Developmental Science 7:2 (2004), pp 201–211

PAPER

Origins of a stereotype: categorization of facial attractiveness by 6-month-old infants

Jennifer L. Ramsey,[1,2] Judith H. Langlois,[1] Rebecca A. Hoss,[1,3] Adam J. Rubenstein[1,4] and Angela M. Griffin[1]

1. *Department of Psychology, The University of Texas at Austin, USA*
2. *Now at Department of Psychology, University of Nevada Las Vegas, USA*
3. *Now at College of Saint Mary, USA*
4. *Now at Department of Psychology, College of William and Mary, USA*

Abstract

Like adults, young infants prefer attractive to unattractive faces (e.g. Langlois, Roggman, Casey, Ritter, Rieser-Danner & Jenkins, 1987; Slater, von der Schulenburg, Brown, Badenoch, Butterworth, Parsons & Samuels, 1998). Older children and adults stereotype based on facial attractiveness (Eagly, Ashmore, Makhijani & Longo, 1991; Langlois, Kalakanis, Rubenstein, Larson, Hallam & Smooth, 2000). How do preferences for attractive faces develop into stereotypes? Several theories of stereotyping posit that categorization of groups is necessary before positive and negative traits can become linked to the groups (e.g. Tajfel, Billig, Bundy & Flament, 1971; Zebrowitz-McArthur, 1982). We investigated whether or not 6-month-old infants can categorize faces as attractive or unattractive. In Experiment 1, we familiarized infants to unattractive female faces; in Experiment 2, we familiarized infants to attractive female faces and tested both groups of infants on novel faces from the familiar or novel attractiveness category. Results showed that 6-month-olds categorized attractive and unattractive female faces into two different groups of faces. Experiments 3 and 4 confirmed that infants could discriminate among the faces used in Experiments 1 and 2, and therefore categorized the faces based on their similarities in attractiveness rather than because they could not differentiate among the faces. These findings suggest that categorization of facial attractiveness may underlie the development of the 'beauty is good' stereotype.

Introduction

Stereotypes are generalizations or beliefs about a group of people that are based on cognitive categories perceivers use in social situations when processing information about individuals (Allport, 1954; Hamilton & Trolier, 1986). Adults know much about the content of many stereotypes, such as what traits and behaviors are associated with which groups (e.g. Devine, 1989), but how is this knowledge acquired? For most people, it is unlikely that knowledge about stereotype content was explicitly taught. Rather, knowledge about stereotype content is most likely accumulated through social interaction, observation, and the cognitive processing of those interactions and observations. As a result, stereotypes can become so automatic that people are unaware of them and their activation (e.g. Banaji & Greenwald, 1994; Bargh & Chartrand, 1999). How does the stereotyping process become so automatic and what mechanisms underlie how stereotypes are formed? This article will explore the early development of facial attractiveness stereotypes and suggest that categorization facilitates the development of this stereotype.

Attractiveness stereotypes may originate from visual preferences for attractive faces. Several studies have documented preferences for attractive faces in infants ranging in age from a few days to 6 months (Langlois *et al.*, 1987; Langlois, Ritter, Roggman & Vaughn, 1991; Samuels, Butterworth, Roberts, Graupner & Hole, 1994; Samuels & Ewy, 1985; Slater, Bremner, Johnson, Sherwood, Hayes & Brown, 2000; Slater, Quinn, Hayes & Brown, 2000; Slater *et al.*, 1998). Infants look longer at faces adults rate as attractive than at faces adults rate as unattractive. This finding is robust in that it has been

Address for correspondence: Judith H. Langlois, The University of Texas at Austin, Psychology Department, 1 University Station, A8000, Austin, TX 78712, USA; e-mail: langlois@mail.utexas.edu

replicated with faces of Caucasian female and male adults, African-American female adults and Caucasian infants (Langlois *et al.*, 1991).

By 3 years of age, full-blown stereotypes favoring attractive peers and adults are evident. Young children attribute positive behaviors and traits to attractive children and negative behaviors and traits to unattractive children, and select attractive children as playmates more than unattractive children (Dion, 1973; Langlois & Stephan, 1977, 1981). Stereotyping based on physical attractiveness is known in the social psychology literature as the 'beauty is good' stereotype and continues throughout adulthood (Dion, Berscheid & Walster, 1972; Eagly *et al.*, 1991; Langlois *et al.*, 2000).

When and how do infants' visual preferences for attractive faces develop into physical attractiveness stereotypes? Results from research with 12-month-old infants suggest that 'beauty is good' stereotype development may begin during the first year of life. In one study, 12-month-olds approached and showed more positive affect toward an attractive female stranger than a less attractive female stranger (Langlois, Roggman & Rieser-Danner, 1990). In a second study, 12-month-olds also played longer with attractive rather than less attractive dolls (Langlois *et al.*, 1990). Thus, by 12 months, infants behave differently toward attractive and unattractive people and objects. In addition, Rubenstein (2000) found that sometime after 9 months of age, but by 12 months of age, infants match pleasant voices with attractive faces and unpleasant voices with unattractive faces, suggesting that infants associate one type of preferred stimuli (e.g. attractive faces) with other types of preferred stimuli (e.g. pleasant voices), and less preferred stimuli (e.g. unattractive faces) with other types of less preferred stimuli (e.g. unpleasant voices). These findings demonstrate that the 'beauty is good' stereotype is evident in development much earlier than previously supposed. The transition from simple visual preferences for attractive individuals to stereotypic associations about attractive and unattractive individuals seems to develop sometime during infancy.

