Prosocial Development in Late Adolescence: A Longitudinal Study

Nancy Eisenberg, Gustavo Carlo, Bridget Murphy, and Patricia Van Court
Arizona State University

EISENBERG, NANCY; CARLO, GUSTAVO; MURPHY, BRIDGET; and VAN COURT, PATRICIA. Prosocial Development in Late Adolescence: A Longitudinal Study. CHILD DEVELOPMENT, 1995, 66, 1179-1197. Change in prosocial moral reasoning over 15 years, gender differences in prosocial reasoning, and the interrelations of moral reasoning, prosocial behavior, and empathy-related emotional responses were examined with longitudinal data from 17-18- and 19-20-year-olds and data from adolescents interviewed for the first time. Hedonistic reasoning declined in use until adolescence, and then increased somewhat in early adulthood. Needs-oriented and stereotypic reasoning increased until mid-childhood or early adolescence and then declined in use. Direct reciprocity and approval reasoning, which appeared to be on the decline in mid-adolescence in previous follow-ups, showed no decline into early adulthood. Several modes of higher-level reasoning increased in use across adolescence and early adulthood. Females' overall reasoning was higher than males'. Scores on interview and objective measures of prosocial moral reasoning were positively correlated. Consistent with expectations, there was some evidence of relations among prosocial reasoning, prosocial behavior, sympathy, and perspective taking.

Most research on moral judgment has concerned reasoning about moral dilemmas in which rules, laws, authorities' dictates, and formal obligations are central (Kohlberg, 1969, 1981; Rest, 1979). There has been much less research on positive morality (Eisenberg, 1986; Gilligan & Attanucci, 1988). To eliminate this gap in the research, Eisenberg and her colleagues have studied age changes in, and correlates of, prosocial moral reasoning (reasoning about moral dilemmas in which one person's needs or desires conflict with those of others in a context in which the role of prohibitions, authorities' dictates, and formal obligations is minimal; Eisenberg, 1986). The primary purpose of the present study was to extend the existing longitudinal study on this topic into late adolescence/early adulthood—that is, to age 17 to 18 years and age 19 to 20 years—and to examine the relations of prosocial moral reasoning to prosocial behavior and empathy-related characteristics. A secondary goal was to examine the relations among empathy-related characteristics and prosocial behavior.

The age-related changes in prosocial moral reasoning found in childhood and early to mid-adolescence generally have been consistent with Kohlberg's (1981) view that the capacities for complex perspective taking and for understanding abstract concepts are associated with advances in moral reasoning. Specifically, researchers have found that young children tend to use primarily hedonistic reasoning or needs-oriented (primitive empathic) prosocial reasoning. In elementary school, children's reasoning begins to reflect concern with approval and enhancing interpersonal relationships as well as the desire to behave in stereotypically "good" ways, although such reasoning appears to decrease in use somewhat in high school (Eisenberg, 1986). Direct reciprocity reasoning, which reflects an orientation to self-gain, has been found to increase in the elementary school years, perhaps because of the cognitive sophistication involved in thinking about reciprocity over time (Eisenberg et al., 1987). However, according to longitudinal data, such reasoning drops off in use somewhat in mid-
Although hedonistic reasoning reflects lower-level concerns and is most common in late adolescence (Eisenberg, Miller, Shell, McNalley, & Shea, 1991). Beginning in late elementary school or thereafter, children begin to express reasoning reflecting abstract principles, internalized affective reactions (e.g., guilt), and self-reflective sympathy and perspective taking. Nonetheless, at all ages individuals sometimes verbalize less mature modes of reasoning, although hedonistic reasoning decreases with age in childhood (Eisenberg, 1986). Of particular interest is the finding that in mid-adolescence (particularly at age 15 to 16 years), hedonistic reasoning, which dropped off sharply until age 11 to 12, increased somewhat, particularly for boys (Eisenberg et al., 1991). The latter finding is consistent with Eisenberg’s (1986) view that levels of prosocial moral reasoning are not hierarchical, integrated structures (so an individual’s reasoning is not necessarily primarily at one stage) or invariant in sequence and universal.

At this time, to our knowledge there are no data on prosocial moral reasoning beyond high school. Longitudinal data on prosocial moral reasoning have been collected from children from 4 to 5 years of age to 15 to 16 years of age (Eisenberg et al., 1991). Because moral development in regard to justice-related issues continues into adulthood (Colby, Kohlberg, Gibbs, & Lieberman, 1983), and advances in logical and sociocognitive capabilities occur during adolescence and early adulthood (Kuhn, Ambsel, & O’Loughlin, 1988; Selman, 1980), there are theoretical reasons to expect further development of prosocial moral reasoning in late adolescence and beyond. Moreover, changes in the complexity of the social environment as individuals move into late adolescence and college might be expected to stimulate perspective taking and, consequently, the development of higher-level moral reasoning (see Mason & Gibbs, 1993; Rest & Narvaez, 1991). Thus, we expected greater use of higher-level, internalized modes of reasoning with age. In contrast, stereotypic and approval/interpersonal-oriented reasoning, mid-level modes of prosocial moral reasoning, were expected to continue to decrease slightly with age (as they do in justice-oriented reasoning during late adolescence and adulthood; Colby et al., 1983).

It was more difficult to make predictions regarding changes in hedonistic (and, to some degree, direct reciprocity) reasoning. Although hedonistic reasoning reflects lower-level concerns and is most common in young children, some investigators have noted relative increases in such lower-level reasoning in late adolescence/early adulthood (Kohlberg & Kramer, 1969). Whether or not one finds evidence of an increase in lower-level reasoning at this age appears to depend on the method of coding of factors such as relativity in moral reasoning and structure of reasoning (see Candee & Kohlberg, 1987; Murphy & Gilligan, 1980). In the older coding system used with justice-oriented moral reasoning data—a system that tapped content of reasoning and structure of reasoning—regression in reasoning was noted in early adulthood (Kohlberg & Kramer, 1969). When the coding system was modified to assess primarily structure of justice-oriented moral reasoning and moral competence (rather than actual use of all types of moral reasoning), little evidence of moral regression was found (Colby et al., 1983). However, the more recently developed coding system involves several procedures that result in the discarding of lower-level reasoning (e.g., if such reasoning occurs infrequently or if a higher-level reason is stated for the same issue; see Colby & Kohlberg, 1987; Eisenberg, 1986). In the scoring of prosocial moral reasoning, content (e.g., an emphasis on egoistic issues) as well as structure, and performance rather than competence, are assessed; thus, a decrease in level of reasoning in late adolescence similar to that found with the older Kohlberg coding system might be expected. Such a “regression” might be due to the relativism, accompanied by seemingly egoistic reasoning due to an emphasis on individualism and individual rights. According to some theorists, relativistic reasoning and a seemingly individualistic focus are part of the process of rejecting conventional moral reasoning (moral reasoning oriented toward the arbitrary maintenance of social order, authority, and societal rules) and shifting to principled, internalized moral reasoning that emphasizes individual rights (Kohlberg & Gilligan, 1971; Turiel, 1974).

From another perspective, Gilligan (1977) noted a shift in early adulthood from a focus on goodness as self-sacrifice to an emphasis on responsibility for the self as well as others. In prosocial moral dilemmas, the conflict is between one’s own and another’s needs, wants, and desires; story protagonists can assist another, but at a cost. Thus, some older adolescents and young adults may express more self-oriented reasoning in response to prosocial moral dilem-
mas as they realistically weigh the relative costs, and their responsibilities, to self and other in moral dilemmas. Given all the aforementioned factors, it was difficult to predict age-related changes in hedonistic moral reasoning in late adolescence.

As in prior follow-ups of the longitudinal study, prosocial moral reasoning was assessed with Eisenberg’s interview procedures. However, at the 15-year follow-up (at age 19 to 20 years), an objective pencil-and-paper measure of prosocial moral judgment (the PROM) also was used (Carlo, Eisenberg, & Knight, 1992). This relatively new instrument had never been used with individuals past tenth grade; thus, we wanted to examine its relation to empathy, perspective taking, prosocial behavior, and moral reasoning as assessed in interviews for this age group.

As noted previously, a goal of this study was to examine the relations among prosocial moral reasoning, prosocial behavior, and empathy-related reactions in late adolescence/early adulthood. Few researchers have studied adolescents’ empathy-related reactions, and even studies of adolescents’ prosocial behavior are rare (Eisenberg, 1990).

