Impressions of Babyfaced Individuals Across the Life Span

Leslie A. Zebrowitz
Brandeis University

Joann M. Montepare
Tufts University

Four questions were addressed concerning perceptions of babyfaced individuals from infancy to older adulthood: (a) Do perceivers make reliable babyface judgments at each age; (b) does a babyface have the same effects on trait impressions at each age; (c) are the effects of a babyface independent of the effects of attractiveness; and (d) what facial maturity features are associated with babyface ratings, and do these features predict trait impressions? Ratings of portrait photographs revealed that perceivers reliably detect variations in babyfacedness across the life span. Facial measurements revealed that large eyes, a round face, thin eyebrows, and a small nose bridge characterized a babyface. Trait impressions showed a babyface overgeneralization effect at each age: Babyfaced individuals were perceived to have more childlike traits than their maturefaced peers, and this effect was independent of attractiveness.

Considerable evidence indicates that appearance has a significant impact on social perceptions and social interactions, which may have important implications for personality development. Not only is appearance one of the initial qualities that children and adults mention in their descriptions of people (Fiske & Cox, 1979; Livesley & Bromley, 1973), but it also influences the traits they ascribe to others (see Alley, 1988; Bull & Rumsey, 1988; and McArthur, 1982, for reviews of pertinent research). Moreover, people’s physical appearance influences the desire to become acquainted with them (Lyman, Hatelid, & Macurdy, 1981) as well as the nature of their social interactions (e.g., Langlois, 1986; Reis, Nezlek, & Wheeler, 1980; Reis, Wheeler, Spiegel, Kernis, Nezlek, & Perri, 1982; Snyder, Tanke, & Berscheid, 1977). Developmental theorists have argued that these effects of appearance on social perceptions and social interactions may yield self-fulfilling prophecy effects, whereby people with certain appearance qualities develop the traits that are expected (Adams, 1977; Langlois, 1986; Sorrell & Nowak, 1981).

One appearance quality that has been shown to influence social perceptions and social interactions is facial attractiveness. Impressions of attractive people have been characterized as a positive halo effect, whereby those who are physically attractive are perceived more positively than the unattractive on a variety of dimensions (see Adams, 1977; Berscheid & Walster, 1974, for reviews). Moreover, the positive halo effect for facial attractiveness has been demonstrated for reactions to targets ranging from infants to older adults (e.g., Berkowitz & Frodi, 1979; Dion, 1972, 1974; Hildebrandt & Fitzgerald, 1983; Johnson & Pittenger, 1984; Langlois, 1986; Stephan & Langlois, 1984). Finally, facial attractiveness has been shown to have a variety of social interaction consequences (see Bull & Rumsey, 1988, for a review of pertinent research).

A second appearance quality that influences social perceptions and social interactions is facial babyishness. Impressions of babyfaced people have been characterized as an overgeneralization effect, whereby adults with facial structures that ethnologists identify as babyish (e.g., Guthrie, 1976; Lorenz, 1943) are perceived to have childlike traits. More specifically, adults with the babyish facial features of large eyes, thin and high eyebrows, a large cranium, a small chin, and a curved rather than an angular face are perceived as more socially dependent, intellectually naive, physically weak, honest, and warm than their maturefaced peers (Berry & McArthur, 1985; Keating, 1985; McArthur & Apatow, 1983–1984; Zebrowitz-McArthur & Montepare, 1989). A babyface has also been shown to have numerous social interaction consequences. (See Alley, 1988, for relevant reviews and also Berry & Zebrowitz-McArthur, 1988; Zebrowitz, Kendall-Tackett, & Fafel, 1991; Zebrowitz & McDonald, 1991; Zebrowitz, Tenenbaum, & Goldstein, 1991.)

