

Developmental Trajectories in Moral Reasoning Across the Life Span

CHERYL ARMON

Antioch University Southern California at Los Angeles, USA

THEO L. DAWSON

University of California at Berkeley, USA

ABSTRACT *This long-term study found that moral reasoning as conceptualised by Kohlberg (1981, 1985) can develop into adulthood. Predominantly white, well-educated, middle-class participants were interviewed four times at 4-year intervals (N = 44). Stage development was sequential and continued throughout the life span, although its occurrence decreased with advancing age in a curvilinear fashion. Post-conventional reasoning was demonstrated by seven adults. Stage of moral reasoning correlated with age strongly in children and moderately in adults, and was moderately correlated with education in all age groups. Additionally, advance in moral reasoning stage was correlated with increase in education in adults. Although no systematic gender differences were found across age groups, men in the younger adult group had significantly higher scores than women.*

This paper reports results from a longitudinal and cross-sectional study of the development of moral reasoning across the life span [1]. Investigated were the sequence and trajectory of development, particularly in adulthood, and the relationships between education, gender, and moral reasoning stage change. Of particular interest was the investigation of Piagetian-type, structural moral reasoning development during adulthood and the impact of education on such development during that period.

The conception of moral reasoning used here relies primarily on Kohlberg's structural-developmental model of moral reasoning and its development, the stage descriptions, criteria and many assumptions of which have been thoroughly articulated elsewhere (Kohlberg, 1969, 1981, 1984; Colby *et al.*, 1983; Kohlberg, Levine & Hower, 1983; Colby & Kohlberg, 1987a, 1987b).

For the most part, Piaget posited that stage development of reasoning ended during adolescence (Inhelder & Piaget, 1958; but see Vuyk, 1981, for exceptions). Kohlberg (1990), however, eventually acknowledged that the development of post-conventional moral reasoning (Stage five), which includes the reciprocal relationship

between systems of rights and systems of duties, was probably a phenomenon of adulthood. The assertion that structural stages of moral development extend into adulthood remains controversial and there have been no long-term studies to date that support or refute it. This study was designed, in part, to examine the occurrence of adult moral reasoning development by studying an "élite" sample over a relatively long period of time.

Recently, there has been more interest in the empirical study of adult moral judgement in the developmental literature (Colby & Damon, 1992, 1995; Walker *et al.*, 1995). At the same time, however, the contemporary fascination with post-modern contextualism is not supportive of the structural approach; this perspective places the validity of the paradigm in question, often pointing to not only the limits of generalisability, but also the inability of abstract reasoning about hypothetical dilemmas to tell us anything meaningful about human moral experience. While sometimes extremist, this perspective is useful; it can expose weaknesses of bias and preference and provide a fresh facet through which to view the complex moral world.

One example of the post-modern influence in moral development is the "narrative" approach, which focuses on text analysis of subjects' moral "stories" or told experience instead of their abstract moral reasoning in response to hypothetical dilemmas. The unique, individual features of moral experience kept intact by a contextual approach are more highly valued than the type of reasoning sought by Kohlberg's approach (see e.g., Brown *et al.*, 1992; Day & Tappan, 1996). A kind of dualism has emerged in the professional literature between what is characterised as this more emotional, relational, and relativistic experience of morality, on one hand, and the rational, autonomous, and normative reasoning of morality on the other hand.

Over time, this chasm may be narrowed. Eventually, the two approaches may be used side by side, each making meaningful contributions that the other would have ignored (Pratt & Arnold, 1995). Alternatively, the division itself may be unfounded (Lourenco, 1996; Puka, 1996). Moral development and experience has a complex array of dimensions, some of which are emotional, others rational, but all are "relational" by definition. In any event, the current study, conducted from 1978 to 1991, would have benefited from the addition of some of these methods of data collection [2]. Its weaknesses in this regard, however, must be balanced against the returns of long-term longitudinal data. Thus, the present study continues a particular research tradition—the structural–developmental approach.

One of the fundamental postulates of this approach is the invariant sequence criterion (Kohlberg, 1969; 1981) which stipulates that stage development, whenever it occurs, should occur in order. Accordingly, no stages should be skipped, nor should development revert to earlier stages except possibly in cases of brain injury or illness. Although individuals' growth may cease at any stage in the sequence, development that occurs should do so toward the next ordered stage. It is acknowledged that this is not the only model of development. For example, it can be contrasted to much of gerontological adult development research, which reports forms of regression (e.g. Charness, 1985; Salthouse, 1993, 1994), as well as lifespan

theories that indicate multidirectional change in some domains (e.g. Baltes, 1987). Moreover, there are contemporary developmentalists (e.g. Case, 1985) who reject many of the assumptions of Piagetian-type stage models. Few of the alternative models, however, have been subjected to long-term longitudinal studies.

In the meantime, longitudinal research within a single domain with the same task sequence should be a reliable test of the invariant sequence criterion of structural stages (Kohlberg & Armon, 1984; Cerella *et al.*, 1993). Since Colby and Kohlberg (1987a, 1987b) completed the current, more reliable and valid system for scoring moral judgement interviews (MJIs), the results of several other longitudinal studies have supported this criterion in both short-term and longer-term designs (Page, 1981; Walker, 1982; Nisan & Kohlberg, 1982; Colby, *et al.*, 1983; Snarey, Reimer & Kohlberg, 1985).

The research reported in the following review is restricted to studies of moral reasoning development produced with this current scoring system and its associated dilemmas for two main reasons. First, the new system is substantially improved. Interrater, test-retest, and internal consistency data for the Standard Issue System indicate that the measure is well above acceptable limits of reliability. Although some problems remain concerning discriminant and construct validity, the Standard Issue System remains the most reliable and valid scoring system available for the assessment of moral reasoning. (see Colby & Kohlberg, 1987a, for a detailed discussion of these issues). Secondly, it is not advisable to compare findings generated from different scoring systems, which use different methods and maintain different assumptions. For this reason, the results of research using Rest's (1975) Defining Issues Test (DIT), for example, or Gibbs' Sociomoral Reflection Measure (SRM; Gibbs, Basinger & Fuller, 1992), are not reported here.

Development of Moral Reasoning Across the Lifespan

Most of the research reported on the development of moral reasoning is cross-sectional. One-time examinations of moral reasoning development within different age groups can identify associations between age and moral reasoning stage, but they are of limited usefulness on questions of *development*, that is, movement from stage to stage during the lifespan. The few longitudinal studies published to date have had two shortcomings when considered in light of this question: (1) they examine development over a relatively short period of time (1-4 years, e.g., Walker, 1989; Pratt *et al.*, 1994); and/or (2) they examine development within a narrow segment of the life span (e.g., Colby *et al.*, 1983; Pratt *et al.*, 1994). In most studies, adolescents and young adults are particularly well-represented (e.g., Page, 1981; Colby *et al.*, 1983), whereas younger children and mature adults are largely neglected (for exceptions, see Walker, 1989 and DeVries & Zan, 1994 for studies in childhood; Walker, 1989 and Pratt *et al.*, 1994 for studies in adulthood). Findings from studies limited in duration or age-range have difficulty addressing whether positive moral reasoning development continues through adulthood. This may help to explain their somewhat contradictory findings regarding adult moral development.

Stages of Moral Reasoning and Age

Some studies report increases in stage development with age, while others do not. Pratt *et al.* (1994) found no change in moral reasoning stage scores over a 4-year period in 64 adults between the ages of 30 and 80 years. Similarly, with a cross-sectional sample, Commons *et al.* (1989) found no relationship between age and stage of moral reasoning in a group of 81 Mensa Club members (who are distinguished by their high IQ scores), aged 18–83 years. In another cross-sectional study, Pratt *et al.* (1983) reported no statistically significant differences between the moral reasoning stage scores of 30–50- and 60–75-year-olds.

