

Infant Preferences for Attractive Faces: A Cognitive Explanation

Adam J. Rubenstein, Lisa Kalakanis, and Judith H. Langlois
University of Texas at Austin

Research on infant face perception has shown that infants' preferences for attractive faces exist well before socialization from parents, peers, and the media can affect these preferences. Four studies assessed a cognitive explanation for the development of attractiveness preferences: cognitive averaging and infant preferences for mathematically averaged faces, or prototypes. Studies 1 and 2 demonstrated that both adults and 6-month-old infants prefer prototypical, mathematically averaged faces. Studies 3 and 4 demonstrated that 6-month-olds can abstract the central tendency from a group of naturalistic faces. Taken together, the studies suggest that infants' preferences for attractive faces can be explained by general information-processing mechanisms.

Conventional wisdom asserts that standards of attractiveness are gradually learned, subjective, and largely a product of the media. However, empirical research conducted over the past decade has contradicted this widely held belief. Langlois et al. (1987) showed that even young infants prefer attractive to unattractive female faces and that these preferences are similar to preferences that adults have for attractive faces. Infants' preference for attractiveness extends to Caucasian and African American female faces, infant faces, and male faces (Langlois, Ritter, Roggman, & Vaughn, 1991; Samuels & Ewy, 1985). Why should young infants prefer attractive faces and have the same standards of attractiveness as adults? After all, young infants have had only minimal exposure to the cultural standards of beauty portrayed by the media. The studies on young infants' preferences for attractive faces seem to rule out the likelihood that extensive socialization accounts for our standards and preferences for attractiveness.

What mechanism might explain such early preferences for attractiveness? Perhaps the answer lies in how infants process information generally. Several cognitive and perceptual mechanisms have been posited that allow infants to make sense of their world, including faces (Bruner, 1990; Mandler, 1992; Quinn & Eimas, 1997; Strauss, 1979). One such mechanism underlying early preferences for attractiveness might be prototype formation, a cognitive ability common to both infants and adults (Langlois & Roggman, 1990; Rhodes & Tremewan, 1996). A prototype can be defined as the mathematical average or mean value of the at-

tributes of a category and represents the averaged members of a class (e.g., Homa, 1978; Komatsu, 1992; Posner & Keele, 1970; Reed, 1972; Rosch, 1978; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Rosch, Simpson, & Miller, 1976). Both infants and adults are capable of abstracting a prototype after viewing exemplars of a class or category. For example, Strauss (1979) showed simple schematic drawings of faces to adults and 10-month-old infants. The drawings were composed of facial dimensions (length of face, length of nose, width of nose, eye separation), each with a specific value representing a continuum of size from 1 to 5. Strauss created a prototype face by using the average values of each facial dimension. Following familiarization to the schematic exemplar faces, adults and 10-month-olds treated the previously unseen averaged prototype as if it were familiar, suggesting that they averaged the exemplars. Others have also shown that infants average category members to form prototypes in a variety of domains, such as dot patterns (e.g., Bomba & Siqueland, 1983; Quinn, 1987; Younger & Gottlieb, 1988) and schematic animals (Younger, 1985).

The prototype of a category, because of its standing as a unique and representative member of a category, is typically the most preferred member of a category. Research has demonstrated that prototypes of color categories (Martindale & Moore, 1988), object categories (Whitfield & Slatter, 1979), and musical categories (Smith & Melara, 1990) are often preferred to less prototypical exemplars. On the basis in part of this research regarding the preferential status of prototypes and on prototype formation research, Langlois and Roggman (1990) proposed that attractive faces might be preferred by infants and adults alike because attractive faces represent the central tendency or average of the population of facial configurations and are thus prototypical. To examine empirically whether averaged faces are indeed preferred just as attractive faces are preferred, Langlois and Roggman (1990) created naturalistic averaged faces by combining individual faces to create a composite image using digital averaging procedures. These averaged faces as well as the original individual faces were rated for attractiveness by adult judges. Judges found the averaged faces to be more attractive than almost all of the individual faces, confirming the hypothesis that averaged faces are judged to be attractive. Rhodes and Tremewan (1996) replicated these findings using line drawings and facial caricatures.

Adam J. Rubenstein, Lisa Kalakanis, and Judith H. Langlois, Department of Psychology, University of Texas at Austin.

This research was supported by National Institute of Child Health and Human Development Grant HD21332-11A1. Portions of this article were presented at the meeting of the Society for Research in Child Development, Washington, DC, in April 1997. We thank Andrea Larson, Jennifer Ramsey, David Rakison, Janie Childers, and Nate Marti for their helpful comments on earlier drafts of this article. We also thank Andrea Larson and Monica Smoot for their invaluable help in the creation of stimuli for the studies.

Correspondence concerning this article should be addressed to Judith H. Langlois, Department of Psychology, University of Texas at Austin, Mezes Hall 330, Austin, Texas 78712. Electronic mail may be sent to langlois@psy.utexas.edu.

Thus, the prototype hypothesis is a plausible explanation for adult attractiveness preferences. However, to be a sufficient and parsimonious explanation of attractiveness preferences in general, the prototype account must also adequately address young infants' attractiveness preferences. The purpose of the studies in this report was to determine if infants (a) prefer averaged faces and (b) can average across complex, naturalistic faces to form facial prototypes. If, in fact, infants did both prefer averaged faces, as adults do, and were capable of averaging across real faces, we would have evidence for a simple information-processing mechanism as a parsimonious account of both infants' and adults' attractiveness preferences.

Study 1

Study 1 had two goals. First, we wanted to replicate, with a different sample of faces and raters, Langlois and Roggman's (1990) finding that adults judge averaged faces to be attractive. Second, we wanted to obtain adult ratings of faces to select stimuli for Study 2, which investigated whether infants prefer averaged faces. To achieve these goals, we had adults rate the attractiveness of individual and mathematically averaged faces.

