Journal of Abnormal Psychology
1986, Vol. 95, No. 1, 81-87

0021-843X/86/$00.75

Children's Judgments and Attributions in Response to the "Mentally Retarded" Label: A Developmental Approach

Richard Bromfield, John R. Weisz, and Tim Messer
University of North Carolina at Chapel Hill

Labels implying deviance or handicap can have stigmatizing effects. However, developmental theory and research suggest that such label effects may depend on cognitive processes that only develop in later childhood—processes such as trait inference and the logical linkage of attributions, expectancies, and behavior prescriptions. To test this possibility, we had third, sixth, and ninth graders watch a videotape in which a child failed a puzzle task; all of the students had identical information about the child's achievement scores and puzzle ability, but half were told that the child was mentally retarded. Consistent with key person perception and social cognition findings, the traitlike *retarded* label had little impact on younger children but strong effects on older ones. Like adults in earlier research, sixth and ninth graders saw low effort as a less important cause of failure for the retarded than for the unlabeled child, and they saw less need to urge the retarded child to persist. Correlations showed no evidence of logical linkages among attributions, expectancies, and persistence-urging among third graders, but strong linkages among sixth and ninth graders. The findings suggest that label effects are mediated by cognitive processes dependent on developmental level.

There is ample evidence that labels can activate sets that shape person perception. Labels such as *warm* versus *cold* can shape people's first impressions of a lecturer (Kelley, 1950) and their more carefully considered judgments about people in simulated job interviews (Huguenard, Sager, & Ferguson, 1970). The simple information that a child is feeling "under par" can significantly color the descriptions people write of that child's behavior (Rapp, 1965).

Labels implying deviance, deficit, or a handicap may be especially important because their effects are often stigmatizing. Langer and Abelson (1974) showed clinicians a videotaped interview with a man labeled as either a *job applicant* or a (mental) *patient*. Clinicians given the patient label evaluated the interviewee as significantly more disturbed than did clinicians given the job applicant label. Piner and Kahle (1984) found that people performed better on a highly salient memory test when they thought their partner was a former mental patient than when they had no such information; the people were presumably attempting to cope with, and compensate for, perceived disability of their partners.

We are grateful to the principals, staff, and students of the following schools for participating in this study: Carolina Friends School, Cresset Christian Academy, Immaculata Academy, and St. Mary's Country Day School, all in Durham, North Carolina; Butner-Stem School, Creedmoor Elementary School, Hawley School, and South Granville High School, all of the Granville County, North Carolina, school system; and St. Thomas More School in Chapel Hill, North Carolina. We also thank Debbie Bromfield and Allison Weisz for their Oscar-worthy roles in the videotape, and Geraldine Dawson and Thomas Pyszczynski for their helpful comments on the manuscript.

Correspondence concerning this article should be addressed to John R. Weisz, Department of Psychology, Davie Hall 013-A, University of North Carolina, Chapel Hill, North Carolina 27514.

Like the patient label, the *mentally retarded* label seems to have potentially stigmatizing effects. Several studies have shown that people form quite different attributions, expectancies, and behavior prescriptions for children labeled mentally retarded than for unlabeled children, even in the face of information placing the children at the same level of ability (Bromfield & Weisz, 1983; Weisz, 1981; Yeates & Weisz, in press). Despite such equal ability information, the retarded label can make people more likely to attribute failure to low ability, less likely to attribute failure to insufficient effort, and less likely to believe that the failure can be reversed with increased effort. It is not surprising that people also indicate that the attributions and expectancies they tend to form for retarded children would discourage them from urging the child to persist after failure.

Findings on the stigmatizing effects of labels can be understood from several theoretical perspectives. A key notion suggested by Goffman (1963) is that in stigmatization, people who possess a certain trait, particularly if it is unusual, are denigrated by those who do not possess the trait. Goffman's analysis of stigmatization has been followed by several other theoretical developments (see a thoughtful review by Katz, 1979). For example, one attribution theory perspective (e.g., Weisz, 1981) emphasizes the impact of labels on processes of causal reasoning—as when the mentally retarded label highlights the salience of low ability attributions and triggers discounting of low effort attributions (cf. Kelley, 1973). Ambivalence-response amplification theory suggests that labels can have stigmatizing effects by stimulating ambivalence about the labeled person and thus provoking exaggerated responses to that person (Katz, 1979). Other theories suggest that labels can be stigmatizing because they stimulate uneasiness or discomfort (Crano & Messe, 1982), or because they convert the labeled person into a novel stimulus (Langer, Taylor, Fiske, & Chanowitz, 1976).