In contrast, a follow-up study performed by Walker (1989) examined sequential development over a 2-year period in a group of 6–15-year-olds and their parents and found significant positive change in moral reasoning stage in both age groups—one-third to one-half-stage for children, and less than one-quarter stage for adults.

Similarly, in a cross-sectional analysis of data from an early phase of the present study, Armon (1984b) reported a curvilinear relationship between age and moral reasoning stage with a steep rise in scores during the course of childhood, a more moderate rise in adulthood, and relative stability in old age. In another cross-sectional study, Pratt and his colleagues (Pratt *et al.*, 1983) also identified a curvilinear relationship between age and moral reasoning stage with rising scores through middle-age, although this relationship was not modelled statistically. Bielby and Papalia (1975) and Bakken and Ellsworth (1990) report similar cross-sectional findings. In most of these studies, average moral reasoning stage scores continue to rise until middle adulthood and then level off or decline in late adulthood. Declines in late adulthood, however, are also not a consistent finding. Recall, for example, the Pratt *et al.* (1983) finding of no difference in average stage scores between younger and older adults.

In sum, findings to date suggest three possible developmental paths of moral reasoning development in adulthood: (1) a plateau effect in adulthood; (2) ongoing development throughout adulthood; or (3) on-going development in earlier adulthood with decline in later adulthood.

Post-Conventional Moral Reasoning in Adulthood

Researchers have had difficulty investigating the existence, nature, and onset of post-conventional moral reasoning. These problems have naturally led to discussions as to whether post-conventional moral Stages Five and Six [3] exist or are “true” or “hard” stages (discussed in Kohlberg, *et al.*, 1983; Kohlberg & Armon, 1984; see also Alexander & Langer, 1990). These concerns can be explained in part by the dearth of data scored at these stages. In the last 10 years, however, studies that include both adolescents and adults indicate that when post-conventional reasoning appears, it does so in the late 20s (Colby *et al.*, 1983), in middle adulthood rather than in adolescence or early adulthood (Armon, 1984a) or after age 37 (Bakken & Ellsworth, 1990). Similarly, in research with adolescents (Higgins, 1980; Kohlberg & Higgins, 1984) and young adults (Armon, 1984a) no incidence

of post-conventional reasoning is found. Indeed, only a few incidences of Stage Four in adolescence are reported.

Adult Development of Moral Reasoning and Higher Education

Kohlberg (1969, 1984) believed that an important prerequisite of moral development was socio-moral experience, that is, direct and repeated experience with moral conflict. A regularly investigated source of this in adulthood is higher education. For example, several researchers have reported a moderate to strong positive relationship between years spent in college or graduate school and stage of moral reasoning in cross-sectional analyses of adults (e.g., Candee, Graham & Kohlberg, 1978; Colby *et al.*, 1983; Pratt *et al.*, 1983, 1984, 1991, 1994; Armon, 1984a; Walker, 1984, 1986; Bakken & Ellsworth, 1990; Markoulis, 1990), even in late adulthood (Pratt *et al.*, 1991, 1994). Moreover, Stage Four and (post-conventional) Stage Five rarely appear in individuals without post-secondary education. For example, Walker (1986), Markoulis (1990), and Armon (1984a) found Stage Four reasoning only among adults who had obtained some post-secondary education, and with Armon's (1984a) and Kohlberg's (Colby *et al.*, 1983; Kohlberg & Higgins, 1984) longitudinal samples, Stage Five was found only in individuals with at least some graduate study. Nucci and Pascarella (1987) report similar findings in their review of research on the relationship between attending college and the development of moral reasoning.

All the above-mentioned studies demonstrate, however, that even among individuals who have completed advanced education, post-conventional reasoning remains elusive. Moreover, some studies have found no higher education effects on moral reasoning at all. For example, Commons *et al.* (1989) found no correlation between moral reasoning stage and educational level with Mensa Club members (although this may be due to unique features of that group). In an attempt to explain such inconsistent findings regarding the effect of education on development, Nucci and Pascarella (1987) cite a range of studies that suggest that individual students may have quite different experiences in college, and that certain types of social and educational experience may be more salient than others. Similarly, in a review of studies that examine the relationship between scores on the Defining Issues Test (DIT) of moral reasoning and years of education, Rest and Narvaez (1991) discuss research that goes beyond correlations toward a deconstruction of "education". These authors persuasively argue that the same *education* is never the same experience for any two individuals, and that the relationship between the development of moral reasoning and education can be meaningfully examined only when the nature and quality of the student-education relationship is taken into account. Some ideas about the nature and quality of educational experiences and their potential influence on development will be addressed further at the conclusion of this paper.

The Development of Moral Reasoning and Gender

One well-known criticism of Kohlberg's model, first made by Gilligan (1982), concerns gender bias. Gilligan made two claims: (1) that there are two primary

orientations to moral reasoning, one involving justice and the other involving care; and (2) that women are more likely to be (incorrectly) scored at Stage Three because concern with caring is stressed in Kohlberg's descriptions of Stage Three reasoning, while it is lacking in descriptions of earlier and later stages (for a thorough discussion of the justice/care distinction see Puka, 1989). Although this paper does not take up this debate, it does address the question of differences in the moral reasoning stage scores of men and women.

Neither Pratt *et al.* (1984) nor Walker's (1984) surveys of the literature on gender differences in moral reasoning expose any systematic differences between males and females on the Moral Judgement Interview (MJI, Colby & Kohlberg, 1987a, 1987b) after differences in education are taken into account. However, the question of gender bias has not been laid to rest. Some recent studies of adult moral reasoning development continue to report gender differences. For example, Bakken and Ellsworth (1990) reported that average moral reasoning stage scores were about one third of a stage higher for males than for females in their cross-sectional study of 94 middle-class adults, ages 28–55 years. Unfortunately, participants' education was not controlled for in that analysis. Another cross-sectional study, this time of 40 undergraduate students, found average moral reasoning stage scores about one-quarter of a stage higher for males than for females (Krebs *et al.*, 1994). Nevertheless, the bulk of studies reported average male and female scores to be virtually identical (e.g. Pratt *et al.*, 1988a, 1988b, 1991; Commons *et al.*, 1989; Walker, 1989; Walker *et al.*, 1995).

The present study examines moral development across the lifespans of men and women. Its long-term (13-year) longitudinal design, combined with the wide age-range of participants (5–72 years at the first interview) offers an opportunity to examine the development of moral reasoning from both longitudinal and cross-sectional perspectives. The primary research questions were: (1) does Piagetian-type moral reasoning development take place throughout the lifespan, including later adulthood and, if so, does its occurrence decrease with advancing age? (2) Does post-conventional moral reasoning (Stage Five) occur only in adulthood? (3) Can educational experience be related meaningfully to the development of moral reasoning throughout the lifespan? (4) Does moral reasoning development proceed through the stage sequence without stage-skipping or reversals in all age groups? (5) Are there systematic gender differences in moral reasoning development, and is gender related to the development of post-conventional reasoning?

Method

Participants

At the first interview time, Time 1 (1977), 50 individuals, ranging in age from 5–72 years, responded to a flyer seeking volunteers. The flyer sought "people interested in development through the life span, or their children", and was distributed at a state university in the Los Angeles area. The first 50 individuals to respond were accepted. Of the 37 who participated in the follow-up in 1981 (Time 2), four were

Latino, one was Asian, two were mixed race and the rest were Caucasian. Average years of education for adult males at Time 1 was 17 years; for females it was 15.5 years. This difference was not found to be significant. Four adults had completed doctorates, nine had completed some post-undergraduate work, six had completed college, and nine had completed some college. Individuals' annual incomes ranged between \$18,000 and \$140,000, with an average of \$30,000. Average years of education and annual income rose throughout the study. Income levels were not found to be significantly related to any other variables.