Method

Participants. Fifty-six undergraduate students (28 men and 28 women) participated. All students received course credit for participation and were recruited from introductory psychology classes at the University of Texas at Austin. Thirty participants were Caucasian, 14 were Asian, 8 were Hispanic, and 3 were African American. One participant did not report ethnicity.

Stimuli. Thirty-two photographs of Caucasian female undergraduates at the University of Texas at Austin were randomly selected from a large database. All slides depicted individuals facing directly toward the camera with a neutral facial expression. Clothing cues were masked with a white drape. Using the methodology of Langlois and Roggman (1990), we combined all 32 faces on a computer to create a mathematically averaged face. The images of the individual faces and the averaged face were equalized for root mean square contrast, a measure of the average deviation of pixel in the image from the average image pixel value. Using Adobe PhotoShop (1998), we filtered and modified images so that brightness, color balance, sharpness, and clarity of the photographs were equivalent. Thus, there were no differences in the photographic quality between the averaged and individual faces.

Procedure and design. Groups of participants viewed the 33 stimulus slides on a projector screen and judged their facial attractiveness on a 5-point scale ranging from 1 (*very unattractive*) to 5 (*very attractive*). A random sample of 10 stimuli faces was presented for approximately 1 s each before the rating session began to allow the judges to see the range of attractiveness in the sample. Slides of faces were then presented in random order for 10 s each, with an interval of 2 s between each slide.¹ After every 11 slides, a slide of a scenic landscape was projected to allow the judges a brief break from the task.

Results and Discussion

Reliability of the attractiveness ratings was calculated using coefficient alpha (Tinsley & Weiss, 1975). Reliability was high, $\alpha = .97$. No significant differences were found between the mean ratings of male and female judges, $F(1, 54) = 0.02, ns$. Therefore, all analyses were conducted on the combined data of the male and female judges. We analyzed attractiveness ratings of the 32-face

averaged face and the mean of the individual faces with a repeated measures analysis of variance (ANOVA). Type of face (averaged face vs. individual face) was the within-subjects factor; gender was the between-subjects factor. We found a significant difference in the ratings of the different types of faces: Participants rated the 32-face averaged face ($M = 3.38, SD = 0.93$) as significantly higher in attractiveness than the average rating of the individual faces ($M = 2.42, SD = 0.48$), $F(1, 54) = 53.77, p < .001$.² In addition, a post hoc *t* test indicated that there was no significant difference between the most attractive individual face and the 32-face averaged face, $t(55) = 0.74, ns$, a finding that provides further evidence for the link between averageness and attractiveness. We would not expect a significant difference in attractiveness ratings between the most attractive individual faces and the 32-face averaged face because, by definition, the most attractive individual faces should be the closest to the mathematical average of the population in configuration.

These findings replicate precisely those of Langlois and Roggman (1990) using a different set of both judges and faces. These results support the hypothesis that averaged faces are both attractive and preferred members of their category. Consistent with prior definitions of prototypes as representing the central tendency of a category and as preferred members of a category (Light, Hollander, & Kayra-Stuart, 1981; Martindale & Moore, 1988; Whitfield & Slatter, 1979), we showed that adults consider prototypical faces to be attractive.

Study 2

The purpose of Study 2 was to determine if infants, like adults, prefer mathematically averaged faces. We examined infants' preferences for averaged faces by replicating the methodology of previous studies showing that young infants prefer attractive faces over unattractive faces (Kramer, Zebrowitz, San Giovanni, & Sherak, 1995; Langlois et al., 1987, 1991; Samuels & Ewy, 1985). In these studies, a visual preference paradigm allowed infants to compare attractive with unattractive faces. In all cases, infants looked longer at the attractive faces than at the unattractive faces, indicating a preference for attractiveness. In this study, we used the same visual preference methodology as these earlier studies had used. However, instead of pairing individual attractive faces with individual unattractive faces, we paired averaged faces with unattractive faces to determine if infants would demonstrate a preference for averaged faces that is similar to their preference for attractive faces. Thus, a preference for averaged faces would allow the inference that infants, like adults, consider averaged faces to be attractive and preferred to unattractive faces.

Method

Participants. Names of infants were obtained from birth announcements in the local newspaper. Parents were initially contacted by a letter

¹ Attractiveness ratings were gathered in three separate rating sessions. The slides were presented in a different order for each session.

² Note that these are mean attractiveness ratings, averaged across 56 raters. Mean ratings of this magnitude for attractive faces are quite typical in the attractiveness literature (e.g., see Langlois & Roggman, 1990).

that explained the experiment and were later contacted by telephone to schedule an appointment. Infants received a T-shirt for participation.

The initial sample consisted of 73 6-month-old infants (35 boys, 36 girls). The data from 35 infants were deleted from the analyses, 15 because of infant fussiness, 9 because of experimenter error, 1 because of equipment failure, and 10 because the infant was off-task and failed to look at both faces in a test comparison, indicating that he or she was not able to compare the faces. Thus, we obtained a final sample of 38 (17 boys, 21 girls) healthy, full-term infants.³ The ethnicity of the infants was predominantly Caucasian ($n = 30$). The remainder were Hispanic ($n = 6$) and "other" ($n = 2$; 1 = Caucasian-Hispanic; 1 = Caucasian-Asian). Mean age of the final sample was 26 weeks (range = 24 weeks 3 days to 27 weeks 6 days). The demographics of the final sample were fairly homogeneous. The majority of parents worked in professional, technical, and managerial occupations ($n = 31$), whereas the remainder worked in clerical and sales occupations ($n = 2$) and machine trades ($n = 3$).⁴ Two families did not report occupational data. In addition, 26 parents had received at least a bachelor's degree from a 4-year college; 8 had some college education; 1 had received a high school diploma; and 1 had not completed high school. Two families did not report educational data.⁵

Stimuli. We chose five color images of Caucasian female faces from Study 1. One face was the 32-face averaged face (attractiveness = 3.38 on a 5-point scale). The remaining four faces consisted of individual, nonaveraged faces of low attractiveness (mean attractiveness = 1.74 on a 5-point scale) from Study 1.