Theorists such as Candee and Kohlberg (1987) and Rest (1979) have argued that moral reasoning influences individuals’ moral decisions and behavior. In fact, moral reasoning, including prosocial moral judgment, in general is modestly correlated with the performance of prosocial behaviors (Eisenberg, 1986; Underwood & Moore, 1982). Children’s and early- to mid-adolescents’ prosocial behavior generally has been positively related to need-oriented reasoning, negatively related to hedonistic reasoning (Eisenberg, 1986; Eisenberg et al., 1987), and occasionally positively correlated with a composite measure of adolescents’ prosocial moral reasoning (e.g., at age 15 to 16, but not 13 to 14, in the longitudinal study; Eisenberg et al., 1991). Researchers have hypothesized that the relation between moral reasoning and behavior increases with age because higher-level reasoning is associated with the “progressive stripping away of bases for justifying behavior that are extrinsic to principle” (p. 104, Rhoses & Bailey, 1983), resulting in stronger motivation to maintain consistency between attitudes and behaviors at higher stages of development (Kohlberg & Candee, 1984). Thus, measures of the longitudinal subjects’ prosocial behavior and inhibition of aggression in late adolescence/early adulthood were expected to be positively correlated with other-oriented and perhaps higher-level modes of reasoning, and negatively correlated with hedonistic reasoning.

Some types of prosocial moral reasoning explicitly reflect cognitive perspective taking and sympathetic tendencies; further, perspective taking is viewed as underlying advances in moral judgment (Colby et al., 1983; Eisenberg, 1986) and some types of prosocial behavior (Underwood & Moore, 1982). Nonetheless, there are very few studies of the relations of empathy-related dispositional characteristics to moral reasoning or adolescents’ prosocial behavior, particularly studies in which researchers have differentiated among various empathy-related responses. Investigators have argued that sympathy (concern for others based on the apprehension of another’s state) and empathy (an emotional reaction elicited by and congruent with another’s state) stimulate the development of internalized moral reasoning reflecting concern for others’ welfare (Hoffman, 1987) and prime the use of pre-existing other-oriented moral cognitions (Eisenberg, 1986). Further, sympathy, which frequently may stem from empathy and perspective taking (Hoffman, 1987), is viewed as resulting in other-oriented, altruistic motivation and has been associated with higher levels of prosocial behavior (Eisenberg & Fabes, 1991). In contrast, personal distress (a self-focused, aversive reaction to cues regarding another’s negative state or emotion; Davis, 1983) has been theoretically linked to egoistic motives and empirically associated with low levels of prosocial behavior (Batson, 1991; Eisenberg & Fabes, 1991). Thus, cognitive perspective taking and sympathy would be expected to be positively correlated with higher-level, other-oriented modes of prosocial moral reasoning and altruism, whereas an inverse pattern of relations would be expected for personal distress.

Consistent with these expectations, at age 15 to 16, cognitive perspective taking was positively related to higher-level reasoning (on the moral reasoning composite score) and negatively related to hedonistic reasoning, whereas dispositional sympathy was positively correlated with primitive empathic (i.e., needs-oriented) reasoning and negatively correlated with hedonistic reasoning (Eisenberg et al., 1991). Similarly,
Carlo et al. (1992) found a positive relation between seventh and tenth graders' internalized prosocial moral reasoning on the PROM and their sympathy and perspective taking. Findings in regard to personal distress are less consistent. Personal distress was unrelated to adolescents' prosocial moral reasoning in the longitudinal study (Eisenberg et al., 1991); in contrast, it was negatively related to internalized and stereotypic reasoning, and positively related to approval-oriented reasoning, on the PROM (Carlo et al., 1992). In the present study, we further examined the relations of prosocial moral reasoning with various empathy-related dispositional characteristics.

Consistent with theory (Hoffman, 1987), some evidence of relations between prosocial behavior and empathy-related characteristics also has been found, primarily in studies involving younger children and global measures of empathy (see Eisenberg & Fabes, 1991; Eisenberg & Miller, 1987; Underwood & Moore, 1982). In a prior follow-up of the longitudinal study, measures of helping behavior and mothers' reports of children's prosocial behavior at age 15 to 16 were unrelated to adolescents' sympathy, perspective taking, or personal distress, whereas adolescents' self-reported prosocial behavior was positively correlated with both sympathy and perspective taking (Eisenberg et al., 1991). Thus, findings in regard to the relations of specific empathy-related characteristics to adolescents' prosocial behavior are inconsistent. Consequently, another goal of this study was to examine the relations between empathy-related characteristics and prosocial behavior during late adolescence/early adulthood.

Also of interest was the consistency of measures of both empathy-related responding and prosocial behavior across time. Although relevant research is quite limited, there is evidence of some consistency in prosocial behavior across 2 or more years in childhood (Eisenberg et al., 1987; Radke-Yarrow & Zahn-Waxler, 1984) and across 2 years (Eisenberg et al., 1991) or several weeks (Small, Zeldin, & Savin-Williams, 1983) in early to mid-adolescence. To our knowledge, consistency of prosocial behavior across longer periods of time in adolescence, and into adulthood, has not been examined. The research on consistency of empathy-related reactions is even more sparse, although sympathy, personal distress, and perspective taking appear to be fairly consistent across 2 or 3 years' time in the junior high and high school years (Eisenberg et al., 1991). Davis and Franzoi (1991) noted that the magnitude of these correlations increased somewhat (albeit nonsignificantly) with age; thus, we expected the consistency of empathy-related characteristics to be fairly high from age 17–18 to 19–20.

Finally, it is possible that self-report measures of moral reasoning and behavior are contaminated by self-presentational concerns. Thus, in the present study we examined the relation of social desirability to other moral indexes. However, because measures of social desirability may partially tap characteristics associated with moral functioning, such as the ability to inhibit oneself (King, Emmons, & Woodley, 1992), as well as the desire to present oneself in a positive light, controlling for social desirability may be a conservative analytic strategy.

**Method**

**Subjects**

Two groups of middle-class children participated in this study. The primary longitudinal cohort (C1) consisted of 16 girls and 16 boys (all Caucasian except two) who had been interviewed seven times previously, at ages 4–5, 5½–6½, 7–8, 9–10, 11–12, 13–14, and 15–16 years (at 156, 138, 120, 96, 72, 48, and 24 months prior to the first assessment in this study); the nine testing sessions henceforth will be referred to as T1 to T9. The mean ages of the children at T8 and T9 were 211 months (range = 202–219 months; approximately age 17–18 years) and 235 months (approximately age 19–20 years). No children were lost since T3 (in 12 years); five have been lost over the 15-year period (three boys, two girls). Mean years of maternal and paternal education for this sample (as reported at T8) were 16.0 and 17.0, respectively (range = 12 to 20 years for both). At T8, all subjects except one were living at home; at T9, all but five were attending college, and only 11 were living with their parents.

The second sample (C2) consisted of 34 twelfth graders from middle-class, predominantly (85%) Caucasian families (20 girls, 14 boys; M age = 213 months). These children attended a public school in the suburban city in which the longitudinal subjects lived at the beginning of the study. They were interviewed only at T8. Means years of education for subjects' mothers and fathers were 14.6 and 15.6 years, respectively.
Moral reasoning.—Children's prosocial moral reasoning was assessed in two ways. First, at both T8 and T9, students were interviewed with the same four moral reasoning stories used in prior follow-ups (see Eisenberg, Lennon, & Roth, 1983; Eisenberg et al., 1987, 1991), although a few words were changed at adolescent follow-ups to make the stories sound less childlike (e.g., "birthday party" was changed to "birthday celebration"). An additional story previously used with adolescents in another study (concerning going into the hospital to donate blood at a cost to the self; Eisenberg-Berg, 1979) also was used at T6 to T9. This story was added because the costs of helping in the giving blood story likely would be substantial for adolescents (losing time at work and school). Story protagonists (if specified) were the same sex as the subject. At T8, the moral reasoning interview for a C2 male was lost due to experimenter error. At T9, four C1 participants did not complete the interview portion of the study because they had moved from the state (2), could not be reached (1), or refused (1).

Second, at T9 only (and, consequently, for only C1), participants were administered the PROM (Carlo et al., 1992). Participants were mailed a version of the PROM containing six stories, five of which were very similar in content to the vignettes used in the moral reasoning interview. The PROM is modeled after Rest's (1979) Defining Issues Test (DIT); subjects are presented with six stories, five of which are similar in content to the vignettes used in the moral reasoning interview. The PROM composite scores (for the five types of reasoning) by the sum of the five PROM composite scores. The hedonistic response (the tendency to feel unease and discomfort in tense interpersonal settings involving others' needs or emotions; .71, .76, and .76). Items on each scale were averaged to compute a composite score. SD was missing for one subject at T9.

