the group of 9-month-old infants.

To further investigate novelty preferences within each age group, we conducted a series of two-tailed $t$ tests to determine whether the time spent looking at novel stimuli differed from the

![Table 1](image)

<table>
<thead>
<tr>
<th>Age and face ethnicity</th>
<th>Mean time looking at the novel face (%)</th>
<th>$t(15)$</th>
<th>$p$</th>
<th>$p_{corr}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African</td>
<td>60.88 (16.52)</td>
<td>2.635</td>
<td>.019*</td>
<td>.942</td>
</tr>
<tr>
<td>Middle Eastern</td>
<td>57.31 (11.37)</td>
<td>2.572</td>
<td>.021*</td>
<td>.937</td>
</tr>
<tr>
<td>Chinese</td>
<td>58.72 (14.07)</td>
<td>2.479</td>
<td>.026*</td>
<td>.929</td>
</tr>
<tr>
<td>Caucasian</td>
<td>63.71 (13.47)</td>
<td>4.072</td>
<td>.001*</td>
<td>.988</td>
</tr>
<tr>
<td>6 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African</td>
<td>55.35 (11.40)</td>
<td>1.380</td>
<td>&gt; .05</td>
<td>.840</td>
</tr>
<tr>
<td>Middle Eastern</td>
<td>56.70 (12.89)</td>
<td>2.079</td>
<td>&gt; .05</td>
<td>.371</td>
</tr>
<tr>
<td>Chinese</td>
<td>56.42 (7.79)</td>
<td>3.295</td>
<td>.005*</td>
<td>.965</td>
</tr>
<tr>
<td>Caucasian</td>
<td>58.27 (8.88)</td>
<td>3.725</td>
<td>.002*</td>
<td>.979</td>
</tr>
<tr>
<td>9 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African</td>
<td>51.33 (10.53)</td>
<td>0.505</td>
<td>&gt; .05</td>
<td>.414</td>
</tr>
<tr>
<td>Middle Eastern</td>
<td>53.51 (8.47)</td>
<td>1.658</td>
<td>&gt; .05</td>
<td>.799</td>
</tr>
<tr>
<td>Chinese</td>
<td>48.23 (13.31)</td>
<td>0.530</td>
<td>&gt; .05</td>
<td>.642</td>
</tr>
<tr>
<td>Caucasian</td>
<td>59.70 (11.16)</td>
<td>3.476</td>
<td>.003*</td>
<td>.971</td>
</tr>
</tbody>
</table>

Note. Standard deviations are given in parentheses. Asterisks highlight conditions in which the infants viewed novel faces significantly more often than predicted by chance.
chance level of 50% (see Table 1). The results showed that 3-month-old infants demonstrated significant novelty preferences in all four face-ethnicity conditions, 6-month-old infants demonstrated significant novelty preferences in two of the four conditions (Chinese and Caucasian), and 9-month-old infants demonstrated a novelty preference for Caucasian faces only.

**DISCUSSION**

The aim of the current study was to investigate the onset of the ORE during the first months of life, following up on previous findings that 3-month-olds already show a preference for own-race faces (Bar-Haim et al., 2006; Kelly et al., 2005, 2007). The results reported here do not provide evidence for the ORE (as measured by differential recognition capabilities for own- and other-race faces) in 3-month-old infants, but they do indicate that the ORE emerges at age 6 months and is fully present at age 9 months.

Our results are consistent with the notion of general perceptual narrowing during infancy (e.g., Nelson, 2001). Our findings are also consistent with those of Pascalis et al. (2002), further demonstrating that the face-processing system undergoes a period of refinement within the 1st year of life. Collectively, these findings lend weight to the concept of a tuning period between 6 and 9 months of age. However, differences between the present study and the work by Pascalis et al. should be noted. For example, there is the obvious difference that Pascalis et al. found between-species effects, and our study focused on within-species effects. It should not be assumed that identical mechanisms necessarily underlie the reductions in recognition accuracy observed in the two cases. In addition, once the ability to discriminate between nonhuman primate faces has diminished, it apparently cannot be recovered easily (Dufour, Coleman, Campbell, Petit, & Pascalis, 2004; Pascalis et al., 2002), whereas the ORE is evidently modifiable through exposure to other-race populations (Sangrigoli et al., 2005) or simple training with other-race faces (Elliott, Wills, & Goldstein, 1973; Goldstein & Chance, 1985; Lavrakas, Buri, & Mayzner, 1976).

