CHAPTER 15

The Good Life

A Longitudinal Study of Adult Value Reasoning

CHERYL ARMON AND THEO LINDA DAWSON

Philosophers and social critics have promoted different conceptions of the good human life for some 2000 years. Such philosophical conceptions always included, or relied entirely on models of good psychological functioning or mental health. In contrast, psychologists have only recently entered the debate about the Good Life. It was not until the 19th and 20th centuries that theorists such as Baldwin (1906), James (1890), Freud (1961), Horney (1937), and Erikson (1963) began to articulate models of mental and psychological health. These models can be understood as attempts to define (in part) a good human life. Contemporary television and song lyrics reveal an interest in the nature a good human life. Yet, many believe that there are as many conceptions of the Good Life as there are persons who seek it (e.g., Nosick, 1974; Rawls, 1971). The findings of this study, however, indicate that although the sources of our conceptions differ widely across time and culture, the number of actual views of the Good Life may be finite. Moreover, despite the 2000 years that passed between ancient philosophers and early psychologists, these groups, too, produced some strikingly similar ideas about the Good Life. Further, there is also dramatic similarity between the Good Life concepts of many philosophers and the work of contemporary developmental psychologists. Finally, the results presented here demonstrate that educated adults, who have studied neither philosophy nor psychology, also construct similar good life concepts.

This study provides a general, developmental model of value reasoning about the Good Life and presents empirical findings from a 13-year study of young and older adults. It has been difficult to appreciate many of the substantial theoretical
and empirical commonalities in this area between philosophical and psychological studies, as well as among developmental studies themselves. Researchers working in different disciplines and subdisciplines are often unfamiliar with one another’s work. There are so many models, findings, and assertions about the development of reasoning about values, it is often difficult to see the forest for the trees. The empirical work and philosphical justification for the analysis of ethical judgments provided here represents an advance in understanding some core commonalities. A general, developmental model of value reasoning about the Good Life can incorporate many of these typically separate findings.

Although ethical philosophers describe the Good Life for adults, not children, there has been little research interest in adult value reasoning or adult morality. While it is adults who make social policy, establish a community’s quality of life, decide to wage war, and determine methods of parenting and educating the young, there is a paucity of research on how adults reason about values and morality, and how they might develop or change such reasoning. In a democratic society in which adult citizens are expected to participate intelligently, it is particularly bewildering that interest in adults’ cognitive, philosophical, and psychological capacities for this sort of intellectual work is so lacking. This phenomenon may be partly a result of psychology’s longstanding rejection of the notion of adult development, that is, the idea that adults are capable of changing their reasoning after their early 20s. Some contributions to this volume begin to fill this barren landscape (e.g., Brabeck and Schrader, this volume).

This chapter provides a partial report on the four assessments of a longitudinal study of developing conceptions of the Good Life begun in 1977, which included both children and adults. After the first follow-up in 1981, a developmental model of reasoning about the Good Life was constructed. The model relied on both structural–developmental and ethical theory for its theoretical framework. At that time, the Good Life Scoring Manual was completed, and it was proposed that developing conceptions of the Good Life could be effectively modeled with an invariant sequence of five stages (Armon, 1984b). There was also an emphasis on adulthood, particularly the investigation of structural–developmental change during the second half of the lifespan. The purpose of this chapter is to explore the nature of this reasoning and development of the adult subjects.

## REVIEW OF THE LITERATURE

### The Developmental Psychology of the Good Life: A Theoretical Framework for the Stage Model

*Structural–Developmental Approaches.* The structural–developmental approach (Colby & Kohlberg, 1987; Kohlberg, 1969, 1984; Piaget, 1960, 1968) has been used to investigate the structural organization of concept and value development in many different domains. The focus of the approach is on the hierarchical, structural organization of reasoning, and development is typically modeled in stages or levels. For example, Loevinger (1976) describes stages of ego development, Selman (1980) defines stages of good friendship concepts, Damon and Hart (1990) delineate stages of self-understanding, Gilligan (1981) describes the
development of care and responsibility, Fowler (1981) studies faith development, and Piaget (1932) and Kohlberg (1981, 1984) construct developmental models of justice and fairness. From a philosophical viewpoint, each of these areas fall within the ethics of human experience, and most could be said to fall within the Good Life. My own work has focused explicitly on the development of value reasoning about the Good Life, generally, and about good work, good relationships, and the good person, in particular (Armon, 1984a, 1984b, 1988, 1993, 1998; Armon & Dawson, 1997).

While philosophical and methodological differences among these models exist, commonalities are more striking, and they reside within the evaluative dimension. The models all describe similar paths of value development within one or more domains. At the earliest, or lowest, stage of development, subjects tend toward egoistic, impulsive, and undifferentiated value concerns. Next develops instrumental, controlled gratification and a focus on power. This is followed by group or interpersonal norms and affective emphasis, which is succeeded by societal or systematically constructed, self-authored values. The final level is always some form of autonomy, interdependence, or dialectical construction of values. (Table 15.1 presents developmental stage theories for comparisons.)

Where significant differences between developmental models do exist, they appear at the highest stages. It is unclear whether these differences reflect actual empirical differences, theoretical or ideological differences between researchers (e.g., liberal bias), or simply an inability to produce definitive constructions owing to the rarity of performances at these stages. The paucity of higher stage performances may be due, in part, to the fact that most developmental studies do not include older adults who are more likely to produce them.

Hard and Soft Stage Theories. The good life stage model is general and can be seen to incorporate some of the more domain-specific stage models within it. Generalizability is dependent, in part, on the extent to which the theory can withstand a “structural analysis.” Theorists regularly debate the structural integrity and/or exclusivity of various stage models. Notable is Kohlberg and Armon’s (1984) discussion of “hard” and “soft” stage models.

For some time, Kohlberg argued that “hard” stage theories meet Piagetian criteria for a stage,1 and are more generalizable (perhaps universal). In contrast, “soft” stage models may meet some of these criteria in a general way, but no serious attempt is made to meet them all. Kohlberg claimed that his stage model of moral judgment and Piaget’s stage model of cognitive development represented the only stage models that have met the necessary criteria. Yet, despite his more than 20-year effort, Kohlberg fell short of adequately demonstrating all four Piagetian criteria for the moral judgment stages. This is particularly evident in the area of defining the logic (hierarchical integration) of the stage sequence (Habermas, 1979; L. Kohlberg, 1979).

