

The Influence of Children's Facial Maturity on Parental Expectations and Punishments

LESLIE A. ZEBROWITZ, KATHLEEN KENDALL-TACKETT, AND JODIE FAFEL

Brandeis University

It was hypothesized that the perception of maturefaced children as more able to follow complicated instructions, more likely to know right from wrong, more shrewd, and physically stronger than their babyfaced peers would lead parents to assign more demanding tasks to these children and to judge their misbehavior more harshly. Study 1 revealed that parents allocated more cognitively demanding, but not more physically demanding, chores to maturefaced 11 year olds depicted in photographs than to babyfaced children of the same age and attractiveness. Study 2 revealed that parents perceived the misbehaviors of maturefaced 4- and 11-year-old children as more intentional than those of their babyfaced peers, an effect that was significant only when parents judged children of the opposite sex. Study 2 further revealed that, with perceived intentionality held constant, a baby-face mitigated the severity of punishment recommended for relatively serious infractions by preschoolers, while increasing it for older children. The latter finding was discussed in light of other evidence that people react negatively to the disconfirmation of their benign expectations regarding babyfaced individuals, and that parents perceived the misbehaviors as more unexpected for 11 year olds than 4 year olds. © 1991 Academic Press, Inc.

Considerable evidence indicates that children's appearance may have a significant impact on their development. A reliance on physical appearance in descriptions of other people occurs at early age (Livesley & Bromley, 1973) and continues into adulthood (Fiske & Cox, 1979; Liggett, 1974). Moreover, the appearance of children, like that of adults, influences impressions of their traits and abilities as well as their social interactions. (See Adams, 1977; Alley, 1988; Bull & Rumsey, 1988; Berscheid & Walster, 1974; Langlois, 1986; and Sorrell & Nowak, 1981, for reviews of pertinent research.

Most of the research concerning children's appearance has focused on

This research was supported by NIMH Grant BSR 5 RO1 MH42684 to the first author. Portions of this research were presented at the meeting of the Society for Research on Child Development, Kansas City, April, 1989. Correspondence regarding this work may be addressed to Leslie A. Zebrowitz, Department of Psychology, Brandeis University, Waltham, MA 02254.

physical attractiveness, which has been shown to have a positive influence both on impressions of children and on their social interactions. For example, children as young as preschool age tend to rate more attractive peers as friendlier, nicer, smarter, less scary, and less likely to start fights (Dion, 1973; Dion & Berscheid, 1974; Langlois & Stephan, 1977; Vaughn & Langlois, 1983). Adults, including teachers, also attach a positive halo to attractive children (Clifford & Walster, 1973; Dion, 1972; Stephan & Langlois, 1984). Attractive children not only are perceived more positively, but also they are likely to experience more positive social interactions. For example, they are more apt to be chosen as playmates than their less attractive peers (Adams, Hicken, & Salehi, 1988; Dion & Berscheid, 1974) and their misdeeds are punished less harshly by adults (Berkowitz & Frodi, 1979).

A second appearance quality that may affect children's development is their facial maturity. It has been shown that children with facial structures that ethologists identify as babyish (e.g., Lorenz, 1943) are perceived to have more childlike traits than their more maturefaced peers. In particular, more babyfaced children ranging in age from 6 months to early adolescence are perceived as more socially dependent, intellectually naive, physically weak, honest, and warm than their more maturefaced peers. They are also judged as less likely "to be able to follow complicated instructions" and "to know right from wrong the way an adult does." Moreover, all of these effects held true when the children's age was held constant and made known to perceivers, and when their attractiveness was statistically controlled (Zebrowitz & Montepare, 1991).

Like variations in facial attractiveness, variations in children's facial maturity may have social consequences beyond impressions. However, most of the research investigating social consequences of facial maturity has focused on adults. This work has revealed that babyfaced and maturefaced adults are given jobs which require traits that are stereotypically associated with their appearance (Zebrowitz, Tenenbaum, & Goldstein, 1991), and that maturefaced defendants are more apt than the babyfaced to be perceived as committing intentional offenses (Berry & Zebrowitz-McArthur, 1988; Zebrowitz & McDonald, 1991). The latter finding is consistent with evidence that babyfaced people are perceived as more honest and naive than the maturefaced—and thus less motivated or able to plan intentional harm (Berry & McArthur, 1985; McArthur & Apatow, 1983–1984; McArthur & Berry, 1987; Zebrowitz & Montepare, 1991). It also has been found that when defendants admit to an offense, the more babyfaced are punished less severely for negligent acts, presumably because they looked like they "couldn't help it," whereas they are punished more severely for premeditated acts, perhaps because judges were particularly disturbed by behaviors that violated their benign expectations (Berry & Zebrowitz-McArthur, 1988).

Only one study has documented social interaction consequences of facial maturity for children. McCabe (1984) found that physically abused children between the ages of 2 and 15 had more mature craniofacial proportions than nonabused controls matched for age, sex, and ethnic background. There are, of course, many contributing factors to child abuse, and a 'mature' facial appearance by itself will not lead to abuse. Moreover, the effects of facial attractiveness and facial maturity cannot be differentiated in McCabe's results. Nevertheless, differences in perceptions of the traits and capabilities of mature- versus babyfaced children may have been one contributing factor. In particular, one can postulate the following sequence of events. First, parents may assign more demanding tasks to maturefaced children just as they do taller children (Brackbill & Nevill, 1981; Eisenberg, Roth, Bryniarski, & Murray, 1984). Second, parents may become more upset when maturefaced children fail to perform as expected, if, like maturefaced adults, their actions are more apt to be perceived as intentional (Berry & Zebrowitz-McArthur, 1988; Zebrowitz & McDonald, 1991). And, finally, parents may punish maturefaced children more severely for their misbehavior.