In addition to explaining label effects per se, an important task for theory and research is to explain situational and individual variations in the impact of labels (Piner & Kahle, 1984). A key finding of several studies has been that not all individuals are equally susceptible to label effects. For example, Langer and Abelson (1974) found that behaviorally oriented clinicians were relatively unaffected by the label *patient,* but that traditional psychodynamically oriented clinicians were quite strongly influenced. Yeates and Weisz (in press) found that effects of the mentally retarded label could be predicted from adults' theoretical beliefs about mental retardation and from their experience with mentally retarded children. For example, graduate students and teachers in the field of special education were relatively uninfluenced by the label, but undergraduates and regular classroom teachers were significantly susceptible to label effects. It seems likely that a number of individual difference factors that have not yet been studied systematically may influence people's susceptibility to stigmatizing effects of labels.

One such factor that may well be of primary importance is developmental level. The theories cited above (see especially Goffman, 1963, and Weisz, 1981) suggest that label effects are mediated in large part by the attribution of stable traits to labeled individuals and by a series of logically linked attributional and expectancy judgments that follow from that trait attribution. There is growing evidence that these cognitive processes only take shape gradually across years of development and that they may not be in place until late childhood (cf. Rholes, Blackwell, Jordan, & Walters, 1980). Consider, for example, the developmental literature on social cognition, particularly the research on person perception. This research (e.g., Barenboim, 1981; Livesley & Bromley, 1973; Peevers & Secord, 1973) indicates that it is only over the course of the elementary school years, and sometimes beyond, that children come to verbalize and use the kinds of stable, personal constructs or traitlike inferences that adults use.[1] Prior to late childhood or early adolescence, children are more likely to describe people in terms of concrete, outwardly observable, and often transitory characteristics such as physical appearance.

It is, evidently, during the middle years of childhood that the concept of stability, or invariance across time and setting, is consolidated and applied to such quantitative domains as mass and number (Inhelder & Piaget, 1958). Similarly, some have argued (e.g., Flavell, 1977), it is only with continuing cognitive development that the notion of stability or constancy is applied to such qualitative domains as gender and traits. Building on this notion, Rholes and Ruble (1984) tested the ability of younger and older elementary school children to predict behavior based on information about another person's internal attributes. Their findings suggest that younger children, in contrast to older ones, do not regard internal factors such as ability "as stable, abiding characteristics of other persons" (Rholes & Ruble, 1984, p. 50). Focusing specifically on the concept of ability, Nicholls and Miller (1984) found that, through most of the elementary school years, children do not appear to construe ability as a stable, underlying capacity, or to comprehend correctly relations between ability, effort, and performance outcomes (for a related point, see Kun, 1977).

This developmental literature strongly suggests that there may be age differences in children's responses to labels that have implications for traits such as ability. One such label, of course, is *mentally retarded.* The preceding findings suggest that younger children may be less likely than older ones to respond to the mentally retarded label with stable, adultlike attributions, and less likely to modify their causal judgments and their expectancies in logically consistent ways. In short, young children should be less susceptible than older ones to potentially stigmatizing label effects because they lack the cognitive prerequisites for such effects. Analogous reasoning might be applied to labels that imply other kinds of deviance or handicap (e.g., "mental patient"). However, the substantial exposure of children to mentally retarded youngsters, in school settings and elsewhere, makes this particular label especially appropriate as a starting point for the study of developmental effects.

To probe for such effects in the present study, we followed a procedure patterned, in part, after that of Langer and Abelson (1974). We asked third, sixth, and ninth graders to view a videotape in which a child worked at a problem-solving task in the presence of an adult. All of the children viewed exactly the same videotape, and all were given exactly the same information about the target child, her achievement level, her level of general ability, and her level of ability at the specific task shown on the videotape. However, half of the children were also told that the child was mentally retarded. This allowed us to pose the question of whether the label alone would produce effects on children's causal attributions, ability assessments, performance expectancies, behavior prescriptions, and ratings of desirability as a friend—and whether the effects would differ as a function of grade level. The inclusion of both boys and girls in our sample allowed us to test for generality of effects across sex. A meaningful analysis of sex effects per se was not possible given the experimental requirements of this study. Our objectives made it essential that all children view precisely the same videotape; this, in turn, meant that we had to choose a single child, in this case a girl, to be taped. Thus, any sex effect that might emerge from our analyses could also be interpreted as a same-sex versus cross-sex effect. For this reason, any interpretation of such "sex effects" should focus only on their implications for the generality of our findings, not on their possible meaning as sex effects per se.

Method

Subjects and Experimental Design

Originally, a total of 504 students participated in this study. However, we set very strict standards for inclusion in the sample, with youngsters required to correctly complete numerous questions (described below) tapping comprehension of the experimental instructions, questionnaire scale format, instructions regarding the child shown in the videotape, and identification of their own gender (see the Method section for details). We dropped 79 students from the original sample for failure to meet one or more of these criteria, or for other errors or omissions on the questionnaires.