What abilities are required for such a transition? Theory and prior research posits that categorization of people into different groups is a cognitive prerequisite of stereotyping (e.g. Tajfel *et al.*, 1971; Zebrowitz-McArthur, 1982). Thus, according to this view, infants must identify attractive and unattractive faces as belonging to different groups of faces based on the physical characteristics of the face. Once infants categorize based on physical appearance, the process of categorization directs perception so that members of a particular group are perceived as having similar traits and behaviors (assimilation effects) whereas members of the other group are perceived as having traits and behaviors different from (contrast effects) those of the first group (e.g. Tajfel *et al.*, 1971; Zebrowitz-McArthur, 1982). If infants categorize individuals based on attractiveness, such categorization should guide their perception so that characteristics of attractive individuals are generalized to other attractive individuals, but not to unattractive individuals, and vice versa. There is evidence that adults categorize female faces according to physical attractiveness (Miller, 1988) and that infants group faces into other social categories, such as gender (Leinbach & Fagot, 1993). Thus, perhaps infants also group faces into attractiveness categories.

Purpose

We investigated whether young infants are capable of categorizing faces as attractive or unattractive. If such categorization abilities are a prerequisite to forming stereotypes based on attractiveness, these abilities should be present prior to stereotypic behaviors and associations. Because behaviors and associations characteristic of the 'beauty is good' stereotype are present at 12 months of age, but not 9 months of age (Langlois *et al.*, 1990; Rubenstein, 2000), stereotypic attributions likely develop between 9 and 12 months of age. Therefore, categorization abilities should be present even earlier in development. But when? At 6 months of age, infant preferences for attractive over unattractive faces are particularly robust, regardless of age, gender or ethnicity (Langlois *et al.*, 1987, 1991; Samuels *et al.*, 1994; Samuels & Ewy, 1985), suggesting strong discrimination between these categories of attractiveness. Given the likelihood that strong preferences for groups may facilitate or result from categorization of groups, we investigated if 6-month-old infants could successfully form categories of attractive and unattractive faces.

Experiment 1

Method

Participants

We recruited 93 infants from birth announcements in the local newspaper and obtained written consent from parents for their infant's participation in the study. The data from 50 infants were not analyzed for the following a priori reasons: born prematurely (6); parent interaction during the study (9); side preference (18); equipment error (1); fussiness, crying and off task (11); and experimenter error (5). Our a priori criterion for side prefer-

ences excluded the data from any infant who spent 90% or more of their time looking at one side of the screen during the test trials (Haaf & Diehl, 1976). Our a priori criterion for being off task excluded the data from any infant who looked around the room or at their parent during the majority of the study rather than looking at the slides on the screen. The 43 infants (24 female) represented in the data analysis were 6 months of age, plus or minus 2 weeks, and were born within 3 weeks of their due date. The majority of the infants who participated were Caucasian (72%), but Hispanic (19%), Asian/Pacific Islander (4.5%) and infants of other or mixed racial backgrounds (4.5%) also participated.

Materials

The stimuli consisted of photographic, color slides of 28 adult, Caucasian, female faces. We obtained these slides from our photo database of faces containing approximately 1000 female adults who posed with neutral facial expressions and masked their clothing cues with a drape. At least 40 adults (20 female, 20 male) previously rated the faces for attractiveness using a 1 to 5 Likert scale (1 = *very unattractive*; 5 = *very attractive*), and there was high interrater reliability (alphas = .95 or greater). From the large sample of faces, we chose 24 faces rated as unattractive and four rated as attractive. Mean ratings of the unattractive faces ranged from 1.33 to 2.12 ($M = 1.69$); mean ratings of the attractive faces ranged from 3.86 to 4.19 ($M = 4.01$). We used 20 of the unattractive faces as familiarization stimuli. For generalization purposes, hair varied in color, length and style. We used the remaining four unattractive female faces plus the four attractive female faces as test stimuli. To control for any artifactual effects of hair color, half of the attractive and unattractive faces used in the test trials had light hair and the other half had dark hair. We standardized all slides for size, brightness, contrast, color balance and background using Adobe Photoshop. We projected life-size images (approximately 24 cm high) of the slides using two carousel projectors, and ran the study using a MacIntosh computer and Habit software (a program that Leslie Cohen and Judith Langlois developed at the University of Texas at Austin for presenting stimuli to infants during various types of testing procedures).

Procedure

We rear-projected the slides in pairs onto a screen with side partitions that blocked the rest of the room from the infant's view. The facial images were approximately 30.5 cm away from one another, so that both images could be easily seen by the infants, but looks to a particular image were discernible by the experimenter. Infants sat on their parent's lap in a darkened room facing the screen, which was approximately 91.4 cm away. Parents wore opaque sunglasses that occluded their view of the faces to ensure that infant looking was not influenced or guided by parental interest in particular faces. To prevent experimenter bias, the experimenter sat on the other side of the screen facing away from it and could not see the faces. The experimenter observed the infant on a television monitor attached to a video camera focused on the infant's eyes, and recorded onset and duration of looking times by pressing and holding down keys on a computer keyboard. The computer interfaced with the projectors and advanced the slides. We videotaped all infants during the study so that we could evaluate the looking time data for reliability. At least two individuals other than the original experimenter assessed reliability for each infant. Intraclass reliability ranged from .86 to .99 ($M = .96$).