The purpose of the present research was to test the foregoing hypotheses regarding the effects of children's facial maturity on the reactions of a random sample of parents, as opposed to abusive parents. More specifically, Study 1 investigated the influence of children's facial maturity on the difficulty level of the chores assigned to them, and Study 2 investigated its influence on reactions to their misbehavior.

STUDY 1

Method

Subjects

Thirty-two white middle class parents of 11 year olds (16 mothers, 16 fathers) were recruited from a University Sunday School and other community groups. Subjects of each sex were randomly assigned to one of two orders of target sex and to one of two orders of chore allocations.

Targets

Black and white slides depicting school yearbook photographs of 24 fifth grade boys and 24 fifth grade girls were rated by 12 male and 12 female undergraduates on 7-point bipolar scales assessing perceived facial maturity (babyface/matureface) and attractiveness (very unattractive/very attractive) as well as several other impressions. Each slide depicted only the child's face. All of the children were dark-haired Caucasians, and those with glasses, birthmarks, or any other distinctive facial attribute were excluded. Based on these ratings, which were highly reliable ($\alpha = .96$ for facial maturity and $.92$ for attractiveness), two maturefaced

and two equally attractive babyfaced children of each sex were selected to serve as targets in the present study. Craniofacial measurements made by two independent judges with an average interrater reliability of .82 further established that the babyfaced targets had more infantile craniofacial proportions. More specifically, a composite of craniofacial features that were found by Zebrowitz & Montepare (1991) to predict facial maturity ratings across the lifespan (eye size, eyebrow thickness, facial roundness, and nosebridge size) showed more mature values for the targets designated as maturefaced than for those designated as babyfaced (mean z scores for features in the composite were .43 and $-.74$ for babyfaced and maturefaced targets, respectively, with a standard deviation of .60 for the entire set of 48 faces).

Dependent Measures

Forty-four tasks taken from the Expected Achievement Index created by Brackbill and Nevill (1981) were employed in the present study. Brackbill and Nevill (1981) had developed age norms for each task by asking parents to indicate at what age a child would be able to carry out the task at a summer camp without the supervision of a counselor. They had also established whether or not each task was perceived to require physical strength. Based on these data, three groups of tasks were selected. A group of eight 'physical' tasks had been judged to require physical strength; a group of eight 'cognitive' tasks had been judged to require varying levels of cognitive ability, but not physical strength; and a group of 28 'mixed' tasks had been judged as not specifically cognitive or physical (see Table 1).

Procedure

Subjects were told that the research was concerned with the factors that influence decisions made by parents, teachers, camp counselors, and other adults concerning the appropriateness of certain chores to a child's capabilities. They were further told that sometimes there is very little information to use when assigning chores, and that the present study sought to determine whether or not parents show agreement in their chore assignments in a minimal information situation, when only a child's photograph and age is available.

Each parent assigned chores to the four male targets and to the four female targets, with order of target sex counterbalanced across subjects. The four photographs within each sex were arranged in a square formation with the position of the babyfaced and maturefaced children counterbalanced across subjects. In the center of the square was the label '11 year olds.' Parents allocated the three groups of chores to the four children in a set by writing the child's identification letter (A, B, C, D) next to the chore description with the constraint that an equal number of chores

TABLE 1
 RATED AGE APPROPRIATENESS OF TASKS ASSIGNED TO TARGETS IN STUDY 1*

Tasks	Age level
Cognitive	
1. Edit camp newspaper	14.05
2. Be cashier in camp store	13.52
3. Write articles for camp newspaper	11.30
4. Gather information for camp newspaper	11.24
5. Answer camp telephone and deliver messages	10.44
6. Help plan menus	10.44
7. Lead grace at table	8.34
8. Make simple foods (cereal, sandwich)	8.13
Physical	
1. Vacuum swimming pool	12.66
2. Launch canoes	12.38
3. Put blankets and saddles on horses	11.59
4. Mow the lawn	11.58
5. Scrub coping around edge of pool	11.53
6. Scrub out showers	10.90
7. Mop the kitchen floor	9.84
8. Rake the leaves	8.15
Mixed	
1. Act as lifeguard for younger children	15.28
2. Operate kiln	14.18
3. Spray bushes for insects	13.72
4. Varnish the canoe	13.44
5. Sharpen knives for wood whittling	13.11
6. Prepare all of a meal	12.92
7. Sharpen kitchen knives	12.93
8. Put clay objects in kiln	11.95
9. Pack equipment for trip	11.13
10. Comb the horses	11.06
11. Skim swimming pool to remove leaves	10.40
12. Wash windows	10.41
13. Heat up prepared foods	9.96
14. Hammer a nail to put up decorations for party	9.91
15. Hang wash on line	9.51
16. Water the yard	8.77
17. Clean up dishes after evening snacks	8.71
18. Fold laundry	8.64
19. Put dishes in dish washer	8.64
20. Prepare simple refreshments (tea)	8.17
21. Set dining room tables	8.02
22. Assist in preparation of simple foods	7.86
23. Dry the dishes	7.82
24. Sweep the floor	7.81
25. Clear the dishes off table after meal	7.51
26. Feed camp pets	7.38
27. Set table	7.17
28. Dust furniture in cabin	7.15

* Tasks and mean age appropriateness are taken from Brackbill and Nevill (1981).

from the group be assigned to each child. The order of allocation of the 8 physical and the 8 cognitive tasks was counterbalanced across subjects while the group of 28 mixed tasks was always the last to be allocated.

