



HAL
open science

REORGANIZATION OF FACE-RACE CATEGORIES IN INFANCY Reorganization in the Representation of Face-Race Categories From 6 to 9 Months of Age: Behavioral and Computational Evidence

Paul C Quinn, Benjamin J Balas, Olivier Pascalis

► To cite this version:

Paul C Quinn, Benjamin J Balas, Olivier Pascalis. REORGANIZATION OF FACE-RACE CATEGORIES IN INFANCY Reorganization in the Representation of Face-Race Categories From 6 to 9 Months of Age: Behavioral and Computational Evidence. *Vision Research*, 2021, 179, pp.34-41. <10.1016/j.visres.2020.11.006>. <hal-03042904>

HAL Id: hal-03042904

<https://hal.science/hal-03042904v1>

Submitted on 7 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Reorganization in the Representation of Face-Race Categories
From 6 to 9 Months of Age: Behavioral and Computational Evidence

Paul C. Quinn¹, Benjamin J. Balas², & Olivier Pascalis³

¹Department of Psychological and Brain Sciences, University of Delaware

²Department of Psychology, North Dakota State University

³Laboratoire de Psychologie et Neurocognition, Université Grenoble Alpes

Author Note

We thank Anna Franklin and an anonymous reviewer for their comments on the initial submission.

Correspondence should be sent to Paul C. Quinn, Department of Psychological and Brain Sciences, 108 Wolf Hall, University of Delaware, Newark, DE 19716, email: pquinn@udel.edu

Abstract

Prior research has reported developmental change in how infants represent categories of other-race faces (Quinn, Lee, Pascalis, & Tanaka, 2016). In particular, Caucasian 6-month-olds were shown to represent African versus Asian face categories, whereas Caucasian 9 month-olds represented different classes of other-race faces together in one category, inclusive of African and Asian faces but exclusive of Caucasian faces. The current investigation sought to provide stronger evidence that is convergent with these findings. In Experiment 1, an experimental group of Caucasian 6-month-olds was familiarized with African (or Asian) faces and then given a novel category preference test with an Asian (or African) face versus a Caucasian face, while a control group of Caucasian 6-month-olds viewed the test faces without a prior familiarization period. Infants in the experimental group divided attention between the test faces and infants in the control group did not manifest a spontaneous preference. Experiment 2 used the same procedure as Experiment 1, but was conducted with Caucasian 9-month-olds. Infants in the experimental group displayed a robust preference for Caucasian faces when considered against the finding that infants in the control group displayed a spontaneous preference for other-race faces. The results provide further evidence that during the period between 6 and 9 months, infants transition to representing own-race versus other-race face categories, with the latter inclusive of multiple other-race face classes with clear perceptual differences. Computational modeling of infant responding suggests that the developmental change is rooted in the statistics of experience with majority versus minority group faces.

Reorganization in the Representation of Face-Race Categories

From 6 to 9 Months of Age: Further Evidence

Laboratory investigations have shown that when presented with multiple instances from a given class, preverbal infants are able to group these instances into a common representation (Quinn, 2011). This ability has been demonstrated in studies that use infant looking time as a dependent measure and rely on visual selectivity, i.e., novelty preference (Fantz, 1964). In a typical procedure, infants are familiarized with exemplars from the same category, and subsequently tested with novel exemplars of the familiarization category and novel members of a novel category. Category formation is inferred if during the test infants generalize looking time responsiveness to novel instances of the familiarization category and display differential responsiveness to novel instances of the novel category. The ability to form category representations in the initial months of life is viewed as adaptive in that it provides infants with a means to organize memory and respond to novel entities from familiar categories as familiar.

While much past research documents infant ability to use perceptual similarity to form category representations for various object classes, newer research has focused on infant ability to form category representations for the three major social categories of faces, namely, race, gender, and age (e.g., Anzures, Quinn, Pascalis, Slater, & Lee, 2010; Damon, Quinn, Heron-Delaney, Lee, & Pascalis, 2016; Quinn, Yahr, Kuhn, Slater, & Pascalis, 2002). This work has been spurred by the belief that asymmetries favoring own-race, female, and adult faces (Rennels & Davis, 2008; Sugden, Mohamed-Ali, & Moulson, 2014) will allow one to better understand the contribution of experiential (i.e., frequency) information to category formation by infants (Quinn, Lee, & Pascalis, 2020).

How infants represent the race category membership of faces is of interest for both theoretical and translational reasons. Theoretically, faces of different races vary in skin tone (Balas, Westerlund, Hung, & Nelson, 2011), physiognomic features (Anzures, Pascalis, Quinn, Slater, & Lee, 2011), and experience, with many infants across the world encountering own- versus other-race faces in greater than a 9 to 1 ratio (Rennels & Davis, 2008; Sugden et al., 2014), thereby allowing investigators to inquire as to which difference or combination of differences contribute to category formation. Translationally, given that racial bias is present in children as young as 3 years of age (Qian et al., 2016), one can ask if evidence of such bias is manifest even earlier in development in the manner in which infants respond to different races of faces. While one would not want to equate differential looking in infancy with attitudinal biases in early childhood, it may be that how attentional mechanisms respond to differential experience early in development sets the stage for subsequent biases (Kinzler & Spelke, 2011; Quinn, Lee, & Pascalis, 2018, 2019; Rhodes, 2020), a point that we return to in the General Discussion.

Nevertheless, to our knowledge, there are only two studies in the literature investigating whether infants form distinct category representations for own- versus other-race faces and for different classes of other-race faces (Anzures et al., 2010; Quinn, Lee, Pascalis, & Tanaka, 2016). In the first of these studies, Anzures et al. (2010) reported that Caucasian 9-month-olds formed distinct categories of Asian and Caucasian faces, although the basis for such a category distinction was unclear. The infants could have relied on perceptual differences between the categories or on differential experience with own- versus other-race faces.