The experimenter captured the infant's attention with a beeping noise and a blinking light controlled from the computer before each trial started. Once the infant looked at the center of the screen, the experimenter began the trial. The procedure was infant-controlled during both the familiarization and test trials, so the computer advanced the projectors to the next trial when the infant looked away from the faces for more than 1 second, regardless of whether infants looked at one or both faces within the face pair during the trials. We used an infant-controlled procedure because past research has shown that infants can discriminate among similar looking faces if they are given a sufficient amount of time to look at and learn about the face (Barrera & Maurer, 1981). Thus, the infant-controlled procedure ensured that infants could view a face for as long as they were interested in learning about that face.

We familiarized infants to the unattractive category by showing them 15 different, unattractive, Caucasian, female faces randomly selected from the 20 familiarization faces. This selection procedure allowed us to present different faces in different orders to each infant, thus increasing the generalizability of our results. During familiarization, we presented each face in pairs, so that the same face was projected side by side. The purpose of presenting same face pairs during the familiarization trials was to make it similar to the test trials in which infants also saw two faces projected side by side. Following familiarization, infants saw two pairs of novel faces during the test trials. One of the faces in each pair was a new face from the familiarized, unattractive category and the other was from the novel, attractive category. Infants saw each pair of faces twice, presented in blocks, with their left-right positions reversed to control for possible side preferences. For example, one infant saw novel

Face A from the familiarized, unattractive category on the left side and novel Face B from the novel, attractive category on the right side during the first test trial and then the positions of these two faces reversed during the second test trial. This infant then saw a new pair of novel faces, one from each category, on the third test trial, and then saw these same faces with their left-right positions reversed during the fourth test trial. We randomly chose the two attractive and two unattractive faces used during the test trials from the four attractive and four unattractive test faces for each infant and alternated the sides on which infants saw the attractive and unattractive faces across trials and across infants.

Data analyses

All analyses in this study and subsequent studies were conducted using SAS software and the PROC MIXED repeated measures analysis. This analysis is similar to running a repeated measures ANOVA, but allowed us to more easily examine the effects of within- and between-participants variables in the same analysis. An alpha level of .05 was used for all statistical tests.

Results and discussion

Following sufficient familiarization to the unattractive faces, infants should look longer at the novel, attractive faces than the novel, unattractive faces during the test trials if they categorize the faces based on their attractiveness. Even though both the attractive and unattractive test faces are new to the infant, the unattractive faces should seem familiar whereas the attractive faces should be perceived as novel if the infants have categorized the faces based on their attractiveness.

Infant total looking times during the familiarization trials varied widely and ranged from 21 to 209 seconds ($M = 64.57$, SD = 33.80). To determine if infants became sufficiently familiarized to the unattractive faces, we compared their looking times during the first three familiarization trials to their looking times during the last three familiarization trials. Infants significantly decreased their looking across the familiarization trials, $F(42) = 11.79$, $p = .001$. Average looking time of all the infants during the first three familiarization trials was 5.73 seconds (SD = 3.39), and average looking time of all the infants during the last three familiarization trials was 3.45 seconds (SD = 2.72), suggesting that they had become familiar or bored with the category of unattractive faces.

To compute infants' total looking time to the attractive and unattractive test faces, we summed the looking times in tenths of seconds to each of the attractive and unattractive faces across the four test trials. Because the procedure was infant-controlled, the looking times during the test trials were highly variable and the distribution was skewed. Thus, we converted the looking times to proportions of the total looking time (PTLT) to better compare the infants' data. PTLT served as the dependent variable with face type (attractive/unattractive) as the within-participants variable and participant gender as the between-participants variable. Any interaction involving participant gender and face type would suggest that male and female infants differed from one another in their categorization of unattractive faces. There was, however, no significant interaction between participant gender and face type. Thus, we collapsed the data across all the infants and reran the analysis to assess the effect of face type on infant looking time during the test trials. The effect of face type was significant, $F(42) = 22.38$, $p < .001$. Infants spent an average of 58% of their time looking at the attractive faces ($M = 10.37$ seconds, SD = 12.94) and an average of 42% of their time looking at the unattractive faces ($M = 6.85$ seconds, SD = 7.89) during the test trials. A non-parametric binomial test showed that 77% of the infants looked longer at the attractive than the unattractive faces during the test trials and that this proportion was significantly different from chance, $p < .001$.

After being familiarized to a series of unattractive faces, infants looked longer at novel, attractive faces when they were paired with novel, unattractive faces. Because infants spontaneously prefer attractive to unattractive faces (Langlois *et al.*, 1987, 1991; Samuels *et al.*, 1994; Samuels & Ewy, 1985; Slater, Bremner *et al.*, 2000; Slater, Quinn *et al.*, 2000; Slater *et al.*, 1998), these results are ambiguous. Perhaps infants were showing simple preferences for attractive faces and did not categorize faces based on their attractiveness. Therefore, in Experiment 2, we familiarized infants to attractive faces and tested them with novel attractive and unattractive faces. If infants categorize faces according to attractiveness, then they should look longer at the novel unattractive faces than the novel attractive faces following familiarization to attractive faces, despite their pre-existing preference for attractive faces.

Experiment 2

Method

Participants

We recruited 88 infants from birth announcements in the local newspaper and obtained written consent from parents for their infant's participation in the study. The data

from 42 infants were not analyzed for the following a priori reasons: too old (2); born prematurely (7); parent interaction during the experiment (5); side preference (10); equipment error (2); fussiness, crying and off task (8); and experimenter error (8). The 46 infants (21 females) represented in the data analysis were 6 months of age, plus or minus 2 weeks, and were born within 3 weeks of their due date. Again, the majority of infants who participated were Caucasian (76%), but Hispanic (13%) infants and infants of other or mixed racial backgrounds (11%) also participated.