---

1 Piagetian criteria for a hard stage theory are: (1) A qualitative difference in structures, or modes of thinking, that serve the same basic function at various points in development; (2) the different structures form an invariant sequence in individual development; (3) each of these different and sequential modes of thinking forms a “structured whole,” in which structures appear as a consistent cluster of responses in development; and (4) stages form an order of increasingly differentiated and integrated structures to fulfill a common function. Accordingly, higher stages integrate the structures found at lower stages (Kohlberg & Armon, 1984).
### Table 15.1. Comparison of Developmental Sequences

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-Operational</td>
<td>Stage 1:</td>
<td>Stage 1:</td>
<td>Level 1:</td>
<td>Stage 1:</td>
<td>Stage 1:</td>
<td>Stage 1:</td>
<td>Stage 1:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Heteronomous</td>
<td>Differentiated and subjective</td>
<td>Symbolic</td>
<td>Categorical identifications</td>
<td>Impulsive</td>
<td>Intuitive-projective</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>morality</td>
<td>Level 2: Impulsive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concrete</td>
<td>Stage 2:</td>
<td>Stage 2:</td>
<td>Level 2:</td>
<td>Stage 2:</td>
<td>Stage 2:</td>
<td>Stage 2:</td>
<td>Stage 2:</td>
<td></td>
</tr>
<tr>
<td>Operational</td>
<td>Individualism,</td>
<td>Self-reflective</td>
<td>Survival</td>
<td>Survival</td>
<td>Survival</td>
<td>Survival</td>
<td>Mythic-literal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>instrumental</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>purpose, and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>exchange</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early Formal</td>
<td>Stage 3:</td>
<td>Stage 3:</td>
<td>Level 3:</td>
<td>Stage 3:</td>
<td>Stage 3:</td>
<td>Stage 3:</td>
<td>Stage 3:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mutual</td>
<td>Third-person</td>
<td>Interpersonal</td>
<td>Interpersonal</td>
<td>Goodness</td>
<td>Synthetic- conventional</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>interpersonal</td>
<td>mutuality</td>
<td>implications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>expectations;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>conformity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consolidated</td>
<td>Stage 4:</td>
<td>Stage 4:</td>
<td>Level 4:</td>
<td>Stage 4:</td>
<td>Stage 4:</td>
<td>Stage 4:</td>
<td>Stage 4:</td>
<td></td>
</tr>
<tr>
<td>Formal</td>
<td>Social system</td>
<td>In depth and</td>
<td>Systematic</td>
<td>Truth</td>
<td></td>
<td>Individuative-reflective</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>maintenance;</td>
<td>societal;</td>
<td>beliefs and plans</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>conscience</td>
<td>Symbolic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stage 4/5:</td>
<td>Stage 5:</td>
<td>Level 5:</td>
<td>Stage 5:</td>
<td>Stage 5:</td>
<td>Stage 5:</td>
<td>Stage 5:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subjective</td>
<td>Second-order</td>
<td>Autonomous</td>
<td>Interindividual</td>
<td>Non-violence</td>
<td>Universalizing</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>relativism</td>
<td>reciprocity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stage 5:</td>
<td>Stage 6:</td>
<td>Level 6:</td>
<td>Stage 6:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prior rights;</td>
<td>Second-order</td>
<td>Integrated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Social contract</td>
<td>mutuality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>or utility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

(a) Gilligan may not agree with this characterization of her stages.

personal communication, June 3, 1986; Sonnert & Commons, 1994; cf. Kohlberg, Levine, & Hewer, 1983). Nor is it clear that Piaget’s model met his own criteria for generalizability.² Piaget's stage of formal operations, for example, has been criticized repeatedly for not providing a full account of its transformational laws, and many have questioned the logic of the stages and substages (e.g., Broughton, 1984; Ennis, 1978; Feldman, 1980).

It is unlikely that this form of structural analysis will ever be completed or even attempted for the many stage models currently in use, at least to the degree demanded by Piaget or Kohlberg. Nevertheless, clearly articulated analytic models, applied in longitudinal studies, using valid and reliable scoring schemes, and including cross-cultural tests can help immensely in distinguishing between content-oriented, phase-type theories (“soft-stage models,” Erikson, 1963; Levinson, Darrow, Klein, Levinson, & McKee, 1978) and sequential, structural theories.

Newer forms of analyses, however, continue to support fundamental structural–developmental constructs (e.g., the General Stage Model, Commons & Richards, 1984; Lam, 1994; Rasch Analyses, Dawson, 1997).

Structural Development in Adulthood. During the 1960s and 1970s, when humanistic psychologists were celebrating adults’ self-realization and capacities for change (e.g., Erikson, 1963; Maslow, 1971; Rokeach, 1973), most developmental psychologists continued to accept Piaget’s (1972) assertions that (1) formal operations developed in early adolescence and (2) formal operations was the final stage of cognitive development. In the last two decades, however, many researchers have posed alternatives to these claims. These alternatives focus on the idea that formal operations is not necessarily the last stage in cognitive development, or that current definitions of formal operations do not include all of the cognitive functions of that period, or that the endpoint of cognitive development does not necessarily occur during adolescence, but rather in adulthood, or combinations of these ideas (e.g., Alexander, Drucker, & Langer, 1990; Richards, Armon, & Commons, 1984). During the late 1980s and early 1990s, an array of books, textbooks, articles, and a professional journal devoted exclusively to theory and research about progressive development in adulthood appeared (e.g., Alexander & Langer, 1990; Commons, Armon, Kohlberg, Richards, Grotzer, & Sinnott, 1990; Commons, Richards, & Armon, 1984; Commons, Sinnott, Richards, & Armon, 1989; Demick, 1999; Rybash, Roodin, & Santrock, 1991; Stevens-Long & Commons, 1990). It should no longer be necessary to defend the study of development during adulthood. Today, progressive research programs are less likely to investigate whether there is adult development but, rather, its prevalence, form, and nature.

Lifespan Developmental Approaches. While not focused so narrowly on structural development, lifespan developmental psychology supported early studies of progressive adult development and aging (Baltes, 1983; Baltes, Hayne, &

² For Piaget (1970), “stages” represent both organizations of structures and the processes and contingent phenomena involved in structural development. Thus, it was primarily the proposed structures that must meet criteria, not the stages, which are more general. His criteria for structures concern: (1) wholeness; (2) transformations; and (3) self-regulation. A structure is a system of transformations, or a system of laws of transformations. Thus, a structural analysis will identify the transformations, the transformational system, and the laws that govern them, of each structural organization or “stage.”
Lipsitt, 1980). In its formative years, lifespan psychology was defined both in terms of universals (Birren, 1964) and in terms of the particular historical and personal events that influence an individual life (Baltes, 1979). Researchers in this area have by now amassed a large body of work on adult development and aging. Although acquired and discussed differently than those of the structural developmental approach, these results also help describe ideals of the Good Life in adulthood. For example, there is extensive literature on psychological well-being in the second half of life that provides theoretical models of positive functioning, life satisfaction, happiness, adjustment, and the like, which can be seen as different ways of talking about the Good Life in adulthood (e.g., Cutler, 1979; Lawton, 1984; Stock, Okun, & Benin, 1986).