Social desirability.—To assess social desirability (SD), participants in C1 and C2 completed 25 items from the Marlowe-Crowne (Crowne & Marlowe, 1964) social desirability scale (alphas = .84 and .89 for C1 at T8 and T9, respectively, and .74 for C2 at T8). Items were averaged to compute a composite score. SD was missing for one subject at T9.

Empathy-related responding.—At T8 and T9, three subscales of Davis's (1983) Interpersonal Reactivity Scale were administered (definitions and alphas for C1 at T8 and T9, and C2 at T8, are in parentheses): sympathy (the tendency to experience feelings of warmth and concern for others; alphas = .71, .83, and .72, respectively; one item was dropped at T8), perspective taking (the tendency to adopt the point of view of others; .79, .91, and .84), and personal distress (the tendency to feel uneasy and discomfort in tense interpersonal settings involving others' needs or emotions; .71, .76, and .76). Items on each scale were averaged (after reversing items if appropriate). Scores for these scales were missing for one C1 subject at T8.

Moral behavior.—At T8, students filled out a 23-item adapted version of Rushton, Chrisjohn, and Fekken's (1981) self-report altruism scale (alphas = .87 for both C1 and
C2). This measure (henceforth called reported prosocial behavior) also was used at T6 and T7. Students indicated on a five-point scale (from “never” to “very often”) how frequently they engaged in 23 behaviors, such as giving money to charity or volunteer work.

At T8, mothers of children in C1 reported on their children’s prosocial behavior using the same measure (with the wording slightly modified). Because they were given the additional option of “Don’t know,” alphas could not be computed (items with this response were coded as missing, so few mothers had scores for all items). One mother did not complete this measure. For both child and mother versions of the scale, scores on the items were averaged.

Students at T8 in both C1 and C2 also were given an opportunity to assist the experimenter by filling out some additional questionnaires and returning them in a stamped, addressed envelope (henceforth called the measure of helping). Students were given a score indicating whether they returned the measures (1 = not returned; 2 = returned).

An additional donating task was administered only to C2 participants at T8. At the end of the interview session, C2 students were paid $5 (4 $1 bills and 4 quarters) for their time, and were told that the experimenter belonged to an organization at the university that was collecting money for a widely publicized child that needed a liver transplant (a newspaper article was presented to them alongside a covered donation box with money in it). Students also were told that they could donate money if they wished to do so and were left alone for 1 min. The amount of money donated was the measure of donating. The donating task was not used for C1 because subjects remembered similar tasks used when they were younger.

At T9, self-reported moral behavior was assessed for C1 with portions of Weinberger’s Adjustment Inventory (WAI; Weinberger, 1991; also see Feldman & Weinberger, 1994). Items from the longer version of two restraint subscales that concerned moral behavior were used (rated 1 = false to 5 = true). The subscales used were as follows (sample items, numbers of items, and alphas for this sample are in parentheses): consideration of others (“I often go out of my way to do things for other people,” seven items, .79) and suppression of aggression (“I lose my temper and let people have it when I’m angry” [reversed], seven items, .87). Items within each subscale were averaged.

Friends’ reports.—At T9, subjects were asked to provide the names of one or two friends to fill out a questionnaire about the subject. Reports from at least one friend were obtained for 25 subjects; reports from two friends were obtained for 11 participants. Mean length of time that friends reported knowing the subjects was 66.42 months. The friends’ questionnaire packet included the seven-item sympathy and perspective-taking subscales (modified slightly for an other-report format). In addition, friends responded to items from Weinberger’s WAI Rating Items, which contains three items per scale. Items on these subscales were similar to those used in the subject report measures. If reports from two friends were obtained, they were averaged (correlations between reports for the two friends for sympathy, perspective taking, consideration for others, and suppression of aggression were .85, .18, .59, .21). Alpha coefficients for the three-item scales were .77 and .62 for consideration of others and aggression.

Procedures
C1 at T8.—Interviews for C1 at T8 usually took place at the university. Mother and child were interviewed individually in different rooms, the mother by a woman and the children by a man who had not been involved in any previous follow-ups. For the adolescents, the prosocial dilemmas were presented in random order; they were read to the adolescents while the adolescents read along (responses were taped). Participants repeated dilemmas to check for comprehension, and a standard sequence of questioning was followed (Eisenberg et al., 1987).

Either subsequent or prior to the moral interview (sequence was counterbalanced across subjects), the children completed the measures of empathy-related constructs, social desirability, and self-reported prosocial behavior (presented in random order). The students were told that their responses were confidential. Next the children were paid for their participation ($20) and the experimenter told the students that he would appreciate them filling out a few more forms at home, but that they need not do so. If the student agreed to take the questionnaires (all did), they were given the forms and a stamped, addressed envelope.
C2 at T8.—Adolescents in C2 were administered the same interview and questionnaires as C1 at their schools by either a male or female experimenter (the male experimenter was the same as used for C1 at T8; approximately half the subjects of each sex were interviewed by each experimenter). Subjects' mothers were not involved. At the end of the session, the helping and donating tasks described previously were administered.

C1 at T9.—At T9, the subjects in C1 were first sent a packet including the PROM and the sympathy, personal distress, perspective taking, and WAI questionnaires (order of the PROM and questionnaires was counterbalanced). Subjects were also asked to supply names and addresses of friends if they were willing to do so (all but three did so). Completed questionnaires were returned by mail, and subjects were sent $25 for participation. After completing the questionnaires, subjects were asked if they would be willing to complete the prosocial moral reasoning interview. Twenty-eight subjects were available for the interview, and they were individually interviewed approximately 1 to 4 months after completing the questionnaires (depending on their availability). Participants were paid an additional $20 for the interview.

Scoring of the Prosocial Reasoning Stories (Interviews)

Reasoning was coded into the categories of reasoning outlined by Eisenberg et al. (1983, 1987, 1991; Eisenberg-Berg, 1979). Those used with any frequency were as follows:

Hedonistic reasoning: (a) hedonistic gain to the self (orientation to gain for oneself), (b) direct reciprocity (orientation to personal gain because of direct reciprocity or lack of reciprocity from the recipient of an act), (c) affectional relationship (orientation to the individual's identification or relationship with another or liking for the other);

Pragmatic (orientation to practical concerns that are unrelated to selfish considerations);

Needs-oriented (orientation to the physical, material, or psychological needs of the other person, e.g., “He needs blood,” or “He’s sad”);

Stereotypes of a good or bad person (orientation to stereotyped images of a good or bad person);

Approval and interpersonal orientation (orientation to others' approval and acceptance in deciding what is the correct behavior);

Self-reflective empathic orientation: (a) sympathetic orientation (expression of sympathetic concern and caring for others), (b) role taking (the individual explicitly takes the perspective of the other or has the story protagonist do so), (c) internalized positive affect related to consequences (orientation to internal positive affect as a result of a particular course of action because of the consequences of one's act for the other person), (d) internalized negative affect related to consequences (the same as [c] but for negative affect);

Internalized affect because of loss of self-respect and not living up to one's values: (a) positive (orientation to feeling good as a consequence of living up to internalized values), (b) negative (concern with feeling bad as a consequence of not living up to internalized values);

Internalized law, norm, and value orientation (orientation to an internalized responsibility, duty, or need to uphold the laws and accepted norms or values);

Other abstract and/or internalized types of reasoning: (a) generalized reciprocity (orientation to indirect reciprocity in a society, that is, exchange that is not one-to-one but eventually benefits all or a larger group), (b) concern with the condition of society (orientation to improving the society or community as a whole), (c) concern with individual rights and justice (orientation to protecting individual rights and preventing injustices that violate another's rights), (d) equality of people (orientation to the principle of the equal value of all people).

Subjects were assigned scores indicating the frequency with which they used each of the various types of reasoning when discussing both the pros and cons of helping the needy other in the story dilemma (1 = no use of category; 2 = vague, questionable use; 3 = clear use of a mode of reasoning; 4 = a major type of reasoning used). Next, the scores for each category were summed across the stories. At T8, a second scorer coded half the data for reliability; at T9, two additional people co-scored the data for reliability. Interrater reliabilities for T1 to T7 are presented in previous papers (e.g., Eisenberg et al., 1983, 1987, 1991). For all time periods, the primary coder was the same per-
son, whereas seven persons have served as reliability coders over the nine time periods. To prevent bias in scoring, the primary coder (as well as the secondary coder at T8 and one secondary coder at T9) was blind to the identity of the children. The coders also were blind to any information regarding the subjects’ scores on other measures. Interrater reliabilities (Pearson correlations) computed for each reasoning category at T8 and T9 (using data for all subjects at T9, and for half the subjects at T8) ranged from .70 to 1.00 at T8, with most being above .85, and from .77 to 1.00 at T9 (with all but two being above .83). (These reliabilities are for five stories; those for four stories were very similar.)