Facial attractiveness and babyishness are most likely to have an impact on personality development if perceivers can reliably identify these qualities in targets across the life span and if an attractive face or a babyface elicits the same trait expectations at various ages. Research has demonstrated that both of these conditions are met for attractiveness (e.g., Berkowitz & Frodi, 1979; Dion, 1972, 1974; Hildebrandt & Fitzgerald, 1983; Johnson & Pittenger, 1984; Langlois, 1986; Stephan & Langlois, 1984). On the other hand, there has been no systematic investigation of these questions regarding a babyface, although there is some evidence for the cross-age generality of the overgeneralization effect for perceptions of competence. Specifically, not only are babyfaced adults perceived to be naive (Berry & McArthur, 1985; McArthur & Apatow, 1983–1984), but also...
babied infants may be perceived as less competent than those with a more mature appearance (Ritter, Casey, & Langlois, 1991). Because the effects of infants’ babyfaceness, presumed age, and attractiveness were not separated in the Ritter et al. study, the question remains as to whether maturefaced and babyfaced infants will be perceived to have different competencies when they are equally attractive and when their ages are known to perceivers.

The present study had four goals. The first goal was to determine whether perceivers make reliable judgments about the facial babyishness of male and female targets across the life span, including infants, preschoolers, young children, adolescents, young adults, and older adults. The second goal was to systematically investigate whether the babyface overgeneralization effect holds true for targets across the life span. The third goal was to determine whether the effects on impressions of a babyface are independent of the effects of attractiveness which, as noted earlier, has also been shown to influence impressions. A final goal was to identify the facial features associated with a babyface for male and female targets across the life span and to determine whether these features influence trait impressions.

Method

Subjects

Three hundred and forty two students who were enrolled in an introductory psychology class participated for partial course credit. Approximately equal numbers of men and women were randomly assigned to rate targets representing one of six age groups: infants (N = 32), preschoolers (N = 48), fifth graders (N = 48), eighth graders (N = 48), young adults (N = 118), and older adults (N = 48). Subjects assigned to the infant group rated both male and female targets, whereas subjects assigned to the other age groups rated either male or female targets. For each group of targets, subjects viewed one of two orders of faces and completed ratings in one of two orders.

Facial Stimuli

The facial stimuli consisted of black-and-white slides depicting the faces of White male and female targets at each of the six age levels. Slides within each age level were selected with the goal of achieving a range of babyface and attractiveness values, and they consequently do not represent a random sample of faces at each age. The slides of infants were drawn from a set of faces used in research by Hildebrandt and Fitzgerald (1979). They consisted of 12 boys and 12 girls ranging in age from 5 to 9 months. The slides of children were made from yearbook photographs of children attending a private school and included 24 boys and 24 girls representing each of three grade levels: preschool, fifth grade, and eighth grade. The slides of the young adults were made from yearbook photographs of 20 men who were college seniors and 20 women who were high school seniors. The slides of the young adult men were those used by Berry and McArthur (1985). The slides of the older adult women were obtained from the Berkeley Growth and Guidance study archives of the Institute of Human Development at the University of California, Berkeley, and included 24 men and 24 women between 52 and 54 years of age.

Subjects who rated infants, children, or older adults were told a specific age or school grade for each target group so that judgments about targets within each group would not reflect variations in their presumed ages. Infants were described as 6 months old; preschoolers were described as 3½ years old; the school grade was given for fifth and eighth graders; and older adults were described as 60 years old. Although the age of the young adults had not been specified when the faces were rated for previous investigations (Berry & McArthur, 1985; Zebrowitz, Tenenbaum, & Goldstein, 1991), subjects in these studies estimated the targets’ ages so that presumed age could be statistically controlled. The real ages of infants were also statistically controlled (see footnote 5).

Two slide orders were generated for each group of targets. Male and female infant faces were presented in blocks such that all of the male faces were presented before the female faces in one order, whereas the sequence of male and female faces was reversed in the other order. For the remaining target groups, one order of slides was a random sequence and the other was the reverse of the first with one exception: The second order of the eighth grade target faces consisted of a random order different from the first order.

Dependent Measures

Impression ratings. Targets were rated on scales reflecting dimensions that had been influenced by variations in facial babyishness in past research with young adult faces (e.g., Berry & McArthur, 1986). Two types of rating scales were used: behavioral scales and bipolar trait scales. For the behavioral scales, subjects indicated on 7-point scales with endpoints labelled definitely no and definitely yes the extent to which the targets were likely to perform certain behaviors. The bipolar trait scales consisted of 7-point scales along which subjects indicated the extent to which the target possessed a particular trait as opposed to its polar opposite.