After completing the chore allocations, parents were asked to rate the facial maturity and attractiveness of each target on the same 7-point bipolar scales on which they had previously been rated by college undergraduates. These manipulation checks were included to ensure that the children viewed as 'maturefaced' and 'babyfaced' by college students were similarly viewed by the parents in this study, and that there were no differences in parents' judgments of the attractiveness of these two groups of children.

Results

The mean age appropriateness of the cognitive, physical, and mixed tasks assigned to each target were computed and subjected to three separate analyses of variance with exemplar faces (2) nested within facial maturity (2) and target gender (2) as within subjects variables and subject gender (2), task order (2), and target gender order (2) as between subjects variables. Parallel analyses were also performed on ratings of facial maturity and attractiveness.

Manipulation Checks

Targets designated as 'maturefaced' on the basis of undergraduates' ratings were also rated by parents as more maturefaced (5.41) than those designated as 'babyfaced' (2.52), $F(1, 24) = 268.08$, $p < .0001$. In addition, parents' ratings of the attractiveness of the maturefaced (4.34) and the babyfaced (4.07) targets did not differ significantly, $F(1, 24) = 2.28$, $p > .10$, indicating that variations in attractiveness could not account for variations in the assignment of chores to babyfaced and maturefaced targets.

Task Assignments

As predicted, the mean age level of the cognitive tasks assigned to maturefaced children (11.56) was significantly higher than that of the cognitive tasks assigned to babyfaced children (10.28), $F(1, 24) = 32.15$, $p < .0001$. Similarly, parents assignments of the mixed tasks, which were not exclusively cognitive or physical, yielded a higher mean age level for maturefaced children (10.88) than for babyfaced children (9.14), $F(1, 24) = 42.20$, $p < .0001$. Contrary to expectation, the mean age level of the physical tasks assigned to maturefaced children (11.20) did not differ significantly from that of the physical tasks assigned to babyfaced children (10.96), although the means were in the predicted direction, $F(1, 24) = 2.10$, $p = .16$. None of the interaction effects in these analyses was significant.

Discussion

Consistent with past evidence that taller children are perceived as capable of performing more difficult tasks than shorter children of the same age (Brackbill & Nevill, 1981; Eisenberg et al., 1984), maturefaced children were also perceived as more capable than their equally attractive babyfaced peers. Parents were strongly predisposed to assign more demanding tasks to maturefaced than to babyfaced children even when they were informed that the children were all 11 years old. Moreover, the significant effects of facial maturity held true regardless of the children's gender and they generalized across four babyfaced and four maturefaced targets.

The higher expectations for maturefaced children held true for tasks that varied in their cognitive requirements, which is consistent with the tendency to perceive more maturefaced children as less naive and more able to follow complicated instructions (Zebrowitz & Montepare, 1991). However, the effect of facial maturity was not significant for tasks that varied primarily in their physical requirements. This result is somewhat surprising, since research has revealed that maturefaced children as well as adults are perceived as physically stronger than their babyfaced age-mates (McArthur & Apatow, 1983–1984; Zebrowitz & Montepare, 1991). On the other hand, this result is consistent with a recent study by Ritter et al. (1991). Although these authors did not explicitly vary facial babyishness, they did demonstrate that parents overestimated the abilities of older looking infants on communication and cognitive tasks but not on physical tasks.

It should be noted that a crucial difference between the Ritter et al. study and the present investigation was that subjects in the former study were not told the ages of the infants whom they judged, and they perceived the older-looking infants to be 1.5 months older than their actual age as well as significantly less attractive than the younger-looking infants. Thus, the contribution to parental expectations of a child's perceived age, attractiveness, and facial maturity cannot be separated in the Ritter et al. study. Although parents in the present study may have assumed that the maturefaced children were some months older than the babyfaced, an age difference of a few months is less significant at Age 11 than in infancy. Moreover, the present investigation established that the effects of facial maturity were independent of attractiveness.

STUDY 2

Having demonstrated that parents expect more from maturefaced children than their babyfaced age-mates, the question remains as to whether children's facial maturity also influences reactions to their misbehavior. Three lines of research suggest that such effects may indeed occur. First,

as noted above, adults' facial maturity influences such reactions. When responsibility is denied, babyfaced adults are less apt to be viewed as committing intentional offenses, and when responsibility is admitted, babyfaced adults are punished more severely for intentional offenses and less severely for negligent ones (Berry & Zebrowitz-McArthur, 1988, Zebrowitz & McDonald, 1991). Second, research has revealed that children's appearance can in fact influence reactions to their misbehavior: the transgressions of unattractive children are judged as more severe (Dion, 1972), and these children are more harshly punished (Berkowitz & Frodi, 1979). Finally, research has revealed that parental responses to children's misbehaviors vary with the age of the child. Specifically, as children grow older, their negative acts are perceived as increasingly intentional and these misdeeds elicit reports of increasingly negative affect (Dix & Grusec, 1985; Dix, Ruble, Grusec, & Nixon, 1986). While this research has influenced reactions to misbehavior by describing the children as varying in age, the maturity of appearance of children at a given age may have parallel effects. Indeed, Eisenberg et al. (1984) found that mothers assigned more punishment for a hypothetical wrongdoing to taller than to smaller preschool girls, holding constant their perceived age.