More recent lifespan work on “successful aging” even more closely parallels attempts to describe the Good Life. This work has gone beyond theoretical models and investigated “how middle-aged and older adults themselves define positive functioning” (Ryff, 1989, p. 195). Performing a content analysis on adults’ responses to questions such as “What is most important in life?” and “How would you define an ideal person?” Ryff (1989) reports a greater incidence of the type of responses commonly found at the higher, adulthood stages in structural–developmental models. For example, while structural–developmental researchers report work-related accomplishments as central to the conventional level of development in adulthood (e.g., Armon, 1993; Kegan, 1982, 1994), Ryff (1989) reports job-related experiences as the second most important activity of midlife. Similarly, while structural–developmental researchers, such as Loevinger (1976), Armon (1984a), Colby and Kohlberg (1987), and others report the maintenance of interpersonal relationships as central to the earlier (and more commonly found) stage of the conventional level in adulthood, Ryff reports family relationships to be the most important activity of midlife. Thus, the content analyses of the lifespan researcher and the structural analyses of the structural–developmental counterpart come together to enhance our understanding of certain aspects of the Good Life in adulthood.

The Philosophy of the Good Life

In addition to developmental psychology, the developmental model of value reasoning about the Good Life relies on moral philosophy, or ethics, in two ways. First, philosophy provides analyses about the possible relationships between the good and the right, which are described briefly later in this section. Second,
traditional philosophical constructions of the Good Life were used to inform the
collection of the content categories used in part of the data analysis. In the ini-
tial model building of the Good Life stage model (Armon, 1984b), it was found that
the material adult subjects offered when describing their ideals of the Good Life,
particularly at the higher stages, was similar to professional philosophers’ views.
Thus, traditional philosophies of the Good Life were used to categorize adult sub-
jects’ responses. The following section briefly describes the different philosophical
orientations that formed such categories for the present study. (A fuller account of
these categories can be found in Armon, 1984b.)

Philosophies of the Good Life: A Theoretical Framework for
Categorizing Content

Traditionally, ethics has attempted to answer value-related questions of seem-
ingly universal interest such as “What is the Good Life?” Responses to this ques-
tion can be roughly classified as either Perfectionistic or Hedonistic. Perfectionist
theories define good living as the development and expression of inherent human
talents and capacities. (Which capacities to perfect is oft debated.) Hedonist theo-
ries define good living as the successful acquisition and appreciation of pleasure—
the ultimate intrinsic value. In Hedonism, the means to pleasure are secondary.
Achieving the result, drawing pleasure from an object or activity, is key.

For the content categories, this study relied primarily on the works of three
perfectionists: Aristotle, Spinoza, and Dewey. It also drew on two hedonists,
Epicurus and Mill. Although these are leading theorists in their persuasions, others
could have been chosen to exemplify the ethical views. Figure 15.1 illustrates these
five Orientations:

Perfectionism. For Aristotle (Ostwald trans., 1962), the distinctive human
capacities, especially practical and theoretical wisdom, are what we must perfect.
From this perspective, the Good Life consists of experiences that ensure the
development of intellectual, ethical, and even physical abilities that are uniquely
human. (This Orientation is referred to here as Perfectionism–Capacities.) For
Spinoza (1949), the Good Life develops our recognition, through self-knowledge,
of our interdependence with Nature or God (referred to here as the Perfectionism–
Unity Orientation). Instead of a focus on particular capacities or endpoints,
Dewey (1944, 1957, 1980) claims that experiences that contribute to continuous, progressive development are central to the Good Life. (This Orientation is called Perfectionism–Progressive in this study.)

**Hedonism.** In contrast to Perfectionistic views, Hedonists Epicurus and Mill claim that the experience of pleasure or happiness, and the absence of pain, should be the universal objects of human desire. Epicurus (from Laterius, 1925) constructs an essentially self-centered philosophy of the Good Life, concerning himself with individual pleasure. (Here, this Orientation is called Classical Hedonism.) Questioning the delights of sensuality and extravagance, however, he recommends pleasures resulting from peaceful equilibrium—attained through self-knowledge, a virtuous character, and the company of intellectually stimulating friends.

Another form of Hedonism is offered by Mill (1957/1861, 1978/1861). He rejects the egoistic aspect of Epicurus’ thought and argues that the Good Life takes place in a social rather than an individual context in which the consideration and maximization of all persons’ pleasure is paramount. (This orientation is called Social Hedonism in this study.) Like Epicurus, Mill asserts that the pleasures resulting from persons’ higher faculties, including pleasure resulting from a virtuous character, are superior to “sensuous pleasures.”

The good life stage model brings together the philosophy and psychology of the Good Life. There are, of course, cultural, historical, ideological, and methodological differences between these disciplines. Thus, one might expect ideals of the Good Life constructed by (some ancient) philosophers to be radically disparate from those inferred from developmental psychology. Yet, when carefully considered, particularly when focusing on adult reasoning at the higher stages of development, the philosophy and developmental psychology of the Good Life appear more similar than different, containing core elements that have remained central for thousands of years. Bringing together different developmental models and combining them with the typically separate discipline of ethical philosophy reveals commonalities that inform our understanding of a good human life from prescriptive and descriptive, theoretical and empirical perspectives.

**The Good and the Right.** The operational definition of the Good Life is the combined set of values that persons affirm in normative, ideal-evaluative judgments about the Good Life, in general, and about good work, good friendship or relationship, and the good person, in particular.

The domain of the Good Life is conceptualized as broad, including the moral good (e.g., ethical dimensions of persons, relationships) and nonmoral good (e.g., nonmoral aspects of work, family, and community). On occasion, subjects also produce judgments concerning the moral right in describing the Good Life and such judgments must also be accounted for. To classify the ethical quality of responses, categories were developed from the philosophical works of Campbell (1935), Frankena (1973), Hare (1952), and Ross (1930), and are generally consistent with the works of Lewis (1946), Perry (1926/1954), and Rawls (1971, 1980). These categories are illustrated in Table 15.2. In Table 15.2, the first category on the left, the moral right includes judgments concerning duties, human and legal rights, and absolute principles. The second category contains judgment of the
morality and the Good Life. Included in this category are both aretaic judgments (judgments of character, the good person) and judgments of welfare consequences, such as the consequences of actions taken on moral motives. The former resides in the person, and is concerned with the moral worth of individuals, or traits of character. The latter may appear nonmoral, such as a group of children receiving an enriched education. It becomes a morally good consequence, however, if it occurred as a result of moral motives.