A detailed description of the rating scales appears in Table 1, where it can be seen that some scales were varied slightly across the different age groups so that they were age appropriate. Also, young adult targets were rated only on trait scales, because these data had been collected in previous research that did not use the behavioral scales. The scales were grouped into five a priori composites paralleling the dimensions used in past research on impressions of babyfaced adults (e.g., Berry & McArthur, 1986): social autonomy, physical weakness, naivete, warmth, and honesty. The last dimension was not represented for infant targets, because it was assumed that they would be rated uniformly high on the deceitful/straightforward scale. For each composite, the behaviors described in Table 1 were scored in a positive direction, as were the second poles of the trait scales.

In addition to rating the targets’ traits and likely behaviors, subjects rated how maturefaced/babyfaced and how unattractive/attractive each target was on 7-point bipolar scales. Subjects who rated the young adults—for whom ages were not supplied—also estimated the age, in years, of each target. Ratings of the targets on all scales could range between 1 and 7 and were coded so that higher scores represented greater social autonomy, physical weakness, naivete, warmth, honesty, babyfaceness, or attractiveness.

Facial characteristics. Five age-related facial characteristics were

---

1 Eighty of these subjects had previously rated targets’ naivete, warmth, and honesty in research by Berry and McArthur (1985). The remaining 38 subjects rated physical weakness and social autonomy, which had not been included in the previous research.
Table 1
Dependent Measures and Reliability Coefficients Within the Six Age Groups

<table>
<thead>
<tr>
<th>Composites/individual impression scales</th>
<th>Infant</th>
<th>Preschool</th>
<th>5th grade</th>
<th>8th grade</th>
<th>Young adult</th>
<th>Older adult</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>F</td>
<td>M</td>
<td>F</td>
<td>M</td>
<td>F</td>
<td>M</td>
</tr>
<tr>
<td>Mature face/babyface</td>
<td>.87</td>
<td>.91</td>
<td>.86</td>
<td>.93</td>
<td>.96</td>
<td>.96</td>
</tr>
<tr>
<td>Not at all attractive/very attractive</td>
<td>.86</td>
<td>.75</td>
<td>.87</td>
<td>.83</td>
<td>.90</td>
<td>.93</td>
</tr>
<tr>
<td>Social autonomy*</td>
<td>.94</td>
<td>.90</td>
<td>.90</td>
<td>.96</td>
<td>.98</td>
<td>.95</td>
</tr>
<tr>
<td>Dependent/ independent*</td>
<td>.83</td>
<td>.67</td>
<td>.66</td>
<td>.79</td>
<td>.92</td>
<td>.88</td>
</tr>
<tr>
<td>Submissive/dominating*</td>
<td>.84</td>
<td>.90</td>
<td>.79</td>
<td>.66</td>
<td>.88</td>
<td>.83</td>
</tr>
<tr>
<td>Not homesick at camp</td>
<td>.81</td>
<td>.89</td>
<td>.93</td>
<td>.90</td>
<td>.87</td>
<td>.78</td>
</tr>
<tr>
<td>Not cry when parents leave the room</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.90</td>
<td>.85</td>
</tr>
<tr>
<td>Not give in to friends</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not listen to parents when they say &quot;Don't touch&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naive*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sophisticated/naive</td>
<td>.77</td>
<td>.89</td>
<td>.65</td>
<td>.86</td>
<td>.91</td>
<td>.90</td>
</tr>
<tr>
<td>Not know right from wrong like an adult</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not know it is naughty to do certain things</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not able to follow complicated instructions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fooled into thinking a toy is gone if hidden behind your back</td>
<td>.76</td>
<td>.83</td>
<td>.75</td>
<td>.87</td>
<td>.74</td>
<td>.83</td>
</tr>
<tr>
<td>Physical weakness*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physically strong/physically weak</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not strong enough to take a toy from a baby of the same age</td>
<td>.89</td>
<td>.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Warmth*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not at all affectionate/very affectionate*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enjoy being hugged</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Honestyf</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. M = male; F = female. * Cronbach alpha coefficients for composite measures. b Scales for infants were very dependent/not very dependent. * Scales for infants were very submissive/not very submissive. a Scales for infants were very naive/very naive and for adults were shrewd/naive. c Scales for young adult men were cruel/kind. d Scales for young adults were dishonest/honest; scales for all others were deceitful/straightforward.

determined from facial measurements made independently by two judges who projected each slide onto a wall and measured the facial features depicted in Figure 1 to the nearest \( \frac{3}{16} \) of an inch. The specific facial characteristics included eye area, eyebrow height, eye separation, chin area, and cranium height. All of these features had been shown in previous research to predict babyface ratings (Berry & McArthur, 1985; Zebrowitz-McArthur & Montepare, 1989).