We used four different faces of low attractiveness to ensure that our findings were not specific to a particular unattractive face. All of the individual faces had dark hair and similar skin tone. Recall from Study 1 that all stimuli were standardized for contrast, brightness, and color balance. Only one 32-face averaged face could be created from the set of faces used in Study 1. However, Langlois and Roggman (1990) showed that 32-face averaged faces look remarkably similar even though composed of nonoverlapping individual faces. Therefore, results from the single 32-face averaged face should readily generalize to other 32-face averaged faces of similar gender and ethnicity.

Procedure and design. Each infant viewed pairs of faces in a darkened room while seated on his or her parent's lap facing a projection screen 36 in. away. The presentation of the pairs of faces was controlled by a computer that operated two slide projectors. Faces were rear-projected at approximately life size. The projection screen and its frame divided the room and hid equipment, the experimenter, and other possible distractions from the infant.

The experimenter captured the infant's attention with a beeping noise and a blinking light, both centered below the projection screen and connected to the computer. Once the infant's attention was secured, the experimenter began the study. Timing of each trial began once the infant looked at one of the faces, after which the infant's looks were recorded by the experimenter pressing keys on the computer, which timed and recorded total looking time to each face. The experimenter observed the infants on a television monitor connected to a video camera placed directly in front of the infant and below the center of the projection screen. To prevent possible experimenter bias, the experimenter faced away from and could not see the stimuli being displayed. For a similar reason, parents wore a sleep mask while the slides were presented so they could not see the stimuli and potentially influence the preference of their infants.

Each infant was presented with two trials lasting 10 s each. The first trial consisted of the 32-face averaged face paired with one of the nonaveraged faces of low attractiveness. The second trial was the left-right reversal of the first trial to control for infant side preferences. Each session was videotaped for later coding by other judges for reliability purposes.

Results and Discussion

Looking times to the averaged face and the individual face were summed across the two presentations. To determine reliability

among experimenters, we computed intraclass correlations for 32% of the infants (Hunter & Koopman, 1990). At least three observers in addition to the original experimenter recorded looking time data for each infant from the videotape. The average intraclass correlation coefficient was .91, indicating high agreement between experimenters. We inspected the data for outliers prior to analysis; no outliers were found.

We analyzed data with a repeated measures ANOVA to test the a priori prediction that the averaged face would be preferred over the individual faces. Type of face (averaged vs. individual) was the within-subjects factor; gender of infant and individual face (which of the four individual faces shown to the infant) were between-subjects factors. No main effect was found for the gender of infant factor, $F(1, 30) = 0.20$, *ns*. As predicted, we found a significant main effect for the type of face factor, $F(1, 30) = 5.32$, $p < .05$, indicating that infants looked longer at the averaged face than the individual faces (7.83 and 6.33 s, respectively). A significant interaction was found between the type of face and individual face factors, $F(3, 30) = 5.24$, $p < .01$. Post hoc *t* tests indicated that infants looked longer at the averaged face in three of the four comparisons; there was no significant preference for either face in the fourth comparison, $t(9) = 1.80$, *ns*. A visual examination of the individual face in the fourth comparison showed that it differed from the other individual faces on one dimension: The hair line was high so that the forehead of this individual appeared much larger than any of the individual faces and the averaged face. Other characteristics of the faces, such as feature size and symmetry, did not appear to differ from one another.

These results demonstrate that infants look longer at averaged faces than at unattractive faces, and the results parallel previous findings in which infants looked longer at attractive versus unattractive faces (Kramer et al., 1995; Langlois et al., 1987, 1991; Samuels & Ewy, 1985). Furthermore, previous research demonstrated that these results are not likely to be the consequence of other characteristics of the face. For example, Langlois, Roggman, and Musselman (1994) demonstrated that adult preferences for averaged faces could not be explained by symmetry, blur, youthfulness, or artifacts of the averaging process itself. Kramer et al. (1995) showed that infants clearly preferred attractiveness over youthfulness, and Samuels, Butterworth, Roberts, Graupner, and Hole (1994) ruled out symmetry as an alternative explanation of infants' preferences for attractive faces. Therefore, we conclude that infants prefer averaged faces because they regard them as attractive, just as adults regard averaged faces as attractive.

³ On the basis of previous research in infant face perception (e.g., Langlois et al., 1987), we determined a minimum sample size for each infant study using a priori power analyses, which ensured power of at least .80 in each study.

⁴ All occupations were classified according to the occupational categories found in the U.S. Department of Labor's (1991) *Dictionary of Occupational Titles*.

⁵ In all cases, occupational and educational data were reported for the father unless the infant lived in a single-parent household.

Study 3

The results from the first two studies indicated that 6-month-old infants, like adults, prefer mathematically averaged faces. However, these studies did not examine the cognitive averaging explanation suggested by Langlois and Roggman (1990). To accept the averaging hypothesis as a potential explanatory mechanism for infants' attractiveness preferences, we had to demonstrate that 6-month-old infants can average very complex, naturalistic faces. Although Strauss (1979) showed that infants can average across schematic facial exemplars, the drawings of facial stimuli were simple and varied on only four dimensions. Furthermore, the infants in Strauss's study, at 10 months of age, were considerably older than the 6-month-olds we used in the studies of infant preferences for attractive faces. Thus, it was important to determine if 6-month-olds can average across and from a prototype of very complex, naturalistic stimuli that vary on hundreds of dimensions (Farkas, 1981).