In light of the foregoing findings, Study 2 investigated parents' reactions to the misbehaviors of babyfaced and maturefaced boys and girls in two different age groups—4 and 11 year olds. It was predicted that parents would perceive the misbehavior of maturefaced or 11-year-old children as more intentional and more upsetting than the same behavior by babyfaced or 4-year-old children, respectively. Predictions regarding punishment severity were less clear cut. Some research (e.g., Eisenberg et al., 1984; McCabe, 1984) suggested that parents would recommend more severe punishment for the misbehavior of older or maturefaced children, whereas other research (e.g., Berry & Zebrowitz-McArthur, 1988) suggested that the maturefaced might be punished more severely for some behaviors and less severely for others.

Method

Subjects

Sixty-four white middle class parents of 10 to 12 year olds (32 mothers, 32 fathers) were recruited from a University Sunday School program, and from other community groups. Subjects were randomly assigned to one of two target ages (4 or 11) and to one of 8 questionnaire forms representing the counterbalanced face-misbehavior combinations.

Targets

The babyfaced and maturefaced 11-year-old targets were the same as those employed in Study 1. Four-year-old targets were selected in a par-

allel fashion. Specifically, black and white slides depicting school yearbook photographs of 24 preschool boys and 24 preschool girls were rated by 12 male and 12 female undergraduates on 7-point bipolar scales assessing perceived facial maturity (babyface/matureface) and attractiveness (very unattractive/very attractive). Each slide depicted only the child's face. All of the children were dark-haired Caucasians, and those with distinctive facial attributes were excluded.

Based on the undergraduates' ratings, which were highly reliable ($\alpha = .90$ for facial maturity and $.84$ for attractiveness), two maturefaced and two equally attractive babyfaced children of each sex were selected to serve as targets. Craniofacial measurements made by two independent judges with an average interrater reliability of $.82$ further established that the babyfaced targets had more infantile craniofacial proportions as reflected in a composite of features (eye size, eyebrow thickness, facial roundness, and nosebridge size) that predicted facial maturity ratings across the lifespan (mean z scores for features in the composite were $.42$ and $-.70$ for babyfaced and maturefaced targets, respectively, with a standard deviation of $.60$ for the entire set of 48 faces).

Vignettes

Eight brief vignettes describing children's misbehaviors were patterned after those developed by Dix et al. (1986). Four vignettes described relatively serious acts of commission (two involved hurting another child, two involved destroying property), and four described less serious acts of omission (not helping, not sharing). The vignettes were deliberately worded in an ambiguous fashion so that it was difficult to tell whether the act described was intentional or accidental. Each vignette used a "gender-neutral" child's name (e.g., Chris, Randy, Gerry), and the vignettes, pictures, and names were counterbalanced so that each picture was matched with each vignette and each name (see Table 2).

Dependent Measures

Each of the eight vignettes was followed by a one-page questionnaire with four questions. Question 1 asked parents to rate on a 7-point scale how they would feel if they witnessed the act described, with the poles labeled "not at all upset" and "very upset." Question 2 asked them to rate on a 7-point scale how likely or unlikely it was that the action was intentional, with the poles labeled "very unlikely" and "very likely."

Questions 3 and 4 asked parents to give a global rating of the punitiveness they recommended on a 7-point scale with "the most severe punishment I would use" and "no punishment at all" as the poles. These questions were identical except that one asked them to assume the actions were intentional, whereas the other asked them to assume they were accidental. This procedure was adopted so that the measure of punishment

TABLE 2
VIGNETTES DESCRIBING CHILDREN'S MISDEEDS

Acts of commission	
1.	Gerry was playing soccer. Gerry went for the ball and, in doing so, kicked another child in the stomach.
2.	A child was playing with an electronic game. Pat sat down next to the child and grabbed the game, cutting the child's hand.
3.	Jamie asked to borrow a nature book from a bunkmate. When Jamie returned the book, some of the pages were torn.
4.	Several children were working with ink markers and paper in the art room. When the teacher gathered the materials, she noticed ink scribbles written on the table where Terry had been sitting.
Acts of omission	
1.	Lee was supposed to feed the rabbits and fish in the Nature Hut. At feeding time, Lee went to another cabin to play a game.
2.	Chris and another child were making ornaments with bread dough. Chris left for lunch without cleaning up, leaving bread dough all over the table and chairs.
3.	A group of children were playing with Legos in the rec room. Randy took most of the Lego plates from the kit and the other children couldn't build what they wanted to.
4.	After lunch, the campers were served cookies for dessert. Andy took several cookies from the dessert plate and some other campers didn't get any.

severity would not be confounded with perceived intentionality. Whether parents were asked to assume accidental or intentional first was counter-balanced across conditions.