Quinn et al. (2016) therefore asked how Caucasian 6- and 9-month-olds represent different classes of other-race faces. Both groups of infants were familiarized with African or Asian faces, and were then tested on novel African versus novel Asian faces. Whereas Caucasian

6-month-olds preferred the novel category, indicating that they had represented African and Asian faces as distinct categories, Caucasian 9-month-olds distributed their attention evenly between the familiar and novel category. However, in an additional experiment in the same investigation, Caucasian 9-month-olds familiarized with African or Asian faces and tested with novel exemplars from the familiar category and novel exemplars from the novel Caucasian category, preferred the Caucasian faces. This combined results that the Caucasian 9-month-olds indicate that they had formed a broader other-race category inclusive of both African and Asian faces, but exclusive of Caucasian faces. Moreover, the findings from both age groups in the Anzures et al. (2010) and Quinn et al. (2016) studies suggest that whereas 6-month-olds form face-race categories on a perceptual basis (i.e., contrasting different classes of other-race faces), the representation of face race at 9 months centers on the more social distinction between own- and other-race faces, in effect, de-emphasizing the perceptual differences between different classes of other-race faces.

The present study aimed to provide stronger evidence for the conclusion that 6-month-olds represent the perceptual differences between face race categories, whereas 9-month-olds represent the contrast of own-race versus other-race face categories with the latter inclusive of multiple other-race faces classes. One limitation of the Quinn et al. (2016) study in terms of supporting the conclusion that 9-month-olds represent different classes of other-race faces in a broad “other” category is that the procedure did not require the Caucasian infants to generalize looking time responsiveness from African to Asian faces or vice versa. The present investigation therefore used a procedure in which an experimental group of Caucasian infants was familiarized with African (or Asian) faces and then tested with a novel Asian (or African) face versus a novel Caucasian face. The rationale is that if the infants represent African, Asian, and Caucasian faces

as distinct categories, then both the face from the novel other-race class and the Caucasian face should be perceived as novel and neither should be preferred. However, if the infants represent African and Asian faces together as other-race faces and separately from own-race Caucasian faces, then infants should generalize from the familiarization other-race class to the novel other-race class, and the Caucasian face should be preferred. Experiment 1 was conducted with 6-month-olds, and Experiment 2 was performed with 9-month-olds.

One additional consideration concerns the baseline or spontaneous preferences for own- versus other-race faces by infants. Two studies have now reported that infants transition from an own-race preference at 3 months, to a null preference at 6 months, followed by an other-race preference at 9 months (Fassbender, Teubert, & Lohaus, 2016; Liu et al., 2015). Because of this consideration, in Experiments 1 and 2 of the current study, baseline preferences for the own- versus other-race test faces were assessed in 6- and 9-month-old control groups, respectively. The measurement of such preferences allows one to determine if any preference observed in the experimental group was a response to the category information presented during familiarization rather than a reflection of a pre-existing preference.

Finally, to probe whether infant responding in the two experiments is related to the statistics of face experience for infants encountering faces from majority and minority categories, we performed computational modeling. Since the mid 1990s, computational modeling has been used as a technique to understand (in principle) the representations infants may form based on learning from experience (Elman et al., 1996), and has been used in particular in the study of infant categorization (Mareschal, French, & Quinn, 2000), inclusive of investigations seeking to explain how infant categorization performance is affected by asymmetrical real-world experience (Mermillod, French, Quinn, & Mareschal, 2003). The modeling in the current paper follows from

this tradition and relied on a principal components analysis (PCA) and training with majority and minority group faces in a ratio corresponding to what infants are experiencing in their demographic location, and used a task design comparable to that presented to the infants in Experiments 1 and 2.

Experiment 1

In Experiment 1, an experimental group of Caucasian 6-month-olds was familiarized with faces from one other-race category and then presented with a novel exemplar from a novel other-race class and a novel exemplar from the own-race (Caucasian) class. For half of the infants, the familiarization category was African and the novel other-race class was Asian, and for the other half of the infants, the familiarization category was Asian and the novel other-race class was African. In addition, to discount a possible role for spontaneous preference, same-aged Caucasian infants in a control group were presented with the same faces as the infants in the experimental group, but without first being presented with a class of other-races during a familiarization period.

Method

Participants. Participants were 64 6- to 7-month-old infants (28 females), mean age = 190.53 days, $SD = 16.24$ days. Four additional infants were tested, but 3 did not complete the procedure due to fussiness, and 1 was excluded from the data analysis because of failure to compare the test stimuli. They were all Caucasian (as reported by their parents), and from predominantly middle-class backgrounds in Newark, DE, where the local racial demographics include 87% Caucasian, 6% African, and 4% Asian.

Stimuli. The stimuli consisted of photographs of African, Asian, and Caucasian female faces that were taken from the NimStim face set (Tottenham et al., 2009). There were 4 faces for

each of the categories, all depicting a neutral expression. The particular stimulus faces used from the Nimstim set were: 11, 12, 13, and 14 (African), 15, 16, 17, and 18 (Asian), and 7, 8, 9, and 10 (Caucasian). These were the same faces used by Quinn et al. (2016). Following the procedure in the earlier paper, each photograph was cropped lightly (so as to eliminate hairstyle differences as a category cue), centered, and pasted onto a white 17.7 x 17.7 cm poster board for presentation. As was reported in Quinn et al. (2016), there were no differences between the three categories in the mean height and width of the faces, and comparison of low-level properties (i.e., luminance and contrast) of the images across the categories revealed no difference between the Caucasian faces and both classes of other-race faces that was greater than the difference between the two classes of other-race faces. Finally, as noted in the prior paper, analysis of salience (Walther & Koch, 2006) indicated that there was no differential salience of the various regions of the face images across the African, Asian, and Caucasian categories.