For some of the analyses, all participants were divided into three age groups, consisting of a *Child and Adolescent Group* of twelve 5–14-year-olds (6 females and 6 males), a *Younger Adult Group* of twenty-one 23–45-year-olds (11 females and 10 males), and an *Older Adult Group* of eight 50–72-year-olds (6 females and 2 males). Categorisation into these groups was based on (1) the view that different developmental trends may occur in childhood/adolescence, adulthood and late adulthood; (2) clear differences in developmental trends among these age groups in this sample; and (3) the younger and older adult groups had more variation in their levels of education.

Procedure

The Standard Form Moral Judgment Interview (MJI, Colby & Kohlberg, 1987b) was administered individually to each participant, tape-recorded and transcribed. Participants were interviewed four times at approximately 4-year intervals from 1977 to 1990. Form A (three dilemmas: Heinz, Judge, Joe) was used on all participants except three who were familiar with it. For these participants, Form B (three dilemmas: Dr Jefferson, Judge, Judy) or C (three dilemmas: Korean, Valjean, Karl) were used. The same form was used on each participant at each time.

Of the participants who were interviewed on two or more occasions, 40 were interviewed at Time 1, 37 were interviewed at Time 2, 34 were interviewed at Time 3 (1985), and 33 were interviewed at Time 4 (1989–90). Seven participants missed two of four interviews. Of these, two entered the study for the first time in 1985, an 11-year-old who entered the study late and a 39-year-old who was not administered the MJI at Time 1 and 2 due to time constraints. Six participants dropped out after Time 2. Their ages at that time were 12, 32, 34, 37, 66 and 68. Three of them were unable to complete their interviews due to health problems. The other three participants could not be contacted at either Time 3 or Time 4. Participants who completed only two interviews were not significantly different from those who completed three or four interviews on any of the measures used in this study.

Scoring and Analysis

The MJIs were scored by the first author, who had obtained a 0.95 reliability rating with the first author of the *Standard Issue Scoring Manual* (Colby &

Kohlberg, 1987a, 1987b). To check for rater bias, half of the interviews from Time 1 and 2, and one-quarter of the interviews from Time 3 and 4 were scored by independent raters (trained graduate and undergraduate students). Different raters were used at each test time. The perfect agreement rates between the independent raters and the first author were never less than 80% and there was 100% agreement within a half stage. Interrater, test-retest, and alternate form reliability estimates for the MJI indicate an overall error in the range of 1/3 Stage (Colby & Kohlberg, 1987a).

The scoring system produces scores that can be reported either as Global Stage Scores (GSS) ranging from 1 to 5 in 1/2 stage increments, or as Weighted Average Scores (WAS), ranging from 100 to 500 in single digit increments. A WAS of 100-124 is equivalent to a Global Stage Score (GSS) of 1, 125-174 equals a GSS of 1.5, 175-224 equals a GSS of 2, 225-274 equals a GSS of 2.5, 275-324 equals a GSS of 3, 325-374 equals a GSS of 3.5, 375-424 equals a GSS of 4, 425-474 equals a GSS of 4.5, and 475-500 equals a GSS of 5.

Several analyses in this report were conducted on the pooled longitudinal and cross-sectional data. This practice has two major advantages. First, power is increased when all measurements at all test times are treated as independent observations, which is permissible when a study design involves more than two test-times separated by relatively long intervals (Willet, 1989). Secondly, the longitudinal information incorporated into the analysis in this way adds valuable information about growth trends. Specifically, by including longitudinal data for individuals, it can be shown that age differences reflect actual trends at the individual level rather than artifacts of statistical averaging. In order to eliminate concerns about the possible introduction of error with this approach, all analyses were also run separately on the data for each test-time. The trends found at each test time were consistent with the trends reported for the pooled sample.

Results

Overview of Developmental Change

Development of moral reasoning occurred for many people in this study, including adults. This can be seen in Fig. 1, which depicts moral reasoning stage scores (WAS) for each participant by age for each test time. To evaluate development in the whole sample and in the three age groups, a series of parallel straight-line regressions of WAS on age were run in which the intercepts for each person were allowed to vary but the slopes on age were all forced to be the same. The common slope was interpreted as the mean slope across individuals. It was found that in the group as a whole, scores increased by an average of 69.20 WAS points over the 13-year course of the study, $t = 10.51$, $P < 0.001$. In the *Child and Adolescent Group*, scores increased by an average of 160.68 points, $t = 11.96$, $P < 0.001$. In the *Younger Adult Group* scores increased by an average of 49.53 points, $t = 8.92$, $P < 0.001$. There was no significant change in scores in the *Older Adult Group* over the course of the study.