Because one cannot directly observe cognitive averaging, one must infer it from the infant's behavior. Both cognitive averaging and prototype formation in infants is traditionally assessed through a paired-comparison novelty preference procedure (e.g., Bomba & Siqueland, 1983; Quinn, 1987; Strauss, 1979). This procedure takes advantage of the fact that following familiarization with a series of stimuli, infants typically look longer at novel than at familiar stimuli (see Hunter & Ames, 1988, for a review). In the paired-comparison novelty procedure, infants can be familiarized to a set of exemplars and tested on familiar and novel exemplars and the prototype of the familiar exemplars. If cognitive averaging and prototype formation have occurred, infants should treat the prototype stimulus as familiar even though it was never actually seen by the infant. Thus, infants should look longer at previously unseen novel stimuli than at the previously unseen prototype. If prototype formation and cognitive averaging have not occurred, infants should consider the prototype to be novel just like any other previously unseen stimulus and look as long at it as at the novel stimulus.

We designed three test comparisons to determine if infants are capable of abstracting a prototype following familiarization to photographs of real faces. First, as a manipulation check, one of the tests compared a familiar face (from the familiarization procedure) with a novel, previously unseen individual face. This comparison served to affirm that infants would exhibit novelty preferences following familiarization. We would be sure our familiarization procedure was successful if infants looked longer at the novel face in this pair. Furthermore, on the basis of this manipulation check, we could assume that when infants look longer at a particular face in the other test comparisons, they are exhibiting a novelty preference.

A second comparison paired a novel face with the averaged face created by combining and averaging the familiarization faces on a computer (see Study 1 section and Langlois & Roggman, 1990). If infants form a prototype of the familiarization faces, they should consider the averaged face, even though never previously seen, as familiar. Therefore, infants should look longer at the previously unseen novel face than at the averaged face.

A final and strong test of prototype formation compared one of the familiar exemplars with the averaged face. If prototype formation has occurred, infants should look longer at the familiar ex-

emplar than at the averaged face (e.g., Bomba & Siqueland, 1983; Quinn, 1987). This seemingly counterintuitive prediction is based on the logic that the prototype of the category is the best and most familiar example of the category (Komatsu, 1992). Thus, the previously unseen averaged face should actually seem more familiar to the infant than the previously seen familiar face. Thus, in this test comparison, we predicted the "familiar" face would be "novel" (Strauss, 1979).

Method

Participants. The initial sample consisted of 92 6-month-old infants (44 boys, 48 girls). The data from 35 infants were deleted from the analyses, 8 because of infant fussiness, 6 because of premature birth, 1 because of postmature birth, 6 because of experimenter error, 6 because of outside distraction, 6 because of an infant side preference greater than 90% to one side, and 2 because of equipment failure. Thus, we obtained a final sample of 57 (30 boys, 27 girls) healthy, full-term infants. The ethnicity of the infants was predominantly Caucasian ($n = 50$). The remainder were Hispanic ($n = 3$), African American ($n = 1$), and "other" ($n = 3$; all Caucasian-Hispanic). Mean age of the final sample was 25 weeks 5 days (range = 23 weeks 5 days to 27 weeks 6 days). The demographics of the final sample were fairly homogenous. The majority of parents worked in professional, technical, and managerial occupations ($n = 43$), and the remainder worked in clerical and sales occupations ($n = 9$), service occupations ($n = 1$), machine trades ($n = 2$), and structural work occupations ($n = 1$). One family reported that the head of the household was a student. In addition, 38 parents had received at least a bachelor's degree from a 4-year college; 12 had some college education; and 6 had received a high school diploma. One family did not report educational data.

Stimuli. Because averaged faces are attractive, it was necessary to ensure that the individual exemplars were also attractive to prevent attractiveness rather than novelty preferences from influencing the results. Thus, we selected eight different individual female Caucasian faces of equal and high attractiveness (as determined by adult ratings) as familiarization stimuli. Three of these familiarization faces were selected to serve as the familiar faces in the test trials. Four additional individual faces, equal in attractiveness to the eight familiarization stimuli, were used as the novel faces. We used multiple familiar and novel faces to ensure that results were not limited to specific faces. Finally, we created an averaged face from the eight familiarization faces. All possible combinations of the novel versus familiar, novel versus averaged, and familiar versus averaged comparisons were tested. Thus, 19 possible pairs of faces (12 novel vs. familiar, 4 novel vs. averaged, and 3 familiar vs. averaged comparisons) were used in this study. As in the previous studies, all faces were standardized to eliminate any quality differences (e.g., image contrast, blurriness).

Procedure. The testing room and stimulus presentation was similar to that of Study 2. The infants were presented with eight familiarization faces in 16 5-s trials, divided into two blocks of 8 trials each. Within each block, faces were presented in a random order, with two images of the same face displayed simultaneously during each trial. The infants then were presented with the three test comparisons (novel vs. familiar, novel vs. averaged, and familiar vs. averaged) and their left-right reversals, totaling six 10-s trials. For each infant, the same novel, familiar, and averaged faces were used (e.g., the three comparisons were Novel #1 vs. Familiar #2, Novel #1 vs. averaged, and Familiar #2 vs. averaged), giving the infants multiple exposures to each face. Because each face was seen multiple times in the test trials, it was possible that familiarity with the faces could develop and subsequently affect the infants' preferences. In order to eliminate any possible familiarity effect, the order of the test trials between participants was random, so that some infants would have become familiar with the novel or familiar tests earlier in the test trials than with the averaged face.