Apparatus. Infants were tested in a visual preference apparatus, modeled after that of Fagan (1970). The apparatus has a gray display panel that includes two compartments to hold the stimuli. The stimuli were illuminated by a fluorescent lamp that was shielded from the infant's view. Center-to-center distance between compartments was 30.5 cm. There was a 0.62 cm peephole located midway between the compartments that permitted an observer to record infant visual fixations. A second peephole, 0.90 cm in diameter, located directly below the first peephole, permitted a Pro Video CVC-120PH pinhole camera and Magnavox DVD recorder to record infant gaze duration.

Procedure. The University of Delaware Institutional Review Board approved all study procedures (study title: Development of Face Processing Expertise, protocol number: 151954-

11). Infants were tested individually and brought to the laboratory by a parent. The parent provided informed consent for their infant.

Each infant was first seated in a reclining position on the parent's lap. An experimenter then moved the apparatus into position over the infant, keeping the infant's head centered with respect to the midline of the display panel. To begin a trial, the experimenter would load the stimuli into the compartments on the display panel and close the panel, thereby exposing the stimuli to the infant. When the panel was closed, it was positioned approximately 30.5 cm above and in front of the infant, and the infant could see only the two stimuli and the gray surround of the viewing chamber. During a trial, the experimenter observed the infant through the upper peephole and recorded fixations to the left and right stimuli with a 605 XE Accusplit stopwatch (Pleasanton, CA, USA) held in each hand. The criterion for fixation was observing corneal reflection of the stimulus over the infant's pupil. Between trials, the experimenter opened the panel, recorded the looking time data, changed the stimuli, and reclosed the panel. For the experimental groups, two experimenters were used to record fixations, one during familiarization and another during test trials. Both experimenters were naive to the hypotheses under investigation. The test for the infants in the control condition required only a single experimenter (who was again naive to the hypotheses being investigated) to record fixations. The experimenter recording during the two preference test trials of the experimental and control groups was kept away from the experimental room until the point at which those trials were to begin (i.e., the test stimuli were positioned in the apparatus, and the infant was ready for them to be presented). This experimenter was thus also naive to the testing condition (experimental vs. control) of a particular infant.

Infants from each of the two age groups were randomly divided into an experimental group and a control group with 32 infants in each. Half of the infants in the experimental group were familiarized with three African faces and the other half with three Asian faces. The three faces were randomly chosen for each infant. During familiarization, following Quinn et al. (2016), the three faces were presented twice over the course of six 15 s familiarization trials. Infants were presented with one face (i.e., two identical copies of that face in the left and right stimulus compartments) on the first trial, another face on the second trial, and the third face on the third trial, with the sequence repeated on trials 4 through 6. The order of presentation of the three faces was randomly determined for each infant.

Two 10 s test trials immediately followed the familiarization trials and paired a novel face from the novel other-race class (i.e., an Asian face for infants familiarized with African faces or an African face for infants familiarized with Asian faces) with a novel own-race Caucasian face. The stimulus from the novel other-race class was randomly selected for each infant. The own-race Caucasian faces were also randomly selected and assigned to one infant who had seen African faces and one infant who had seen Asian faces during familiarization. The Caucasian members of each pairing were identical for both groups of infants. Left-right positioning of the novel face from the novel other-race category was counterbalanced across infants on the first test trial and reversed on the second test trial.

The control group was not presented with a familiarization stimulus; the infants from this group only took part in the two preference test trials that were administered in the same manner as described for the experimental group. For each test pairing presented to an infant in the experimental group, there was a corresponding pairing presented to an infant in the control group. Inter-observer agreement, as determined by comparing looking times measured by the

experimenter using the center peephole, and an additional naive observer measuring looking times offline from DVD records, was calculated for the test trial preference scores of 16 infants (25% of the sample), 8 from each group (experimental vs. control). Average level of agreement was 98.36% ($SD = 1.53$). In addition, the average of the difference scores for the original and reliability-check preference scores was -0.10 , $SD = 1.21$, which was not significantly different from 0, $t(15) = 0.21$, $p = .836$, two-tailed, indicating that the differences between observers were randomly distributed, rather than having one observer with consistently higher or lower scores relative to the other observer.

Results and Discussion

Familiarization Trials. Individual looking times were summed over the left and right copies of the stimulus presented on each trial to the experimental group and then averaged across the first three and last three trials. Mean looking times are shown in Table 1. An analysis of variance (ANOVA), Face Category (African vs. Asian) x Trials (1-3 vs. 4-6), performed on the individual scores revealed only a significant Trials effect, $F(1, 30) = 4.96$, $p = .034$, partial $\eta^2 = .14$. Neither the effect of Face Category, $F(1, 30) = .02$, $p = .888$, nor the interaction of Face Category x Trials, $F(1, 30) = 2.86$, $p = .101$, were significant. Using the standard operational definition of habituation as a decline in responsiveness with repeated stimulation (Cohen & Gelber, 1975), the decrement in looking time from the first to the second half of familiarization for infants presented either with African or Asian faces indicates that both groups habituated to the stimuli.

Preference Test Trials. Each infant's looking time to the Caucasian test stimulus was divided by the looking time to both test stimuli and converted to a percentage score. Mean preference scores for the Caucasian faces in the experimental and control groups are shown in

Table 1. An ANOVA performed on the individual preference scores, with factors of Group (experimental vs. control) and Test Comparison (Asian-Caucasian vs. African-Caucasian) did not yield any significant effects, $F(1, 60) < 0.92, p > .34$. The null results take on greater meaning from the comparisons of the mean preferences to the chance value of 50%, which revealed that neither of the test comparisons in either the experimental group or control group yielded evidence of reliably longer looking to the Caucasian test stimuli (Table 1). The pattern of performance of individual infants relative to chance paralleled those observed at the group level: only 17 (8 for the Asian-Caucasian test and 9 for the African-Caucasian test) of 32 infants in the experimental group displayed individual preference scores greater than 50% (binomial probability = .43), and just 16 (7 for the Asian-Caucasian test and 9 for the African-Caucasian test) of 32 infants in the control group displayed individual preference scores greater than 50% (binomial probability = .57). The pattern of both the mean and individual preference scores suggests that the infants in the experimental group did not have a preference for the novel Caucasian faces (whether they were paired with novel Asian faces for infants familiarized with African faces or novel African faces for infants familiarized with Asian faces). Likewise, the infants in the control group did not spontaneously prefer Caucasian faces to either Asian or African faces.