In addition to the foregoing linear facial measures, three qualitative ratings were made of other facial features that had been shown to be associated with babyface ratings. Specifically, two independent judges rated the following facial characteristics of each target on 7-point bipolar scales: thin eyebrows/thick eyebrows, small nosebridge/large nosebridge, and round face/angular face. Faces were coded as round to the extent that face length and width appeared equal, and the jaw was rounded. They were coded as angular to the extent that face length appeared greater than width, and the jaw was angular. A neutral rating was given to faces that appeared equal in length and width and had an angular jaw or to faces that had greater length than width and a rounded jaw. Ratings of smiling were also obtained for use in the data analyses as a control variable. Specifically, the two judges rated the amount of smiling displayed by each target on 7-point scales with endpoints labeled little smile/big smile.

Procedure

Subjects were informed that the study dealt with how people's physical appearance influences perceivers' impressions of them and that their task was to rate faces of different people on a series of scales. Each face was presented for 8 s with a 2-s interstimulus interval, during which time subjects made their ratings on one scale. After the first rating of all of the faces was completed, subjects viewed the set of target faces a second time and completed the second rating. This procedure was repeated until all of the ratings had been made.
Figure 1. Facial measurements. (The methods for calculating physiognomic measures were as follows: Eyesize was the product of eye height [4; distance between top and bottom of eye at center] and eye width [5; distance between inner and outer corners of the eye]. Eyebrow height was the distance from the pupil center to the eyebrow [3]. Eye separation was the distance between the inner corners of each eye [6]. Chin area was the product of the width at the point halfway between the bottom lip and the base of the chin [9] and the distance from the lower lip center to the base of the chin [10]. Cranium height was the distance from the middle of the pupil to the top of the head minus half of hair length [2]. To correct for variations in face size, all vertical measures were divided by face length [1], horizontal measures [5] and [6] were divided by face width [7], and horizontal measure [9] was divided by face width [8].)

Two orders of rating scales were generated for each target group. For infant targets, half of the subjects rated the targets on a random order of the behavioral and trait scales, and this rating order was reversed for the remaining subjects. For the child and older adult targets, half of the subjects rated the targets on a random order of the set of trait scales followed by a random order of the set of the behavioral scales. The random order of scales within each set was reversed for the remaining subjects, but the trait scales still preceded the behavioral scales. For the young adult targets, for whom only trait scales were used, half of the subjects rated the targets on a random sequence of the scales, and this rating order was reversed for the remaining subjects. For every target group, ratings of attractiveness and babyfaceness were made after all trait and behavioral ratings had been completed, with half of the subjects making babyface ratings first and half of the subjects rating attractiveness first. These ratings were made last so that subjects would not be explicitly aware of variations along these appearance dimensions when making their impression ratings.

Results

Data Base

Facial measures. Pearson correlation coefficients between the two judges' measurements of facial features revealed high reliability except for the size of the infants' nose bridge. The low reliability for this measure (.33 for boys and .44 for girls) reflected a tendency for the judges' ratings to be clustered at the floor of the scale owing to the absence of a noticeable nose bridge in most infants. The average reliability across all target age and sex groups for each facial measure was .80 for eye area, .84 for eyebrow height, .83 for eye separation, .80 for eyebrow thickness, .87 for cranium height, .84 for chin area, .82 for facial roundness, .74 for smiling, and .84 for nosebridge size, excluding infants. The reliable facial measurements and ratings made by the two judges were averaged to yield a single score for each target for use in subsequent analyses.

Cronbach's coefficients computed across subjects' babyface and attractiveness ratings for male and female targets at each age revealed high reliability (see Table 1). Thus, facial babyishness and facial attractiveness reliably differentiate individuals of both sexes and all ages. Having demonstrated acceptable interrater reliability, mean babyface and attractiveness ratings for each face were computed across subjects. These means and standard deviations are presented in Table 2.