Results and Discussion

The dependent variable for all analyses was the sum of the looking time to each face in a pair and its left–right reversal. To determine experimenter reliability, we computed intraclass correlations for 30% of the infants. The average intraclass correlation coefficient was .95, indicating high agreement between experimenters. The data from the original experimenter were averaged with that of the experimenters who conducted reliability checks for the final data analysis.

We analyzed test trial data with a repeated measures ANOVA. Comparison (novel vs. familiar, novel vs. averaged, familiar vs. averaged) and direction of preference (novelty preference vs. familiarity preference) were the within-subjects factors; gender of infant was the between-subjects factor. The main factor of interest was direction of preference: Obtaining a novelty preference in all three comparisons would support the hypothesis that infants can average across faces to form a prototype of complex real faces. Although the comparison factor is largely uninformative on its own, a lack of Comparison \times Direction of Preference interaction would indicate that the infants' preferences for novelty or familiarity are consistent across all the comparisons.

No main effect was found for the infant gender factor, $F(1, 54) = 0.26$, *ns*. As predicted, we found a significant effect for the direction of preference factor, $F(1, 54) = 13.10$, $p < .001$. The means indicated that infants demonstrated consistent novelty preferences, indicating that they formed a prototype of the familiarization faces. Furthermore, neither a significant main effect for comparison, $F(2, 108) = 0.19$, *ns*, nor a significant Comparison \times Direction of Preference interaction, $F(2, 108) = 0.05$, *ns*, was obtained, indicating that all three comparisons yielded the predicted pattern of infant looking behavior (see Table 1). Separate analyses of the pattern of preferences for each of the familiar and novel faces indicated that the findings generalized to all of the faces used in this study, $F(2, 54) = 1.20$, *ns*, and $F(3, 53) = 0.74$, *ns*, respectively.

Exposure to the eight familiarization faces led to the predicted preference in all three test comparisons. These data suggest that infants average across faces to form facial prototypes that appear familiar to them. This evidence is the first we know of to indicate that young infants can average across complex, naturalistic faces, and it suggests that cognitive averaging is a plausible mechanism underlying infants' preferences for attractive faces. This pattern of results is consistent with other studies that have demonstrated that infants are capable of averaging across exemplars from artificial

Table 1
Study 3: Infant Novelty Preferences Following Familiarization

Comparison	Face	<i>M</i>	<i>SD</i>
Novel vs. familiar	Novel	5.65	2.68
	Familiar	4.80	2.45
Novel vs. averaged	Novel	5.79	2.98
	Averaged	4.83	2.58
Familiar vs. averaged	Familiar	5.86	3.04
	Averaged	4.86	2.45

Note. The predicted face in the comparison appears in boldface type. Values represent average number of seconds and standard deviation of infant looking time to faces.

categories, such as dot patterns (Bomba & Siqueland, 1983; Quinn, 1987), schematic animals (Younger & Cohen, 1986), and schematic faces (Strauss, 1979).

Note that in Study 2, in which visual preferences were examined, our prediction was that infants would look longer at the face they preferred—the averaged face. In contrast, we made the opposite prediction in this study, which used the paired-comparison novelty preference procedure: Infants should not prefer to look at the averaged face if cognitive averaging has occurred. These are not inconsistent or mutually exclusive predictions. Although averaged, prototypical stimuli are typically preferred, the procedure in this study takes advantage of the fact that infants will prefer to look at novel stimuli following familiarization to a set of stimuli (Hunter & Ames, 1988), overriding typical visual preference behavior. Note, for example, that infants prefer to look at pictures of their familiar mothers over pictures of novel strangers in a visual preference paradigm (Barrera & Maurer, 1981), but on the basis of the familiarization procedure logic, they prefer the face of a stranger following familiarization to the face of their mother (Field, Cohen, Garcia, & Greenberg, 1984). Thus, the predictions of Study 2 and this study are not contradictory because infants' looking behavior is dependent on the demands of the methodology.

Study 4

What if some unknown preexisting preferences of the infants rather than the familiarization process in Study 3 led to the results we obtained? It was important to establish that our results in Study 3 were due to the familiarization procedure before inferring that prototype formation occurred. To establish that the results we obtained in Study 3 were due to our familiarization procedure and the resulting formation of a prototype, and not to preexisting preferences for some of our stimuli, we evaluated in Study 4 the preferences of infants for the faces in the novel vs. familiar, novel vs. averaged, and familiar vs. averaged comparisons used in Study 3.

Method

Participants. Infants were recruited in the same manner as in Studies 2 and 3. The initial sample consisted of 123 6-month-old infants (58 boys, 65 girls). The data from 34 infants were deleted from the analyses, 5 because of infant fussiness, 11 because of experimenter error, 10 because of low experimenter reliability, 4 because of outside distraction, 3 because of equipment failure, and 1 because of an infant side preference. Thus, a final sample of 89 (39 boys, 50 girls) healthy, full-term infants was obtained. This large number of participants was necessary to test effectively the preferences within all 19 test comparisons used in Study 3. The ethnicity of the infants was predominantly Caucasian ($n = 66$). The remainder were Hispanic ($n = 10$), Asian ($n = 2$), and "other" ($n = 11$; 5 = Caucasian–Hispanic; 3 = Caucasian–Asian; 2 = Hispanic–African American; 1 = Caucasian–African American). Mean age of the final sample was 26 weeks 2 days (range = 24 weeks 3 days to 28 weeks 6 days). The demographics of the final sample were fairly homogenous. The majority of parents worked in professional, technical, and managerial occupations ($n = 63$), whereas the remainder worked in clerical and sales occupations ($n = 7$), service occupations ($n = 5$), benchwork occupations ($n = 6$), structural work ($n = 3$), and miscellaneous occupations ($n = 3$; 2 = radio and television occupations; 1 = motor freight occupations). Two families reported that the head of the household was a student. In addition, 66 parents had received at least a bachelor's degree from a 4-year college; 10

had some college education; and 9 had received a high school diploma. Four families did not report educational data.