To determine if there was change in the primary coder’s scoring over the years (and to prevent the primary coder from knowing the age of subjects being coded), four protocols from each of the previous follow-ups were mixed together with the various protocols from T8 and T9 and rescored by the primary coder to determine if there was any change in her scoring over the years (the coder was blind to which protocol was from which follow-up). Scoring of the data from earlier sessions was highly similar to the original scores for the same data (agreement on codings within 1 point was 90% or higher on all categories).

The categories of reasoning are viewed as representing components of developmental levels of prosocial moral reasoning; these levels were derived from the results of cross-sectional research (Eisenberg-Berg, 1979; see Eisenberg, 1986). Briefly, the levels are as follows: Level 1, hedonistic, self-focused orientation; Level 2, needs-of-others orientation; Level 3, approval and interpersonal orientation and stereotyped orientation; Level 4, self-reflective, empathic orientation; and Level 5, strongly internalized orientation. Based on these levels, a score representing level of moral judgment was computed for each child. The level score was constructed in a manner similar to that used to score Kohlbergian moral reasoning; subjects were assigned composite scores by weighing the proportion of the child’s reasoning at each level (see Eisenberg et al., 1983, for more detail). Because it is debatable whether Level 5 is more moral than Level 4, and because they were weighted equally in previous follow-ups, they were weighted equally in the analyses presented in this article (although the data changed little if Level 5 was weighted higher).

Results

Age Changes in Moral Judgment

To examine age changes in moral reasoning for C1 over the 15 years, multivariate and univariate trend analyses of variance were computed with one within-subjects factor (time; adjusted for unequal time gaps when appropriate) and one between-subjects factor (sex). Only the 28 subjects interviewed at T9 were included in the trend analyses (recall that four C1 subjects were not interviewed at T9). Different multivariate analyses had to be computed for groups of reasoning that emerged at different ages because linear dependencies in the data occur if a particular mode of reasoning is not used at more than one time period (and because quadratic trends could occur if a type of reasoning was not used in childhood and then emerged in adolescence). Only categories of reasoning used with some frequency during at least one time period were included in the analyses. Because types of reasoning that were used infrequently tended to be positively skewed, a logarithmic transformation was performed on the data (although the means presented in Table 1 and in the text are nontransformed means). Linear, quadratic, and cubic trends were examined for early emerging modes of reasoning because from early childhood through late adolescence, some categories of reasoning were expected to show both increases and decreases in usage, sometimes with a period of relative stability in use (which could result in a cubic trend analysis).

In the first analysis, the categories of reasoning were those that had been used with some frequency (by at least one sex) at eight or more time periods (i.e., hedonistic, needs-oriented, pragmatic, direct reciprocity, approval-oriented, and stereotypic; see Eisenberg et al., 1987). Scores were computed from the four stories used at all seven follow-ups. The multivariate Fs for the linear, quadratic, and cubic effects of time were highly significant, \( F(7, 20) = 36.47, 13.55, \) and 3.34, \( ps < .001, .001, \) and .016. For hedonistic reasoning, the univariate Fs for the linear and quadratic trends were highly significant; the cubic trend was somewhat lower in significance. Hedonistic reasoning decreased sharply with age until 11–12 years, and then increased somewhat in adolescence and early adulthood (with a slight drop at age 17 to 18; see Table 1 for means and Table 2 for Fs). Interestingly, perusal of the means indicated that scores on hedonistic reasoning for girls changed little in adolescence until age 19 to 20 years (Ms for T5,
### TABLE 1

**Moral Reasoning Categories: Means for Cohort 1**

<table>
<thead>
<tr>
<th>Reasoning Category</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hedonistic</td>
<td>12.11</td>
<td>8.89</td>
<td>6.46</td>
<td>5.39</td>
<td>4.57</td>
<td>4.86</td>
<td>5.32</td>
<td>4.75</td>
<td>5.43</td>
</tr>
<tr>
<td>Direct reciprocity</td>
<td>4.00</td>
<td>4.07</td>
<td>4.11</td>
<td>4.25</td>
<td>5.32</td>
<td>5.71</td>
<td>4.89</td>
<td>5.29</td>
<td>5.50</td>
</tr>
<tr>
<td>Affectional relation</td>
<td>4.04</td>
<td>4.43</td>
<td>4.00</td>
<td>4.18</td>
<td>4.21</td>
<td>4.11</td>
<td>4.61</td>
<td>4.36</td>
<td>4.18</td>
</tr>
<tr>
<td>Pragmatic</td>
<td>4.07</td>
<td>4.11</td>
<td>4.18</td>
<td>4.86</td>
<td>5.36</td>
<td>6.00</td>
<td>6.46</td>
<td>5.75</td>
<td>6.50</td>
</tr>
<tr>
<td>Stereotypic</td>
<td>4.57</td>
<td>4.36</td>
<td>4.68</td>
<td>5.14</td>
<td>5.71</td>
<td>6.79</td>
<td>6.46</td>
<td>6.43</td>
<td>6.25</td>
</tr>
<tr>
<td>Approval/interpersonal</td>
<td>4.00</td>
<td>4.07</td>
<td>4.21</td>
<td>4.36</td>
<td>4.82</td>
<td>4.96</td>
<td>5.00</td>
<td>4.61</td>
<td>5.00</td>
</tr>
<tr>
<td>Sympathetic</td>
<td>4.00</td>
<td>4.03</td>
<td>4.00</td>
<td>4.21</td>
<td>4.39</td>
<td>4.07</td>
<td>4.14</td>
<td>4.04</td>
<td>4.32</td>
</tr>
<tr>
<td>Role taking</td>
<td>4.00</td>
<td>4.00</td>
<td>4.07</td>
<td>4.68</td>
<td>4.68</td>
<td>5.14</td>
<td>6.00</td>
<td>5.57</td>
<td>6.07</td>
</tr>
<tr>
<td>Positive affect/simple or related to consequences</td>
<td>4.00</td>
<td>4.00</td>
<td>4.11</td>
<td>4.57</td>
<td>4.79</td>
<td>5.21</td>
<td>5.39</td>
<td>4.79</td>
<td>6.04</td>
</tr>
<tr>
<td>Negative affect/simple or related to consequences</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.18</td>
<td>4.21</td>
<td>4.11</td>
<td>4.29</td>
<td>4.61</td>
<td>4.68</td>
</tr>
<tr>
<td>Positive affect/values and self-respect</td>
<td>4.00</td>
<td>4.00</td>
<td>4.11</td>
<td>4.04</td>
<td>4.00</td>
<td>4.14</td>
<td>4.25</td>
<td>4.68</td>
<td>4.39</td>
</tr>
<tr>
<td>Negative affect/values and self-respect</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.07</td>
<td>4.00</td>
<td>4.07</td>
<td>4.07</td>
<td>4.07</td>
<td>4.29</td>
</tr>
<tr>
<td>Internalized law, norm, or value orientation</td>
<td>4.00</td>
<td>4.00</td>
<td>4.04</td>
<td>4.00</td>
<td>4.00</td>
<td>4.18</td>
<td>4.39</td>
<td>4.46</td>
<td>4.86</td>
</tr>
<tr>
<td>Generalized reciprocity</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.04</td>
<td>4.29</td>
<td>4.18</td>
<td>4.25</td>
<td></td>
</tr>
<tr>
<td>Condition of society</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.04</td>
<td>4.18</td>
<td>4.14</td>
<td>4.14</td>
</tr>
<tr>
<td>Individual rights</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.14</td>
<td>4.14</td>
<td>4.04</td>
<td>4.11</td>
<td>4.14</td>
</tr>
<tr>
<td>Equality of individuals</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.00</td>
<td>4.07</td>
<td>4.00</td>
<td>4.00</td>
<td></td>
</tr>
</tbody>
</table>