Furthermore, event-related potential (ERP) studies have shown that in 6-month-olds, the putative infant N170 (a face-selective ERP component elicited in occipital regions) is sensitive to inversion for both human and monkey faces, whereas the N170 recorded in adults is sensitive to inversion only for human faces (de Haan, Pascalis, & Johnson, 2002). An adultlike N170 response is not observed in subjects until they are 12 months of age (Halit, de Haan, & Johnson, 2003). The ERP response for other-race faces has not yet been investigated during infancy, but studies with adults have revealed no differences in the N170 response to own- and other-race faces (Caldara et al., 2003; Caldara, Rossion, Bovet, & Hauert, 2004).

Our findings differ from those reported by Sangrigoli and de Schonen (2004b) in the only other study to have investigated the emergence of the ORE during infancy. In their initial experiment, Sangrigoli and de Schonen found that 3-month-old infants discriminated own-race faces, but not other-race faces, as measured by the VPC task. However, numerous methodological differences between our study and theirs (e.g., color stimuli in our study vs. gray-scale stimuli in theirs) could have contributed to these contrasting results. Furthermore, Sangrigoli and de Schonen were able to eliminate the ORE with only a few trials of exposure to multiple exemplars, which suggests that even if the ORE is already present in 3-month-olds, it is weak and reversible. Between Sangrigoli and de Schonen's work and our own, there are now three VPC experiments (one here, two in Sangrigoli & de Schonen) that have been conducted with 3-month-old infants, yet only one has yielded evidence for the ORE. The weight of the evidence thus suggests that a strong and sustainable ORE may not be present at 3 months of age, but rather develops later.

One might ask whether the ORE arises from differences in the variability of faces from different ethnic groups. However, the available evidence indicates that no category of faces has greater homogeneity than any other (Goldstein, 1979a, 1979b). Moreover, the data suggest that the ORE does not exclusively reflect a deficit for non-Caucasian faces: Individuals from many ethnic groups demonstrate poorer recognition of other-race than own-race faces (Meissner & Brigham, 2001). Evidently, a full account of the ORE will involve factors other than heterogeneity.

We have argued elsewhere (Kelly et al., 2007) that the ORE may develop through the following processes: First, predominant exposure to faces from one's own racial group induces familiarity with and a visual preference for such faces. Second, a preference for faces within one's racial group produces greater visual attention to such faces, even when faces from other racial groups are present in the visual environment. Third, superior recognition abilities develop for faces within one's racial group, but not for faces from groups that are infrequently encountered. Although supporting evidence for the first two processes has been obtained previously (Bar-Haim et al., 2006; Kelly et al., 2005, 2007), the data reported here provide the first direct evidence for the third. According to our account, the ORE can be explained by a modifiable face prototype (Valentine, 1991). If each person's face prototype is an average of all faces that person has encoded during his or her lifetime, then one may assume that it will resemble the race of the faces most commonly encountered. Furthermore, one would expect that individuating dimensions will be optimized for recognition of own-race faces, but not other-race faces.

An alternative to the single-prototype account is that people may possess multiple face-spaces that represent different face categories (e.g., gender, race) separately within a global space. In this contrasting scheme, rather than individuating...
dimensions being unsuitable for recognition of other-race faces, a face-space for other-race faces (e.g., Chinese faces) either does not exist or is insufficiently formed because of a general lack of exposure to those face categories. In both accounts, recognition capabilities improve through exposure to other-race faces. In the case of the single-prototype account, individuating dimensions acquire properties of newly encountered other-race faces that facilitate recognition. Alternatively, in the multiple-face-spaces account, a relevant space for other-race faces develops through similar exposure.

In summary, this is the first study to investigate the emergence of the other-race effect during infancy by comparing three different age groups’ ability to recognize faces from their own race and a range of other races. The data reported here support the idea that very young infants have a broad face-processing system that is capable of processing faces from different ethnic groups. Between 3 and 9 months of age, this system gradually becomes more sensitive to faces from an infant’s own ethnic group as a consequence of greater exposure to such faces than to faces from other racial groups. This shift in sensitivity is reflected in the emergence of a deficit in recognition accuracy for faces from unfamiliar groups. Future research should address whether the pattern of results we obtained with Caucasian infants is universal, or whether the ORE emerges at different ages in other populations.

Acknowledgments—This work was supported by National Institutes of Health Grants HD-46526 and HD-42451 and by an Economic and Social Research Council studentship awarded to David J. Kelly.

REFERENCES


(RECEIVED 1/16/07; REVISION ACCEPTED 4/2/07; FINAL MATERIALS RECEIVED 4/10/07)
This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.