Finally, as a check on the facial appearance manipulation, parents were shown all of the faces they had seen and asked to rate their attractiveness and babyfacedness on 7-point scales. Attractiveness was rated first by half of the parents and babyfacedness was rated first by the remainder. The faces were presented in one of two random orders in a separate mini-album, where each page had only the child's photograph and an identification number.

Procedure

Each parent was given a notebook containing a total of eight vignettes, either about 4-year-old children (4 females, 4 males) or 11-year-old children (4 females, 4 males). Within each age group, there were pictures of two babyfaced and two maturefaced children of each sex. The vignettes for the two age groups were identical. In the notebook, each vignette with a photograph attached was on the left-hand page, while the dependent measure questionnaire for that vignette was on the right.

Prior to responding to the vignettes, parents were given written instructions which stated that the purpose of the study was to compare parents' judgments about children's actions with those of young adults who are not parents. They were further told that, to make these comparisons,

descriptions of children's behavior had been taken from the journals of junior counselors at a day camp who were instructed to keep a log of the children's worrisome behaviors in order to get advice on how to deal with them. Parents were informed that it was not known whether the actions were accidental or intentional, and that they would be asked to indicate how they would handle each incident both if it were accidental and if it were purposeful. Finally, they were told that all of the children about whom they would read were 4 (or 11) years old. After responding to all of the vignettes, parents rated the children's faces. They were then debriefed as to the purposes of the study, and thanked for their participation.

Results

Data Analysis

Ratings of babyfacedness, attractiveness, intentionality, and upset were subjected to separate analyses of variance with facial maturity (2), misbehavior type (2), and target gender (2) as within subjects factors and target age (2) and parent gender (2) as between subjects factors. The severity of recommended punishment was analyzed by a six-factor analysis of variance which added the within subjects variable of assumed intentionality (2) of the misbehavior (accidental or intentional). The Latin Square counterbalancing of specific faces with specific misbehaviors across subjects precluded incorporating exemplar face as an independent factor in the design, since it was not crossed with type of misbehavior for any given subject. However, the results of Study 1 indicated that effects of facial maturity do generalize across the exemplars of 11-year-old faces. Only those main effects and interactions involving target facial maturity or age will be reported, since other effects are not germane to the hypotheses under investigation.

Manipulation Checks

Perceived babyfacedness. As predicted, there was a main effect for facial maturity, with maturefaced children ($M = 5.16$) perceived as significantly more maturefaced than babyfaced children ($M = 3.05$), $F(1, 60) = 224.66$, $p < .0001$. Although there was a target age \times facial maturity \times target gender interaction, $F(1, 60) = 9.53$, $p < .01$, t test comparisons revealed that the predicted facial maturity effect held true for boys and girls at each age level, all $ps < .001$. Similarly, although there was a parent gender \times target gender effect, $F(1, 60) = 6.29$, $p = .01$ with fathers rating boys as more maturefaced ($M = 4.09$) than mothers did ($M = 3.98$) and mothers rating girls as more maturefaced ($M = 4.34$) than did fathers ($M = 4.01$), this did not qualify the effect of the facial maturity manipulation, which was significant when mothers and fathers rated both boys and girls.

Perceived attractiveness. A significant target age \times facial maturity \times target gender interaction, $F(1, 60) = 6.94, p = .01$, revealed that the only significant facial maturity effect was for 4-year-old girls: maturefaced 4-year-old girls ($M = 5.50$) were judged to be significantly more attractive than babyfaced girls ($M = 4.97$), $t(63) = 6.04, p < .001$. The babyfaced and maturefaced children in all other groups did not differ in attractiveness, all $ps > .25$. Since in no groups were maturefaced children perceived as less attractive than babyfaced children, the predicted effect of less punitive treatment of the babyfaced cannot be attributed to an attractiveness halo. On the other hand, a failure to find more punitive treatment of maturefaced 4-year-old girls could reflect their greater attractiveness.

Perceived Intentionality

Although the trend was in the predicted direction, the misbehaviors of 11 year olds were not perceived as significantly more intentional ($M = 3.51$) than those of 4 year olds ($M = 3.16$), $F(1, 60) = 2.45, p = .12$. A subject gender \times target gender \times facial maturity interaction, $F(1, 60) = 5.47, p = .02$, and a subject gender \times target age \times facial maturity interaction, $F(1, 60) = 4.36, p = .04$, revealed that the predicted tendency to perceive the misbehaviors of maturefaced children as more intentional ($M = 3.49$) than those of babyfaced children ($M = 3.18$), $F(1, 60) = 7.89, p < .01$, depended on the sex of the parent as well as the age and sex of the child.

Mothers showed a cross-sex effect. Although they did not differentiate between girls who were babyfaced ($M = 3.36$) and maturefaced ($M = 3.41$), $t < 1$, they did perceive the misbehaviors of maturefaced boys ($M = 3.88$) as more intentional than those of babyfaced boys ($M = 2.88$), $t(63) = 4.65, p < .001$, as predicted, and these effects held true for both 4 and 11 year olds, both $ps < .01$. Fathers also showed a cross-sex effect, but it was significant only for 11 year olds. They perceived the misbehaviors of 11-year-old maturefaced girls ($M = 3.41$) as more intentional than those of 11-year-old babyfaced girls ($M = 2.59$), $t(31) = 2.89, p = .01$, whereas they did not differentiate between baby- and maturefaced 11-year-old boys or between baby and maturefaced 4 year olds of either sex, all $ts < 1$.