Stimuli. The stimuli for this study were the Caucasian female faces used in the test phase of Study 3. Thus, eight different faces (four novel, three familiar, and one averaged face) were used. The faces were presented in the 19 different combinations used in Study 3 (i.e., 12 novel vs. familiar comparisons, 4 novel vs. averaged comparisons, and 3 familiar vs. averaged comparisons).

Procedure and design. The general procedure was similar to that of Study 2. Because familiarity of the faces had to be controlled (e.g., an infant could only see the averaged face in one pair), this reduced the possible number of pairs that each infant could see. Thus, pairs of faces from Study 3 were randomly assigned to the participants, with the exception that infants were not allowed to view the same face in more than one pair. Infants were presented with 3 of the possible 19 pairs of faces and their left-right reversals, for a total of six trials. Each trial lasted 10 s. At least 11 infants viewed each possible pair to ensure sufficient power ($\text{power} > .80$).

Results and Discussion

The dependent variable for all analyses was the total looking time to each face. To determine experimenter reliability, we computed intraclass correlations for all of the infants. The average intraclass correlation coefficient was .93, indicating high agreement between experimenters.

Paired-sample *t* tests indicated that no significant preference existed in 18 of the 19 test comparisons. One comparison between a familiar and a novel face did produce a significant preference, $t(22) = 2.14$, $p < .05$, for the familiar over the novel face. However, the infants in Study 3 exhibited a preference for the novel over the familiar face, a pattern opposite that of the preference found in this study. Thus, this preexisting preference could not have accounted for our findings in Study 3.

To directly address if the results from Studies 3 and 4 led to different looking patterns, we combined the data from Studies 3 and 4 and analyzed looking times using a repeated measures ANOVA. Direction of preference (novelty preference vs. familiarity preference) was the within-subjects factor; comparison (novel vs. familiar, novel vs. averaged, familiar vs. averaged) and experiment (Study 3 vs. Study 4) were the between-subjects factors. We obtained a significant Experiment \times Direction of Preference interaction, $F(1, 482) = 4.33$, $p < .05$, indicating that the pattern of novelty preference found in Study 3 was significantly different from the infants' preferences in Study 4. Thus, we concluded that preexisting preferences for the stimuli could not have accounted for the findings in Study 3 and that the familiarization process and presumably cognitive averaging actually changed infants' looking patterns.

General Discussion

According to conventional wisdom, attractiveness standards and preferences are a product of ever-changing cultural norms to which infants are oblivious and immune; these preferences are thought to be learned gradually only through extensive experience. In contrast, our studies suggest that conventional wisdom has failed to appreciate the considerable cognitive competencies of young infants. Our research indicates that rather than being the result of slow acculturation, attractiveness preferences are the result of a basic cognitive process that is present extremely early in life

(Walton & Bower, 1994). Study 1 confirmed that averaged faces are attractive and are preferred by adults. Study 2 showed that infants prefer averaged faces over unattractive faces just as they prefer attractive faces over unattractive faces (Langlois et al., 1987, 1991; Samuels & Ewy, 1985). Studies 3 and 4 together demonstrated that young infants are capable of averaging across complex, naturalistic faces. Our finding that infants are capable of forming prototypes of naturalistic facial stimuli is consistent with previous research demonstrating that infants form prototypes of dot patterns (Bomba & Siqueland, 1983), schematic animals (Younger & Cohen, 1986), schematic faces (Strauss, 1979), and speech sounds (Kuhl, 1991). Furthermore, our studies used naturalistic faces that varied on a large number of dimensions and were thus much more complex than schematic faces or dot patterns.

Taken together, the four studies suggest a reason for the unusual amount of agreement about who is and is not attractive:⁶ If prototypes and cognitive averaging are used by infants and adults to organize and consolidate incoming information (Bruner, 1957, 1990), people may form a common prototype of faces representing the central tendency of the population very early in life. This representative facial prototype then may be the comparison by which other faces are evaluated, just as the prototypical dot pattern, or sound, or color, or piece of furniture is used to judge other members of their respective categories (Bomba & Siqueland, 1983; Kuhl, 1991; Light et al., 1981; Martindale & Moore, 1988).

Although the ability to average across stimuli to form prototypes is probably innate (Bruner, 1957, 1990; Walton & Bower, 1994), the particular exemplar faces that are encountered and averaged are surely environmentally determined. This fact implies that, within a single culture, only the youngest neonates will have highly idiosyncratic standards of attractiveness because they have not yet formed a prototype. How much experience with faces is necessary before infants form a facial prototype representing the central tendency of the population that they can use to make attractiveness judgments similar to the attractiveness judgments of adults? In computer modeling research, an attractive prototype face emerges when between 16 and 32 faces are averaged together (see Langlois & Roggman, 1990). Furthermore, an averaged face made of 32 faces looks very similar to any other 32-face averaged face even when they are created from completely different individual faces (Langlois & Roggman, 1990; Langlois et al., 1994). Thus, it may be the case that the average of only 32 facial exemplars is sufficient to approximate the population mean and that each person's experience with different faces yields a consolidated and equivalent prototype. However, because there is no way of knowing if human mental processing parallels digital processing, the number of faces necessary to create an attractive prototype in the human mind is unknown. Given that both infants and adults prefer 32-face averaged faces, it may be that human processing is similar to digital processing and that exposure to only 32 faces is sufficient to create a mental prototype by which one judges attractiveness. Because of the number of faces to which even newborns

⁶ Contrary to conventional wisdom that beauty is in the eye of the beholder, meta-analysis showed that the average correlation of attractiveness judgments among 1,368 people was .88 (Langlois et al., 1999).

are exposed,⁷ and because of newborns' early ability to form prototypes (Walton & Bower, 1994), they are likely to form a prototype within days, if not hours, of birth.