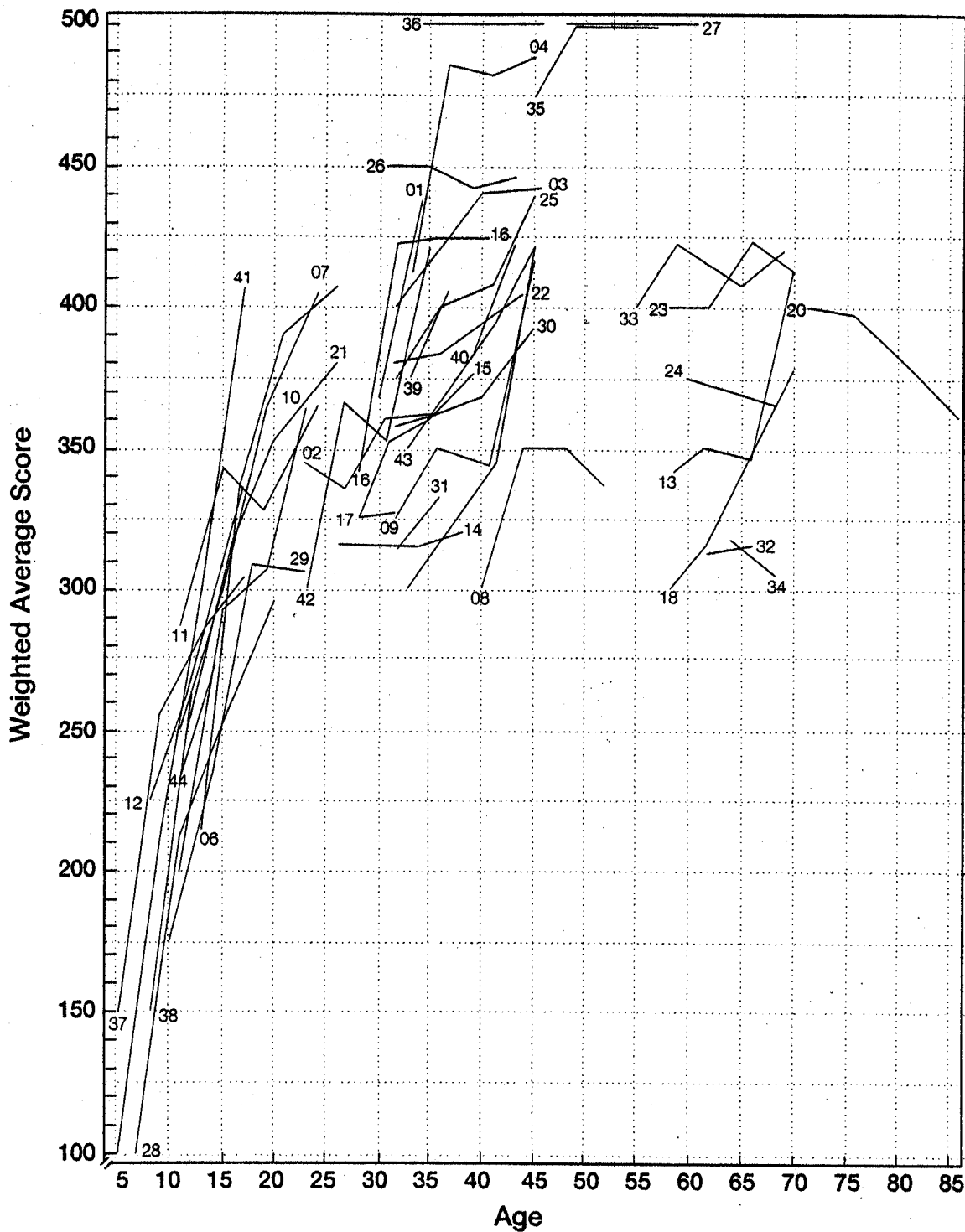


FIG. 1. The Relationship Between Weighted Average Score and Age

The Development of Moral Reasoning and Age

As shown in Fig. 2, moral stage scores increase with age in a curvilinear fashion. In a regression of the log of age [4] with weighted average score (WAS), 46% of the variance in WAS is explained by the log of age, $F(1,166) = 141.62, P < 0.001$. $WAS = 54.70 + 203.31\log(\text{age})$.

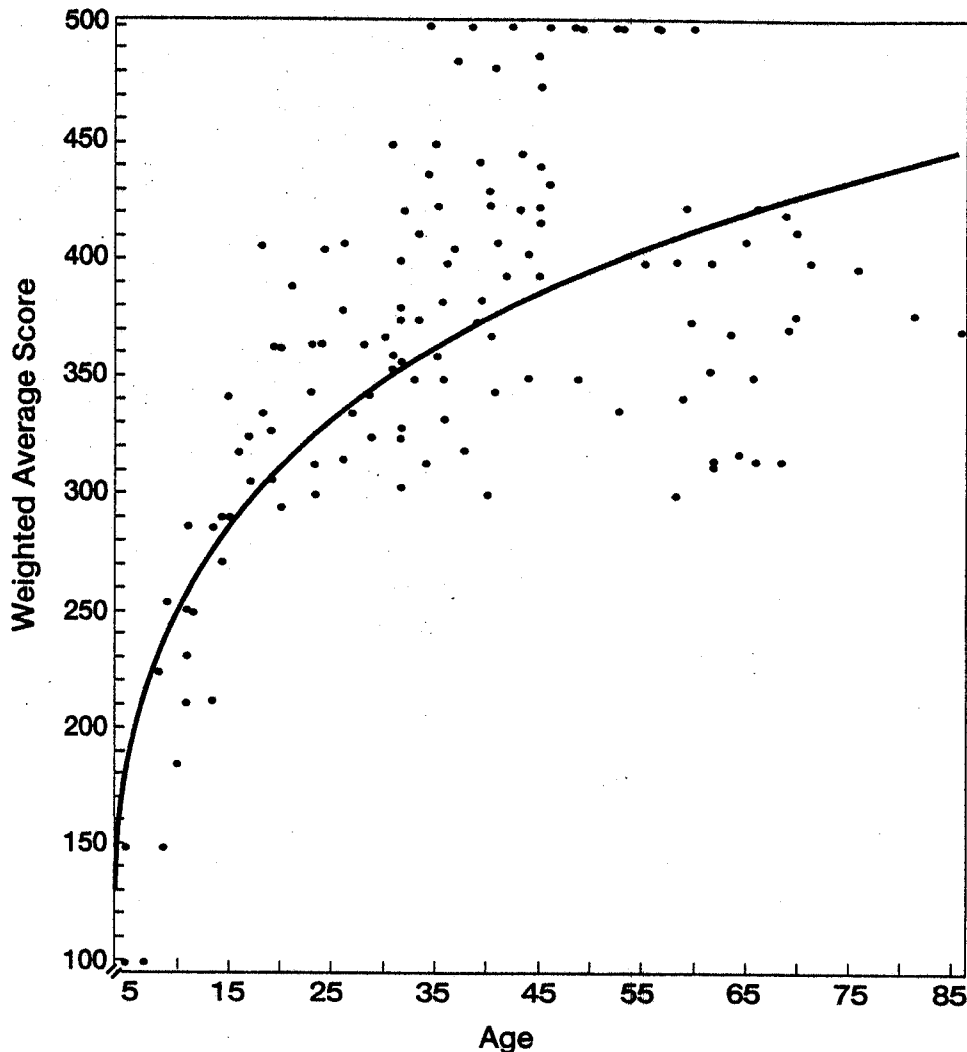


FIG. 2. Plot of Weighted Average Score With the Log of Age in 42 Individuals Over 4 Test-Times

This curvilinear trend suggests different incident rates of development during different periods of the lifespan, with diminished occurrence as adulthood advances. Further analysis of the age cohorts confirms this pattern. (Recall that scores at all test times for each participant are used.) Figure 3 depicts regressions of WAS by age for each participant. Cohort distinctions stand out clearly in this figure, as does the increasing variability of weighted average scores and individual developmental trends with advancing age. Superimposed over the individual slopes are regressions of weighted average scores by age for each age group.

Age group analyses. In the *Child and Adolescent Group* (5–13 years), mean weighted average score increases by 12.27 WAS points with every 1-year increase in age (constant = 97.94) and age explains 77% of the variance in WAS ($F(1,44) = 148.91$, $P < 0.0001$). In the *Younger Adult Group* (23–50 years) mean weighted average score