The third category is intrinsic, nonmoral good, and contains judgments about generally accepted human values, for example, knowledge, sociality, or artistic expression, which in themselves are nonmoral. These are sometimes referred to as “end-values” (Rokeach, 1973). The final category is extrinsic, nonmoral good, which contains judgments about “goods” that people value because of what they bring or do, not because of what they are in themselves. This category would include cars, pencils, or houses. These are sometimes referred to as “means values” or “instrumental values.”

Distinctions Between the Good and the Right. As described, judgments about the Good Life often concern the moral good and, sometimes, the moral right. Some developmental psychologists, notably Kohlberg (1981), have gone to great lengths to keep these categories of judgments distinguishable. Others, for example, Gilligan (1981) and Selman (1980), have ignored these distinctions.

Ethical philosophers typically expend great effort in defining the relationship between the good and the right. For example, Rawls’ (1971, 1980) deontic perspective holds that conceptions of the good are pluralistic. He focuses on the primacy of universal principles of right action (moral right, justice) for the distribution of the conditions for the attainment of any good life. Thus, conceptions of the good are always subordinate to conceptions of the right. In contrast, Mill’s (1957/1861, 1978/1861) utilitarian view requires that the Good first be defined since the moral obligation, or first principle, is to maximize the Good.

For the present study, the debate concerning the primacy of the Right or the Good need not be resolved. It is proposed here only that they can be distinguished and that a consistent and generalizable psychological theory of reasoning about the Good Life may be possible.

This report of the longitudinal findings briefly covers the first and second assessments (described in detail in Armon, 1984b) and provides more information
on the results of the third and fourth assessments of the study of good life reasoning of a group of adults over a period of 13 years. The study examines stage development during the lifespan in four domains: Good Life, Good Work, Good Friend, and Good Person. Findings from the broad good life domain are reported here. The main purpose of this follow-up report is to validate the sequentiality of the stage model and to continue investigating adult structural development. The following hypotheses were specified for the analyses of follow-up data: (1) Conceptions of the Good Life can be represented by a developmental, sequential stage model, ordered by qualitative increase in cognitive complexity (inclusiveness) and social perspective-taking capacities; (2) structural development (stage change) occurs in adulthood; and (3) the higher stages of the model will appear only in adulthood, even with a “privileged” sample.

METHOD

Study Design

The design of the original cross-sectional study, which was later followed up longitudinally, was determined by a number of theoretical concerns. There was a direct attempt to create a design that would be most likely to identify development throughout the lifespan, particularly in adulthood. Three variables were most important in the design: age, interest, and higher education. In accordance with Dewey’s (1944) notion that an individual’s interests are key factors in his or her learning and development, individual interest in development or change was hypothesized to be positively associated with continued development. It was also thought that interest would affect the subjects’ willingness to perform the practical duties of a long-term longitudinal subject (e.g., notifying the experimenter when relocating). Participation in higher education was also expected to be associated with continued development, as shown by other studies (e.g., Colby, Kohlberg, Gibbs, & Lieberman, 1983; Rest & Thoma, 1985; Walker, 1986). The study was first executed cross-sectionally in 1977. In 1980, most of the subjects were located and interviewed again. In total, subjects were interviewed four times—in 1977, 1981, 1985, and 1989–90.

Participants

At time 1 (1977), 50 individuals, ranging in age from 5 to 72, responded to a flyer seeking volunteers. The flyer sought “people interested in development through the lifespan, or their children,” and was distributed to friends and strangers around a state university in the Los Angeles area. The first 50 individuals to respond, who ranged in age from 5 to 72, were accepted. 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 in this study.

Only adult data are presented here. Adults were divided into two age groups, consisting of a younger adult group of twenty-one 23- to 45-year-olds (11 females and 10 males), and an older adult group of eight 50- to 72-year-olds (6 females and 2 males). Categorization into these groups was based on the view that different developmental trends may occur in adulthood and late adulthood.

**Procedures**

Thirty adults were interviewed at time 1, 27 were interviewed at time 2, 25 were interviewed at time 3 (1985), and 23 were interviewed at time 4 (1989–90). Five participants missed the last two interviews due to health problems or 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.

*Training of Interviewers and Scorers.* The training of two adult, advanced undergraduate students was performed by the author in 1985 and 1989. First, the two students studied the rules and procedures outlined in the Good Life Scoring Manual (Armon, 1984b). Second, they conducted practice interviews with volunteer subjects not associated with this study. The practice interviews were tape-recorded and transcribed. Third, each of the two students practice-scored the other student’s interview independently, using the Good Life Scoring Manual, while the author scored both. Finally, the three of us met and reviewed the protocols and scoring to discuss errors in both the interview methods and the scoring, and differences between the three of us in the scoring. This process involved approximately 10 interviews (over 3 months) until a level of reliability and validity was reached that was acceptable to the three scorers. (See Reliability, below.)

**Administration**

The structural–developmental research approach (Colby & Kohlberg, 1987) was used to collect and analyze the interview data at each test time. (See Armon, 1984, for a complete explanation of how data were collected and analyzed.) The semistructured Good Life Interview (Appendix A), administered at each test time, consists of open-ended questions in many domains beginning with “What is the Good Life?” (combined with multiple probe questions to elicit subjects’ underlying reasoning, e.g., “Why is that good?”), and including questions about good work, good friendship or relationship, and good person. Demographic and life history information was also collected each time. It was not uncommon for interviewers to travel out of state, to other countries, and to federal prisons to locate and interview subjects. Thus, the environmental milieu of the interviews was variable. Interviews were always conducted, however, in privacy, with only the
interviewer and the subject present in a room. The interviews were tape-recorded and transcribed. The average interview time was two and a half hours.

RESULTS

Data Analysis and Results of 1977 and 1981 Assessments

This section provides a very brief summary of the main findings from the first follow-up study. Detailed explanations can be found for all analyses and results mentioned here in Armon (1984b).