To determine whether the variability in babyface ratings provided equivalent potential for the prediction of trait impres-
Table 2
Means and Standard Deviations for Babyface and Attractiveness Ratings Within the Six Age Groups

<table>
<thead>
<tr>
<th>Age group</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Babyface ratings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infant</td>
<td>4.62</td>
<td>0.74</td>
</tr>
<tr>
<td>Preschool</td>
<td>4.15</td>
<td>0.81</td>
</tr>
<tr>
<td>5th grade</td>
<td>4.45</td>
<td>1.32</td>
</tr>
<tr>
<td>8th grade</td>
<td>4.29</td>
<td>0.98</td>
</tr>
<tr>
<td>Young adult</td>
<td>3.93</td>
<td>0.97</td>
</tr>
<tr>
<td>Older adult</td>
<td>3.65</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>$F_{max}(6, 20)^a = 3.22$</td>
<td>$F_{max}(6, 20) = 2.77$</td>
</tr>
<tr>
<td>Attractiveness ratings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infant</td>
<td>4.02</td>
<td>0.77</td>
</tr>
<tr>
<td>Preschool</td>
<td>4.07</td>
<td>0.76</td>
</tr>
<tr>
<td>5th grade</td>
<td>3.88</td>
<td>0.81</td>
</tr>
<tr>
<td>8th grade</td>
<td>3.75</td>
<td>0.74</td>
</tr>
<tr>
<td>Young adult</td>
<td>3.54</td>
<td>0.76</td>
</tr>
<tr>
<td>Older adult</td>
<td>3.37</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>$F_{max}(6, 20) = 1.20$</td>
<td>$F_{max}(6, 20) = 1.65$</td>
</tr>
</tbody>
</table>

Note. For infants, $n = 12$; for young adults, $n = 20$; for all others, $n = 24$.
*df* = the harmonic mean of the cell ns.

Sions across the different age groups, we performed tests of heterogeneity of variance. Hartley's test revealed that the average variance in babyface ratings did not differ significantly across the six age groups for male or for female targets, both $p > .05$ (see Table 2).2

**Impression measures.** Standardized Cronbach's coefficients computed across subjects' impressions ratings for male and female targets at each age revealed high reliability (see Table 1). Having demonstrated acceptable interrater reliability, mean impression ratings for each face were computed across subjects. Next, composite scores for social autonomy, physical weakness, naivete, warmth, and honesty were computed by averaging across mean impression ratings included in the five a priori groupings of the rating scales. As shown in Table 1, the alpha coefficients calculated for each composite were highly significant.3

**Impact of a Babyface on Impressions**

To assess whether previously documented effects of a babyface held true for male and female targets at different age levels, we computed Pearson correlation coefficients between babyface ratings and the composite impression scores within each target group. These coefficients appear in Table 3 and provide strong evidence for cross-age and cross-gender generality in impressions elicited by a babyface. With few exceptions, babyfaced male and female targets in all age groups were perceived as less socially autonomous, physically weaker, more naive, and warmer than their more mature-looking peers. A babyface was also positively correlated with perceived honesty, although this effect did not always achieve significance. Table 3 also includes correlations between babyface and attractiveness ratings, as well as correlations between attractiveness and trait composites for the interested reader.4

Although the correlations shown in Table 3 indicate that facial babyishness and attractiveness were significantly correlated only in the infant age group, we performed a series of forced-entry regression analyses to establish that the effects of a babyface on impressions were independent of any linear attractiveness effects. Degree of smiling was also entered as a predictor in these analyses along with the babyface and attractiveness ratings, because smiling was not held constant in the facial photographs and it has been shown to affect social perceptions (Cunningham, 1986; Keating, Mazur, & Segall, 1981).

The resultant partial correlations between babyface and trait ratings controlling for the linear effects of attractiveness and smiling are presented in Table 3. A babyface continued to predict impressions of low social autonomy, high physical weakness, high naivete, high warmth, and high honesty.3 An exception to the general tendency for the effects of a babyface to hold up in the partial correlations was obtained for ratings of infants' attractiveness,

---

2 Note that the means in Table 2 reveal a linear decrease in babyface ratings with increasing age for both male and female targets, linear $F(1, 122) = 6.60$ and 6.66, respectively, both $p < .01$, as well as marginally higher babyface ratings for female targets, $F(1, 244) = 3.26, p < .07$.