Degree of Upset

Contrary to prediction, there was no significant main effect for target age on parents' ratings of how upset they would be by a misbehavior ($Ms = 3.49$ and 3.58 for 4 and 11 year olds, respectively), $F < 1$, nor was there a significant main effect for target facial maturity ($Ms = 3.53$ and 3.54 for baby and maturefaced, respectively), $F < 1$. There was, however, a significant target age \times facial maturity \times type of misbehavior interaction, $F(1, 60) = 4.93, p = .03$. Contrary to prediction, parents

indicated that they would be less upset by the commissions of maturefaced ($M = 3.53$) than babyfaced ($M = 3.88$) 11 year olds, $t(63) = 2.31, p = .02$, whereas there was no significant effect for 4 year olds, although the means were in the predicted direction ($M_s = 3.81$ and 3.55 for mature and babyfaced targets, respectively), $t(63) = 1.78, p = .08$. For acts of omission, the effects of facial maturity on degree of upset were also not significant, $ps > .10$.

Recommended Punishment Severity

There was no significant effect for target age on the recommended severity of punishment ($M_s = 3.13$ and 3.05 for 4 and 11 year olds respectively), $F < 1$. There was also no main effect for target facial maturity ($M_s = 3.07$ and 3.11 for baby and maturefaced respectively), $F < 1$. However, there was a significant subject gender \times facial maturity effect, $F(1, 60) = 4.82, p < .03$. Fathers recommended more severe punishment for maturefaced ($M = 3.14$) than babyfaced ($M = 2.97$) children, $t(127) = 2.90, p < .01$, whereas mothers did not ($M = 3.07$ and 3.16 for maturefaced and babyfaced, respectively), $t(127) = 1.53, p > .10$.

A target age \times facial maturity \times type of misbehavior interaction, $F(1, 60) = 4.76, p < .03$, revealed that when the misbehaviors involved acts of commission, parents recommended more severe punishment for maturefaced ($M = 3.44$) than babyfaced ($M = 3.16$) 4 year olds, $t(63) = 3.17, p < .04$, while recommending *less* punishment for maturefaced ($M = 3.11$) than babyfaced ($M = 3.45$) 11 year olds, $t(63) = 3.85, p < .01$. For acts of omission, more severe punishment was recommended for maturefaced than babyfaced children of both ages, although none of these differences was significant, all $ps > .10$. Finally, it should be noted that the foregoing effects held true both when parents assumed that the misbehavior was intentional and also when they assumed that it was accidental.

Discussion

As predicted, the misdeeds of maturefaced 4- and 11-year-old children were judged as more intentional than the same behaviors by their babyfaced peers, though this effect unexpectedly held true only for mothers' judgments about boys and fathers' judgments about 11-year-old girls. Although restricted to these cross-sex perceptions, the intentionality judgments are consistent with the tendency to perceive more maturefaced children as less naive and more likely to know right from wrong the way an adult does (Zebrowitz & Montepare, 1991) as well as with the finding that more maturefaced adults are more apt to be judged as engaging in intentional misconduct (Berry & Zebrowitz-McArthur, 1988; Zebrowitz & McDonald, 1991).

The fact that parents' intentionality judgments showed more evidence of facial maturity stereotypes for children of the opposite sex may reflect the fact that parents tend to interact less with opposite sex children (Lamb, 1977; Weinraub & Frankel, 1977), which may yield more stereotyped judgments. The absence of a cross-sex effect for fathers judging 4-year-old girls may be due to a countervailing attractiveness halo: the greater attractiveness of the maturefaced girls in this age group may have negated the tendency to perceive their misbehaviors as more intentional than those of the babyfaced.

The effects of facial maturity on punishment severity were qualified by parent gender as well as by the children's age and the nature of their misbehavior. Fathers, but not mothers, showed the predicted overall tendency to recommend more severe punishment for maturefaced than babyfaced children. Fathers' greater responsiveness to appearance stereotypes parallels evidence for a greater impact of sex stereotypes on men, and both of these effects may reflect men's lesser experience with children (Barry, 1980; Block, 1973; Rubin, Provenzano, & Luria, 1974; Sobieszek, 1978).

Neither mothers nor fathers recommended significantly more punishment for maturefaced than babyfaced children when the misbehaviors were relatively innocuous omissions, such as not helping or sharing. In the case of more serious actions, such as hurting another child or destroying property, parents recommended significantly more severe punishment for maturefaced than babyfaced 4 year olds, while recommending less severe punishment for maturefaced than babyfaced 11 year olds. The more negative reactions to baby- than maturefaced 11 year olds committing relatively serious offenses parallels the tendency for subjects in an earlier study to recommend more severe punishment for babyfaced defendants who had confessed to premeditated crimes (Berry & Zebrowitz-McArthur, 1988).