Materials

All materials were the same as those used in Experiment 1, except for the familiarization slides. The familiarization slides consisted of 20 attractive, rather than unattractive, faces. Mean ratings of the attractive faces used as familiarization or test slides ranged from 3.41 to 4.19 ($M = 3.78$); mean ratings of the unattractive faces used as test slides ranged from 1.33 to 1.49 ($M = 1.38$).

Procedure

The procedure was the same as Experiment 1, except that infants saw 15 different attractive Caucasian female faces during the familiarization trials. For each infant, we randomly chose the 15 faces from the 20 familiarization faces. Again, we videotaped all infants during the study so that we could evaluate the looking time data for reliability. At least two individuals other than the original experimenter assessed reliability for each of the infants. Intraclass reliability ranged from .88 to .99 ($M = .96$).

Results and discussion

Following sufficient familiarization to the attractive faces, infants should look longer at the novel, unattractive faces than the novel, attractive faces during the test trials if they categorize the faces based on their attractiveness. Even though both the attractive and unattractive test faces are new to the infant, the attractive faces should seem familiar whereas the unattractive faces should be perceived as novel if the infants have categorized the faces based on their attractiveness.

Infant total looking times during the familiarization trials varied widely and ranged from 23.1 to 196.9 seconds ($M = 69.01$, SD = 38.00). As in Experiment 1, infants significantly decreased their looking across the familiarization trials, $F(45) = 4.54$, $p = .04$, suggesting that they had become familiar or bored with the category of attractive faces. Average looking time during the first three familiarization trials was 5.97 seconds (SD = 3.03) and average looking time during the last three familiarization trials was 4.27 seconds (SD = 4.49). Thus, infants decreased their looking across the familiarization trials, although the decrease in looking to the attractive faces was not as large as the decrease in looking to the unattractive faces during the familiarization trials in Experiment 1.

As in Experiment 1, we converted the looking times to PTLT and ran the analysis with PTLT as the dependent variable, face type as the within-participants variable and participant gender as the between-participants variable. Again, there was no significant interaction between participant gender and face type. Thus, we collapsed the data across all the infants and reran the analysis to assess the effect of face type on infant looking time during the test trials. The effect of face type was significant, $F(45) = 20.00$, $p < .001$. As predicted, infants spent an average of 58% of their time looking at the unattractive faces ($M = 9.45$ seconds, SD = 7.17) and an average of 42% of their time looking at the attractive faces ($M = 7.16$ seconds, SD = 6.11) during the test trials. A non-parametric binomial test showed that 70% of the infants looked longer at the unattractive than the attractive faces during the test trials and that this proportion was significantly different from chance, $p = .006$.

After being familiarized to a series of attractive faces, infants looked longer at novel, unattractive faces when they were paired with novel, attractive faces. Given that infants have pre-existing preferences for attractive faces (e.g. Langlois *et al.*, 1987), this finding is particularly noteworthy because the effect of categorization needed to be strong enough to overcome infants' general tendency to look longer at attractive than unattractive faces when paired together. This finding does not contradict previous research demonstrating that infants prefer attractive faces (Langlois *et al.*, 1987, 1991; Samuels *et al.*, 1994; Samuels & Ewy, 1985; Slater, Bremner *et al.*, 2000; Slater, Quinn *et al.*, 2000; Slater *et al.*, 1998) because in those studies the researchers tested infant interest in equally novel faces and did not manipulate infants' experience or familiarity with the faces. In contrast, in the present studies, we tested infant interest in faces following experience with a particular category of faces. The results from Experiment 2 suggest that by 6 months of age, infants categorize attractive faces as belonging to one group and recognize that unattractive faces are different from that group.

Experiment 3

The results from Experiment 1 suggest that 6-month-olds categorize unattractive female faces as belonging to

one group and do not include attractive female faces within that group. One possibility, however, is that the infants simply could not distinguish between the unattractive female faces because of their similarity in attractiveness. Thus, infants may have treated the unattractive faces as the same face rather than as different faces from the same category. To ensure that infants were treating the unattractive faces as individual faces belonging to the same category and not simply as one face, it is important to demonstrate that 6-month-olds can discriminate among faces belonging to the same attractiveness category. The purpose of this study was to investigate infants' ability to discriminate among the unattractive faces used during the familiarization trials of Experiment 1 and the test trials of Experiments 1 and 2. Unlike Experiments 1 and 2, which tested 6-month-olds' ability to discriminate between attractiveness categories (i.e. discriminate between attractive and unattractive faces), this study investigates 6-month-olds' ability to discriminate within attractiveness categories (i.e. discriminate between unattractive faces). If infants demonstrate significant interest in the novel relative to the familiar unattractive face during the test trials, then this finding provides evidence that infants are able to discriminate between the faces even though they are from the same category.

Method

Participants

Using birth certificate information provided by the Texas Department of Health, we recruited 41 infants to participate in the study. The data from 17 infants were not analyzed for the following a priori reasons: born prematurely (4); parent interaction during the study (2); side preference (2); equipment error (1); fussiness, crying and off task (8). The 24 infants (15 females) represented in the data analysis were 6 months of age, plus or minus 2 weeks, and were born within 3 weeks of their due date. The ethnicity of the infants who participated was as follows: 71% Caucasian, 8% Hispanic and 21% mixed or other race.