Preliminary Good Life Reasoning Stages. In 1977, working from the original cross-sectional data of 50 subjects, five preliminary stages of good life reasoning were proposed (Erdynast, Armon, & Nelson, 1978). In 1982, a construction sample of 12 follow-up cases was created by randomly selecting four cases from each of the three age groups—children, young adults, and older adults. Based on the clinical analyses of these data, which is described briefly below, the “value reasoning” construct was refined, the Philosophical Orientations were operationalized with the use of Norms and Elements (see below), and the Good Life Scoring Manual (Armon, 1984b) was constructed. The manual’s scoring system converts stage scores to a continuous scale, for example, stage 1 = 100 EMS (Ethical Maturity Score).

Method of Analysis. Units of analysis were those statements subjects used to affirm something as good, worthy, or ideal and the values and reasons offered in the justification. This modeling of value judgments was adapted from Nowell-Smith (1954). The contents of the scorable units were categorized using an Issues, Norms, and Value Elements system similar to that of Colby and Kohlberg (1987), but additional Norms and Value Elements had to be added from Rokeach (1979) and Maslow (1971) to cover all the good life judgments. In a fully elaborated good life judgment, an Issue, Norm, Modal Element, and Value Element is present.

Issues, Norms, and Elements. In assessing good life judgments, the Issue is merely the topic under discussion, usually identified by the question that the subject is asked (e.g., “What is a good person?”) or sometimes by spontaneous response of the subject. The Issue categorizes the content that the subject is expressing an evaluation about (e.g., good person, good work). The Norm categorizes the general value that the subject assigns to the issue (e.g., honesty, caring, love, society). In other words, the Norm canotes the value area that the subject begins discussing in relation to the Issue (e.g., in the response, “A good person is trustworthy,” “good person” is the Issue and “trustworthiness” is the Norm). There are moral and non-moral Norms. Modal Elements express the type of judgment; they express the “mood” or modality of ethical language, such as “having a right,” or “character,” or “blaming/approving.” Finally, Value Elements are values used to support and justify Norms. Value Elements provide the final aspect of a judgment. They are terminal or end-values (Rokeach, 1973) for which the Norm serves as object.

To give an example of the Issue, Norm, and Value Element system, let us say a subject’s response to “What is a good person?” and “Why?” is: “A good person
works for the community.” When questioned, “Why is that good?” the subject replies, “Because it shows a certain character.” Questioned further: “What do you mean a certain character?” subject replies, “Well, I think a good person gets his personal satisfaction from working in community with and for others.” In this judgment, the Issue is Person. The Norm is Community. The Modal Element is Character. And the Value Element is Satisfaction/Fulfillment.

Once content was categorized in this way, systems of reasoning, that is, methods of organizing the Norms and Elements, were identified in subjects’ protocols independent of the particular Norms or Elements used. Similar systems were grouped and then groups were hierarchically ordered by increasing cognitive complexity, or inclusiveness, and by increasing levels of social perspective-taking (Armon, 1984a, 1984b; Selman, 1980).

Final Good Life Reasoning Stages. The resulting five-stage, hierarchical model confirmed the general construction of the initial stages derived from the cross-sectional data (Erdynast, Armon, & Nelson, 1978). It begins in early childhood with an egocentric conception of the Good Life derived primarily from pleasure-seeking fantasy (e.g., “The Good Life is having my birthday party every day”) and culminates with a complex conception of the good that encompasses complex criteria, including a preeminent societal dimension (“The Good Life is the worthy life. It is the integrated life—bringing the various facets of experience into balance with my interests and talents. It is also constructed in social context. To be good, it must move the society forward in some way”).

The stages are most easily observed in individuals’ constructions of their evaluative criteria; that is, the criteria the subject uses to decide whether a person, idea, state, or activity is good. Table 15.3 summarizes the stages of value reasoning about the Good Life.

Reliability of the Scoring Manual for Stage Assessment. The interrater reliability of the scoring manual, including stage and Philosophical Orientation assessment, and the long- and short-term test–retest reliability of the good life reasoning construct were substantially above acceptable limits. Detailed tests were performed using the Good Life Scoring Manual in 1983 and are reported in Armon (1984b). Compared to the reliability estimates of other, similar models (e.g., the Standard Form Moral Judgment Scoring System, Colby & Kohlberg, 1987), the present model and scoring scheme is well within acceptable limits. In addition, the Good Life Scoring Manual has been used in other studies with acceptable results (e.g., Commons, Armon, Richards, & Schrader, 1989; Lam, 1994).

CURRENT METHODS OF ANALYSIS

Several analyses in this report are 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). Second, the longitudinal information incorporated into the analysis in this way adds valuable
<table>
<thead>
<tr>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Stage 4</th>
<th>Stage 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egoistic hedonism</td>
<td>Instrumental hedonism</td>
<td>Affective mutuality</td>
<td>Individuality</td>
<td>Autonomy/community</td>
</tr>
</tbody>
</table>

The individual child does not possess conscious value criteria. Nor is the rational distinguished from the irrational; possible and impossible occurrences are not distinguished. Only ends are considered, not the means for their attainment, nor their possible consequences. Perceived to be good are those material objects and physical activities that provide pleasure to the self. The Good is synonymous with the desired. No distinction is made between physical pleasure, happiness, or contentment. Other persons, including family members, are rarely included.

The individual thinks instrumentally about achieving the Good Life for him-or herself. In considering the means, individuals at this stage contemplate others’ interests, motives, and intentions, as well as external physical and social conditions. What characterizes these means, however, is their concrete, instrumental quality. Others are considered as separate persons with their own interests, but the focus is on how others serve the self’s ends. Other people are important aspects of the good life because they are a means to the self’s ends.

The Good is shared by the self and others. Reasoners see the mutual quality of interaction with others as an integral part of the Good Life. Mutuality in relationships and consensus in valuing is sought. Beyond the general distinction between happiness and pleasure, individuals attend to the form of happiness itself. Happiness has a distinct meaning, grounded in affective contentment. There is also a sense in which happiness, or the good life in general, can be defined as the absence of certain negative affective states or experiences, such as loss, crisis, loneliness, fear, anxiety, worry, and stress.

The origin of value lies within the individuated self. A central feature is a concern with individualism. This is an orientation toward self-chosen values and the freedom to go against consentual norms, if necessary, to make choices about, and to pursue, one’s particular vision of the Good Life. Meaning, worthiness, and value are criteria for satisfaction, independent of others’ beliefs and desires. Wide variability in individuals’ values is acknowledged and tolerated. This awareness can be coupled with a form of relativism in reasoning about the good, particularly in the Classical Hedonism orientation. In the Perfectionistic Orientation, relativism is less prevalent. The conception of good tends to be generalized to other persons. The focus remains, however, on the fulfillment and realization of the self’s chosen values.