**NOTE.** Means presented are for the nontransformed data for the 28 subjects with interviews at all time periods.

### TABLE 2

**Trend Analyses for the Individual Moral Reasoning Categories: Univariate Effects**

<table>
<thead>
<tr>
<th>Moral Category</th>
<th>Linear</th>
<th>Quadratic</th>
<th>Cubic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hedonistic</td>
<td>63.13****</td>
<td>85.03****</td>
<td>10.72**</td>
</tr>
<tr>
<td>Direct reciprocity</td>
<td>49.84****</td>
<td>2.60</td>
<td>1.75</td>
</tr>
<tr>
<td>Affectional relationship</td>
<td>8.98**</td>
<td>.40</td>
<td>2.61</td>
</tr>
<tr>
<td>Pragmatic</td>
<td>64.36****</td>
<td>4.12*</td>
<td>3.30*</td>
</tr>
<tr>
<td>Needs-oriented</td>
<td>1.69</td>
<td>51.97****</td>
<td>17.76***</td>
</tr>
<tr>
<td>Stereotypic</td>
<td>35.96***</td>
<td>3.52*</td>
<td>7.17*</td>
</tr>
<tr>
<td>Approval</td>
<td>22.21****</td>
<td>2.71</td>
<td>.31</td>
</tr>
<tr>
<td>Sympathetic</td>
<td>.48</td>
<td>.01</td>
<td>. .</td>
</tr>
<tr>
<td>Role taking</td>
<td>17.26**</td>
<td>.57</td>
<td>. .</td>
</tr>
<tr>
<td>Positive affect/simple or related to consequences</td>
<td>41.20****</td>
<td>.12</td>
<td>. .</td>
</tr>
<tr>
<td>Negative affect/simple or related to consequences</td>
<td>8.98**</td>
<td>1.35</td>
<td>. .</td>
</tr>
<tr>
<td>Positive affect/values and self-respect</td>
<td>10.59**</td>
<td>.88</td>
<td>. .</td>
</tr>
<tr>
<td>Negative affect/values and self-respect</td>
<td>2.08</td>
<td>.51</td>
<td>. .</td>
</tr>
<tr>
<td>Internalized law, norm, or value orientation</td>
<td>19.36</td>
<td>2.68</td>
<td>. .</td>
</tr>
<tr>
<td>Generalized reciprocity</td>
<td>3.82*</td>
<td>. .</td>
<td>. .</td>
</tr>
<tr>
<td>Condition of society</td>
<td>2.66</td>
<td>. .</td>
<td>. .</td>
</tr>
<tr>
<td>Individual rights</td>
<td>.01</td>
<td>. .</td>
<td>. .</td>
</tr>
<tr>
<td>Equality of individuals</td>
<td>.86</td>
<td>. .</td>
<td>. .</td>
</tr>
</tbody>
</table>

* *Indicate that an analysis was not computed (see text).

*p < .10.

*p < .05.

**p < .01.

***p < .001.

****p < .0001.
T6, T7, T8, and T9 were 4.60, 4.60, 4.67, 4.27, and 5.00), whereas such reasoning started to increase in frequency at age 13 to 14 (T6) for boys (Ms = 4.54, 5.15, 6.08, 5.31, and 5.92). According to an additional analysis, the increase in hedonistic reasoning from T8 to T9 was significant, F(1, 26) = 4.46, p < .045. For needs-oriented reasoning, there was a highly significant quadratic trend and a weaker (but highly significant) cubic trend. Needs-oriented reasoning increased with age until 7–8 years, was relatively stable from 7–8 to 11–12 years, and declined considerably in adolescence, particularly at age 19 to 20 years (see Table 1). According to a highly significant linear trend, direct reciprocity reasoning increased in use with age; specifically, it was used with little frequency until age 9–10, increased in use until early adolescence (13–14 years), started to decline slightly in mid-adolescence, and then shot up in use in late adolescence. For stereotypic reasoning, there was a highly significant linear trend and a cubic trend. Stereotypic reasoning was used infrequently until mid- to late elementary school, increased in use until age 13–14, and then decreased slightly in use in mid- to late adolescence. For approval-oriented reasoning, only the linear trend was significant; such reasoning increased until age 15 to 16, dropped slightly at age 17 to 18, and then rose slightly again at age 19 to 20. Finally, pragmatic reasoning increased in a linear fashion with age.

A second 2 (sex) × 5 (time) trend analysis was computed for those higher-level categories of reasoning used with any frequency at T3 or T4 (sympathetic, role taking, internalized positive affect about consequences, internalized negative affect about consequences, internalized positive affect about values, internalized negative affect about values, and internalized law, norm, or value orientation reasoning). The multivariate Fs for sex and the linear trend for time were significant, Fs(7, 20) = 2.60 and 16.38, p < .044 and .001. Females scored higher than males (at p < .059 or better) on role taking, positive affect/values, and negative affect/consequences reasoning, Fs(1, 26) = 3.91, 5.73, and 4.20, ps < .059, .024, and .051 (Ms = 5.47, 4.35, and 4.43 for females; 4.84, 4.09, and 4.14 for males). Role taking, positive affect/consequences, positive affect/values, negative affect/consequences, and internalized norm, rule, and law reasoning increased with age (see Table 2).

In a third 2 (sex) × 3 (time) analysis, we examined linear and quadratic age changes in the use of categories of reasoning that emerged only in adolescence (generalized reciprocity, concern with society, rights and justice, and equality of people reasoning). These categories of reasoning were used quite infrequently (see Table 1). The multivariate effect for the sex × quadratic interaction was significant, p < .041; however, follow-up analyses were not. The multivariate linear trend was only marginally significant, p < .10.

In a summary analysis, we examined change in C1 adolescents’ moral reasoning composite scores from late childhood (T5) through late adolescence. According to a 2 (sex) × 5 (time) trend analysis, there were main effects of sex and a strong linear trend, Fs(1, 26) = 18.84 and 30.09, ps < .001. Females scored higher than males, and scores increased with age (Ms = 231, 241, 251, 252, and 270 for T5, T6, T7, T8, and T9, respectively).

Correlations also were computed to examine the relations between moral reasoning at earlier periods of adolescence (T6 and T7) and moral reasoning at T8 and T9. Reasoning at T6 was at least marginally positively related to reasoning at T7 and T8, rs(30) = .37 and .32, ps < .035 and .078. Reasoning at T7 was positively correlated with reasoning at T8, r(30) = .46, p < .001. In contrast, there was a discontinuity in reasoning at T9; T9 composite scores were not significantly related to analogous scores at T6, T7, or T8, rs(26) = .20, .22, and .07, N.S.

If the age trends in the children’s reasoning were primarily the result of repeated testing, one would not expect the reasoning for C1 to be similar to that of children of the same age interviewed for the first time at T8 (C2). At T8, the only difference in reasoning

---

1 For the entire sample of longitudinal and cross-sectional subjects, there was not a significant sex difference in the moral reasoning composite scores at T8, whereas at T9 females scored higher than males on the composite scores composed of both five stories and marginally higher on that for the primary four stories, ps < .016 and .08. In regard to individual moral reasoning categories, for C1 and C2 at T8, males scored higher on affectional relationship reasoning, ps < .044, whereas females scored higher on pragmatic reasoning at T9, p < .007 (ps are for four stories; those for five stories were similar).
between C1 and C2 was that C1 used more approval and needs-oriented reasoning, \( F(1, 62) = 5.33 \) and \( 7.11, p < .024 \) and .011. Neither of these types of reasoning is higher level; thus, if there was any effect of repeated testing, it was to encourage greater expression of relatively low-level reasoning.

**Correlation of Interview Moral Reasoning with Scores on the PROM**

Mean scores for adolescents' PROM subscales are presented in Table 3; according to an analysis of variance for repeated measures, the means differed significantly, \( F(4,124) = 24.14, p < .001 \). Internalized reasoning was most preferred; scores on stereotypical and needs-oriented reasoning were moderately high; hedonistic reasoning was the second least preferred category; and approval-oriented reasoning were least preferred. All categories differed significantly from one another at \( p < .01 \), except needs-oriented reasoning did not differ from either internalized or stereotypic reasoning, and hedonistic and stereotypic reasoning differed at \( p < .05 \) (Tukey's tests). This pattern is highly consistent with that obtained by Carlo et al. (1992).