Materials

The stimuli consisted of the same 24 unattractive female faces used during the familiarization trials in Experiment 1 and test trials in Experiments 1 and 2. Using the original Adobe Photoshop files from which the color slides were created for Experiment 1, we presented images of life-size female faces (approximately 24 cm high) on two computer monitors using a MacIntosh computer and Habit 2000 software.

Procedure

Because of a move to new facilities and a concurrent change in equipment and lab set-up (e.g. using video monitor presentations rather than slide projections for stimulus presentation), we provide specific details about this study's procedure, most of which is relatively similar to the procedure in Experiments 1 and 2, but includes minor changes. Infants viewed pairs of faces with each face presented on one of two computer monitors 28 cm apart, so that both images could be easily seen by the infants, but looks to a particular image were discernible by the experimenter. A large black wooden structure framed the two computer monitors so that the monitors were equidistant from the center of the structure and the structure blocked the rest of the room and equipment from the infant's view. Infants sat on their parent's lap in a darkened room facing the monitors, which were approximately 91.4 cm away. To prevent experimenter bias, the experimenter sat in a separate control room and could not see the faces. The experimenter observed the infant on a television monitor attached to a low-light video camera focused on the infant's face, and recorded onset and duration of looking times by pressing and holding down keys on a computer keyboard. We videotaped all infants during the study trials so that we could evaluate the looking time data for reliability. At least two individuals assessed reliability for each infant. Intraclass reliability ranged from .92 to .99 ($M = .97$).

The experimenter captured the infant's attention with a whistling noise accompanied by green circles that grew larger and smaller on the computer monitors before each trial started. Once the infant looked toward the monitors, the experimenter began the trial. We familiarized each infant to an unattractive face by showing them the same face presented in same face pairs across 20 familiarization trials that each lasted 10 seconds. We randomly chose one of the 24 unattractive faces used in Experiment 1 to be used as the familiar face and we chose two novel faces from the remaining faces for which each infant was to evidence discrimination. Each infant saw a different unattractive face during the familiarization trials than the other infants saw. To illustrate, Infant A viewed Face 1 during the familiarization trials and then saw Face 1 paired with Faces 2 and 3 during the test trials. Infant B viewed Face 2 during the familiarization trials and then saw Face 2 paired with Faces 3 and 4 during the test trials. This method allowed us to investigate if each infant could distinguish between three of the unattractive faces used in Experiment 1. Logically, it also permitted us to infer that if Infant A could distinguish Faces 2 and 3 from Face 1 and Infant B could distinguish Face 2 from Face 3 and Face 4, then they

likely could both distinguish Face 1 from Face 4 without directly testing this discrimination. Such a procedure was necessary because running one infant for every possible comparison would have required an incredibly large sample of infants ($N = 276$). Infants saw the familiar and novel face pairs twice, presented in blocks, during the test trials, with their left-right positions reversed to control for possible side preferences during the second presentation. We randomized the sides on which the infants saw the familiar and novel faces presented across infants. Because the study lasted over 4 minutes and required multiple presentations of the same stimulus during familiarization, infants saw a colorful, animated butterfly presented on both computer monitors during a final test trial in order to assess fatigue.

Results and discussion

Infant total looking times during the familiarization trials ranged from 61.7 to 146.1 seconds ($M = 107.74$, SD = 23.85). Infants significantly decreased their looking across the familiarization trials, $F(23) = 17.97$, $p < .001$, suggesting that they had become familiar or bored with the face. Average looking time during the first three familiarization trials was 6.45 seconds (SD = 1.72) and average looking time during the last three familiarization trials was 4.28 seconds (SD = 1.83).

To compute infants' total looking time to the novel and familiar test faces, we summed the looking times in tenths of seconds to each of the faces across the four test trials. To be consistent with Experiments 1 and 2, we converted infant looking times to the novel and familiar faces during the test trials to PTLT. PTLT during the test trials served as the dependent variable with face type (novel/familiar) as the within-participants variable and participant gender as the between-participants variable. There was no interaction between gender and face type. Thus, we collapsed the data across all the infants and reran the analysis to assess the effect of face type on infant looking time during the test trials. The effect of face type was significant, $F(23) = 23.47$, $p < .001$. As predicted, infants spent an average of 60% of their time looking at the novel face ($M = 12.68$, SD = 4.92) and an average of 40% of their time looking at the familiar face ($M = 8.67$, SD = 3.80). A non-parametric binomial test showed that 79% of the infants looked longer at the novel than the familiar faces during the test trials and that this proportion was significantly different from chance, $p = .007$.

Despite the length of the study, the infants were not fatigued. Infants looked longer when the butterfly was shown at the end of the study ($M = 5.73$ seconds, SD = 1.46) than they did during the average of their last three familiarization trials ($M = 4.28$ seconds, SD = 1.83), $F(23) = 9.30$, $p = .006$.

Infants' significant interest in the novel face during the test trials provides evidence that they could discriminate between the unattractive faces paired together. Thus, the results from Experiment 1 that demonstrate that infants categorize unattractive faces as a group are not due to infants' inability to discriminate among the faces. Rather, infants categorize unattractive faces as belonging to a group that does not include attractive faces, despite their noticing the individual differences among the unattractive faces.