Individuals attempt to construct what is of value independent of social and historical norms. Thus, value is something to be perceived and constructed by each individual. Individuals employ generalizability, universality, and/or intrinsicality, rather than individualism, as criteria of value. The emphasis is placed not on the choosing of values, but on the perception and construction of the worth of the values themselves, for both the self and others. The focus is on those traits, objects, processes, or states that possess intrinsic value and the obligation to uphold those values, once recognized. A principled, ethical view of an ideal human world is constructed in which justice is a precondition for goodness.
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 of 1985 and 1990/91 Assessments

Manual Reliability. To continue to test the reliability of the Good Life Scoring Manual, one third of the 1985 and one third of the 1989–1990 protocols were each scored independently by the author and two trained undergraduate students using the unaltered Good Life Scoring Manual (Armon, 1984b) in 1990. There was a 100% perfect agreement rate within one stage, 93% agreement within a half stage (50 EMS points), and 88% agreement within a fourth stage (25 EMS points).

Invariant Sequence of Stages. To meet the criteria for invariant sequence, it must be demonstrated that, beyond measurement error, no subject’s stage score at \( T+1 \) is less than their score at \( T \) (which would constitute regression) and that no subject skipped a stage while developing progressively through the sequence. These criteria were strongly supported by the follow-up data. While significant stage change did not occur in all subjects at each test time, for those whose reasoning did change, it changed toward the next, successive stage. Assuming a conservative measurement error of \( \pm 31 \) EMS points (just over \( 1/4 \) stage), no regression was identified. In the 4 (sometimes a bit more) years between test times it might have been possible for subjects to develop beyond the next ordered stage, thereby possibly “skipping” a stage. In this sample, however, this did not occur. Figure 15.2 plots all adult subjects’ scores at the four test times with age. There it can be seen that subjects’ reasoning either remains at the same stage or changes toward the next, ordered stage.

Internal Consistency of Stage Scores. The internal consistency of stage scores was tested in 1981 and 1985 by measuring the proportion of reasoning at each stage in the construction protocols (12 cases, two interviews each, from three age groups, see above). The 1981 results can be found in Armon (1984b) and are almost identical to those of 1985. In 1985, on average, 74% of the scorables

---

4To find an upper limit for the standard error of measurement in the 77–81 analyses, the highest variation was coupled with the lowest correlation. A 95% confidence interval was computed around the long-term, test–retest correlation,.95. The SD of 100 was used with the lower limit of that interval,.90. The standard error of measure was then estimated with the following equation:

\[
\varepsilon = 100 \sqrt{1 - .90} = 36.97
\]

This results in a standard error of 31 EMS points.
responses in each protocol was at a single stage—the “modal stage” (Colby & Kohlberg, 1987). The mean percent of the next most often used stage, always adjacent to the modal stage, was 21%. The mean percent of the third most often used stage, also always adjacent to the modal stage, was 5%. The number of scorable responses in a given protocol was widely variable, particularly with the young children in the sample. A 6-year-old might produce as few as two scorable responses for an Issue, while an older adult might produce 20 or more. Nevertheless, in the construction sample, no subject produced a single scorable response (even a “guess score”) that was assessed at a stage other than the modal or adjacent stage.  

Rasch Analysis

Rasch analysis was used to further examine the construct validity of the Good Life measure. Though well-known in psychometric circles, Rasch models have been employed by developmentalists only recently (Andrich, 1986; Andrich & Styles, 1986; 1987). Another test of internal consistency investigates the stability of stage scores across the four different Issues. Although each Issue is scored separately, independence was not guaranteed. Correlations between pairs of Issues ranged from .82 to .94 in 1977 and 1981 (see Armon, 1984b).
A Longitudinal Study of Adult Value Reasoning

1994; Bock, 1991; Bond, 1994, 1995; Dawson, 1996; Demetriou, Efklides, Papadaki, Papantoniou, & Economou, 1993; Draney, 1996; Goodheart, Dawson, & Commons, 1996a, 1996b; Hautamäki, 1989; Müller, Reene, & Overton, 1994; Müller, Winn, & Overton, 1995; Noeltling, Coudé, & Rousseau, 1995; Noeltling, Rousseau, & Coudé, 1994; Wilson, 1985, 1989a, 1989b; Wilson & Draney, 1997). One function of these models is to examine behavior on measures intended to capture hierarchies of difficulty, which makes them highly suitable for developmental applications. The Rasch model tests the assumption that performances and items (or levels of items) form an invariant, hierarchical sequence (within probabilistically determined constraints) that can be successfully modeled along a single continuum (Andrich, 1989; Fisher, 1994; Masters, 1988; Wright & Linacre, 1989).

In their raw form, little can be said about the relative distances between stage scores (Duncan, 1984; Michell, 1990; Thurstone, 1959; van der Linden, 1994). A Rasch analysis transforms these scores into interval form through a log transformation (Wright & Linacre, 1989). The result is a common metric along which both stage difficulty and respondent ability estimates are arranged. The metric is referred to as a logit scale. (For more information on logits and algorithms for estimating them, see Wright and Masters, 1982, or Ludlow and Haley, 1995.) The distance between logits has a particular probabilistic meaning. An ability estimate for a given individual, say of 1.0 logits, means that the probability of that individual performing accurately on an item at the same level is 50%. There is a 73% probability that the same individual will perform accurately on an item whose difficulty estimate is 0.0, an 88% probability that he or she will perform accurately on an item whose difficulty estimate is −1.0, and a 95% probability that he or she will perform accurately on an item whose difficulty estimate is −2.0. (For general discussion of the properties of the Rasch model and related models, see Andrich, 1988; Fisher, 1994; Masters, 1982, 1997.)

One advantage of Rasch models is that they provide estimates of both participant abilities and item difficulties. Each of these includes an error term, which makes it possible to establish confidence intervals for all item and person estimates. Fit statistics are also provided for each estimate, so both items and persons can be examined for their conformity with the requirements of the model. Both individual items or persons and subgroups of items or persons can be examined in this way. As is demonstrated below, the results of the analysis provide a great deal of useful information about the psychometric properties of one's instrument, patterns of behavior, and the latent variable being investigated. (The software used to run this analysis was Quest [Adams & Khoo, 1993].)