To compare reasoning on the PROM with the composite index of interview prosocial moral reasoning, a composite score for the PROM was computed. Based on the means in this study and the findings in Carlo et al. (1992), a weighted score was computed in which percent of internalized reasoning was multiplied by 3, percent of needs-oriented and stereotypic reasoning was multiplied by 2, and percents of hedonistic and approval-oriented reasoning were multiplied by 1. These weighted values were then summed. Although approval-oriented reasoning is considered of moderate level in studies of spontaneously elicited moral reasoning (e.g., Eisenberg et al., 1991), it is clear that the approval-oriented items on the PROM, a preference measure of moral judgment requiring merely the endorsement or rejection of options, reflected relatively low-level moral reasoning. Students tended to reject the bluntly worded approval-oriented items, and use of these items decreased with age (Carlo et al., 1992). When the composite PROM and interview moral reasoning composites (computed for all five vignettes at T9) were correlated, the relation was moderate and comparable to findings for the DIT, \( r(26) = .57, p < .002 \). Moreover, adolescents' interview moral reasoning composite scores at T8 were positively correlated with their PROM composite score at

**TABLE 3**

**MEANS AND STANDARD DEVIATIONS FOR MEASURES OF SOCIAL DESIRABILITY, PROSOCIAL BEHAVIOR, EMPATHY-RELATED CHARACTERISTICS, AND SCORES ON THE PROM**

<table>
<thead>
<tr>
<th>MEASURE</th>
<th>T8 M</th>
<th>T8 SD</th>
<th>T9 M</th>
<th>T9 SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social desirabilityb</td>
<td>1.42</td>
<td>.18</td>
<td>1.39</td>
<td>.24</td>
</tr>
<tr>
<td>Helpingb</td>
<td>1.45</td>
<td>.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Donatingc</td>
<td>1.82</td>
<td>1.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-reported moral behaviorde</td>
<td>3.17</td>
<td>.52</td>
<td>.00</td>
<td>1.80</td>
</tr>
<tr>
<td>Mothers' reports of prosocial behaviord</td>
<td>3.13</td>
<td>.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sympathyf</td>
<td>3.95</td>
<td>.64</td>
<td>5.51</td>
<td>.99</td>
</tr>
<tr>
<td>Perspective takingf</td>
<td>3.37</td>
<td>.76</td>
<td>4.64</td>
<td>1.29</td>
</tr>
<tr>
<td>Personal distressg</td>
<td>3.17</td>
<td>.52</td>
<td>3.00</td>
<td>.94</td>
</tr>
<tr>
<td>Scores on PROM:g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hedonistic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approval</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Needs-oriented</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stereotypic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internalized</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite score</td>
<td>1.95</td>
<td>.10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(a\) T8, data are from C1 and C2 combined when possible.

\(b\) Scores could range from 1 to 2.

\(c\) Scores were in dollars and available only for C2.

\(d\) Scores could range from 1 to 5.

\(e\) Modified Rushkin et al. (1981) scale at T8; standardized WAI subscales at T9.

\(f\) Scores could range from 1 to 5 at T8 and 1 to 7 at T9.

\(g\) The scores presented are the proportional scores.
T9, r(30) = .38, p < .03. Note that two-tailed correlations are reported here and throughout, even though specific hypotheses frequently were formulated.

Consistency of Social Desirability, Prosocial Behavior, and Empathy-Related Responding across Time

Similar scales of social desirability, prosocial behavior, and empathy-related responding were administered at more than one time period. In general, there was consistency across time for these measures. The Marlowe-Crowne Social Desirability Scale (administered at T7, T8, and T9) was highly correlated from T7 to T8 and T9, rs(30 and 29) = .86 and .76, ps < .001, and also between T8 and T9, r(29) = .69, p < .001. Children’s reported prosocial behavior (on the modified Rushton et al., 1981, scale) at T8 was positively correlated with their reports on the same measure at T6 and T7, both rs(30) = .43 and .50, ps < .013 and .004. Mothers’ reports of children’s prosocial behavior at T8 also were significantly correlated with their reports at T6 and T7, r(27) = .54 and .65, ps < .003 and .001. Helping by returning questionnaires at T8 was positively correlated with similar helping at T6 and T7, rs(30) = .35 and .45, ps < .049 and .011.

In regard to empathy-related constructs, reports of sympathy at T8 were positively related to scores on Bryant’s (1982) empathy scale at T5 and T6, both rs(29) = .55, p < .001, as well as scores on the same sympathy scale at T7, r(29) = .53, p < .002. Sympathy at T9 also was positively correlated with empathy at T5 and T6, rs(30) = .45 and .49, ps < .011 and .005, as well as sympathy at T7, r(30) = .56, p < .001, and T8, r(29) = .77, p < .001. Similarly, perspective taking (PT) scores at T8 and T9 were highly related, r(29) = .74, p < .001, and were positively correlated with PT at T7, rs(29 and 30) = .62 and .62, ps < .001, respectively. Finally, personal distress (PD) was positively correlated from T8 to T9, r(29) = .58, p < .001, and PD at T7 was positively correlated with PD at both T8 and T9, rs(29 and 30) = .49 and .45, ps < .005 and .01.

Relations of Friends’ Reports at T9 to Subjects’ Reports

Recall that for 25 subjects, friends reported (often with shortened scales) on friends’ perspective taking, sympathy, consideration for others, and aggression. Correlations between subjects’ and friends’ reports on these measures were as follows (degrees of freedom generally were 22 or 23): perspective taking, .05, N.S.; sympathy, .59, p < .002; consideration for others, .43, p < .032; suppression of aggression, .39, p < .057. These correlations provided some evidence of the validity of subjects’ self-reports of sympathy, consideration for others, and suppression of aggression. Perspective taking is an internal process that is not easily observed by others; thus, scores on this variable were retained in the analyses even though friends’ and subjects’ reports were not related. Because of the small sample size for friend data (the degrees of freedom would be only 20 for some analyses), the friend data were not used in the primary analyses.

Interrelations of Measures of Morally Relevant Behavior

The measures of prosocial behavior were helping at T8 (C1 and C2), mothers’ reports of offsprings’ prosocial behavior at T8 (C1), adolescents’ reports of prosocial behavior at T8 (C1 and C2), donating at T8 (C2), and adolescents’ reports of consideration for others and aggression at T9. Relations among prosocial behaviors at each time period were examined with correlations. In these and all subsequent analyses, data from C2 as well as C1 were used at T8 when possible.

At T8, there were no significant relations among the prosocial behavior measures, although r(29) = .35, p < .057 between mother and child report measures and r(31) = .31, p < .076 between C2 donating and reported prosocial behavior.

At T9, consideration for others (M = 3.82, SD = .64) and suppression of aggression (M = 3.60, SD = .90) were highly correlated, r(30) = .63, p < .001. Thus, these two measures were standardized and aggregated for further analyses (henceforth called the moral behavior aggregate). This aggregate score was moderately positively related to the same composite score computed from the data from friends, r(23) = .52, p < .026. This composite score was not significantly correlated with helping at T8, although it was positively correlated with self-reported prosocial behavior at T8, r(29) = .46, p < .008, and with mothers’ reports of prosocial behavior at T8, r(29) = .44, p < .013.

Relation of Social Desirability to Moral Judgment, Moral Behavior, and Empathy

In these and subsequent analyses involving moral judgment, results for the composite scores from all five stories are re-
ported items because composite scores based on more reliable (Rushton, Brainerd, & Pressley, 1983), and the extra story was considered to be more age-appropriate than some of the other four stories. However, the findings based on these composite scores generally were very similar to those based on data from four stories.

There were numerous relations between adolescents' scores on social desirability (SD) and other measures obtained from the adolescents. At T8, SD was positively correlated with sympathy, perspective taking, \( r_s(63) = .37 \) and \( .54, p < .002 \) and \(.001, \) and reported prosocial behavior, \( r(64) = .35, p < .004, \) and negatively correlated with personal distress, \( r(63) = -.27, p < .027. \) At T9, SD was not significantly correlated with any measure of reasoning (including PROM scores) except stereotypic reasoning during the interview, \( r(25) = .45, p < .02. \) However, SD was positively correlated at T9 with perspective taking, sympathy, and the moral behavior aggregate, \( r_s(29) = .54, .69, \) and \(.63, p < .002, .001, \) and \(.001. \)

Because of the aforementioned relations between SD and some of our measures and the socially desirable nature of both empathy-related and prosocial responding, auxiliary partial correlations controlling for SD were computed in addition to zero-order correlations in subsequent analyses. In most cases, the partial correlations were similar to the zero-order correlations. Thus, zero-order correlations are reported in all subsequent analyses, although partial correlations are also mentioned in text if they differ considerably from the zero-order correlations.

The Relation of Moral Reasoning to Prosocial Behavior

The relations of measures of prosocial/moral behavior to moral reasoning were examined with correlational analyses. Data are presented for the interview composite score, as well as the two types of interview reasoning that most frequently have been associated with moral behavior (i.e., hedonistic and needs-oriented reasoning; Eisenberg, 1986; Eisenberg et al., 1991). Correlations for other individual categories that were used with some regularity (i.e., \( Ms \) of at least 4.45) were less frequent, and were always consistent with the findings for the composite scores and/or theoretical expectations. For the PROM, correlations for the five level scores and the composite score are presented.