The punitiveness of subjects sentencing a babyfaced defendant for a premeditated offense was attributed to a tendency for people to react negatively to the disconfirmation of their benign expectations regarding babyfaced individuals. Similarly, parents may also react very punitively when their benign expectations for the behavior of babyfaced children are violated. The fact that parents reacted very punitively to the serious wrongdoings of babyfaced 11 year olds but not to similar acts by babyfaced 4 year olds may reflect the fact that these acts were less surprising in 4 year olds, regardless of their facial appearance. To assess the validity of this argument, a separate sample of parents rated on 7-point scales how unexpected each misbehavior would be for a 4 year old or an 11 year old. The results revealed a strong main effect for age, with the misbehaviors being more unexpected for 11 year olds ($M = 4.92$) than for 4 year olds ($M = 2.33$), $F(1, 20) = 97.46$, $p < .0001$.

Additional evidence to support the argument that the commissions of

11-year-old babyfaced children were more unexpected is provided by parents' ratings of how upset they would be by a particular action. The pattern of results paralleled the severity of recommended punishments, with parents reporting that they would be more upset by the commissions of babyfaced than maturefaced 11 year olds and less upset by the commissions of babyfaced than maturefaced 4 year olds.

A caveat is necessary regarding the generalizability of the effects of facial maturity on punishment severity. Specifically, it should be recalled that parents were asked to assume that a misbehavior was intentional (or accidental) when making their punishment recommendations so that the measure of punishment severity would not be confounded with perceived intentionality. Thus, manipulated intentionality superseded the initial perceptions of intentionality which would ordinarily influence punishment severity. Indeed, parents recommended much more severe punishment when they had been instructed to assume that the misdeeds were intentional ($M = 4.10$) than when they had been instructed to assume that they were accidental ($M = 2.08$), $F(1, 60) = 279.70$, $p < .0001$, an effect which held true across target age and facial maturity and for omissions as well as for commissions.

The foregoing pattern of results indicates that when intentionality is in question, the misdeeds of maturefaced children will be perceived as more intentional than those of the babyfaced and, since intentional misdeeds are punished more severely, maturefaced children will receive harsher punishments. On the other hand, when parents know that an act is intentional or when they know that it is accidental, then the tendency to punish maturefaced children more severely varies with the age of the child and the nature of the misbehavior. In the case of relatively innocuous actions, maturefaced 4- and 11-year-old children are punished more severely than their babyfaced peers, albeit not significantly so. In the case of more serious actions, a babyface mitigates the severity of punishment for preschool age children while increasing its severity for older children.

Interestingly, and contrary to prediction, there were no significant main effects for target age in this study. The most likely explanation for this finding is that age was a between-subjects variable whereas facial maturity was manipulated within subjects. Thus, parents' use of the rating scales may not have been as sensitive to differences in their judgments about 4- and 11-year-old children as they would be if they were making direct comparisons between these age groups. Moreover, the error variance was typically higher when testing the between subjects age effects than the within subjects facial maturity effects.

CONCLUSIONS

The present findings reveal that children's facial maturity exerts a significant impact on the kinds of behaviors they are expected to be capable of performing as well as on reactions to their misbehavior. Compared

with their equally attractive babyfaced peers, maturefaced children were given more cognitively demanding tasks and their misbehavior was more apt to be viewed as intentional. Since intentional misbehavior is punished more severely, maturefaced children will receive harsher punishments than the babyfaced when the intentionality of an act is ambiguous, although this effect may be restricted to parents of the opposite sex. On the other hand, when intentionality is known, young maturefaced children receive more severe punishment than the babyfaced for serious infractions while older maturefaced children are treated more leniently.

The foregoing effects have both positive and negative implications for the development of maturefaced children. On the positive side, the early experiences of maturefaced children may yield self-fulfilling prophecies, such that they live up to the higher expectations that parents have for their ability to perform cognitive tasks and to exhibit proper social behavior. On the negative side, maturefaced children may be unable to live up to parents' expectations. Because their unsatisfactory behavior is perceived as more intentional than equivalent behavior by babyfaced children, they may be more likely to receive harsh discipline in their preschool years and even to become the victims of abusive parents.

REFERENCES

- Adams, G. R. (1977). Physical attractiveness research: Toward a developmental social psychology of beauty. *Human Development, 20*, 217-239.
- Adams, G. R., Hicken, M., & Salehi, M. (1988). Socialization of the physical attractiveness stereotype: Parental expectations and verbal behaviors. *International Journal of Psychology, 23*, 137-149.
- Alley, T. R. (1988). *Social and applied aspects of perceiving faces*. Hillsdale, NJ: Erlbaum.
- Barry, R. J. (1980). Stereotyping of sex role in preschoolers in relation to age, family structure, and parental sexism. *Sex Roles, 6*, 795-806.
- Berkowitz, L., & Frodi, A. (1979). Reactions to a child's mistakes as affected by her/his looks and speech. *Social Psychology Quarterly, 42*, 420-425.
- Berry, D. S., & McArthur, L. Z. (1985). Some components and consequences of a babyface. *Journal of Personality and Social Psychology, 48*, 312-323.
- Berry, D. S., & Zebrowitz-McArthur, L. A. (1988). What's in a face? Facial maturity and the attribution of legal responsibility. *Personality and Social Psychology Bulletin, 14*, 23-33.
- Berscheid, E., & Walster, E. (1974). Physical attractiveness. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 157-215). New York: Academic Press.
- Block, J. H. (1973). Conceptions of sex-roles: Some cross-cultural and longitudinal perspectives. *American Psychologist, 28*, 512-526.
- Brackbill, Y., & Nevill, D. D. (1981). Parental expectations of achievement as affected by children's height. *Merrill-Palmer Quarterly, 27*, 429-441.
- Bull, R., & Rumsey, N. (1988). *The social psychology of facial appearance*. New York: Springer-Verlag.
- Clifford, M., & Walster, E. (1973). Research note: The effects of physical attractiveness on teacher expectations. *Sociology of Education, 46*, 248-258.
- Condry, J., & Condry, S. (1976). Sex differences: A study of the eye of the beholder. *Child Development, 47*, 812-819.