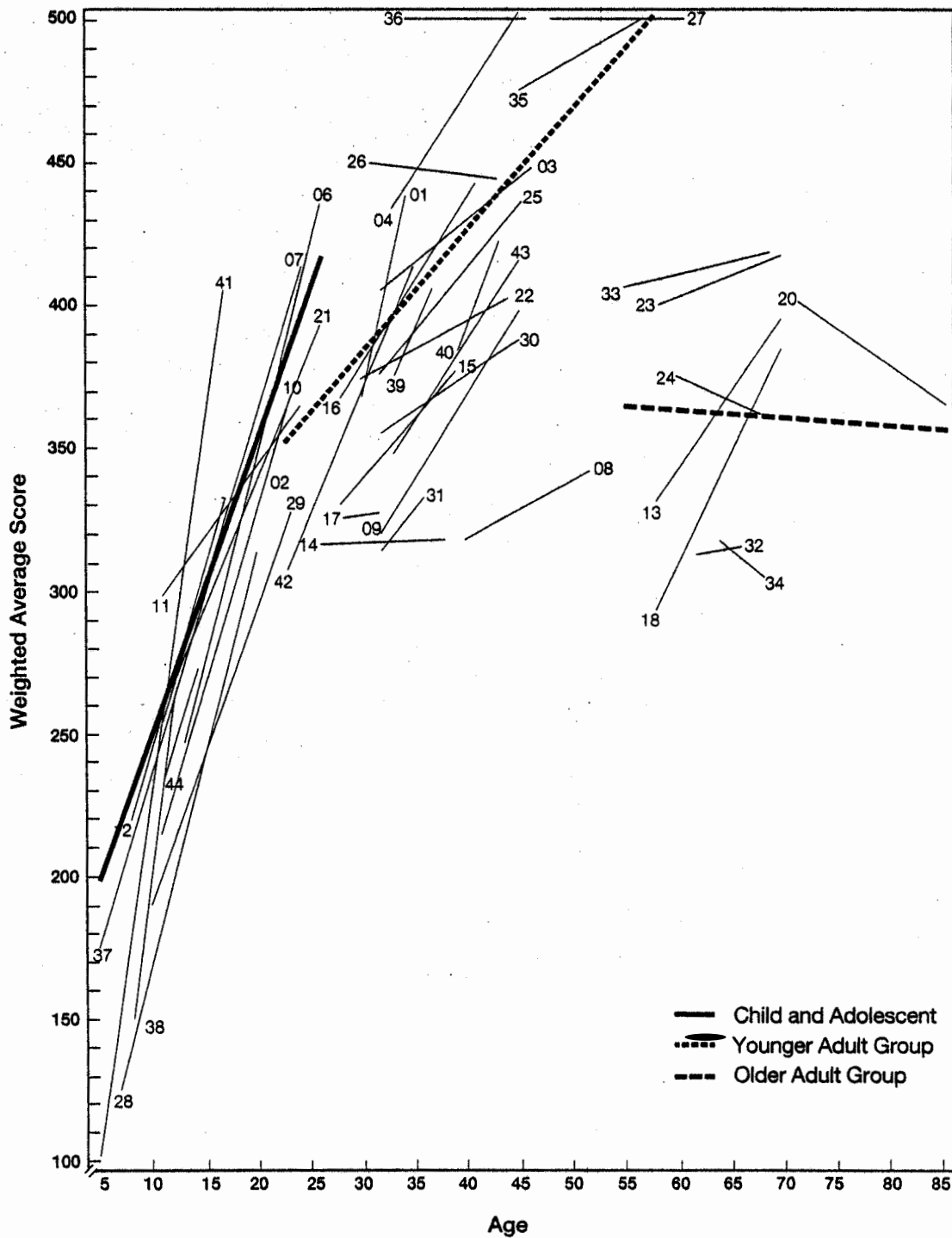


FIG. 3. Regressions of Weighted Average Scores With Age for Each Subject And for Each of Three Age Groups

increases by 4.87 points with every 1-year increase in age (constant = 213.00), and age explains 37% of the variance in WAS ($F(1,86) = 51.51, P < 0.001$). Finally, in the *Older Adult Group* (50–72 years), WAS does not increase significantly with age ($F(1,29) = 0.04, R^2 = -0.03, NS$).

TABLE I. Distribution of highest attained weighted average scores (WAS) and global stage scores (GSS) of all subjects who were older than 20 in 1989-90

WAS/ GSS	Men	Women	Total
225-274	0	0	0
275-324	0	4 (18%)	4 (11 %)
325-374	2 (14%)	6 (27%)	8 (22 %)
375-424	7 (50%)	10 (45%)	17 (47 %)
425-474	3 (21%)	1 (05%)	4 (11 %)
476-500	2 (14%)	1 (05%)	3 (09 %)

Adult Stages of Moral Reasoning

No participant reached post-conventional Stage Five before the age of 35 and with fewer than 21 years of education, and no participant achieved Stage Four before the age of 24 and with fewer than 13 years of education with one exception: a 17-year-old with 11 years of education was scored at Stage Four. As shown in Table I, Stages Four and Five were represented by both men and women. Ten women (45%) and seven men (55%) who were older than 20 at the first time of testing developed to Stage Four. In addition, two women (10%) and five men (35%) received stage scores above Stage Four.

The Development of Moral Reasoning and Education

As mentioned earlier, age groups differed significantly on the education variable so analyses were performed separately with each group.

Child and adolescent group. In addition to the strong relationship between age and WAS in this group (shown in Fig. 3), a strong relationship was found between WAS and educational attainment. Mean WAS increases by 15.94 points with every additional year of education (constant = 135.11), and education explains 76% of the variance in WAS in that group ($F(1,44) = 147.30, P < 0.000$). Not surprisingly, age and educational attainment are highly correlated, $r = 0.95, P < 0.0001$. This almost perfect correlation, however, obscures the differential effects of age and education on WAS. A multiple regression of age and educational attainment on WAS reveals non-significant t s for both age ($t = 1.97, P < 0.06$) and educational status ($t = 1.87, P < 0.07$). In keeping with their high intercorrelation, however, 78% of the variance in WAS is explained by these factors in combination, approximately the same percentage of the variance each explains on its own ($F(2,43) = 80.38, P < 0.0001$). Thus, the effect of education is indistinguishable from the effect of age in this group, rendering further analysis of education effects meaningless.

Younger adult group. In the younger adult group, age explains 37% of the variance in weighted average score (see Fig. 3). A regression of educational attainment on WAS demonstrates a somewhat milder positive relationship. Mean WAS increases by 11.39 points for each year of additional education (constant = 206.14) and educational attainment accounts for 25% of the variance in WAS ($F(1,86) = 29.36$, $P < 0.0001$). To assess the combined effects of age and education, a multiple regression was conducted on WAS. Both age and education have significant effects, $t = 5.56$, $P < 0.0001$ and $t = 3.48$, $P < 0.001$, respectively. Together, they explain 44% of the variance in WAS for this group, ($F(2,85) = 35.21$, $P < 0.0001$, $WAS = 133.94 + 3.88age + 6.90ed$).

Because several of the adults in this group participated in one or more years of formal education during the course of the study, the longitudinal effects of education were examined. To assess whether educational level increased significantly over time, a regression was conducted on education with times of testing. Education increased an average of 0.55 years during each interval ($F(1,81) = 4.51$, $R^2 = 0.04$, $P < 0.05$). The relationship between weighted average scores and this change in educational attainment was then assessed by comparing the change in WAS of those who experienced a change in education between Time 1 and 4 with the change in WAS of those who did not participate in further education. (The two participants who scored the maximum score of 500 at every test time were dropped from these change analyses.) A moderate, positive relationship between change in WAS and change in education was found. Weighted average scores of those participants who advanced their education during the course of the study increased an average of 38.06 points more than did those of participants with no additional education (constant = 28.27). Change in educational attainment accounted for 31% of the variance in change in WAS ($F(1,18) = 9.65$, $P < 0.01$). These analyses demonstrate that although an increase in education heightens the probability of moral reasoning development, it is apparently neither sufficient nor even necessary for stage change to occur.

Older adult group. The relationship between education and weighted average score is moderately positive in the older adult group. Mean WAS increases 7.06 points for each year of added education (constant = 250.08) and educational attainment accounts for 38% of the variance in WAS ($F(1,29) = 19.77$, $P < 0.001$). The lower overall educational attainment in the older cohort may partially explain the group's lower mean Weighted Average Score. (Mean educational attainment in this group is 15.88, $SD = 3.71$, $n = 32$, whereas it is 16.93, $SD = 2.65$, $n = 88$, in the younger adult cohort.)

Stage Sequence

The vast majority of changes in moral reasoning stage scores were increases in all age groups. As can be seen in Fig. 1, reversals are neither common nor large in any age

group, although decreases in stage scores occurred with some adults. These apparent reversals are difficult to interpret because, except for one, they are smaller in magnitude than measurement error (1/3 stage). The only incidence of a significant reversal (one-half stage over 12 years) occurred in the oldest participant who was 71 years old at Time 1. There was no "stage skipping" demonstrated, that is, all increases (or decreases) were to the next stage in the sequence.

The Development of Moral Reasoning and Gender

No systematic gender differences were found when all participants were studied as a group. Therefore, age group analyses were performed.