3 There was also a linear decrease in attractiveness ratings with increasing age for male and female targets, linear $F(1, 122) = 13.07$ and 10.59, respectively, both $p < .001$, but no sex differences in attractiveness, $F < 1$.

4 A factor analysis revealed that the five composites could be reduced to two: power, comprised of high social autonomy, low naivete, and high physical strength; and social goodness, comprised of high warmth and honesty. However, all five composites were retained for two reasons: (a) They parallel the five rating dimensions used in past research with adult targets; and (b) although high social autonomy, low naivete, and high physical strength can be statistically grouped, each may relate differently to the appearance variables. For example, past research has shown that attractiveness creates impressions of high competence (i.e., low naivete), but there is no evidence that it creates impressions of physical strength.

5 Although various relationships have been hypothesized between facial babyishness and attractiveness in adults, past research has not yielded consistent effects, and the present findings are therefore not at variance with the existing literature (see Berry, 1991; Cunningham, 1986, 1990; Keating, 1985; McArthur & Apatow, 1983–1984).

6 It should be recalled that subjects were told a specific age or grade level for all targets except young adults, so that babyface effects would not be confounded with variations in presumed age. To assess the effects of presumed age for young adults, the regressions were repeated, adding subjects' mean age estimates for each target as a predictor. Although presumed age occasionally predicted impressions, the babyface effect lost significance in only one instance, as a predictor of men's physical weakness. Because infants' actual ages were known to the researchers, and because small variations in age may have contributed to impression effects, the regression analyses were also repeated for infants, controlling for actual age. The predictive power of babyfacedness remained highly significant for all impressions, and actual age emerged as a significant predictor only for perceived warmth, with which it was negatively related. Thus the effects of babyfacedness cannot be accounted for by variations in presumed or actual age.
warmth. Whereas the zero-order correlations revealed a positive relationship between a babyface and the perceived warmth of infants, this relationship was not significant in the partial correlations because of the positive correlation between attractiveness and babyfaceness. The regression analyses also revealed some independent effects on impressions of attractiveness when the linear effects of babyfaceness and smiling were controlled. Specifically, attractiveness tended to predict impressions of high warmth, high honesty, and low naiveté (see Table 3).

### Structural Components of a Babyface

Having demonstrated that a babyface influences impressions across the life span, we attempted to identify the facial features that gave rise to the perception of a babyface. Because a babyface is a configurual quality that cannot be captured by any single structural feature, we performed regression analyses to determine the predictive power of a constellation of features. Specifically we entered the eight facial structure measures of each infant into hierarchical regression analysis.
ses within target sex using the combined data for preschoolers through older adults. Infants were omitted from these analyses, because one of the facial predictors, nosebridge size, had not been reliable for that group. In these analyses, targets' age group was entered on the first step as a control variable, and the facial measures were entered as predictors using a stepwise procedure on the second step.

The subset of significant predictors that emerged for male faces included a round as opposed to an angular face, large eyes, and thin eyebrows, which together accounted for 32% of the adjusted variance in babyface ratings over and above that explained by age group alone, \( F(4, 111) = 16.01, p < .0001 \). The subset of significant predictors for female faces included a small nose bridge and a round face, accounting for 25% of the adjusted variance in babyface ratings over and above that explained by age group alone, \( F(3, 112) = 15.24, p < .0001 \). Stepwise regressions performed within each target age and sex grouping yielded results quite similar to those from the cross-age regressions. In particular, a rounder face was a significant predictor for male targets in every age group except eighth graders, and a smaller nosebridge was a significant predictor for female targets in every age group except older adults. We incorporated the physiognomic predictors of babyface ratings into structural babyface composites for use in subsequent analyses by converting their values to z scores within their respective target age and sex group and then computing the average of the transformed scores.

Because the measure of nosebridge size was unreliable for infants, a separate regression analysis was performed for these targets in which the seven facial measures excluding nosebridge were entered as predictors of babyface ratings within each target sex. A single predictor emerged for male infants—larger eyes—accounting for 42% of the adjusted variance in babyface ratings, \( F(1, 10) = 8.86, p < .05 \), and eye size was used as a structural babyface index for male infants. The regression equation was not significant for female infants, and consequently the one facial component that was most highly correlated with babyface ratings—a large cranium, \( r(10) = .55, p < .10 \)—was used as a structural babyface index for female infants. These structural babyface composites were more highly correlated with infant babyface ratings than the composites derived from the other age groups with nosebridge deleted. The average intercorrelation between babyface ratings and the structural babyface composite scores across all age groups was .60 for both male and female targets (see Table 4).