- Dion, K. (1972). Physical attractiveness and evaluation of children's transgressions. *Journal of Personality and Social Psychology*, *24*, 207-213.
- Dion, K. (1973). Young children's stereotyping of facial attractiveness. *Developmental Psychology*, *9*, 183-188.
- Dion, K., & Berscheid, E. (1974). Physical attractiveness and peer perception among children. *Sociometry*, *37*, 1-12.
- Dix, T., & Grusec, J. E. (1985). Parent attribution process in the socialization of children. In I. E. Sigel (Ed.) *Parental belief systems: The psychological consequences for children* (pp. 201-233). Hillsdale, NJ: Erlbaum.
- Dix, T., Ruble, D. N., Grusec, J. E., & Nixon, S. (1986). Social cognition in parents: Inferential and affective reactions to children of three age levels. *Child Development*, *57*, 879-894.
- Eisenberg, N., Roth, K., Bryniarski, K. A., & Murray, E. (1984). Sex differences in the relationship of height to children's actual and attributed social and cognitive competencies. *Sex Roles*, *11*, 719-734.
- Fiske, S. T., & Cox, M. G. (1979). Person concepts: The effect of target familiarity and descriptive purpose on the process of describing others. *Journal of Personality*, *47*, 136-161.
- Lamb, M. E. (1977). The development of parental preference in the first two years of life. *Sex Roles*, *3*, 495-497.
- Langlois, J. H. (1986). From the eye of the beholder to behavioral reality: Development of social behaviors and social relations as a function of physical attractiveness. In C. P. Herman, M. P. Zanna, & E. T. Higgins (Eds.), *Physical appearance, stigma, and social behavior. The Ontario symposium* (Vol. 3, pp. 23-51). Hillsdale, NJ: Erlbaum.
- 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.
- Liggett, J. C. (1974). *The human face*. New York: Stein & Day.
- Livesley, W. J., & Bromley, D. B. (1973). *Person perception in childhood and adolescence*. London: Wiley.
- Lorenz, K. (1943). Die angeborenen Formen möglicher Erfahrung [The innate forms of potential experience.] *Zeitschrift für Tierpsychologie*, *5*, 234-409.
- McArthur, L. Z., & Apatow, K. (1983-1984). Impressions of babyfaced adults. *Social Cognition*, *2*, 315-342.
- McArthur, L. Z., & Berry, D. S. (1987). Cross-cultural agreement in perceptions of babyfaced adults. *Journal of Cross-Cultural Psychology*, *18*, 165-192.
- McCabe, V. (1984). Abstract perceptual information for age level: A risk factor for maltreatment? *Child Development*, *55*, 267-276.
- Meyer, J. W., & Sobieszek, B. (1972). Effect of a child's sex on adult interpretations of its behavior. *Developmental Psychology*, *6*, 42-48.
- Ritter, J. M., Langlois, J. H., & Casey, R. J. (1991). Adults' responses to infants varying in appearance of age and attractiveness. *Child Development*, *62*, 68-82.
- Rubin, J. Z., Provenzano, F. J., & Luria, Z. (1974). The eye of the beholder: Parents' views on sex of newborns. *American Journal of Orthopsychiatry*, *44*, 512-519.
- Sobieszek, B. I. (1978). Adult interpretations of child behavior. *Sex Roles*, *4*, 579-588.
- Sorrell, G. T., & Nowak, C. A. (1981). The role of physical attractiveness as a contributor to individual development. In R. M. Lerner & N. A. Busch-Rossnagel (Eds.), *Individuals as producers of their development: A life span perspective* (pp. 389-446). London: Academic Press.
- Stephan, C., & Langlois, J. H. (1984). Baby beautiful: Adult attributions of infant competence as a function of infant attractiveness. *Child Development*, *55*, 576-585.

- Vaughn, B. E., & Langlois, J. H. (1983). Physical attractiveness as a correlate of peer status and social competence in preschool children. *Developmental Psychology*, *19*, 561-567.
- Weinraub, M., & Frankel, J. (1977). Sex differences in parent-infant interaction during free play, departure, and separation. *Child Development*, *48*, 1240-1249.
- Zebrowitz, L. A., & McDonald, S. (1991). *The impact of litigants' babyfacedness and attractiveness on adjudications in small claims court*. *Law and Human Behavior*, in press.
- Zebrowitz, L. A., & Montepare, J. M. (1991). *Impressions of males and females across the lifespan as a function of their babyfacedness and attractiveness*. Unpublished manuscript, Brandeis University.
- Zebrowitz, L. A., Tenenbaum, D. R., & Goldstein, L. H. (1991). The impact of job applicants' facial maturity, sex, and academic achievement on hiring recommendations. *Journal of Applied Social Psychology*, *21*, 525-548.

RECEIVED: January 22, 1991; REVISED: April 5, 1991.