Child and adolescent group. No significant relationship was found between weighted average scores and gender in this pooled group, although the non-significant trend favoured females by 14 WAS points. In addition, gender differences in WAS change scores were non-significant and there were no significant differences between males and females on years of education completed.

Younger adult group. In this pooled group, mean WAS for men is 45.64 points higher than for women (constant = 332.57) and gender accounts for 13% of the variance in WAS ($F(1,86) = 14.68, P < 0.001$). There are no significant differences between men and women on amount of education completed. When gender is entered as the last variable in a stepwise regression on WAS with age and education, the advantage for males decreases but remains significant ($t = 4.11, P < 0.01$). Fifty-three per cent of the variance in WAS is explained by age, education and gender, and WAS is greater by 36.84 points for men than for women ($F(3,84) = 33.50, P < 0.0001, WAS = 101.37 + 3.94age + 5.52ed + 36.84male$).

No direct, significant relationship was found between *change* in Weighted Average Score and gender. Nevertheless, women's scores increased significantly over the course of the study while men's did not. Women's scores increased by an average of 24.81 points per testing interval (constant = 313.03) and the passage of time accounted for 26% of the variance in women's scores ($F(1,41) = 15.66, P < 0.001$). Men's scores increased by an average of 13.66 points per time of testing (constant = 389.70), a change that was not statistically significant.

Similarly, women's educational attainment increased significantly over the four times of testing, while men's did not. Women's education changed by an average of 0.8 years per testing interval (constant = 14.69) ($F(1,41) = 6.28, R^2 = 0.11, P < 0.05$). Men's education increased by an average of 0.27 years per testing interval (constant = 16.8, NS). The small n of this sample made it impossible to adequately examine the relationships between change in education and change in WAS by gender.

Older adult group. Because there were only two men in this group, it was not possible to assess gender effects.

Discussion

The data from this study corroborate previous findings on the sequentiality of Piagetian-type moral reasoning stage development. The absence of stage-skipping, coupled with the dramatic infrequency and small magnitude of reversals supports the structural-developmental criterion of sequentiality over virtually the entire lifespan. One significant reversal was evident in only the oldest participant, who was 71 at the onset of the study. Whether this anomaly can be associated with organic changes cannot be ascertained. In contrast, a 66-year-old participant's moral reasoning score *increased* significantly for the first time in 12 years at Time 4. She had been diagnosed with terminal cancer 2 years earlier and decided to forego further surgery and chemotherapy. At her last interview, she reported dramatic changes in her moral beliefs and behaviour. She died only days later.

Whether age beyond a certain point typically brings with it a decline in the structural complexity of moral thought warrants serious investigation. Our data suggest that such research will reveal tremendous variability in the development of moral reasoning or its maintenance after early adulthood. Genetic history, lifestyle, individual interest and activity, competence, access to resources, specific life events and general health may all be important factors.

In the small sample of older adults the relationship between age and development was not significant. This is in keeping with previous cross-sectional and short-term longitudinal research that indicates an absence of moral development (or, in some cases, reversals) in older participants. The results from this study, the first we are aware of that has investigated long-term moral reasoning development in older adults using the Standard Issue Scoring System (Colby & Kohlberg, 1987a, 1987b), indicate that development may be possible for some older adults. More research and experimental intervention in this era of the lifespan is clearly needed.

The lack of a systematic relationship between the development of moral reasoning and gender in this study is noteworthy. No significant relationship was found in the *Child and Adolescent Group* yet, in the *Younger Adult Group*, even after controlling for the effects of education and age, males had significantly higher overall scores than females. Adult women's scores increased significantly, however, over the course of the study, while men's did not. In addition, years of education completed for women changed significantly over the course of the study, while for men it did not, hinting at a partial explanation for the differences in developmental change between genders. It is intriguing that although the women in this study lag behind men in their overall scores, they appear to be developing somewhat more quickly. This trend, although marginal, combined with (1) the lack of a male advantage in the younger group, (2) the fact that eight of 19 women who entered the study during adulthood demonstrated Stage Four or Five reasoning and, finally, (3) the mixed findings of previous research counters any notion that relatively simple, systematic gender bias exists in this model of development. In fact, these findings indicate that the relationship between gender and the development of moral reasoning is highly complex.

For many of the participants in this study, moral development continued into middle age, although its occurrence diminished. This progressive reduction in development provides support for a curvilinear developmental path, and adds important evidence for the possibility of continued, significant development during middle and, perhaps, even late adulthood.

This group of bright, accomplished, high-functioning adults had more Stage Five reasoners than reported in typical samples. It appears to take more than three decades to construct post-conventional reasoning strategies, but time alone is not sufficient. Kohlberg (1981) may have argued for the greater adequacy of post-conventional reasoning, but he did *not* claim that its achievement was universal.

Many of these participants had the social privileges associated with high-stage development but did not achieve it. These findings lead to interesting interpretations, such as the notion that development may get "harder" at the higher stages. It is typically assumed that development from one stage to the next presents equivalent challenges, regardless of rank on the stage sequence, but it makes more sense that as stage constructs increase in complexity, they also increase in difficulty.

The consistent finding that reasoning above Stage Four occurs only during adulthood should encourage researchers to invest more energy in studying post-conventional reasoning development in adulthood. We need to know what distinguishes those adults who do develop post-conventional reasoning from those who do not.

Generally, formal education and the development of moral reasoning were moderately correlated throughout the life span in this study, which is consistent with others' findings. The longitudinal relationship found here between change in education and change in moral reasoning stage scores adds an additional dimension to our understanding of the relationship between education and the development of moral reasoning. Apparently, education that takes place in adulthood can be directly related to development during the same era. Not every participant who engaged in education during the course of the study, however, changed their moral reasoning. Conversely, not every participant who demonstrated an increase in moral reasoning stage score participated in formal education. Thus, formal education does not appear to be necessary for adult moral reasoning development. These findings indicate the possibility of a causal effect of education on the development of moral reasoning for some, and the importance of other influential life-experience factors for others. To "unpack" the impact of education, it is recommended that we take into account not only the educational content and its context, but also individuals' reactions to them (Spickelmier, 1983; Walker, 1986; Deemer, 1987; Nucci & Pascarella, 1987; Rest & Narvaez, 1991).

Elsewhere, "moral events" in adulthood are discussed in terms of their potential impact on development (Armon, 1995). As described earlier in this paper, socio-moral conflict is a core element of a moral event. The moral dimensions of conflict, however, must first be perceived as such and considered salient by the participant before he or she begins to engage in the kinds of moral problem-solving that are hypothesised to affect moral reasoning development. Unfortunately, the moral

nature of an educational experience is often ambiguous to participants. Moreover, morally educative events take place outside school as often as not. Narrative material from the participants of the present study indicates that even events containing moral conflict dramatically evident to the investigator can remain obscure to those living it. An analysis of these stories indicates that many moral and non-moral individual differences affect whether or not a moral conflict is seen as such, including personality variables, the particular content and context of the event, and the timing of the event in one's life. What seems clear is that "moral education" needs to teach its students the skills to be sensitive to the moral dimensions of their experiences, wherever and whenever they take place. Unfortunately, particularly in the United States, most educational institutions at all levels have not deemed these skills relevant.

We need to look at other social institutions as well. We should pay more serious attention to the social, political, and cultural influences on adult moral development. While power and privilege may provide conditions for conventional moral reasoning, we believe there is little demand and few rewards for post-conventional reasoning in the citizens of modern, capitalistic societies. These more contextual, socio-political forces should be included in future investigations of the possibilities and limitations of adult moral development.