To conduct the analysis on EMS scores, it was necessary to convert weighted average scores into stage scores. The translation was conducted as follows:

- 100–124 = Stage 1.0
- 125–174 = Stage 1.5
- 175–224 = Stage 2.0
- 225–274 = Stage 2.5
- 275–324 = Stage 3.0
- 325–374 = Stage 3.5
- 375–424 = Stage 4.0
- 425–474 = Stage 4.5
- 475–500 = Stage 5.0
It should be noted that the half-stage scores represent a mixture of reasoning at adjacent stages. The Good Life scoring system does not include scoring criteria for half-stages. Data from all test times for all three age groups was included in the analysis to maximize the sample size. The larger sample size provided more reliable item estimates. However, only case estimates for adult participants are shown in Fig. 15.3. The partial credit model (Masters, 1982, 1994), an extension of the original Rasch model designed to permit the estimation of multiple levels for each item in an instrument, was utilized for this analysis. The partial credit model is appropriate here because stage of reasoning was assessed in each of four domains (the Good Life, Good Work, Good Friendship, and the Good Person), and

![Figure 15.3. Rasch analysis: item and case map.](image-url)
each of these domains has nine possible levels (five stages, and four half-stages). In the analysis, each domain is equivalent to an item. The proportion of the observed case estimate variance that is considered to be true (reliability of case estimates) is .93. (See Wright & Masters, 1982, for more information about reliability of case estimates.) Given the small sample size ($n = 147$) and the number of levels and items, .93 is remarkably good. Reliabilities for case estimates of .85 and up are considered acceptable in ability testing (Wilson, January 1997, personal communication). All items have acceptable in-fit (weighted) and out-fit (unweighted) fit statistics (Wright & Masters, 1982), indicating that they all belong on the same scale or, in other words, that they measure the same dimension of ability. In-fit and out-fit statistics are expressed as meansquares and $t$s. Meansquares are expected to be close to 1.0 and $t$s are expected to be close to zero.

Figure 15.3 shows the item and case estimates for the Good Life analysis. The case estimates are displayed on the left of the figure, and the item estimates are displayed on the right. Near the center of the figure is the logit scale. Standard errors for individual cases range from .63 to 1.08, with a mean of .78. Standard errors for the item estimates range from .44 to 1.28 logits. Two standard errors are shown in Figure 15.3 as bars around each level of each item for which a standard error could be calculated. These can be thought of as 95% confidence intervals. Standard errors are greater at the lowest and highest stages because the sample sizes on which estimates are based are small at the extremes of the range; hence, we are less confident of the locations of persons and items at these locations.

Cases that fit the model are indicated with Xs, those with negative weighted misfit are indicated with Os, and those with positive weighted misfit are indicated with Xs. Altogether, 11 performances (7.5%) do not fit the model. This is outside the expected 5% that can justifiably be attributed to measurement error. Not all of these performances, however, raise actual contradictions in terms of the theoretical model. Six cases exhibit negative misfit (4.1%). All of these are participants who performed at a single stage. If some performances at a given difficulty level are much more consistent than others at that level, they will have negative misfit. These performances are more consistent than expected, but they are congruous with the theoretical expectation that some individuals will demonstrate performances that are consolidated at a single stage. Five of the negatively misfitting cases are individuals whose performances on all items were scored at stage 3.0. Note that their estimates are located approximately 3.0 logits above the difficulty estimates for the stage 3 items. This means that they, and the other cases at their ability level have about a 88% probability of scoring at stage 3.0 and somewhat more than a 50% probability of exhibiting some stage 4.0 reasoning. Only five cases exhibit positive misfit (3.4%). These are the individuals whose scores span more than 1.5 stages. Although these cases pose problems for both the model and our theoretical expectations, 3.4% is well below the 5% expected error level. Overall, the reliability of case estimates, along with case and item fit statistics, provide good evidence for the construct validity of the Good Life instrument. This analysis also provide support for the invariant sequence of stage development and for the structured whole criterion. Evidence in support of invariant sequence is provided by the relatively small standard errors for most levels of most items. In Figure 15.3, the error bars clearly show gaps between each full stage and the half-stage above it (one exception occurs at the stage 2.0/2.5 transition).
Item separation of this kind is strong evidence of invariant sequence. It is also apparent from Fig. 15.3 that stages tend to cluster together. Note the horizontal gray lines between stage 3.0 and the 3.5/4.0 cluster and stage 4.0 and the 4.5/5.0 cluster. These gray lines mark the areas in which there is no overlap of error ranges between scores at one full stage and those at the next half stage (which is, as stated above, a combination of reasoning at the two adjacent stages). In contrast, note that error ranges between half-stages and the next full stage usually overlap. This pattern is supportive of the idea that, at least within a given domain, individuals tend to reason using the structures of no more than two adjacent stages, and that reasoning tends to consolidate at a given full stage before any reasoning at the next stage appears.

The combination of good case and item fit is evidence that would support the construct validity of any instrument, in the sense that it upholds the notion that a single, underlying dimension of ability is being assessed. In the present case, however, construct validity is additionally supported by evidence for two important postulates of developmental stage theory, that of invariant sequence and the structured whole criterion. These findings are consistent with those reported in other analyses of instruments based on Piagetian principles (Dawson, 1997, 1996; Dawson, Goodheart, Draney, Wilson, & Commons, 1997; Draney, 1996; Goodheart et al., 1996a, 1996b; Müller et al., 1994, 1995; Noelting et al., 1994, 1995).

**Adult Development**

Three questions related to adult development were addressed: (1) What is the relationship between age and stage attainment in adulthood, and are the relationships homogeneous in different age groups? (2) Are there stages that occur only in adulthood? (3) Does structural developmental change occur in adulthood and, if so, what can be said to influence it?

*Adulthood Stages.* The 1985 and 1989 analyses confirmed earlier outcomes about the existence of adult stages (Armon, 1984a, 1984b). No subject attained a stage 4 score prior to age 24, and no subject attained a stage 5 score prior to age 34, which confirmed hypothesis 2 that these stages occur only in adulthood, if they occur at all, even with a “privileged” sample.

*Structural Developmental Change in Adulthood.* The largest structural change during any of the three 4-year periods was three quarters of a stage, demonstrated by only two people (6%). Seven subjects (36%) developed one half stage, while 15 (47%) developed one fourth stage (equivalent to measurement error), and 8 subjects (27%) showed no development at any of the test times. The most change over the whole study period occurred with a single subject and was one full stage. Eight subjects demonstrated no change; 10 demonstrated one quarter stage (not beyond measurement error), 7 one half stage, and 5 three quarters of a stage.