Although the results of the experimental group were null, they take on additional significance when considered against the positive results reported in Experiment 1 of Quinn et al. (2016) where the same procedure was used with the same age group, and the infants were familiarized with African or Asian faces and tested Asian versus African faces. There the result was that the infants preferred the novel category faces, indicating that the infants formed a category for African faces that excluded Asian faces and a category for Asian faces that excluded

African faces. Notably, when the combined mean novel category preference score from Experiment 1 of Quinn et al. (2016) ($M = 58.41$, $SD = 13.59$, $N = 32$) was compared with the combined mean novel category preference for the Caucasian faces from experimental group of the current experiment ($M = 48.88$, $SD = 14.82$, $N = 32$), the difference was reliable, $t(62) = 2.68$, $p = .009$, $d = 0.67$. Thus, in the current experiment, when infants were familiarized with faces from an other-race category and tested with novel faces from a novel other-race category and novel faces from the own-race category, it can be argued that the novelty of the novel face from the novel other-race category would have been perceived. Likewise, in Anzures et al. (2010), same-aged infants familiarized with other-race faces (i.e., Asian) subsequently preferred novel Caucasian faces over novel Asian faces. This finding implies that infants in the present experiment would have also perceived the novelty of the novel own-race Caucasian face. Both of the prior results suggest that infants in the current experiment showed a null preference at test because they perceived both the novelty of the novel other-race category and the novelty of the novel own-race category. Taken together, the findings from all three studies are consistent with an interpretation in which 6-month-old Caucasian infants form distinct categories for African, Asian, and Caucasian faces. Finally, it can be noted that the null spontaneous preference between own-race Caucasian faces and other-race (African or Asian) faces observed in the control group is consistent with the null preferences observed for 6-month-olds tested on own-other race face contrasts in prior papers (Fassbender et al., 2016; Liu et al., 2015).

Experiment 2

Experiment 2 was a replication of Experiment 1, but conducted with 9-month-olds.

Method

Participants. Participants were 64 6- to 7-month-old Caucasian infants (28 females),

mean age = 282.45 days, $SD = 15.43$ days. Two additional infants were tested but did not complete the procedure due to fussiness.

Stimuli. Stimuli were the same as those used in Experiment 1.

Procedure. Procedure was the same as in Experiment 1, except that 9-month-olds were tested. Inter-observer agreement was determined with the same procedure described for Experiment 1 and calculated for the test trial preference scores of 16 infants (25% of the sample), 8 from each group (experimental vs. control). Average level of agreement was 98.65% ($SD = 1.49$). In addition, the average of the difference scores for the original and reliability-check preference scores was .19, $SD = 1.24$, which was not significantly different from 0, $t(15) = 0.61$, $p = .551$, two-tailed, indicating that the differences between observers were again randomly distributed.

Results and Discussion

Familiarization Trials. Mean looking times are shown in Table 2. An ANOVA, Face Category (African vs. Asian) x Trials (1-3 vs. 4-6), performed on the individual scores revealed only a significant Trials effect, $F(1, 30) = 9.22$, $p = .005$, partial $\eta^2 = .24$. Neither the effect of Face Category, $F(1, 30) = 1.63$, $p = .211$, nor the interaction of Face Category x Trials, $F(1, 30) = 0.82$, $p = .372$, were significant. The decrement in looking time from the first to the second half of familiarization suggests that the 9-month-olds in Experiment 2, like the 6-month-olds in Experiment 1, habituated to the African and Asian faces.

Preference Test Trials. Mean preference scores for the Caucasian faces in the experimental and control groups are shown in Table 2. An ANOVA performed on the individual preference scores, with factors of Group (experimental vs. control) and Test Comparison (Asian-Caucasian vs. African-Caucasian), yielded only an effect of Group, $F(1, 60) = 21.22$, $p < .0001$,

partial $\eta^2 = .26$, indicating that the experimental group infants displayed higher preference scores than the control group infants. In addition, comparisons of the mean preferences to the chance value of 50% revealed that the experimental and control groups of infants had reliable preferences in opposite directions. That is, while the experimental groups looked longer to the novel own-race category (i.e., Caucasian faces), the control groups looked reliably longer to the other-race African and Asian faces. The pattern of performance of individual infants relative to chance resembled that observed at the group level: whereas 23 (11 for the Asian-Caucasian test and 12 for the African-Caucasian test) of 32 infants in the experimental group displayed individual preference scores greater than 50% (binomial probability = .01), only 7 (4 for the Asian-Caucasian test and 3 for the African-Caucasian test) of 32 infants in the control group displayed individual preference scores greater than 50% (binomial probability = .001).