The final sample, then, included 425 predominantly white, middle class students drawn from Grades 3 (n = 140), 6 (n = 188), and 9 (n = 97) of five private and four public schools. The 77 girls and 63 boys in Grade 3 averaged 8.13 years of age (SD = 0.50). The 86 girls and 102

[1] MacClennon and Jackson (1985), though, have found some evidence of *nonverbal* traitlike inference even in the early elementary years.

boys in Grade 6 averaged 11.12 years (SD = 0.55). The 45 girls and 52 boys in Grade 9 averaged 14.39 years (SD = 0.73). At each grade level approximately half of the children were randomly assigned to the label condition, in which they were told that the videotaped child was mentally retarded; the remaining children were *not* told that the child was mentally retarded (no label condition). In the 3 × 2 × 2 (Grade × Sex × Condition) design thus created, the three factors were nonorthogonal due to unequal cell sizes.

Procedure

Groups ranging in size from 13 to 35 children were asked to read an information sheet as the experimenter read it aloud. Half of the groups were told that "Tammy," the child they would observe in the videotape, was mentally retarded. The remaining groups were told only that her name was Tammy. All were told "Tammy's achievement test scores show that she can do third grade school work," and "she can also do puzzles that most third graders can do." Thus, general achievement levels and ability to perform the specific task that Tammy was to fail in the film were described identically for the labeled and unlabeled child.

Children were then asked questions designed to check their understanding of the information they had been provided about Tammy (e.g., "Tammy can do puzzles that most _____ graders can do"); all who failed one or more of these questions were dropped from the sample. In addition, children in the label condition were dropped if they failed to answer the following postexperiment question correctly: "What do you know about Tammy? Circle one: (a) She is mentally retarded, (b) she is not mentally retarded."

All of the children viewed the same 4½ min videotape. It showed Tammy working at block design problems in the presence of a student teacher. As the film began, Tammy was pushing a successfully completed block design problem toward the teacher for evaluation. The teacher praised Tammy for coming up with the correct solution and suggested that Tammy try another one. Tammy then spent the remaining, major portion of the film working on the second block puzzle. During this period she alternately appeared distracted and bored, or attentive but unable to complete the problem. Toward the end of the tape, Tammy showed her teacher an incorrectly completed design. The teacher told her that it was not quite right and suggested that she try a little longer. After further effort and another word of encouragement from the teacher, Tammy declared that she really could not do the problem. The teacher responded, "That's all right. Why don't we try another one?" Overall, the tape was intended to provide mixed cues, some suggesting insufficient effort on Tammy's part, other suggesting insufficient ability and the possibility that the task might be too difficult for Tammy. The tape was also intended to be somewhat ambiguous as to whether the teacher had allowed Tammy to give up at the appropriate time or had urged too much or too little persistence.

Dependent Measures: Child Questionnaire

After viewing the tape, children filled in a questionnaire adapted from those used in earlier studies of adults' perceptions (Bromfield & Weisz, 1983; Weisz, 1981). Building on the conceptual and empirical work of Weiner et al. (1972), we asked children to rate the importance of four key factors as causes of the child's failure: insufficient effort (i.e., "She did not try hard enough"), bad luck (i.e., "She had bad luck"), insufficient ability (i.e., "She is not very good at doing puzzles"), and task difficulty (i.e., "The puzzle was too hard"). Five-point rating scales used in the adult studies ranged from *not important* to *very important*. However, pilot data suggested that this format was difficult for many children to comprehend and that scales ranging from *very false* (0) to *very true* (4) would be preferable; thus, we used this latter format.

Using similar 4-point scales, children also rated (a) Tammy's likelihood of success if she tried the same puzzle on the very next day with the same teacher, (b) Tammy's likelihood of success if she tried the same puzzle on the very next day with a different teacher, (c) "how good . . . Tammy is at doing puzzles," and (d) how much they "would like to have Tammy for a friend." Subsequently, subjects indicated whether they thought the teacher, (a) "should have let Tammy quit trying sooner than she did," (b) "did just the right thing," or (c) "should have made Tammy keep trying longer." Similarly, children were asked to imagine that they were the teacher and to indicate whether they would have (a) "let Tammy quit sooner than her teacher did," (b) "done just what her teacher did," or (c) "made Tammy try longer than her teacher did." In an attempt to identify the factors influencing the preceding judgments, children were asked, "Remember when you answered the questions about Tammy. What thing about Tammy did you think about most?" The alternatives listed were her face, her voice, her clothes, how she moved, how she worked, how she felt, or "if something else, what? _____." To test whether young respondents understood the mechanics of the rating scales, we asked them to rate the accuracy of the statements "I eat food" and "Dogs have seven legs" on a scale ranging from *very false* (0) to *very true* (4). Youngsters who circled false or very false for the first question and/or true or very true for the second question were excluded from the sample.