In summary, this study demonstrates support for the structural–developmental model of moral reasoning development and provides new information about adult moral development. While it bolsters support for the structural–developmental approach, it also illuminates some of its shortcomings. Over the years we have learned much about the consistency of structural development. Moreover, newer methods and analyses (such as Rasch, e.g. Bond, 1994; Dawson, 1997) continue to validate many of the basic assumptions of the approach. What we lack is a rich understanding of the impact of that development on lived morality, particularly of adult subjects. It is our hope that future research will complement this reliable dimension of development with more contextual, behavioural and socio-political perspectives in the study of adult morality.

Acknowledgements

Findings and analyses of the first follow-up were included as part of the first author's doctoral dissertation at Harvard University, under the supervision of Lawrence Kohlberg. The balance of the research was completed with support from Antioch University Southern California. We wish to graciously thank all of the participants who have grown (and grown up) with this study for sharing their thoughts with us over 13 years. We are indebted to our interviewers Joyce Friedman, Larry Smiley, Lura Beth Illig and Sandra Nahan who sometimes travelled thousands of miles to unexpected places to meet with participants. We thank Michael Commons, Rhett Diessner, Richard Mattingly, Bill Puka and Lawrence Walker for valuable input on earlier versions of this paper. We also acknowledge Paul Holland and Michael Commons for their assistance with recent statistical analyses.

Correspondence: Dr Cheryl Armon, Antioch University, Southern California at Los Angeles, Liberal Studies Program, 13274 Fiji Way, Marina del Rey, California 90292, USA. e-mail: cheryl_armon@antiochla.edu.

NOTES

- [1] The results reported here were part of a larger study that included the investigation of evaluative reasoning about the good life (Armon, 1984a, 1984b, 1989, 1993, in preparation).
- [2] Some aspects of the narrative approach were incorporated into this study in the later years and the outcomes are reported elsewhere (Armon, 1995).
- [3] The current scoring system (Colby & Kohlberg, 1987a, 1987b) provides scoring methods for protocols only up to Stage Five.
- [4] The use of the log of age is a transformation that helps to correct for the abnormal distribution of that variable. (See, for example, Armon, 1984a, 1993; Brabeck & Wood, 1990; Wood, 1990, for previous uses; see Tabachnick & Fidell, 1989, for a statistical explanation.)

REFERENCES

- ALEXANDER, C. & LANGER, E. (1990) *Higher Stages of Human Development: perspectives on adult growth* (New York, Oxford University Press).
- ARMON, C. (1984a) Ideals of the good life: evaluative reasoning in children and adults. Unpublished doctoral dissertation, Harvard University.
- ARMON, C. (1984b) Ideals of the good life and moral judgment: ethical reasoning across the life-span, in: M. COMMONS, F. RICHARDS & C. ARMON (Eds) *Beyond Formal Operations: Vol. 1. Late adolescent and adult cognitive development* (New York, Praeger).
- ARMON, C. (1989). Individualism and autonomy in adult ethical reasoning, in: M.L. COMMONS, F. RICHARDS & C. ARMON (Eds) *Adult Development. Vol. 1: Comparisons and Applications of Developmental Models* (New York, Praeger).
- ARMON, C. (1993) Developmental conceptions of good work: a longitudinal study, in: J. DEMICK & P.M. MILLER (Eds) *Development in the Workplace*, pp. 21-37 (Hillsdale, NJ, Lawrence Erlbaum Associates, Inc.).
- ARMON, C. (1995) Moral judgment and self-reported moral events in adulthood, *Journal of Adult Development*, 2, 1, pp. 49-62.
- BAKKEN, L. & ELLSWORTH, R. (1990) Moral development in adulthood: relationship to age, sex, and education, *Educational Research Quarterly*, 14, pp. 2-9.
- BALTES, P.B. (1987) Theoretical propositions of life-span developmental psychology: on the dynamics between growth and decline, *Developmental Psychology*, 23, pp. 611-626.
- BIELBY, D. & PAPALIA, D. (1975) Moral development and perceptual role-taking: their development and interrelationship across the life-span, *International Journal of Aging and Human Development*, 6, pp. 293-308.
- BOND, T. (1994) Piaget and measurement II: empirical validation of the Piagetian model, *Archives de Psychologie*, 63, pp. 155-185.
- BRABECK, M.M. & WOOD, P.K. (1990) Cross-sectional and longitudinal evidence for differences between well-structured and ill-structured problem solving abilities, in: M.L. COMMONS, C. ARMON, L. KOHLBERG, F.A. RICHARDS, T.A. GROTZER & J.D. SINNOTT (Eds) *Adult Development, Vol. 2. Models and methods in the study of adolescent and adult thought*, pp. 133-146 (New York, Praeger).
- BROWN, L., DEBOLD, E., TAPPAN, M. & GILLIGAN, C. (1992). Reading narratives of conflict and choice for self and moral voice: a relational method, in: W. KURTINES & J. GEWIRTZ (Eds) *Handbook of Moral Behaviour and Development: Vol. 2, Theory, research, and application*, pp. 25-62 (Hillsdale, NJ, Erlbaum).

- CANDEE, D., GRAHAM, R. & KOHLBERG, L. (1978) *Moral development and life outcomes*. Harvard University: Report submitted to the National Institute of Education.
- CASE, R. (1985) *Intellectual Development: birth to adulthood* (New York, Academic Press).
- CERELLA, J., HOYER, W., RYBASH, J. & COMMONS, M.L. (Eds) (1993) *Adult Information Processing: limits on loss* (San Diego, CA, Academic Press).
- CHARNESS, N. (Ed.) (1985) *Aging and Human Performance* (Chichester, England, Wiley).
- COLBY, A. & DAMON, W. (1992) *Some do Care: contemporary lives of moral commitment* (New York, The Free Press).
- COLBY, A. & DAMON, W. (1995) The development of extraordinary moral commitment, in: M. KILLEN & D. HART (Eds) *Morality in Everyday Life* (Cambridge, Cambridge University Press).
- COLBY, A., KOHLBERG, L., GIBBS, J. & LIEBERMAN, M. (1983) Report on a 20-year longitudinal study of moral development, *Monograph of the Society for Research in Child Development*, 48, pp. 1-124.
- COLBY, A. & KOHLBERG, L. (1987a) *The Measurement of Moral Judgment, Vol. I. Standard issue scoring manual* (New York, Cambridge University Press).
- COLBY, A. & KOHLBERG, L. (1987b) *The Measurement of Moral Judgment, Vol. II. Standard issue scoring manual* (New York, Cambridge University Press).
- COMMONS, M.L., ARMON, C., RICHARDS, F.A. & SCHRADER, D. with FARRELL, E., TAPPAN, M. & BAUER, N. (1989) A multidomain study of adult development, in: M. COMMONS, J. SINNOTT, F. RICHARDS & C. ARMON (Eds) *Adult Development, Vol. 1. Comparisons and applications of developmental models*, pp. 33-56 (New York, Praeger).
- DAWSON, T.L. (1997) New tools, new insights: Kohlberg's moral reasoning stages revisited. Paper presented at the Twenty Seventh Annual Symposium of the Jean Piaget Society, Santa Monica, CA.
- DAY, J.M. & TAPPAN, M. (1996) The narrative approach to moral development: from epistemic subject to dialogic selves, *Human Development*, 39, pp. 67-82.
- DEEMER, D. (1987) Life experiences and moral development. Unpublished doctoral dissertation, University of Minnesota.
- DEVRIES, R. & ZAN, B. (1994) *Moral Classrooms, Moral Children: creating a constructivist atmosphere in early education* (New York, Teachers College Press).
- GIBBS, J.C., BASINGER, K.S. & FULLER, D. (1992) *Moral Maturity: measuring the development of sociomoral reflection* (Hillsdale, NJ, Lawrence Earlbaum Assoc.).
- GILLIGAN, C. (1982) *In a Different Voice* (Cambridge, Harvard University Press).
- HIGGINS, A. (1980) Research and measurement issues in moral education interventions, in: R.L. MOSHER (Ed.) *Moral Education: a first generation of research and development* (New York, Praeger).
- INHELDER, B. & PIAGET, J. (1958) *The Growth of Logical Thinking from Childhood and Adolescence* (New York, Basic Books, Inc.).
- KOHLBERG, L. (1969) Stage and sequence: the cognitive-developmental approach to socialization, in: D. GOSLIN (Ed.) *Handbook of Socialization Theory and Research*, pp. 349-480 (New York, Rand MacNally).
- KOHLBERG, L. (1981) *The Philosophy of Moral Development* (New York, Harper & Row).
- KOHLBERG, L. (1984) *The Psychology of Moral Development* (San Francisco, Harper & Row).
- KOHLBERG, L. (1990) Which postformal levels are stages? in: M.L. COMMONS, C. ARMON, L. KOHLBERG, F.A. RICHARDS, T. GROTZER & J. SINNOTT (Eds) *Adult Development, Vol. 2. Models and methods in the study of adolescent and adult thought*, pp. 263-268 (New York, Praeger).
- KOHLBERG, L. & ARMON, C. (1984) Three types of stage models used in the study of adult development, in: M.L. COMMONS, F. RICHARDS & C. ARMON (Eds) *Beyond Formal Operations: late adolescent and adult cognitive development*, pp. 383-394 (New York, Praeger).
- KOHLBERG, L., LEVINE, C. & HEWER, A. (1983) Moral stages: a current formulation and a response to critics, *Contributions to Human Development, Vol. 10* (New York, Meachum).
- KOHLBERG, L. (with HIGGINS, A.) (1984) Continuities and discontinuities in childhood and adult development revisited—again. *The Psychology of Moral Development* (San Francisco, Harper & Row).
- KREBS, D., VERMEULEN, S., DENTON, K. & CARPENDALE, J. (1994) Gender and perspective differences in moral judgment and moral orientation, *Journal of Moral Education*, 23, pp. 17-26.