Both the mean and individual preferences converge to suggest that whereas infants in the control group had a spontaneous preference for looking at other-race African or Asian faces, infants in the experimental group were able to overcome the spontaneous preference for other-race faces to display a novel category preference for own-race Caucasian faces. The findings take on additional significance when considered alongside those of Quinn et al. (2016). In the earlier study, 9-month-old infants familiarized with either African or Asian faces and then tested with novel African versus novel Asian faces displayed a null preference between the categories. However, 9-month-old infants familiarized with either African or Asian faces and tested with novel instances from the familiarization category and novel Caucasian faces, generalized looking time responsiveness to novel instances of the familiarized category and preferred the novel Caucasian faces. The data from both studies taken together indicate that 9-month-old Caucasian infants form a category for African faces that includes not only novel African faces, but also

novel Asian faces, although it excludes novel Caucasian faces. Likewise, 9-month-old Caucasian infants form a category for Asian faces that includes not only novel Asian faces, but also novel African faces, although it excludes novel Caucasian faces. The present findings thus provide convergent evidence that 9-month-old infants form categories for own-race faces versus other-race faces, with the latter category inclusive of multiple other-race faces with clear perceptual distinctions. In addition, the current control group results showing a spontaneous preference for other-race faces at 9 months of age are consistent with prior reports of other-race preference in the same age group (Fassbender et al., 2016; Liu et al., 2015).

Computational Modeling

The present results, in conjunction with those reported in Quinn et al. (2016), demonstrate that there are important changes in the tuning of infants to face categories defined by race during the first year of life. Specifically, these results are consistent with the suggestion that accumulated experience with faces that is biased towards a majority category leads to changing abilities to categorize faces belonging to majority versus minority groups as defined by experience. To examine the extent to which the behavioral results of the infants are indeed a consequence of the statistics of face experience, we implemented a simple “face space” model based on PCA to determine how increased experience with faces might affect responding to categories.

We implemented our model using a set of faces sampled from the FERET face database (Phillips, Weschler, Huang, & Rauss, 1998) that depicted African, Asian, and Caucasian individuals ($N = 140$ faces). All of the images were grayscale and cropped with a uniform outline that eliminated both the hairline and jawline of each individual. To approximate the biased experience of infants raised in a majority Caucasian environment, each iteration of the model

was trained with Caucasian, African, and Asian faces in a ratio of 18:1:1, which is commensurate with real-world estimates of the exposure of infants to own- and other-race faces (Rennels & Davis, 2008; Sugden et al., 2014) and also with the racial demographics of the locale in which the infants were studied. To simulate increased face experience during the first year of life (i.e., between 6 and 9 months of age or the difference in the age of the infants studied in Experiment 1 vs. Experiment 2), we trained the model with either 80 total faces or 320 total faces, sampled randomly at each iteration from the full set of images.

In each iteration of the model, we used the training images to carry out PCA, determining a low-dimensional ($D = 12$) embedding of the images in a “face space” (Valentine, 1991). Given that Caucasian faces make up the majority of images used to recover the principal components of this space, we expected that the variability across African and Asian faces would be captured less effectively, and that both categories of faces might occupy a similar region of the recovered space (Caldara & Abdi, 2006). To test how their positioning in this face space might affect performance in the categorization task used to test infants, we simulated task performance 64 times per model iteration. Specifically, consistent with the experimental task presented to the infants, on each simulated trial, we randomly sampled either three African or three Asian faces to serve as the familiarization images. Next, a Caucasian face and an Asian or African face were sampled to serve as the test images. With this procedure, the question becomes: Which of the test images is the most novel relative to the familiarization images? We operationalized the answer to this question by calculating the distance between each test image and the centroid of the three familiarization images. Whichever distance was larger was deemed the most novel and trials were scored as either correct or incorrect according to whether or not the Caucasian face was selected.

For each level of face experience (80 faces vs. 320 faces), we carried out 500 iterations of the model. Each iteration was characterized by calculating the proportion of trials that were correct, collapsed across familiarization categories. Overall model performance was estimated by calculating a bootstrapped confidence interval of median accuracy using Matlab's `bootci.m` function. When the model was trained with 80 faces, this 95% confidence interval spanned the values [50% - 53.1%], while training with 320 faces led to a 95% confidence interval spanning the values [53.2% - 55.9%]. This difference is qualitatively consistent with the difference in infant performance observed in Experiment 1 (i.e., null preference between Caucasian faces and Asian [or African] faces after familiarization with African or Asian faces in Caucasian 6-month-olds) versus Experiment 2 (i.e., reliable preference for Caucasian faces over Asian [or African faces] after familiarization with African or Asian faces in Caucasian 9-month-olds), suggesting that infant performance in this task may be a reflection of biased experience with majority versus minority categories and the statistical consequences of that biased experience on low-dimensional representations of faces.

Does the same model also account for the results reported in Quinn et al. (2016)? To examine this question, we used the same model parameters to simulate categorization performance when Caucasian infants were familiarized to the category of African or Asian faces and subsequently tested with a novel face from the familiarization category alongside a novel face from the novel other-race category. As in our first set of simulations, trials were scored as correct if the distance between the novel face from the novel category was further from the centroid of the familiarized faces than the novel face from the familiar category. When trained with 80 faces per model iteration, the 95% confidence interval for accuracy spanned [71.9% - 76.6%]. When trained with 320 faces, the 95% confidence interval spanned [67.2% - 68.9%].

While both of these intervals indicate above-chance performance, they more importantly indicate a decrease in categorization performance with more face experience, which was same result observed with the infants in Quinn et al. (2016). Overall then, this basic face space model of categorization performance reflects some of the key properties of infant behavior in both the current task and in Quinn et al. (2016).

The modeling results are informative in highlighting how the statistics of experience can lead to representations of face appearance that affect a range of recognition behaviors. There are several important caveats that we should acknowledge regarding the model's limitations, however. First, we note that there are several free parameters in the model that are not easily justified given what is known about infant face perception. For example, our model relies on using a set of input faces to derive a dimensional model of facial appearance that is low-dimensional. That is, rather than using hundreds of pixel values per face to describe our images, we use the correlational structure of the input images to come up with a set of just 12 numbers to describe each image. The dimensionality of infants' face space is not well understood, however – there is little experimental data from infant observers to justify using 12 coefficients, or any particular number as the dimensionality of this face space. Thus, the number of components that should be included in low-dimensional face spaces like these is an open question. Second, the precise number of faces used to obtain the principal components that define a face space is also not easy to determine from the literature, and this is another free parameter in our model. We have examined how the model's behavior changes as we provide it with varying numbers of input images, and ideally these numbers would reflect real-world infant face experience. Finally, the actual architecture of the model (distance measurements in a PCA-based space) is similar to many other computational investigations of face recognition (Burton, Kramer, Ritchie, &

Jenkins, 2016; Hancock, Burton, & Bruce, 1996; O'Toole, Deffenbacher, Valentin, & Abdi, 1994), but is certainly not a complete model of infant face recognition abilities. Overall, then, we offer this computational analysis as a way of examining change in some basic face categories as a function of the statistics of experience, which we believe provides insight into the patterns of infant behavior reported here and in Quinn et al. (2016).