Assuming that infants possess the ability to form prototypes at an early age, one might argue that the infants in Study 3 were using a preexisting prototype in their judgments instead of forming a prototype of the specific exemplars presented in the study. However, the differences in patterns of looking times in Studies 3 and 4 argue against this possibility. Together, Studies 3 and 4 show that familiarization with specific faces changed the infants' preferences relative to the preferences exhibited without familiarization. This change indicates that infants were processing the faces "online" in our experiment. Online processing in a specific experimental task is consistent with the results of others (e.g., Jones & Smith, 1993). Note that our claim of online processing in Study 3 differs from our assumptions in Study 2. Because infants were not familiarized with faces in Study 2, we assumed they would exhibit preferences based on attractiveness when making judgments about faces in a preference task (Langlois et al., 1987, 1991; Samuels & Ewy, 1985). In the absence of familiarization, infants should prefer the attractive prototype face rather than less attractive individual exemplars in the same way that infants prefer to look at their mothers over novel strangers in a visual preference task (Barrera & Maurer, 1981) as opposed to looking at a stranger's face following a familiarization procedure (Field et al., 1984). Thus, because of the differences in task demands for Studies 2 and 3, different assumptions about prototype formation versus prototype use are required.

It is important to note that prototype theory is not the only plausible cognitive mechanism explaining preferences for attractive faces. Exemplar-based models of categorization, for example, also predict that the central tendency of a category will be special because the central tendency is optimally similar to all of the individual exemplars of the category (e.g., Heit, 1992, 1994; Nosofsky, 1988, 1991). Because prototype and exemplar models make similar predictions concerning how category members will be encoded and evaluated, it is impossible to make an empirical distinction between the two types of processing models (Barsalou, 1990; Komatsu, 1992). Thus, distinguishing between prototype and exemplar processing of faces was not, nor could it be, the purpose of this article. What is most relevant here, however, is that both exemplar and prototype models assume a special status for the central tendency of a category. Prototype models predict that mathematical averageness in a face is considered attractive because of its similarity to the abstracted central tendency that people store as a category representation. Likewise, exemplar models predict that a mathematically averaged face is favored because of the greater overlap between the averaged face and individual exemplars. Both models are useful for understanding how the central tendency can come to be preferred through cognitive processing of faces.

A limitation of the studies we report here is that we examined primarily the responses of Caucasian infants (80%) to Caucasian female faces. Although there is high agreement about attractiveness within a given ethnic group, agreement across cultures, although higher than might be suspected, is lower than agreement within ethnic groups.⁸ Perhaps the lower levels of agreement across cultures is because different ethnic groups have been exposed to faces from different populations with different structural properties (Brace & Hunt, 1990), resulting in slightly different

facial prototypes. However, because Caucasian infants have shown clear attractiveness preferences for African American faces (Langlois et al., 1991), one might infer that experience with any type of face results in the formation of a single prototype used in evaluating all faces, or that most 6-month-olds have had sufficient experience with African American faces to form a prototype. This inference remains to be directly tested. Research comparing the judgments of faces from different ethnic and gender groups or research that accounts for the experience young infants have with different groups could help evaluate the role of experience in facial prototype formation and subsequent judgments of faces.

In addition, although we demonstrated that infants possess the competence to average across exemplar faces to form facial prototypes, we do not know if in fact infants actually use this mechanism to make their attractiveness judgments. This uncertainty is common in this type of research and represents the standard competence versus performance distinction in developmental psychology (e.g., Carey, 1993). Demonstrating that infants *can* average across faces does not preclude the possibility that other mechanisms may also be involved. For example, according to some, it is possible that there is an innate template of attractiveness operating at birth (Symons, 1979). Future research in the area of infant attractiveness preferences should focus on newborns within the first minutes after birth. Using this extremely young age group would allow researchers to determine if an innate template is present at birth and what, if any, experience is necessary for the development of attractiveness preferences.

The finding from a decade ago that infants prefer attractive faces (Langlois et al., 1987) was surprising at the time in part because no obvious theoretical explanation could account for these counterintuitive results. It was then inconceivable that infants could demonstrate a preference for attractiveness without having extensive exposure to the cultural norms and ideals surrounding attractiveness. We have now identified a simple, parsimonious, and universal mechanism that can explain infants' and adults' preferences for attractive faces. The basic cognitive process of prototype formation appears to be responsible for a seemingly complex social preference and perhaps for the rudiments of stereotyping based on attractiveness.

⁷ On the basis of observations in the hospital (Kalakanis, 1997), newborns see at least 5 to 10 faces before they leave the hospital.

⁸ Cross-cultural reliability from a recent meta-analysis (Langlois et al., 1999) revealed an average correlation of .72.