Results

Data were analyzed via general linear models (GLM) analysis of variance (ANOVA) procedures designed for data sets in which nonorthogonality is produced by unequal cell sizes (Ray, 1982). All F values reported below are based on conservative sums of squares, with each tested effect adjusted for other effects in the model (i.e., via *eliminating* tests—cf. Appelbaum & Cramer, 1974). All marginal and subclass means reported below are least squares means, adjusted to represent the mean values that would be obtained in a completely balanced design (see Ray, 1982, pp. 148–149, 177–178). Finally, we should note that initial analyses showed no significant effects of school type (i.e., private vs. public); thus, for clarity, we have excluded that variable from all the analyses presented below.

Initial Multivariate Analysis of Variance

Because of the large number of dependent variables generated by the post-video questions, it was important to guard against capitalizing on chance findings. Thus, we began by subjecting responses to the 10 questions to a 3 × 2 × 2 (Grade × Condition × Sex) multivariate analysis of variance (MANOVA). The results revealed significant multivariate main effects of grade, $F(20, 784) = 4.06$, $p < .0001$; condition, $F(10, 392) = 5.14$, $p < .0001$; sex, $F(10, 392) = 4.74$, $p < .0001$; and Grade × Condition, $F(20, 784) = 2.40$, $p < .001$. No other effects were significant at the multivariate level; thus no other effects will be discussed in the following report of univariate tests.

Univariate Analysis of Variance Results

A series of 3 × 2 × 2 (Grade × Condition × Sex) ANOVAs revealed a number of significant effects. All of these are summarized in Table 1. On the attributional question of whether Tammy failed because she did not try hard enough, there were significant main effects of condition, $F(1, 401) = 11.92$, $p < .001$, and grade, $F(2, 401) = 6.76$, $p = .001$. Low effort attributions were given more credence by children in the no label condition than by children who thought Tammy was mentally retarded (M = 2.52 and 2.20), and were given more credence by

Table 1
Summary of ANOVA Findings

Dependent variables	Main effects	Interaction
Low effort attributions	Condition (MR < U)***	Condition × Grade
	Grade (3 > 6 < 9)***	Grade 3 (MR > U)
		Grade 6 (MR < U)**
		Grade 9 (MR < U)**
Bad luck attributions	Grade (3 > 6 > 9)****	
Tammy's puzzle ability	Condition (MR > U)**	Condition × Grade
	Grade (3 > 6 > 9)****	Grade 3 (MR < U)
		Grade 6 (MR > U)*
		Grade 9 (MR > U)**
Like Tammy as a friend	Condition (MR > NR)****	Condition × Grade****
	Grade (3 > 6 = 9)**	Grade 3 (MR < U)
	Sex (G > B)****	Grade 6 (MR > U)***
		Grade 9 (MR > U)***
Teacher-urged persistence	Sex (G < B)***	Condition × Grade*
		Grade 3 (MR > U)
		Grade 6 (MR < U)
		Grade 9 (MR < U)**
Self-urged persistence	Sex (G < B)***	Condition × Grade*
		Grade 3 (MR > U)
		Grade 6 (MR < U)*
		Grade 9 (MR < U)*

Note. ANOVA = univariate analysis of variance. MR = mentally retarded label; U = unlabeled condition. 3 = third graders; 6 = sixth graders; 9 = ninth graders. B = boys; G = girls.
* $p < .05$. ** $p < .01$. *** $p < .001$. **** $p < .0001$.

third graders than by sixth graders (Newman-Keuls test, $p < .05$, $M = 2.58$, 2.19, and 2.35, for Grades 3, 6, and 9, respectively). The most important effect, however, was a significant Grade × Condition interaction, $F(2, 415) = 2.93$, $p = .05$. Third graders did not differ significantly in their effort attributions in the retarded versus no label conditions ($M = 2.61$ and 2.59). By contrast, both older groups showed significant label effects ($p < .01$). The retarded label reduced the perceived importance of low effort as a cause of failure among both sixth graders ($M = 1.98$ and 2.44) and ninth graders ($M = 2.06$ and 2.60).

No significant effects were found for attributions of Tammy's failure to low ability ("not very good at doing puzzles").[2] Attributions to bad luck differed only as a function of grade, $F(2, 401) = 17.92$, $p < .0001$; luck attributions were given less credence with increasing grade level ($M = 1.85$, 1.36, and 1.28). No significant effects were found for attributions of Tammy's failure to task difficulty ("puzzle was too hard") or for predictions of Tammy's future success with the same puzzle.