General Discussion

The present results provide further evidence that the category representations infants form for face race undergo reorganization during the period from 6 to 9 months of age. In Experiment 1, Caucasian 6-month-old infants familiarized with African or Asian faces subsequently divided their attention at test when presented with novel instances of the novel, other-race class paired with novel Caucasian faces. Although a null result, it takes on greater significance when considered alongside the findings that same aged Caucasian infants (1) familiarized with African or Asian faces and tested with novel African faces versus novel Asian faces displayed novel category preferences (Quinn et al., 2016), and (2) familiarized with Asian faces showed a subsequent novel category preference for Caucasian faces (Anzures et al., 2010). Collectively, these findings support the contention that Caucasian infants at 6 months of age represent the category distinctions between African, Asian, and Caucasian faces. Given the perceptual differences between the different classes of faces, inclusive of skin tone and facial physiognomy (e.g., Farkas, Katic, & Forrest, 2007; Le, Farkas, Ngim, Levin, & Forrest, 2002), the positive evidence for category distinctions between them should not be surprising.

In Experiment 2, by contrast, Caucasian 9-month-old infants familiarized with African or Asian faces subsequently preferred novel Caucasian faces over novel instances of the novel, other-race class. This preference was observed in spite of Caucasian 9-month-old infants in the control group showing preferences for African and Asian faces over Caucasian faces. The pattern of outcomes can be considered alongside the findings from Quinn et al. (2016) where same-aged Caucasian infants familiarized with African or Asian faces subsequently divided their attention at test between novel African faces paired with novel Asian faces, but showed a novel category preference for Caucasian faces over novel instances of the familiarization category. The set of findings as a whole support the view that the primary category distinction for face race at 9 months of age is that between own-race faces and other-race faces, with the other-race face category inclusive of multiple classes of other-race faces with clear perceptual differences.

Face processing includes not only category formation, but also attention and individuation (Pascalis, Fort, & Quinn, 2020). Each of these other processes has been linked to the differential experience that infants in many parts of the world have with majority versus minority group faces. For example, while African 3-month-olds exposed primarily to African faces look more to African than Caucasian faces, and Caucasian 3-month-olds exposed primarily to Caucasian faces look more to Caucasian faces, African 3-month-olds exposed to both African and Caucasian infants divide their attention between the two categories of faces (Bar-Haim, Ziv, Lamy, & Hodes, 2006). In the case of distinguishing between faces, Caucasian infants exposed primarily to Caucasian faces show increased difficulty distinguishing between other-race faces (e.g., Asian faces) while maintaining differentiation of Caucasian faces (Kelly et al., 2007) during the period between 3 and 9 months of age. However, this perceptual narrowing for Asian

faces can be prevented or undone if Caucasian infants are provided with picture book or video experience with Asian faces (Anzures et al., 2012; Heron-Delaney et al., 2011).

Given the evidence favoring an experiential account of the developmental changes in attention toward and individuation of own- and other-race faces, we used computational modeling to determine if the behavioral data of the infants in the present study could be simulated based on the ratio of majority and minority group faces estimated to be experienced by the infants in their everyday environments. In particular, PCA was performed on a model trained with African, Asian, and Caucasian faces in an 18:1:1 ratio. A “less experienced” model designed to simulate 6-month-old performance was presented with 80 faces, and a “more experienced” model designed to simulate 9-month-old performance was presented with 320 faces. The “more experienced” model showed stronger categorization performance than the “less experienced” model when the task presented to the model matched the one presented to the infants (i.e., familiarization with African or Asian faces and test with novel faces from the novel other-race class and novel Caucasian faces). Moreover, the “more experienced” model showed weaker categorization performance than the “less experienced” model when the task presented to the model matched the one presented to the infants in Quinn et al. (2016) (i.e., familiarization with African or Asian faces and test with novel faces from each category). Remarkably, both of the simulation results corresponded with the direction of developmental change in infant responding to the categories.

The computational results provide support for the view that the reorganization of face race categories between 6 and 9 months of age is rooted in the statistics of experience with majority and minority group categories, and the broader view that infants are statistical learners (Saffran & Kirkham, 2018). In the case of face race, the 9:1 or greater ratio of own- to other-race

experience may lead infants not only to tune out distinctions between individuals within other-race categories (Kelly et al., 2007), but also to tune out distinctions between other-race categories. In this sense, the perceptual narrowing that has been traditionally evidenced by within-category discrimination data may extend to between-category distinctions.

Consistent with theme of this special issue on “Calibrating the Visual System”, the behavioral and computational findings reported in the present paper provide an illustrative example of the role of experience and plasticity in visual development, in particular, highlighting how infants adapt to their native social group. If infants use frequency information as a proxy for determining perceived relevance or importance, it could be that any face that does not correspond with the majority group category is coded as “other”. Given that low frequency of experience is a source of outgroup category formation (Bigler & Liben, 2007), this manner of coding is consistent with the implicit bias that has been observed in children as young as 3 years of age (Qian et al., 2016) and even infants between 7 and 9 months of age in terms of their association of positive and negative valence with own- and other-race categories, as well as selective learning from own- versus other-race category members (Xiao et al., 2018a, 2018b). The overall body of evidence is consistent with the proposal that frequency information may come to over-ride similarity information, and possibly bring about an early transition from perceptual to more social categories surprisingly early in development (Quinn et al., 2020).