At T8, adolescents' reports of their prosocial behavior were unrelated to their composite moral interview scores, but were positively correlated with needs-oriented moral reasoning. Similarly, mothers' reports of greater prosocial responding at T8 were positively correlated with adolescents' needs-oriented reasoning, and negatively correlated with hedonistic reasoning (see Table 4). Helping was unrelated to any of the three measures of reasoning. C2's donating at T8 was significantly positively correlated with higher-level prosocial moral reasoning as assessed by the interview composite score and negatively correlated with hedonistic reasoning.

At T9, scores on the moral behavior aggregate were positively correlated with the interview moral reasoning composite score and needs-oriented reasoning, and negatively correlated with hedonistic reasoning (see Table 4). In addition, they were positively related to the PROM composite score and internalized PROM reasoning, and negatively correlated with PROM hedonistic reasoning (see Table 5).

Relations of Moral Reasoning to Empathy-Related Constructs

To examine the relation of empathy-related constructs to moral reasoning, scores on sympathy, personal distress, and perspective taking at both T8 and T9 were correlated with moral reasoning as assessed contemporaneously (with interviews and, at T9, with the PROM). In general, sympathy and perspective taking were related to moral reasoning, whereas personal distress was not.

Specifically, at T8 and T9, sympathy was negatively correlated with hedonistic reasoning during the interview (Table 4). However, the significant correlation for hedonistic reasoning was not even marginally significant at T9 when social desirability was partialled. At T9, sympathy was also positively correlated with the PROM composite score, as well as internalized and stereotypic PROM reasoning, and negatively correlated with hedonistic reasoning on the PROM (see Table 5).

Perspective taking at T8 was negatively correlated with hedonistic reasoning and positively related to needs-oriented reasoning (see Table 4); the correlation for hedonistic reasoning dropped considerably (to \(-.15\) when social desirability was controlled. At T9, PT was positively correlated with the composite interview measure of

Eisenberg et al. 1191
### TABLE 4

**RELATIONS OF INTERVIEW MEASURES OF PROSOCIAL MORAL REASONING TO CONTEMPORANEOUS PROSOCIAL BEHAVIOR AND EMPATHY-RELATED CHARACTERISTICS**

<table>
<thead>
<tr>
<th>Measure of Moral Reasoning</th>
<th>Hedonistic Needs-Oriented Composite Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prosocial index:</td>
<td></td>
</tr>
<tr>
<td>T8 self-report</td>
<td>-.24*</td>
</tr>
<tr>
<td>T8 maternal report</td>
<td>-.35</td>
</tr>
<tr>
<td>T8 helping</td>
<td>-.07</td>
</tr>
<tr>
<td>T8 donating</td>
<td>-.38*</td>
</tr>
<tr>
<td>T9 composite score</td>
<td>-.52*</td>
</tr>
<tr>
<td>Empathy-related characteristics:</td>
<td></td>
</tr>
<tr>
<td>T8 sympathy</td>
<td>-.33**</td>
</tr>
<tr>
<td>T8 perspective taking</td>
<td>-.28*</td>
</tr>
<tr>
<td>T8 personal distress</td>
<td>-.04</td>
</tr>
<tr>
<td>T9 sympathy</td>
<td>-.40*</td>
</tr>
<tr>
<td>T9 perspective taking</td>
<td>-.31</td>
</tr>
<tr>
<td>T9 personal distress</td>
<td>-.14</td>
</tr>
</tbody>
</table>

**Note.** —Correlations for the T8 helping, self-report prosocial measure, and empathy-related measures were for Cohorts 1 and 2 combined; donating was assessed only for C2 at T8; and maternal reports of prosocial behavior at T8 were assessed for C1. At T9, data were available for only C1.

* * p < .10.
** p < .05.
*** p < .01.

---

moral reasoning. In addition, at T9, PT was positively related to the PROM composite score and negatively correlated with approval-oriented moral reasoning on the PROM (see Table 5).

Because role-taking moral reasoning (during the interview) was conceptually related to perspective taking, its relation to perspective taking also was examined. Such reasoning was significantly correlated with perspective taking at both T8 and T9, rs(62 and 26) = .27 and .37, ps < .032 and .05. Personal distress was not significantly related to any measures of moral reasoning.

**Relation of Prosocial Behavior to Empathy-Related Constructs**

At T8, adolescents’ reported prosocial behavior was positively correlated with both sympathy and perspective taking, but not personal distress (see Table 6). Similarly, mothers’ reports of adolescents’ prosocial behavior were positively correlated with children’s perspective taking. In contrast, neither donating nor helping at T8 were significantly correlated with empathy-related characteristics. At T9, the moral behavior aggregate was highly positively correlated with sympathy and perspective taking, but not personal distress.

**Discussion**

The age-related trends identified in this study are similar to those presented in previous reports on this longitudinal study because only the last two follow-ups in the trend analyses are new to the study. None-
<table>
<thead>
<tr>
<th>Prosocial/Empathy Measures</th>
<th>Composite Index</th>
<th>Hedonistic Reasoning</th>
<th>Needs-Oriented Reasoning</th>
<th>Approval Reasoning</th>
<th>Stereotypic Reasoning</th>
<th>Internalized Reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prosocial aggregate</td>
<td>.51**</td>
<td>-.47**</td>
<td>-.05</td>
<td>-.17</td>
<td>.32†</td>
<td>.35*</td>
</tr>
<tr>
<td>T9 sympathy</td>
<td>.49**</td>
<td>-.49**</td>
<td>-.11</td>
<td>-.24</td>
<td>.42**</td>
<td>.37*</td>
</tr>
<tr>
<td>T9 perspective taking</td>
<td>.44*</td>
<td>-.30*</td>
<td>.03</td>
<td>-.38*</td>
<td>.21</td>
<td>.31†</td>
</tr>
<tr>
<td>T9 personal distress</td>
<td>-.03</td>
<td>-.02</td>
<td>.06</td>
<td>-.16</td>
<td>.16</td>
<td>-.08</td>
</tr>
</tbody>
</table>

* $p < .05$.  
** $p < .01$.  
† $p < .10$.  

---

**TABLE 5**

**RELATIONS OF MORAL REASONING ON THE PROM TO PROSOCIAL BEHAVIOR AND EMPATHY-RELATED CHARACTERISTICS AT T9**


<table>
<thead>
<tr>
<th>Prosocial Index</th>
<th>Empathy-Related Characteristic</th>
<th></th>
<th>Perspective Taking</th>
<th>Personal Distress</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T8 self-report</td>
<td>.49***</td>
<td>.46***</td>
<td></td>
<td>.05</td>
</tr>
<tr>
<td>T8 maternal report</td>
<td></td>
<td>.27</td>
<td>.45**</td>
<td>.23</td>
</tr>
<tr>
<td>T8 helping</td>
<td></td>
<td>.11</td>
<td>-.15</td>
<td>.03</td>
</tr>
<tr>
<td>T8 donating</td>
<td></td>
<td>.04</td>
<td>.11</td>
<td>.17</td>
</tr>
<tr>
<td>T9 prosocial aggregate</td>
<td>.80***</td>
<td>.65***</td>
<td>-.16</td>
<td></td>
</tr>
</tbody>
</table>

Note.—Correlations for the T8 helping, self-report prosocial measure, and empathy-related measures were for Cohorts 1 and 2 combined; donating was assessed only for C2 at T8; and maternal reports of prosocial behavior at T8 were assessed for C1. At T9, data were available for only C1.

**p < .01.
***p < .001.

 Nevertheless, the results of the last two follow-ups provide interesting information on a variety of aspects of prosocial functioning in late adolescence and early adulthood. As predicted, some self-reflective and internalized modes of moral reasoning (role taking, positive affect/consequences, positive affect/values, negative affect/consequences, internalized norm, rule, and law reasoning) increased in use, whereas stereotypic reasoning continued to decrease in use into adulthood. The linear increases in positive affect/values and negative affect/consequences had not been found by mid-adolescence. However, hedonistic reasoning, which had decreased into early mid-adolescence and then started to increase slightly in mid-adolescence, continued to increase in use at age 19–20 (although there was a small drop in its use from age 15 to 16 to age 17 to 18). Moreover, direct reciprocity and approval-oriented reasoning, which had begun to decline in mid-adolescence at the T7 follow-up (when 32 rather than 28 subjects were included), showed little evidence of declining in early adulthood (and even increased somewhat). Although there was a linear increase in overall reasoning throughout adolescence, moral reasoning at age 19 to 20 was not predicted from moral reasoning at earlier points in adolescence, apparently due to substantial declines in reasoning (due to increases in direct reciprocity and hedonistic reasoning) for some people and substantial increases in reasoning due to the use of higher level categories of reasoning for some others. In contrast, there was some continuity in reasoning across age 13–14 to 17–18.