Table 4 reports the correlations between trait impressions and the structural babyface composite for each target group. Male targets who were independently determined to have the babyish features of relatively small nosebridges and round faces (or, for infants, a large cranium) were also perceived as less autonomous, weaker, and more naive, though not consistently warmer or more honest than those with larger nosebridges and more angular faces. This pattern of results replicated across all age groups, although the correlations were not significant for infants or older adult women.6

Discussion

The present findings reveal that perceivers do reliably detect variations in facial babyishness across the life span. There are babyfaced babies, babyfaced children, babyfaced adolescents, babyfaced young adults, and babyfaced older adults of both sexes. The present study also determined that babyface ratings have an objective referent in structural features that generalize across the life span. Specifically, larger eyes, thinner eyebrows, and a rounder face are characteristic of more babyfaced males from infancy through older adulthood, whereas a smaller nosebridge and a rounder face are characteristic of more babyfaced females.

The fact that somewhat different facial features emerged as the best predictors of a babyface for males and females may reflect idiosyncrasies in the faces of the targets used in this research. Evidence to indicate that the same facial features may predict babyface ratings for certain male and female faces is provided by the finding that eye size, a significant predictor for male infants as well as for male targets in the cross-age regressions, did emerge as a significant predictor for preschool and fifth grade girls in within-age regressions. Also, nose bridge size, a significant predictor only for female targets in the cross-age regressions, emerged as a significant predictor for fifth and eighth grade boys in within-age regressions. Thus, there is some evidence to suggest that the components of a babyface do generalize across targets of both sexes.

Although there are similarities in the determinants of a babyface across sex, it is also possible that the most salient determinants differ somewhat for male and female faces. In particular, the finding that larger eyes and thinner eyebrows were more strongly associated with babyface ratings for men could reflect the fact that women in our culture often pluck their eyebrows and apply makeup to enlarge the appearance of their eyes, making large eyes and thin eyebrows less salient markers of a babyface for women. Consistent with this suggestion, female targets in the present study had significantly thinner eyebrows and larger eyes than male targets, and both of these effects were

---

6 We also performed regression analyses to determine which, if any, of the measured features predicted attractiveness ratings. For male preschoolers through older adult males, the only significant predictor was large eyes, which accounted for 3% of the adjusted variance. For female preschoolers and adults, the only significant predictor was high eyebrows, which accounted for 6% of the adjusted variance. The prediction of attractiveness was no better in within-age regression analyses than in these cross-age analyses, and it was no better in separate analyses of the ratings of male and female subjects than in the analyses for all subjects combined.
significant only for adults. There were no significant sex differences in any of the other features that were included in the babyface composites.

Perceivers not only agreed as to who was babyfaced, but they also shared common expectations concerning the traits possessed by these targets: The babyface overgeneralization effect held true for male and female targets across the life span. With the linear effects of attractiveness and smiling held constant, more babyfaced male and female targets from infancy to older adulthood were perceived as less socially autonomous and more naive than their more maturefaced peers. Specifically, the more babyfaced targets were perceived as more dependent, more submissive, more naive and easily fooled, less likely to know right from wrong, and less able to follow complicated instructions. More babyfaced individuals were also perceived as physically weaker, albeit only at a marginal level of significance for infants and older adults.

Babyfaced male and female targets not only were perceived as more childlike in the domains of social, intellectual, and physical power, but they were also perceived as warmer and more honest. Specifically, they were rated as more affectionate and desirous of being hugged, kinder, warmer, more straightforward, and more honest than their maturefaced peers, although the last effect cannot be generalized to infants because their perceived honesty was not assessed. Also, it should be noted that the effects of a babyface and attractiveness on impressions of infants' warmth were not independent of each other. Finally, the strong correlations between trait ratings and the structural babyface composite demonstrate that the correlations between trait and babyface ratings do not merely reflect carryover effects from subjects' trait ratings to their babyface ratings. Rather, these trait ratings are evoked by a specific constellation of infantile facial features that can be measured objectively.