It would be of interest in future computational and behavioral work to determine if there is a “tipping point” in which sufficient experience with less frequently encountered categories does not yield the category reorganization observed in the present paper. In addition, while our approach to explaining the behavioral data of the infants has been based on estimates of the overall ratio of majority versus minority group face members, we acknowledge that the face

environments of infants are likely dominated by a few exemplars from the majority group category, e.g., caregiver faces (Jayaraman, Fausey, & Smith, 2015; Quinn et al., 2002; Rennels & Davis, 2008). It is also the case that computational modeling using a Linear Discriminant Analysis suggests that social categories can emerge from multiple encounters with a small number of people (Kramer, Young, Day, & Burton, 2017). It therefore remains an open question whether computational modeling efforts that incorporate multiple repetitions of such faces may provide an even stronger simulation of the infant looking time data. However this latter question is resolved, the present experimental and computational results are consistent with the view that differential experience with majority versus minority group faces drives infants between 6 and 9 months of age from a representation of face race based on perceptual differences (i.e., African vs. Asian vs. Caucasian) to a representation that contrasts more social groupings (i.e., own vs. other) based on proportional group sizes.

References

- Anzures, G., Pascalis, O., Quinn, P. C., Slater, A. M., & Lee, K. (2011). Minimizing skin color differences does not eliminate the own-race recognition advantage in infants. *Infancy, 16*, 640-654.
- Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., & Lee, K. (2010). Categorization, categorical perception, and asymmetry in infants' representation of face race. *Developmental Science, 13*, 553-564.
- Anzures, G., Wheeler, A., Quinn, P. C., Pascalis, O., Slater, A. M., Heron-Delaney, M., Tanaka, J. W., & Lee, K. (2012). Brief daily exposures to Asian females reverses perceptual narrowing for Asian faces in Caucasian infants. *Journal of Experimental Child Psychology, 112*, 484-495.
- Balas, B., Westerlund, A., Hung, K., & Nelson, C. A. III. (2011). Shape, color and the other-race effect in the infant brain. *Developmental Science, 14*, 892-900.

- Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. *Psychological Science, 17*, 159-163.
- Bigler, R., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's stereotyping and prejudice. *Current Directions in Psychological Science, 16*, 162-166.
- Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. L. (2016). Identity from variation: Representations of faces derived from multiple instances. *Cognitive Science, 40*, 202-223.
- Caldara, R., & Abdi, H. (2006). Simulating the other-race effect with autoassociative neural networks: Further evidence in favor of the face-space model. *Perception, 35*, 659-670.
- Cohen, L. B., & Gelber, E. R. (1975). Infant visual memory. In L. B. Cohen & P. Salapatek (Eds.), *Infant perception: From sensation to cognition* (Vol. 1, pp. 347-403). New York, NY: Academic Press.
- Damon, F., Quinn, P. C., Heron-Delaney, M., Lee, K., & Pascalis, O. (2016). Development of category formation for faces differing by age in 9- to 12-month-olds: An effect of experience with infant faces. *British Journal of Developmental Psychology, 34*, 582-597.
- Elman, J. L., Bates, E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness a connectionist perspective on development*. Cambridge, MA: MIT Press.
- Fagan, J. F. (1970). Memory in the infant. *Journal of Experimental Child Psychology, 9*, 217-226.
- Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science, 164*, 668-670.

- Farkas, L. G., Katic, M. J., & Forrest, C. R. (2007). Comparison of craniofacial measurements of young adult African-American and North American white males and females. *Annals of Plastic Surgery*, *59*, 692-698.
- Fassbender, I., Teubert, M., & Lohaus, A. (2016). The development of preferences for own-race versus other-race faces in 3-, 6- and 9-month-old infants. *European Journal of Developmental Psychology*, *13*, 152-165.
- Hancock, P. J. B., Burton, A. M., & Bruce, V. (1996). Face processing: Human perception and principal components analysis. *Memory and Cognition*, *24*, 21-40.
- Heron-Delaney, M., Anzures, G., Herbert, J. S., Quinn, P. C., Slater, A. M., Tanaka, J. W., Lee, K., & Pascalis, O. (2011). Perceptual training prevents the emergence of the other race effect during infancy. *PLoS ONE*, *6*(5): e19858.
- Jayaraman, S., Fausey, C. M., & Smith, L. B. (2015). The faces in infant-perspective scenes change over the first year of life. *PLoS ONE*, *10*(5): e0123780.
- Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science*, *18*, 1084-1089.
- Kinzler, K. D., & Spelke, E. S. (2011). Do infants show social preferences for people differing in race? *Cognition*, *119*, 1-9.
- Kramer, R. S. S., Young, A. W., Day, M. G., & Burton, A. M. (2017). Robust social categorization emerges from learning the identities of very few faces. *Psychological Review*, *124*, 115-119.