By contrast, the groups differed sharply and in unexpected directions on their judgments as to how good Tammy was at doing puzzles. Children who were told that Tammy was retarded actually rated her higher in puzzle ability than did children in the no label condition, $F(1, 401) = 9.45$, $p < .005$ ($M = 2.29$ and 2.64; lower score = higher ability). Tammy's ability was rated at lower levels with increasing grade level, $F(2, 401) = 11.31$, $p < .0001$ ($M = 2.16$, 2.48, and 2.76). These main effects of condition and grade were accompanied by a Condition × Grade interaction, $F(2, 401) = 3.55$, $p < .05$; third graders did not differ significantly as a function of condition ($M = 2.14$ and 2.20), but sixth graders ($p < .05$; $M = 2.34$ and 2.61) and ninth graders ($p < .01$; $M = 2.40$ and 3.11) both rated Tammy as *more* able in the retarded than in the nonretarded condition.

The question asking how much the children would like to have Tammy for a friend yielded strong group differences as a function of condition, $F(1, 401) = 25.06$, $p < .001$; grade, $F(2, 401) = 4.32$, $p = .01$; sex, $F(1, 401) = 42.26$, $p < .0001$; and Condition × Grade, $F(2, 401) = 14.93$, $p < .0001$. Tammy was rated more desirable as a friend by those who thought she was retarded than by those given no label ($M = 3.43$ and 2.95); she was rated more desirable by third graders than by sixth and ninth graders ($M = 3.35$, 3.11, and 3.12); and, not surprisingly, girls gave Tammy, a fellow girl, higher ratings than did boys ($M = 3.47$ and 2.92). Third graders did not differ significantly as a function of condition ($M = 3.28$ and 3.42), but for sixth graders ($p < .001$; $M = 3.34$ and 2.88) and ninth graders ($p < .001$; $M = 3.68$ and 2.56), the information that Tammy was retarded apparently led to enhanced ratings.

In some respects, the two most important questions were those that asked the children how much the teacher should have urged Tammy to persist and how much they themselves might urge persistence. Both questions yielded significant differences between boys and girls for teacher, $F(1, 401) = 10.88$, $p = .001$, and for self, $F(1, 401) = 4.81$, $p < .05$. In both instances, girls gave lower persistence-urging ratings than did boys (teacher $M = 2.11$ and 2.24; self $M = 2.21$ and 2.32). Both questions also revealed po-

[2] This nonsignificant group difference seems inconsistent with the significant group difference found on the question about how good Tammy is at doing puzzles (see following paragraph). Possibly the earlier question was too complex for some children to understand; to assert that Tammy had adequate ability, a child would have had to select *very false* for the statement "She is not very good at doing puzzles." This amounts to a potentially difficult double negative.

Table 2
Correlations Involving Low Effort and Low Ability Attributions

Attribution	Likelihood of future success	Teacher-urged persistence	Self-urged persistence
Third graders (N = 140)			
Low effort	−.050	.034	−.027
Low ability	−.062	.056	.064
Sixth graders (N = 188)			
Low effort	−.221**	.259***	.233***
Low ability	−.091	−.010	.007
Ninth graders (N = 97)			
Low effort	−.118	.284**	.339***
Low ability	−.272**	.075	−.221*

* $p < .05$. ** $p < .01$. *** $p < .001$.

tentially important Condition × Grade effects: for teacher, $F(2, 401) = 4.00$, $p < .05$, and for self, $F(2, 401) = 4.41$, $p < .05$. On the teacher question, only the ninth graders differed significantly as a function of experimental condition ($p < .01$); they judged persistence-urging by the teacher to be more appropriate for the unlabeled than for the mentally retarded Tammy (M = 2.28 and 2.06). On the self question, both sixth graders ($p < .05$) and ninth graders ($p < .05$) showed this same pattern (sixth grade M = 2.32 and 2.20; ninth grade M = 2.41 and 2.25). Overall, with increasing grade level, children seemed increasingly inclined to go easy on Tammy (i.e., urge less persistence) if she had been labeled mentally retarded than if she had not.

What Children Thought About Most

Another set of analyses dealt with what children said they thought about most when answering the questions about Tammy. Initially, the six response categories (the *other* category yielded no analyzable data) were subjected to an ANOVA-like least squares procedure designed for categorical variables (Grizzle, Starmer, & Koch, 1969). We began with an overall 2 × 3 × 2 × 6 (Condition × Grade × Sex × Response Category) analysis and proceeded to finer-grained tests as appropriate. We found that self-reports of attending to Tammy's face and voice declined marginally with increases in grade level, $\chi^2(2, N = 425) = 8.91$, $p = .02$, and 5.77, $p = .07$, respectively. Attention to "how she moved" showed the reverse grade trend, $\chi^2(2, N = 425) = 5.93$, $p = .06$. Girls reported focusing on "how she felt" more than did boys, $\chi^2(1, N = 425) = 11.53$, $p < .001$. No other effects approached significance.