Experiment 4

The results from Experiment 2 suggest that 6-month-olds categorize attractive female faces as belonging to one group and do not include unattractive female faces within that group. One possibility, however, is that the infants simply could not distinguish between the attractive female faces because of their similarity in attractiveness, particularly because research suggests that attractive faces are similar in configuration (Langlois & Roggman, 1990). Thus, infants may have treated the attractive faces as the same face rather than as different faces from the same category. To ensure that infants in Experiment 2 treated the attractive faces as individual faces within the same category and not simply as one face, we investigated infants' ability to discriminate among the attractive faces used during the familiarization trials of Experiment 2 and the test trials of Experiments 1 and 2. As in Experiment 3, this study investigates 6-month-olds' ability to discriminate within attractiveness categories (i.e. discriminate between attractive faces). We determined that if infants demonstrated significant interest in the novel relative to the familiar attractive face during the test trials, then they could discriminate among the attractive faces used in Experiments 1 and 2.

Method

Participants

Using birth certificate information provided by the Texas Department of Health, we recruited 55 infants to participate in the study. The data from 26 infants were not analyzed for the following reasons: born prematurely (8); parent interaction during the study (4); side preference (3); fussiness, crying and off task (9); and experimenter error (2). The 29 infants (15 females) represented in the data analysis were 6 months of age, plus or minus 2 weeks, and were born within 3 weeks of their due date.

The ethnicity of the infants who participated was as follows: 55% Caucasian, 17% Hispanic, 7% Asian-American/Pacific Islander and 21% of mixed or other race.

Materials

The stimuli consisted of the same 24 attractive female faces used during the familiarization trials in Experiment 2 and test trials in Experiments 1 and 2. Using the original Photoshop files from which the color slides were created for Experiments 1 and 2, we presented images of the female faces on two computer monitors using a MacIntosh computer and Habit 2000 software.

Procedure

The procedure was the same as Experiment 3 except that the faces the infants saw during the familiarization trials were of attractive rather than unattractive females. Also, unlike Experiment 3, the number of infants that participated was larger than the number of attractive faces that could be used as the familiar face (29 vs. 24). As a result, 24 infant participants each saw a different unattractive face during the familiarization trials than the other infants, and the remaining five infants each saw a familiar face that one of the first 24 infant participants saw. We kept the test faces paired with the familiar face the same for the infants who saw the same face during the familiarization trials. Intraclass reliability of the infant looking times ranged from .88 to .99 ($M = .96$).

Results and discussion

Infant total looking times during the familiarization trials ranged from 62.8 to 154.7 seconds ($M = 111.81$, SD = 26.04). Infants significantly decreased their looking across the familiarization trials, $F(29) = 37.56$, $p < .001$, suggesting that they had become familiar or bored with the face. Average looking time during the first three familiarization trials was 6.93 seconds (SD = 1.04) and average looking time during the last three familiarization trials was 4.58 seconds (SD = 1.82).

To compute infants' total looking time to the novel and familiar test faces, we summed the looking times in tenths of seconds to each of the faces across the four test trials. As in all the other experiments, we converted infant looking times to the novel and familiar faces during the test trials to PTLT. PTLT during the test trials served as the dependent variable with face type (novel/familiar) as the within-participants variable and participant gender as the between-participants variable. Gender did not interact with face type preference, so we collapsed the data across all infants in the subsequent analysis. The effect of face type was significant, $F(29) = 45.36$, $p < .001$. Infants spent an average of 62% of their time looking at the novel faces ($M = 14.29$ seconds, SD = 6.25) and an average of 38% of their time looking at the familiar faces ($M = 8.23$ seconds, SD = 3.60) during the test trials. A non-parametric binomial test showed that 80% of the infants looked longer at the novel than the familiar faces during the test trials and that this proportion was significantly different from chance, $p = .002$.

Despite the length of the study, the infants were not fatigued. Infants looked longer when the butterfly was shown at the end of the study ($M = 5.81$ seconds, SD = 1.92) than they did during the average of their last three familiarization trials ($M = 4.58$ seconds, SD = 1.82), $F(29) = 6.45$, $p = .02$.

Infants' significant preference for the novel face during the test trials provides evidence that they could discriminate among the attractive faces used in Experiments 1 and 2. Thus, the results from Experiment 2 that demonstrate that infants categorize attractive faces as a group are not due to infants' inability to discriminate among the faces. Rather, infants categorize attractive faces as belonging to a group that does not include unattractive faces, despite their noticing the individual differences among the attractive faces.

General discussion

Taken together, the results from Experiments 1 and 2 provide evidence that 6-month-old infants categorize attractive and unattractive faces as different types of faces. The results from Experiments 3 and 4 provide support that this categorization occurs not because infants cannot discriminate among the individual unattractive or among the individual attractive faces, but because infants recognize the similarities among faces equal in attractiveness and their differences from faces unequal in attractiveness. According to stereotype theory and research, categorization is a necessary prerequisite of stereotype formation (e.g. Tajfel *et al.*, 1971; Zebrowitz-McArthur, 1982). Thus, infants' ability to categorize faces as attractive or unattractive may serve as a cognitive mechanism that directs infant perception and eventually leads to stereotyping based on facial attractiveness. Given the results from these experiments, this mechanism appears to be well established by 6 months of age.

Model of attractiveness stereotype development

How do attractiveness stereotypes develop? We now turn to a discussion of 'beauty is good' stereotype development based on previous research and theory. Although

speculative, we propose the following sequences and mechanisms as plausible developing parts that culminate in a coherent system of expectations, assumptions and behaviors that together constitute a stereotype. Our goal in such speculation is to stimulate research to either validate or refute our proposal.

Step one: preferences for attractive faces

The evidence for this first stage of stereotype development is robust: Beginning very early in life, infants can discriminate between adult-judged attractive and unattractive faces and show preferences for attractive faces (Langlois *et al.*, 1987, 1991; Samuels & Ewy, 1985; Samuels *et al.*, 1994; Slater, Bremner *et al.*, 2000; Slater, Quinn *et al.*, 2000; Slater *et al.*, 1998).