References

- Adobe Photoshop 5.0 [Computer software]. (1998). San Jose, CA: Adobe Systems Inc.
- Barrera, M. E., & Maurer, D. (1981). Recognition of mother's photographed face by the three-month-old infant. *Child Development, 52*, 714-716.
- Barsalou, L. W. (1990). On the indistinguishability of exemplar memory and abstraction in category representation. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition: Vol. III. Content and process specificity in the effects of prior experiences* (pp. 61-88). Hillsdale, NJ: Erlbaum.
- Bomba, P. C., & Siqueland, E. R. (1983). The nature and structure of infant

- form categories. *Journal of Experimental Child Psychology*, 35, 294–328.
- Brace, C. L., & Hunt, K. D. (1990). A non-racial craniofacial perspective on human variation: A(ustralia) to Z(uni). *American Journal of Physical Anthropology*, 82, 341–360.
- Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, 64, 123–152.
- Bruner, J. S. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.
- Carey, S. (1993). Speaking of objects, as such. In G. Harman (Ed.), *Conceptions of the human mind: Essays in honor of George A. Miller* (pp. 139–159). Hillsdale, NJ: Erlbaum.
- Farkas, L. G. (1981). *Anthropometry of the head and face in medicine*. New York: Elsevier North-Holland.
- Field, T. M., Cohen, D., Garcia, R., & Greenberg, R. (1984). Mother-stranger face discrimination by the newborn. *Infant Behavior & Development*, 7, 19–25.
- Heit, E. (1992). Categorization using chains of examples. *Cognitive Psychology*, 24, 341–380.
- Heit, E. (1994). Models of the effects of prior knowledge on category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1264–1282.
- Homa, D. (1978). Abstraction of ill-defined form. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 407–416.
- Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (pp. 69–95). Norwood, NJ: Ablex.
- Hunter, M. A., & Koopman, R. (1990). Interobserver agreement and reliability of infant visual fixation data. *Infant Behavior & Development*, 13, 109–116.
- Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development*, 8, 113–139.
- Kalakanis, L. (1997). *Newborn preferences for attractive faces*. Unpublished doctoral dissertation, University of Texas at Austin.
- Komatsu, L. K. (1992). Recent views of conceptual structure. *Psychological Bulletin*, 112, 500–526.
- Kramer, S., Zebrowitz, L. A., San Giovanni, J. P., & Sherak, B. (1995). Infants' preferences for attractiveness and babyfacedness. In B. G. Bardy, R. J. Bootsma, & Y. Guiard (Eds.), *Studies in perception and action: III* (pp. 389–392). Hillsdale, NJ: Erlbaum.
- Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, 50, 93–107.
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (1999). *Maxims and myths of beauty: A meta-analytic and theoretical review*. Manuscript submitted for publication.
- Langlois, J. H., Ritter, J. M., Roggman, L. A., & Vaughn, L. S. (1991). Facial diversity and infant preferences for attractive faces. *Developmental Psychology*, 27, 79–84.
- Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science*, 1, 115–121.
- Langlois, J. H., Roggman, L. A., Casey, R. J., Ritter, J. M., Rieser-Danner, L. A., & Jenkins, V. Y. (1987). Infant preferences for attractive faces: Rudiments of a stereotype? *Developmental Psychology*, 23, 363–369.
- Langlois, J. H., Roggman, L. A., & Musselman, L. (1994). What is average and what is not average about attractive faces? *Psychological Science*, 5, 214–220.
- Light, L. L., Hollander, S., & Kayra-Stuart, F. (1981). Why attractive people are harder to remember. *Personality & Social Psychology Bulletin*, 7, 269–276.
- Mandler, J. M. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review*, 99, 587–604.
- Martindale, C., & Moore, K. (1988). Priming, prototypicality, and preference. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 661–670.
- Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 700–708.
- Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 3–27.
- Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, 83, 304–308.
- Quinn, P. C. (1987). The categorical representation of visual pattern information by young infants. *Cognition*, 27, 145–179.
- Quinn, P. C., & Eimas, P. D. (1997). A reexamination of the perceptual-to-conceptual shift in mental representations. *Review of General Psychology*, 1, 271–287.
- Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, 3, 382–407.
- Rhodes, G., & Tremewan, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychological Science*, 7, 105–110.
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–47). Hillsdale, NJ: Erlbaum.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382–439.
- Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 491–502.
- Samuels, C. A., Butterworth, G., Roberts, T., Graupner, L., & Hole, G. (1994). Facial aesthetics: Babies prefer attractiveness to symmetry. *Perception*, 23, 823–831.
- Samuels, C. A., & Ewy, R. (1985). Aesthetic perception of faces during infancy. *British Journal of Developmental Psychology*, 3, 221–228.
- Smith, D. J., & Melara, R. J. (1990). Aesthetic preference and syntactic prototypicality in music: 'Tis the gift to be simple. *Cognition*, 34, 279–298.
- Strauss, M. S. (1979). Abstraction of prototypical information by adults and 10-month-old infants. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 618–632.
- Symons, D. (1979). *The evolution of human sexuality*. New York: Oxford University Press.
- Tinsley, H. E. A., & Weiss, D. J. (1975). Interrater reliability and agreement of subjective judgments. *Journal of Counseling Psychology*, 22, 358–376.
- U.S. Department of Labor, Employment and Training Administration. (1991). *Dictionary of occupational titles*. Lanham, MD: Bernan Press.
- Walton, G. E., & Bower, T. G. (1994). Newborns form "prototypes" in less than 1 minute. *Psychological Science*, 4, 203–205.
- Whitfield, T. W., & Slatter, P. E. (1979). The effects of categorization and prototypicality on aesthetic choice in a furniture selection task. *British Journal of Psychology*, 70, 65–75.
- Younger, B. A. (1985). The segregation of items into categories by ten-month-old infants. *Child Development*, 56, 1574–1583.
- Younger, B. A., & Cohen, L. B. (1986). Developmental change in infants' perception of correlations among attributes. *Child Development*, 57, 803–815.
- Younger, B. A., & Gottlieb, S. (1988). Development of categorization skills: Changes in the nature or structure of infant form categories? *Developmental Psychology*, 24, 611–619.

Received December 22, 1997

Revision received October 16, 1998

Accepted October 20, 1998 ■