Correlation of Attributions With Expectancies and Behavior Prescriptions

Next we focused on relations between the effort and ability attributions children made for Tammy and their expectancies about her likelihood of future success at the puzzle, and their behavior prescriptions on the two persistence-urging questions (teacher and self). As we noted earlier, the label effects found in research with adults seem to depend partly on a logical linkage between attributions and these other cognitive/judgmental processes. For instance in the present case, a logical analysis would suggest that if Tammy failed the puzzle because of low ability, then her likelihood of future success at the puzzle should be estimated at low levels, and there would be little point in urging her to persist. However, if Tammy failed because of low effort, then her likelihood of success might be rated relatively higher, and she might reasonably be urged to persist. The correlation coefficients in Table 2 reveal that such logical linkages were not in evidence at all among third graders. At the sixth-grade level, though, low effort attributions were appropriately linked to the two persistence-urging measures; however, there was a puzzling *negative* correlation between low effort attributions and success expectancies. The more important sixth graders thought low effort was, the less likely they thought it was that Tammy would succeed in the future. Perhaps this reflects a traitlike view of low effort—that is, as a stable trait of laziness. Ninth graders showed the most uniformly logical linkages among attributions, expectancies, and behavior prescriptions. For this oldest group, low effort attributions were linked to increased persistence-urging (both measures), and low ability attributions were linked to lowered success expectancies and lowered persistence-urging (self-measure).[3]

Discussion

The literature on social cognition and person perception (e.g., Barenboim, 1981; Livesley & Bromley, 1973) suggests that there may be developmental differences in the use and effects of labels that refer to relatively stable, underlying personal characteristics. Such traitlike characteristics tend not to have the powerful or pervasive effects on the judgments of younger elementary school children that they can have on the judgments of older children and adolescents (Rholes & Ruble, 1984). Consistent with this general notion, we found exceedingly modest effects of the men-

[3] We also explored whether relations among attributions, expectancies, and behavior prescriptions differed as a function of whether Tammy was thought to be retarded. Correlations computed separately for the label and no-label conditions revealed similar relations, with two exceptions: Low effort attributions and persistence-urging (self-measure) were significantly correlated in the label ($r = .185$, $p < .01$) but not in the no label condition ($r = .091$). And low ability attributions and success expectancies were significantly correlated in the no label ($r = .162$, $p < .05$) but not in the label condition ($r = .092$). Evidently, the information that Tammy was mentally retarded did not alter relations among attributions, expectancies, and behavior prescriptions in any very striking way.

tally retarded label among third graders, but numerous significant effects among older children. These grade differences emerged despite the fact that our sample included only those youngsters who could correctly report *all* of the information we had given them about Tammy and her abilities.

Although developmental differences in the existence of labeling effects could be anticipated, the precise nature of the effects that would materialize was difficult to predict. Some of the effects found in our study resembled earlier findings with adults, whereas other effects did not. Among our older subjects, there were two strong similarities to adult findings: The mentally retarded label reduced the likelihood that failure would be attributed to low effort and also reduced the perceived need for persistence-urging (cf. Bromfield & Weisz, 1983; Weisz, 1981; Yeates & Weisz, in press).

Perhaps the most striking reversal of earlier findings with adults was the fact that the mentally retarded label actually led to *increased* ratings of ability at the target task. In one respect, the findings with regard to perceived ability and low effort attributions can be linked to the finding that the mentally retarded label enhanced Tammy's desirability as a friend. All three effects of the label could be termed *benevolent;* all involve a relatively charitable interpretation of what was seen on the videotape. To put these findings into context, consider the fact that the tape showed a schoollike interaction in which a child failed at a task and gave up—not a particularly flattering picture. When the child who failed was unlabeled, her failure was especially likely to be blamed on low effort (perhaps akin to laziness), her ability at puzzles was disparaged, and she was judged a relatively undesirable candidate for friendship. By contrast, when the child was labeled mentally retarded, her failure was less likely to be blamed on lack of effort, and she was more likely to be viewed as able at puzzles and desirable as a friend. In short, the mentally retarded label appeared to mitigate potentially adverse effects of failure on the judgments and perceptions of other children. In this respect, the findings might actually be called counterstigmatizing.

These findings may also reflect a kind of differential frame-of-reference effect. For example, some in the label condition may have implicitly rated Tammy's ability relative to retarded children, whereas subjects in the no label condition may have used a more demanding standard of comparison (cf. Towne & Joiner's, 1968, sick-role hypothesis).