- Le, T. T., Farkas, L. G., Ngim, R. C. K., Levin, L. S., & Forrest, C. R. (2002). Proportionality in Asian and North American Caucasian faces using neoclassical facial canons as criteria. *Aesthetic Plastic Surgery, 26*, 64-69.
- Liu, S., Xiao, W. S., Xiao, N. G., Quinn, P. C., Zhang, Y., Chen, H., Ge, L., Pascalis, O., & Lee, K. (2015). Development of visual preference for own- versus other-race faces in infancy. *Developmental Psychology, 51*, 500-511.
- Mareschal, D., French, R. M., & Quinn, P. C. (2000). A connectionist account of asymmetric category learning in early infancy. *Developmental Psychology, 36*, 635-645.
- Mermillod, M., French, R. M., Quinn, P. C., & Mareschal, D. (2003). The importance of long-term memory in infant perceptual categorization. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th Annual Conference of the Cognitive Science Society* (pp. 804-809). Mahwah, NJ: Erlbaum.
- O'Toole, A. J., Deffenbacher, K. A., Valentin, D., & Abdi, H. (1994). Structural aspects of face recognition and the other-race effect. *Memory & Cognition, 22*, 208-224.
- Pascalis, O., Fort, M., & Quinn, P. C. (2020). Development of face processing: Are there critical or sensitive periods? *Current Opinion in Behavioral Sciences, 36*, 7-12.
- Phillips, J. P., Weschler, H., Huang, J. & Rauss, P. J. (1998). The FERET database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing, 16*, 295-306.
- Qian, M. K., Heyman, G. D., Quinn, P. C., Messi, F. A., Fu, G., & Lee, K. (2016). Implicit racial biases in preschool children and adults from Asia and Africa. *Child Development, 87*, 285-296.

- Quinn, P. C. (2011). Born to categorize. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development*, 2nd ed. (pp. 129-152). Oxford, UK: Wiley-Blackwell.
- Quinn, P. C., Lee, K., & Pascalis, O. (2018). Perception of face race by infants: Five developmental changes. *Child Development Perspectives*, 12, 204-209.
- Quinn, P. C., Lee, K., & Pascalis, O. (2019). Face processing in infancy and beyond: The case of social categories. *Annual Review of Psychology*, 70, 165-189.
- Quinn, P. C., Lee, K., & Pascalis, O. (2020). Beyond perceptual development: Infant responding to social categories. In J. B. Benson (Ed.), *Advances in Child Development and Behavior* (Vol. 58, pp. 35-61). Cambridge, MA: Academic Press (Elsevier).
- Quinn, P. C., Lee, K., Pascalis, O., & Tanaka, J. W. (2016). Narrowing in categorical responding to other-race face classes by infants. *Developmental Science*, 19, 362-371.
- Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception*, 31, 1109-1121.
- Rennels, J. L., & Davis, R. E. (2008). Facial experience during the first year. *Infant Behavior & Development*, 31, 665-678.
- Rhodes, M. (2020). Are humans born to hate? Three myths and three developmental lessons about the origins of social categorization and intergroup bias. In J. Decety (Ed.), *The social brain: A developmental perspective*. Cambridge, MA: MIT Press.
- Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, 69, 181-203.
- Sugden, N. A., Mohamed-Ali, M. I., & Moulson, M. C. (2014). I spy with my little eye: Typical, daily exposure to faces documented from a first-person infant perspective. *Developmental Psychobiology*, 56, 249-261.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., ... Nelson, C.

A. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*, 242–249.

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

Xiao, N. G., Quinn, P. C., Liu, S., Ge, L., Pascalis, O., & Lee, K. (2018a). Older but not younger infants associate own-race faces with happy music and other-race faces with sad music. *Developmental Science, 21*, e12537.

Xiao, N. G., Wu, R., Quinn, P. C., Liu, S., Tummeltshammer, K. S., Kirkham, N., Ge, L., Pascalis, O., & Lee, K. (2018b). Infants rely more on gaze cues from own-race than other-race adults for learning under uncertainty. *Child Development, 89*, e229-e244.

Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks, 19*, 1395-1407.

Table 1

Mean Fixation Times (seconds) During the Familiarization Trials and Mean Preference Scores (percentages) for the Caucasian Faces During the Preference Test Trials of Experiment 1

	Fixation Time				Novelty Preference	
	Trials 1-3		Trials 4-6		<i>M</i>	<i>(SD)</i>
	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>		
<i>Experimental Groups</i>						
African Familiarization (Asian vs. Caucasian Test)	7.98	(3.16)	7.72	(3.82)	47.66%	(17.83) -0.52
Asian Familiarization (African vs. Caucasian Test)	8.65	(2.84)	6.76	(3.14)	50.10%	(11.54) 0.03

Combined	8.32 (2.97)	7.24 (3.47)	48.88% (14.82)	-0.43
----------	-------------	-------------	----------------	-------

Control Groups

Asian versus Caucasian Test			49.30% (12.24)	-0.23
-----------------------------	--	--	----------------	-------

African versus Caucasian Test			53.22% (10.36)	1.24
-------------------------------	--	--	----------------	------

Combined			51.26% (11.33)	0.63
----------	--	--	----------------	------

^a*t* vs. chance.

Table 2

Mean Fixation Times (seconds) During the Familiarization Trials and Mean Preference Scores (percentages) for the Caucasian Faces During the Preference Test Trials of Experiment 2

	Fixation Time				Novelty Preference		
	Trials 1-3		Trials 4-6		<i>M</i>	<i>(SD)</i>	<i>t</i> ^a
	<i>M</i>	<i>(SD)</i>	<i>M</i>	<i>(SD)</i>			
<i>Experimental Groups</i>							
African Familiarization (Asian vs. Caucasian Test)	8.00	(2.05)	6.11	(3.11)	56.93%	(15.11)	1.83*
Asian Familiarization (African vs. Caucasian Test)	6.68	(2.12)	5.66	(2.09)	54.64%	(8.40)	2.21**

Combined	7.34 (2.16)	5.88 (2.62)	55.78% (12.08)	2.71***
----------	-------------	-------------	----------------	---------

Control Groups

Asian versus Caucasian Test			43.90% (10.13)	-2.41**
-----------------------------	--	--	----------------	---------

African versus Caucasian Test			42.13% (9.50)	-3.31****
-------------------------------	--	--	---------------	-----------

Combined			43.02% (9.70)	-4.07*****
----------	--	--	---------------	------------

^a*t* vs. chance. * $p < .05$, one-tailed, ** $p < .05$, *** $p < .02$, **** $p < .005$, ***** $p < .01$