Step two: categorization of facial attractiveness

As demonstrated in the studies presented in this paper, by 6 months of age, and perhaps even younger, infants group faces into attractive and unattractive categories. The perceptual and categorization abilities of 6-month-olds appear to be advanced enough to allow infants to notice consistencies among faces belonging to the same attractiveness group, which, in turn, may facilitate formation of meaningful associations between categories of attractiveness and other characteristics and attributes. Just as young infants' ability to detect consistencies and differences in facial expressions may allow them to later couple the meaning of an emotion with a particular facial expression (Nelson, 1987), categorization of facial attractiveness might also lead to the rudimentary associations of attractiveness stereotypes.

Step three: association of similarly valenced stimuli

Once infants perceive attractive faces as a group and unattractive faces as a different group, how do they link attractive faces with generally positive information and unattractive faces with more negative information? Although stereotype theory and research posits that categorization, in and of itself, causes assimilation and contrast effects (e.g. Tajfel *et al.*, 1971; Zebrowitz-McArthur, 1982), the question remains as to how *specific* attributes come to be associated with certain attractiveness categories.

By the end of their first year of life, infants are behaving more positively toward attractive people and objects and more negatively toward unattractive people and objects (Langlois *et al.*, 1990). Perhaps infants exhibit differential behaviors based on attractiveness because they associate various types of preferred stimuli with one another, such as attractive faces and pleasant voices, and in turn, associate various types of less preferred stimuli with one another, such as unattractive faces and unpleasant voices (Rubenstein, 2000). Therefore, as a result of these associations, infants also behave in a preferred or positive manner toward attractive people and objects and behave in a less preferred or negative manner toward unattractive people and objects.

Given infants' inability to engage in discussions about traits and behaviors of attractive and unattractive individuals during their first year of development, these associations may stem from observations of social interactions. To illustrate, a recent meta-analysis found that the effect sizes related to differential judgment, treatment and behaviors of attractive and unattractive children and adults are large enough to suggest that activation of the 'beauty is good' stereotype is observable during social interactions (Langlois *et al.*, 2000). Thus, infants may not only witness attractive individuals being treated more favorably than unattractive individuals, they may also observe attractive people acting more positively than unattractive people. Because young infants are quite expert at detecting correlations in their environment and use this ability to learn about such things as the function of objects and the rhythm of language (e.g. Greco, Hayne & Rovee-Collier, 1990; Saffran, Johnson, Aslin & Newport, 1999; Younger, 1992), it is conceivable that infants also use this ability to detect real correlations between the attractiveness of individuals and their behavior within the infant's social environment. Although future research is needed to investigate the specific processes by which social observations affect infant learning and brain development and how these differential judgments, treatments and behaviors initially evolve, it is clear that by the end of the first year of life, the different categories of facial attractiveness and unattractiveness acquire positive and negative meaning.

Step four: elaboration of positive and negative associations into stereotypes

As children develop during their toddler and preschool years, several mechanisms may further strengthen and elaborate initial correlations among positively and among negatively valenced stimuli. First, as previously noted, attractive and unattractive children and adults are treated differently and this differential treatment is large enough to be observable (Langlois *et al.*, 2000). Thus, correlations between attractiveness and positivity are reinforced in the environment.

Second, children acquire the mental capacity to learn and understand more about traits and behaviors as their cognitive and social skills develop, and they may assimilate

this knowledge into existing networks of associations through schematic information processing. There is much evidence to support how schematic processing directs information about other social groups, such as gender (Bigler & Liben, 1992; Koblinsky & Cruse, 1981; Signorella & Liben, 1984), race (Bigler & Liben, 1993) and age (Davidson, Cameron & Jergovic, 1995). Indeed, information about the behavior and traits of attractive and unattractive females is processed schematically. For example, young children have an easier time processing information about females whose appearance and traits are consistent with the 'beauty is good' stereotype than females whose appearance and traits are inconsistent with the stereotype, and these effects do not appear to be moderated by individual differences (Ramsey & Langlois, 2002). As a result of schematic processing, positive traits and behaviors, such as kindness and sharing, continue to be associated with attractiveness whereas negative traits and behaviors, such as being mean or not sharing, continue to be associated with unattractiveness (Dion, 1973; Langlois & Stephan, 1977).

Conclusions

The present studies demonstrated 6-month-olds' ability to: (1) categorize unattractive faces as belonging to a group that does not include attractive faces; and (2) categorize attractive faces as belonging to a group that does not include unattractive faces. This categorization is not due to infants' inability to discriminate among the faces within each group, but rather it is due to infants' ability to detect similarities among the attractive faces and similarities among the unattractive faces, despite noticing individual differences among the faces. Based on the results from these and previous studies, we have proposed a model of attractiveness stereotype development. Obviously, much research is required to confirm or disconfirm our model, but the purpose of the model is to help guide and generate more research in this area. For example, investigating how categories of attractiveness become linked to specific attributes is the least understood step of this model, but we have proposed some ideas for examining this phenomenon. Evidence from such research will help us to learn more about the growth of the 'beauty is good' stereotype and how stereotypical responses are maintained and strengthened throughout development.

Acknowledgements

This work was supported by a grant to Judith H. Langlois from the National Institute of Child Health and Human Development. Parts of this research were previously presented at the 2000 biennial meeting of the International Conference of Infant Studies, Brighton, England.

References

Allport, G.W. (1954). *The nature of prejudice*. Reading, MA: Pegasus.