Two of the most important labeling effects found here involved judgments about how much Tammy should be urged to persist in the face of failure. Older children considered such persistence more appropriate for Tammy when she was unlabeled than when she was labeled mentally retarded—this despite the fact that they had actually rated the retarded Tammy higher in puzzle ability than her unlabeled counterpart. These persistence-urging effects—seen when children were rating their own likely behavior and when they judged what the teacher should have done—might be viewed as another form of benevolence, that is, of going easy on the retarded child.

The effects can also be viewed as an outgrowth of developmental increments in the logical linkage among causal attributions, expectancies, and behavior prescriptions. Our correlational evidence seems to indicate that sixth and ninth graders operated according to a logically appealing implicit principle: If failure results from low effort, then urge persistence. Ninth graders appeared to follow this same reasoning and to use an additional, two-part principle: If failure results from low ability then the likelihood of future success at the task is low and persistence need not be urged. Third graders evidently did not subscribe to any of these principles. What the correlation data suggest is that the logical linkages among trait-related attributions, expectancies, and behavior prescriptions that seem to underlie labeling effects may only emerge over the years from middle childhood to early adolescence.

The group differences in persistence-urging judgments are difficult to construe as good news regarding labeling effects. In fact, they resemble a kind of helplessness-condoning effect discussed in the adult literature, with the mentally retarded label reducing the likelihood that a child *at a given level of ability* will be encouraged to persist in response to failure (cf. Yeates & Weisz, in press). The present evidence suggests that such effects may not be confined to the adults in the retarded child's world, but may extend to the child's nonretarded peers—at least the older ones. This, in turn, suggests that to understand the causes of helpless behavior in retarded children (see Weisz, 1982; Weisz, Bromfield, Vines, & Weiss, 1985), we may need to focus not only on the role of adults, but on that of peers as well. The information that a child is retarded may have variable effects depending on the developmental level of those peers, but one possibility among older children and adolescents may be a set of closely linked attributional and judgmental processes that lead to a condoning of nonpersistence and relinquished control by the labeled child.

In examining our findings with an eye toward future research, it is useful to consider two questions about the data. One is whether the several Condition × Grade interactions we found—with condition effects present in older children but absent in third graders—might mean, in part, that the third graders know little about mental retardation. It does seem quite likely that children learn more about retardation and other handicapping conditions, as they mature. However, it should be noted that our ninth graders were, in some respects, actually further from the adult views seen in earlier research than were our third graders; ninth-graders, for example, unlike third graders, gave higher ability ratings to the retarded than to the unlabeled child. There is considerable evidence that the term *mentally retarded* is meaningful to third graders (see, e.g., Gottlieb, 1969; Gottlieb & Switzky, 1982; Wilkins & Velicer, 1980), but we certainly need to know more about developmental differences in the way children construe the term and the condition. Such knowledge could enrich our understanding of group differences like those reported here.

A second question concerns the role of social desirability in children's responses. It is possible that some in our sample showed the benevolent label effects found here out of a belief that one *should* be nice to retarded children. Some of the Condition × Grade interactions might reflect, at least in part, developmental differences in children's social awareness and their understanding of normative expectations. In future research, it would be useful to explore whether such developmental differences play a role and whether there are parallel effects on actual social behavior. Some evidence (e.g., Goodman, Gottlieb, & Harrison, 1972) suggests that retarded children suffer more real-life social rejection than do nonretarded children. But the effects of the mentally

retarded label per se on social acceptance and rejection have been difficult to isolate for clear-cut, unconfounded scrutiny (MacMillan, Jones, & Aloia, 1974). Such scrutiny, though, remains an important objective in the search for real-world effects of the social cognitions stimulated by labels.

References

Appelbaum, M., & Cramer, E. M. (1974). Some problems in the nonorthogonal analysis of variance. *Psychological Bulletin, 81,* 335–343.

Barenboim, C. (1981). The development of person perception from childhood to adolescence: From behavior comparisons of psychological constructs to psychological comparisons. *Child Development, 52,* 129–144.

Bromfield, R., & Weisz, J. R. (1983). *Effects of the "mentally retarded" label on adults' judgments about a videotaped child's failure.* Unpublished manuscript, University of North Carolina, Chapel Hill.

Crano, W. D., & Messe, L. A. (1982). *Social psychology: Principles and themes of interpersonal behavior.* Homewood, IL: Dorsey Press.

Flavell, J. H. (1977). *Cognitive development.* Englewood Cliffs, NJ: Prentice-Hall.

Goffman, E. (1963). *Stigma.* Englewood Cliffs, NJ: Prentice-Hall.

Goodman, H., Gottlieb, J., & Harrison R. H. (1972). Social acceptance of EMRs integrated into a non-graded elementary school. *American Journal of Mental Deficiency, 76,* 412–417.