- LOURENCO, O. (1996) Reflections on narrative approaches to moral development, *Human Development*, 39, pp. 83-99.
- MARKOULIS, D. (1990) Postformal and postconventional reasoning in educationally advanced adults, *Journal of Genetic Psychology*, 150, pp. 427-439.
- NISAN, M. & KOHLBERG, L. (1982) Universality and variation in moral judgment: a longitudinal cross-sectional study in Turkey, *Child Development*, 53, pp. 865-876.
- NUCCI, L. & PASCARELLA, E. (1987) The influence of college on moral development, in: J. SMART (Ed.) *Higher Education: handbook of theory and research*, Vol. 3, pp. 271-326 (New York, Agathon Press).
- PAGE, R.A. (1981) Longitudinal evidence for the sequentiality of Kohlberg's stages of moral judgment in adolescent males, *Journal of Genetic Psychology*, 139, pp. 3-9.
- PRATT, M. & ARNOLD, M.L. (1995) Narrative approaches to moral socialization across the lifespan, *Moral Education Forum*, 20, pp. 13-22.
- PRATT, M., DIESSNER, R., HUNSBERGER, B., PANCER, S. M. & SAVOY, K. (1991) Four pathways in the analysis of adult development and aging: comparing analyses of reasoning about personal-life dilemmas, *Psychology and Aging*, 6, pp. 666-675.
- PRATT, M., GOLDING, G. & HUNTER, W. (1983) Aging as ripening: character and consistency of moral judgment in young, mature, and older adults, *Human Development*, 26, pp. 227-288.
- PRATT, M., GOLDING, G. & HUNTER, W. (1984). Does morality have a gender? Sex, sex role, and moral judgment relationships across the adult life span, *Merrill-Palmer Quarterly*, 30, pp. 321-340.
- PRATT, M., GOLDING, G., HUNTER, W. & NORRIS, J. (1988a) From inquiry to judgment: age and sex differences in patterns of adult moral thinking and information-seeking, *International Journal of Aging and Human Development*, 27, pp. 115-130.
- PRATT, M., GOLDING, G., HUNTER, W. & SAMPSON, R. (1988b) Sex differences in adult moral orientations, *Journal of Personality*, 56, pp. 373-391.
- PRATT, M., PANCER, M., PRATT, A. & HUNSBERGER, B. (1994). *Perspective-taking and moral reasoning in later life: a longitudinal study*, Poster presented at the Canadian Psychological Association meeting, Penticton, BC, June.
- PUKA, B. (1989) Caring—in an interpretive voice, *New Ideas in Psychology*, 7, pp. 295-314.
- PUKA, B. (1996) Commentary. *Human Development*, 39, pp. 108-116.
- REST, J. (1975) Longitudinal study of the Defining Issues Test: a strategy for analyzing developmental change, *Developmental Psychology*, 11, pp. 738-748.
- REST, J. & NARVAEZ, D. (1991) The college experience and moral development, in: W.M. KURTINES & J.L. GEWIRTZ (Eds) *Handbook of Moral Behavior and Development*, Vol. 2. *Research*, pp. 229-245 (Hillsdale, NJ, Earlbaum).
- SALTHOUSE, T. (1993) The information-processing perspective on cognitive aging, in: R. STERNBERG & C. BERG (Eds) *Intellectual Development* (New York, Cambridge University Press).
- SALTHOUSE, T. (1994) The aging of working memory, *Neuropsychology*, 8, pp. 535-543.
- SNAREY, J.R., REIMER, J. & KOHLBERG, L. (1985) Development of socio-moral reasoning among kibbutz adolescents: a longitudinal cross-sectional study, *Child Development*, 53, pp. 865-876.
- SPICKELMEIR, J. (1983) College experience and moral judgment development. Unpublished doctoral dissertation, University of Minnesota.
- TABACHNICK, B.G. & FIDELL, L.S. (1989) *Using Multivariate Statistics* (New York, Harper Collins).
- VUYK, R. (1981) Overview and Critique of Piaget's Genetic Epistemology 1965-1980, Vols 1-2 (New York, Academic Press).
- WALKER, L. (1982) The sequentiality of Kohlberg's stages of moral development, *Child Development*, 53, pp. 1330-1336.
- WALKER, L. (1984) Sex differences in the development of moral reasoning: a critical review, *Child Development*, 55, pp. 677-691.
- WALKER, L. (1986) Experiential and cognitive sources of moral development in adulthood, *Human Development*, 29, pp. 113-124.
- WALKER, L. (1989) A longitudinal study of moral reasoning, *Child Development*, 60, pp. 157-166.
- WALKER, L., PITTS, R., HENNIG, K. & MATSUBA, M.K. (1995) Reasoning about morality and real-life

moral problems, in: M. KILLEN & D. HART (Eds) *Morality in Everyday Life: developmental perspectives* (New York, Cambridge University Press).

WOOD, P.K. (1990) Construct validity and theories of adult development: testing for necessary but not sufficient relationships, in: M.L. COMMONS, C. ARMON, L. KOHLBERG, F.A. RICHARDS, T. GROTZER & J. SINNOTT (Eds) *Adult Development, Vol. 2. Models and methods in the study of adolescent and adult thought*, pp. 113-130 (New York, Praeger).

WILLET, J.B. (1989) Some results on reliability for the longitudinal measurement of change: implications for the design of studies of individual growth, *Educational and Psychological Measurement*, 49, pp. 587-602.