As just noted, there were increases in some modes of high-level, internalized moral reasoning that were not found in prior follow-ups. It is likely that such advances are due not only to increases in formal operational reasoning in late adolescence and early adulthood, but also to increased opportunities for role-taking experiences (Mason & Gibbs, 1993). Further, college experience has been associated with increases in moral reasoning, perhaps due to the general level of intellectual stimulation in college courses and the extracurricular college milieu (Rest & Narvaez, 1991). However, entry into college did not seem to enhance moral reasoning in regard to some modes of internalized reasoning in this study; for example, the increase in positive affect/values and negative affect/consequences reasoning seemed to occur between ages 15 to 16 years and 17 to 18 years. Most of our subjects were only second semester freshmen, however, so it may have been too early to find any effects of the college experience.

The aforementioned rise in subjects’ use of hedonistic and direct reciprocity moral reasoning during adolescence/early adulthood is reminiscent of Kohlberg and Kramer’s (1969) finding of a regression in moral judgment in early adulthood, and similar to Gilligan’s (1977) data in regard to an increase in focus on responsibility to self rather than merely to others at a late stage of development. However, it is important to note that hedonistic reasoning was relatively infrequent in early adulthood, and was used primarily in response to dilemmas in which the costs of helping were high. Further, sub-
jects’ moral composite scores continued to rise into early adulthood due to increased use of numerous higher-level modes of reasoning. Most of the 19- and 20-year-olds did not verbalize principles such as responsibility to the self and did not use reasoning that could be viewed as relativistic (Murphy & Gilligan, 1980); rather, they simply seemed to have weighed long-term and substantial costs to self versus other (e.g., money for college, class time, possible injury or poor health) more carefully and realistically than when in high school and, perhaps, were relatively self-focused in their perspective. This explanation is consistent with prior work indicating that subjects’ reasoning sometimes plummets in high-cost situations and is not always consistent in terms of level (Eisenberg-Berg, 1979; Rest, 1979). Perhaps young adults are cognitively able and inclined to evaluate the long-term costs for helping in some of our dilemmas (e.g., in terms of getting behind in school or losing money needed for college), whereas younger adolescents, who are less likely to use formal operational reasoning, are less likely to do so. In addition, late adulthood and early adulthood may be a time of life in which concerns about achieving success and independence are heightened, so potential costs in these domains are viewed as substantial. In brief, we suggest that the rise in egoistic modes of reasoning was not due to structural or competence-based factors. However, we have no definitive evidence regarding the reason for this increase. Indeed, whether the increase in egoistic modes of reasoning is interpreted as relativism or merely greater use of lower-level reasoning likely pivots on whether one views moral development as involving invariant, irreversible stages in which higher-level reasoning restructures lower-level reasoning (see Turie!, 1974) or as involving the emergence of new levels of reasoning that do not eliminate or restructure lower-level reasoning (Eisenberg, 1986; Rest, 1979).

In general, the pattern of findings regarding the PROM suggests that the PROM has adequate psychometric properties to use with young adults. Young adults’ scores on their moral reasoning interviews were moderately correlated with their scores on the PROM, even though the two measures usually were administered at least a month apart in time. Further, the alphas for the PROM subscales were adequate. The pattern of correlations between the PROM scores and perspective taking, sympathy, and reported moral behavior also was similar to the pattern between interview scores and these constructs. These findings provide further evidence of the validity of the PROM for use with young adults.

At T9, some interview as well as PROM measures of prosocial moral judgment were significantly correlated with contemporaneous self-reported prosocial behavior and, to some degree, with individual differences in sympathy and perspective taking in early adulthood. Similarly, at T8, moral interview scores were significantly correlated with the more costly mode of prosocial behavior (donating), with subjects’ and mothers’ reports of prosocial behavior, and, to some degree, with sympathy and perspective taking. This pattern of findings is consistent with theoretical assertions and findings in prior follow-ups linking moral judgment with individual differences in perspective taking, sympathy, and moral behavior. However, these links have seldom been demonstrated in adolescence or adulthood.

Of interest is the fact that moral reasoning was more often significantly related to subjects’ and mothers’ reports of prosocial behavior and donating behavior than to helping behavior. The helping task—filling out a few forms and mailing them in a stamped envelope—was relatively low cost and not anonymous. In contrast, donating, which was relatively high cost and anonymous, was significantly correlated with higher-level reasoning and negatively related to hedonistic reasoning. Taken together, these findings are consistent with previous data demonstrating a relation between moral reasoning and costly behaviors such as donating but not low-cost helping behaviors (e.g., Eisenberg et al., 1987; Eisenberg & Shell, 1986). In addition, in the present study, subjects’ and mothers’ reports of prosocial behavior included a variety of prosocial actions; thus, the links between these measures of prosocial behavior and moral reasoning may have occurred because aggregate measures of a construct generally are more reliable than a single index.

The self- and mother-report measures of prosocial behavior, but not donating or helping, were positively related to subjects’ sympathy (for child report only) and perspective taking. This pattern is similar to that found at T7 in mid-adolescence (except mothers’ reports of prosocial behavior were unrelated at T7; Eisenberg et al., 1991). Perhaps this pattern of findings was obtained because the
recipient of donating was not vividly depicted, and the benefits of helping for the experimenter (the recipient) may not have seemed high. Thus, sympathy and perspective taking may not have been relevant to helping or donating (particularly the former).

Measures of prosocial behavior, empathy-related constructs, and social desirability were relatively stable over periods of 4 years and longer. Thus, individual differences in these behaviors and reactions appear to have an enduring quality in adolescence and early adulthood. Moreover, judging from the findings in previous follow-ups, in general the patterns of relations among moral judgment, prosocial behavior, and empathy-related characteristics were fairly similar across adolescence and early adulthood (although, for example, different measures of prosocial behavior may relate to moral reasoning or empathy-related responding at different ages). Further, interrelations, when significant, generally have been consistent with theoretical expectations. Consequently, it appears that at least some of the processes involved in prosocial development are relatively stable across adolescence and into early adulthood. However, research on people other than middle-class Caucasians is needed because it is unclear whether our findings from a small, middle-class, primarily Caucasian sample can be generalized to other groups.

References


development. Cambridge: Cambridge University Press.


You have printed the following article:

**Prosocial Development in Late Adolescence: A Longitudinal Study**
Nancy Eisenberg; Gustavo Carlo; Bridget Murphy; Patricia van Court
Stable URL: [http://links.jstor.org/sici?sici=0009-3920%28199508%2966%3A4%3C1179%3APDILAA%3E2.0.CO%3B2-C](http://links.jstor.org/sici?sici=0009-3920%28199508%2966%3A4%3C1179%3APDILAA%3E2.0.CO%3B2-C)

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

**References**

**An Index of Empathy for Children and Adolescents**
Brenda K. Bryant
Stable URL: [http://links.jstor.org/sici?sici=0009-3920%28198204%2953%3A2%3C413%3AAIOEFC%3E2.0.CO%3B2-9](http://links.jstor.org/sici?sici=0009-3920%28198204%2953%3A2%3C413%3AAIOEFC%3E2.0.CO%3B2-9)

**A Longitudinal Study of Moral Judgment**
Anne Colby; Lawrence Kohlberg; John Gibbs; Marcus Lieberman; Kurt Fischer; Herbert D. Saltzstein
Stable URL: [http://links.jstor.org/sici?sici=0037-976X%281983%2948%3A1%3C1%3AALSOMJ%3E2.0.CO%3B2-9](http://links.jstor.org/sici?sici=0037-976X%281983%2948%3A1%3C1%3AALSOMJ%3E2.0.CO%3B2-9)

**Self-Restraint as a Mediator of Family Influences on Boys' Delinquent Behavior: A Longitudinal Study**
S. Shirley Feldman; Daniel A. Weinberger
Stable URL: [http://links.jstor.org/sici?sici=0009-3920%28199402%2965%3A1%3C195%3ASAAMOF%3E2.0.CO%3B2-X](http://links.jstor.org/sici?sici=0009-3920%28199402%2965%3A1%3C195%3ASAAMOF%3E2.0.CO%3B2-X)