Gottlieb, J. (1969). Attitudes toward retarded children: Effects of evaluator's psychological adjustment and age. *Scandinavian Journal of Educational Research, 13,* 170–182.

Gottlieb, J., & Switzky, H. H. (1982). Development of school-age children's stereotypic attitudes toward mentally retarded children. *American Journal of Mental Deficiency, 86,* 596–599.

Grizzle, J. E., Starmer, C. F., & Koch, G. G. (1969). Analysis of categorical data by linear models. *Biometrics, 25,* 489–504.

Huguenard, T., Sager, E. B., & Ferguson, L. W. (1970). Interview time, interview set, and interview outcome. *Perceptual and Motor Skills, 31,* 831–836.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books.

Katz, I. (1979). Some thoughts about the stigma notion. *Personality and Social Psychology Bulletin, 5,* 447–460.

Kelley, H. H. (1950). The warm–cold variable in first impressions of persons. *Journal of Personality, 18,* 431–439.

Kelley, H. H. (1973). The process of causal attribution. *American Psychologist, 28,* 107–128.

Kun, A. (1977). Development of the magnitude-covariation and compensation schemata in ability and effort attributions of performance. *Child Development, 48,* 862–873.

Langer, E. J., & Abelson, R. P. (1974). A patient by any other name . . . : Clinician group differences in labeling bias. *Journal of Consulting and Clinical Psychology, 42,* 4–9.

Langer, E. J., Taylor, S. E., Fiske, S. T., & Chanowitz, B. (1976). Stigma, staring, and discomfort: A novel stimulus hypothesis. *Journal of Experimental and Social Psychology, 12,* 451–463.

Livesley, W. J., & Bromley, B. D. (1973). *Person perception in childhood and adolescence.* London: Wiley.

MacClennon, R. N., & Jackson, D. N. (1985). Accuracy and consistency in the development of social perception. *Developmental Psychology, 21,* 31–36.

MacMillan, D. L., Jones, R. L., & Aloia, G. F. (1974). The mentally retarded label: A theoretical analysis and review of research. *American Journal of Mental Deficiency, 79,* 241–261.

Nicholls, J. G., & Miller, A. T. (1984). Development and its discontents: The differentiation of the concept of ability. In J. G. Nicholls (Ed.), *The development of achievement motivation* (pp. 185–218). Greenwich, CT: JAI Press.

Peevers, B. H., & Secord, P. F. (1973). Developmental changes in attribution of descriptive concepts to persons. *Journal of Personality and Social Psychology, 27,* 120–128.

Piner, K. E., & Kahle, L. R. (1984). Adapting to the stigmatizing label of mental illness: Foregone but not forgotten. *Journal of Personality and Social Psychology, 47,* 805–811.

Rapp, D. W. (1965). *Detection of observer bias in the written record.* Unpublished manuscript, University of Georgia.

Ray, A. A. (1982). *SAS user's guide: Statistics.* Cary, NC: SAS Institute.

Rholes, W. S., Blackwell, J., Jordan, C., & Walters, C. (1980). A developmental study of learned helplessness. *Developmental Psychology, 16,* 616–624.

Rholes, W. S., & Ruble, D. N. (1984). Children's understanding of dispositional characteristics of others. *Child Development, 55,* 550–560.

Towne, R. C., & Joiner, L. M. (1968). Some negative implications of special placement for children with learning disabilities. *Journal of Special Education, 2,* 217–222.

Weiner, B., Frieze, I., Kukla, A., Reed, L., Rest, S., & Rosenbaum, R. M. (1972). Perceiving the causes of success and failure. In E. E. Jones, D. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 95–120). Morristown, NJ: General Learning Press.

Weisz, J. R. (1981). Effects of the "mentally retarded" label on adult judgments about child failure. *Journal of Abnormal Psychology, 90,* 371–374.

Weisz, J. R. (1982). Learned helplessness and the retarded child. In E. Zigler & D. Balla (Eds.), *Mental retardation: The development–difference controversy* (pp. 27–40). Hillsdale, NJ: Erlbaum.

Weisz, J. R., Bromfield, R., Vines, D. L., & Weiss, B. (1985). Cognitive development, helpless behavior, and labeling effects in the mentally retarded. In F. Morrison & C. Lord (Eds.), *Advances in applied developmental psychology* (Vol. 3, pp. 129–168). New York: Academic Press.

Wilkins, J. E., & Velicer, W. F. (1980). A semantic differential investigation of children's attitudes toward three stigmatized groups. *Psychology in the Schools, 17,* 364–371.

Yeates, K. O., & Weisz, J. R. (in press). On being called mentally retarded: Does a developmental perspective mitigate labeling effects? *American Journal of Mental Deficiency.*

Received May 17, 1985
Revision received September 11, 1985 ■