Although the foregoing effects of a babyface on social perceptions held true with the linear effects of attractiveness controlled, the question remains as to the joint effects of various degrees of facial babyishness and attractiveness. A recent study by Berry (1991) simultaneously manipulated three levels of facial babyishness and attractiveness in young adult faces and found that the perception of highly babyfaced men and women as less powerful than their least babyfaced peers held true at all levels of attractiveness. However, the perception of highly babyfaced men as more sincere than those low in facial babyishness held true at all levels of attractiveness. Similarly, the perception of highly babyfaced women as more sincere did not hold true for women low in attractiveness. Given this evidence of an interactive effect of facial babyishness and attractiveness, it would be interesting to investigate their joint effects on impressions across the life span as Berry (1991) did at a single age level.

In addition to providing evidence for cross-age generality of the babyface overgeneralization effect, the present results reaffirm the cross-age generality of the attractiveness halo effect. With babyface ratings and smiling held constant, more attractive male and female targets from preschoolers to older adults tended to be perceived as warmer and more honest than their less attractive peers. In contrast to these findings, attractiveness had no effect on perceptions of the targets' social autonomy or physical weakness. This could reflect the fact that these traits do not have the strong evaluative connotations of warmth and honesty and consequently do not engage the attractiveness halo effect. On the other hand, being physically strong, dominant,
and independent are certainly positive traits in our society, particularly for men, which suggests that the attractiveness halo may be more specific than previously assumed. Finally, attractiveness, like a babyface, had an independent effect on perceptions of naivete, albeit in the opposite direction. Whereas a babyface increased impressions of naivete, an attractive face diminished such impressions, which is consistent with past evidence for an attractiveness halo effect on judgments of intellectual competence.

The pattern of results on the naivete measures for infants clarifies the contribution of facial maturity and attractiveness to the findings of Ritter et al. (1991). These authors found that 6-month-old infants who were both relatively young looking (i.e., babyfaced) and attractive were rated less positively on specific measures of cognitive skills, such as uncovering a hidden toy, but more positively on global measures of cognitive competency. To explain these findings, Ritter et al. suggested that ratings of specific skills were influenced by infants' facial maturity, whereas global ratings were influenced by their attractiveness. The partial correlations in the present study support this argument: Infant babyfacedness decreased the positivity of competence (naivete) ratings, whereas infant attractiveness increased the positivity of these ratings, at least for male infants. Thus, attractiveness has the same halo effect on the perceived competence of infants as on the perceived competence of individuals from other age groups so long as facial babyishness is controlled.

The present findings clearly demonstrate that facial babyishness, like attractiveness, has the potential to influence personality development across the life span. Individuals of all ages are reliably differentiated along this dimension. Moreover, a babyface elicits the same strong trait expectations, regardless of the target's age, and these expectations may result in their being treated in ways that elicit confirmatory behavior. Indeed, recent research reveals that parents assign less intellectually demanding tasks to babyfaced children and also respond differently to their misdeeds (Zebrowitz, Kendall-Tackett, & Fafel, 1991). Whereas the existent cross-sectional research reveals strong effects of facial babyishness on social perceptions and social interactions, longitudinal research investigating the stability of facial babyishness across the life span is needed to further clarify its likely impact on personality development.

It is noteworthy that Zebrowitz, Kendall-Tackett, and Fafel (1991) selected babyfaced and maturefaced children on the basis of the ratings provided by undergraduates in the present study, and they found that parents' facial maturity ratings showed strong agreement with those of the undergraduates. This demonstrates that the impressions documented in the present study generalize beyond the population of college students. Additional research also indicates that impressions of babyfaced individuals generalize across perceiver's race and age (McArthur & Berry, 1987; Montepare & Zebrowitz-McArthur, 1989). Finally, although the faces investigated in the present study were selected to represent a range of values on the babyface dimension, there is reason to believe that the results would generalize to a more random sample of faces. More specifically, a recent investigation revealed a significant impact of litigant facial maturity on adjudications in small claims court, where over 1,000 faces were represented (Zebrowitz & McDonald, 1991). Thus, the effects of a babyface across the life span that were documented in the present study should hold true across a wide range of perceivers and targets.

References

References


Received June 15, 1990
Revision received August 9, 1991
Accepted October 7, 1991