Banaji, M., & Greenwald, A.G. (1994). Implicit stereotyping and prejudice. In M.P. Zanna & J.M. Olson (Eds.), *The psychology of prejudice: The Ontario symposium on personality and social psychology, 7* (pp. 55–76). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bargh, J.A., & Chartrand, T.L. (1999). The unbearable automaticity of being. *American Psychologist*, **54**, 462–479.

Barrera, M.E., & Maurer, D. (1981). Discrimination of strangers by the three-month-old. *Child Development*, **52**, 558–563.

Bigler, R.S., & Liben, L.S. (1992). Cognitive mechanisms in children's gender stereotyping: theoretical and educational implications of a cognitive-based intervention. *Child Development*, **63**, 1351–1363.

Bigler, R.S., & Liben, L.S. (1993). A cognitive-developmental approach to racial stereotyping and reconstructive memory in Euro-American children. *Child Development*, **64**, 1507–1518.

Davidson, D., Cameron, P., & Jergovic, D. (1995). The effects of children's stereotypes on their memory for elderly individuals. *Merrill-Palmer Quarterly*, **41**, 70–90.

Devine, P.G. (1989). Stereotypes and prejudice: their automatic and controlled components. *Journal of Personality and Social Psychology*, **56**, 5–18.

Dion, K.K. (1973). Young children's stereotyping of facial attractiveness. *Developmental Psychology*, **9**, 183–188.

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, **24**, 285–290.

Eagly, A.H., Ashmore, R.D., Makhijani, M.G., & Longo, L.C. (1991). What is beautiful is good, but . . . : a meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, **110**, 109–128.

Greco, C., Hayne, H., & Rovee-Collier, C. (1990). Roles of function, reminding, and variability in categorization by 3-month-old infants. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, 617–633.

Haaf, R.A., & Diehl, R.E. (1976). Position bias and the paired-comparison procedure in studies of infant attention. *Developmental Psychology*, **12**, 548–549.

Hamilton, D.L., & Trolier, T.K. (1986). Stereotypes and stereotyping: an overview of the cognitive approach. In J.F. Dovidio & S.L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 127–163). Orlando, FL: Academic Press.

Koblinsky, S.A., & Cruse, D.F. (1981). The role of frameworks in children's retention of sex-related story content. *Journal of Experimental Child Psychology*, **31**, 321–331.

Langlois, J.H., Kalakanis, L., Rubenstein, A.J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims of myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, **126**, 390–423.

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., & Rieser-Danner, L.A. (1990). Infants' differential social responses to attractive and unattractive faces. *Developmental Psychology*, **26**, 153–159.

Langlois, J.H., & Stephan, C. (1977). The effects of physical attractiveness and ethnicity on children's behavioral attributions and peer preferences. *Child Development*, **48**, 1694–1698.

Langlois, J.H., & Stephan, C. (1981). Beauty and the beast: the role of physical attractiveness in the development of peer relations and social behavior. In S.S. Brehm, S.M. Kass & F.X. Gibbons (Eds.), *Developmental social psychology: Theory and research* (pp. 152–168). New York: Oxford University Press.

Leinbach, M.D., & Fagot, B.I. (1993). Categorical habituation to male and female faces: gender schematic processing in infancy. *Infant Behavior and Development*, **16**, 317–332.

Miller, C.T. (1988). Categorization and the physical attractiveness stereotype. *Social Cognition*, **6**, 231–251.

Nelson, C.A. (1987). The recognition of facial expressions in the first two years of life: mechanisms of development. *Child Development*, **58**, 889–909.

Ramsey, J.L., & Langlois, J.H. (2002). Effects of the 'beauty is good' stereotype on children's information processing. *Journal of Experimental Child Psychology*, **81**, 320–340.

Rubenstein, A.J. (2000). The ability to form stereotypic associations during infancy: a cognitive look at the basis of the 'beauty is good' stereotype. *Dissertation Abstracts International: Section B: The Sciences & Engineering*, **61** (2-B), 1114. US: University Microfilms International.

Saffran, J.R., Johnson, E.K., Aslin, R.N., & Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, **70**, 27–52.

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.

Signorella, M.L., & Liben, L.S. (1984). Recall and reconstruction of gender-related pictures: effects of attitudes, task difficulty, and age. *Child Development*, **55**, 393–405.

Slater, A., Bremner, G., Johnson, S.P., Sherwood, P., Hayes, R., & Brown, E. (2000). Newborn infants' preference for attractive faces: the role of internal and external facial features. *Infancy*, **1**, 265–274.

Slater, A., Quinn, P.C., Hayes, R., & Brown, E. (2000). The role of facial orientation in newborn infants' preference for attractive faces. *Developmental Science*, **3**, 181–185.

Slater, A., Von der Schulenburg, C., Brown, E., Badenoch, M., Butterworth, G., Parsons, S., & Samuels, C. (1998). Newborn infants prefer attractive faces. *Infant Behavior and Development*, **21**, 345–354.

Tajfel, H., Billig, M.G., Bundy, R.P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, **1**, 149–178.

Younger, B. (1992). Developmental change in infant categorization: the perception of correlations among facial features. *Child Development*, **63**, 1526–1535.

Zebrowitz-McArthur, L. (1982). Judging a book by its cover: a cognitive analysis of the relationship between physical appearance and stereotyping. In A.H. Hastorf & A.M. Isen (Eds.), *Cognitive social psychology* (pp. 149–209). New York: Elsevier.

Received: 17 July 2001
Accepted: 7 March 2003