*Age.* Age was not significantly related to the development of reasoning about the Good Life in adulthood when all adults were combined into one group \( r = .127 \). There was a significant, linear relationship between the two in the younger adult group, however \( r = .35, F = 11.10, p < .001 \).
Education. While neither income nor previous stage were found to be significantly related to adult structural change, years of education were significantly related in both adult groups ($r = .60, F = 33.36, p < .00$). In addition, those who continued in school, or returned to school during the study period, had a significantly higher probability of an increase in their stage scores than those who terminated their education, or did not return to school during the study period ($r = .42, F = 5.82, p < .02$). It is important to note, however, that although increased education affected many subjects’ scores, it was not necessary for stage change in all cases. A few subjects developed without it, while a few others had it and did not advance in stage.

Occupational Status. Hollingshead Occupational Scale scores (Hollingshead, 1965) were significantly correlated with scores on reasoning about the Good Life ($r = .59, F(1,100) = 54.47, EMS = 22.24, HOL + 218.39, p < .001$). (They also accounted for about 5% of the variance over education scores when run on the nonstandardized residuals from education X EMS: $r = .23, F(1,100) = 5.47, RESID$ of EMS by ED $7.45, HOL - 54.42, p < .05$.)

Gender. The overall effect of gender on adult scores was significant, with men attaining an average of almost a half-stage above the women ($r = .41$ 2.40, EMS 45.51 MALE + 319.11, $p < .001$). In a multiple regression with education and occupational status, the salience of gender is reduced by remains significant ($r = .71, F(3,98) = 32.74, p < .001$, EMS = [5.96 HOL + 38.52 GEN + 146.09], tED = 2.86, p < .01; tHOL = 2.59, p < .05; tGEN = 4.74, p < .01).

DISCUSSION

The longitudinal data reported here further validate the invariant sequence stage model of value reasoning about the Good Life. Many subjects, including adults, demonstrated development and, when they did, it was always toward the next stage in the sequence. The Good Life Scoring Manual, developed in 1984, continues to demonstrate high interrater and test–retest reliability. Newer forms of analyses, for example, Rasch, also continue support for the model. The general findings of this study provide robust support for a structural–developmental model of value reasoning as well as some important findings on adult structural development.

Related Variables

Although age variables are useful in providing evidence for the developmental hypotheses of the Good Life Stage model, they provide only indirect evidence. Virtually all studies with developmental outcomes variables show high correlations between age and stage in children, particularly with middle-class samples. Indeed, at first glance, the mere passage of time appears to be salient in individual development, at least to stage 3—the stage of conventional (socially accepted) reasoning about the person, friendship, marriage, morality, and the self. (See, e.g., Colby et al., 1983; Damon & Hart, 1990; Fowler, 1981; Kegan, 1982; Selman, 1980.)
The degree of correspondence between age and stage, however, was observed to decrease to insignificance throughout adulthood.

Although adults’ age and stage scores were not significantly correlated, only adults had reached the two highest stages. Hence, although advancement in age is not a sure indicator of development to higher stages, it appears to be a necessary condition. (See Adult Development, below.)

With this sample, income did not have a significant impact on development. I suspect, however, that the average income was too high to reveal differences. If sufficient income provides for a relatively comfortable lifestyle and continuing educational opportunities, income may cease to have relevance for development.

Colby et al. (1983) reported that in middle-class samples of older adults with average to high IQs, neither income, previous stage, nor age appeared to influence structural change in adulthood. With subjects beyond the age of 30, only the impact of education has been examined and has proved to be modestly significant. The findings reported here are consistent with previous studies that have consistently demonstrated the positive impact of education on stage attainment (e.g., Colby et al., 1983; DeVos, 1983; Walker, 1986; see also Armon & Dawson, 1997). Further research is needed to investigate the relationship between specific forms of education—the actual content and activities—and structural—developmental change.

**Philosophies of the Good Life.** In 1984 it was demonstrated that nonphilosopher adults think much like “expert” philosophers when contemplating the Good Life. Although the Philosophical Orientations were shown to be reliable then, they proved insignificant in terms of their impact on stage development. For example, the Perfectionism Orientation was not shown to be significantly related to stage change scores as was hypothesized in 1984. These findings may be inconclusive since there were too few subjects associated with some of the Orientations, particularly the subgroups of Perfectionism, to explicitly support or reject the hypothesis.

**Gender Differences.** Although statistically significant gender differences were found in good life reasoning stage attainment in the adult group, they had not been found in the younger age groups. This may indicate complex relationships between historical periods and gender role socialization. Alternatively, it may indicate that women’s development is truncated in the adult years. Male and female subjects were approximately equal in their age, educational attainment, occupational status, and socioeconomic status. Thus, the gender differences cannot be reduced to differences in these variables (as has been the case with moral judgment [e.g., Walker, 1984]). Only additional studies with larger samples of males and females can expand our understanding. Previously (Armon, 1984a), no gender differences were found in the distribution of Philosophical Orientations in any age group, which may indicate some interesting directions for future work in gender studies.

**Adult Development**

The findings presented here should encourage researchers to take structural development in adulthood seriously. More research, explicitly designed to study adult structural development needs to be done. If the findings in this study are
replicated with different populations, there are many important implications. For example, we may need to think about including structural–developmental information in the construction of adult education, training programs, and psychological intervention methods in the same ways that we do for child and adolescent populations.

More germane to this study, however, is the issue of why only adults attain higher stages, and only some adults at that. We need to move beyond the question of whether adults develop, and ask the question, “How and why do adults develop?” Why is age a necessary condition for higher stage development? What are the necessary ingredients of continued change? Must one have a life of responsibility and role-taking as Kohlberg (1984) claimed? The antecedents of adult development are almost completely unknown, as are its variable forms. Even in literature where the notion of structural development in adulthood is accepted, it is conceived of as either hierarchical or absent. This dichotomy is unlikely. Other forms of development, such as decalage and consolidation, need to be investigated and distinguished from one another.

As researchers, we need to go beyond delineating static stage models and begin investigating variables that may impact adult development. A few small studies (mostly single-case) have suggested that early religious training and/or confrontation with challenging value conflicts are salient. In Colby and Damon’s (1992) case studies of moral exemplars, or Kohlberg’s (1981) analysis of Andrea Simpson, it seems that some people make growth opportunities out of circumstances they just “fall into,” while others pass by such opportunities unaware.

Although only a few theorists are examining the relationships between personality or temperament and life experiences, findings are promising. For example, Blasi (1980) discusses the possibility of certain self-concept variables having influences on moral development; Kitwood (1988) argues that early psychological wounds inhibit development; and certain temperament variables (such as extroversion) appear to be related to successful and unsuccessful experiences in childhood (Cowen, Wyman, Work, & Parker, 1990; Masten, Best, & Garmezy, 1990). Future individual case analyses of subjects from this study may provide relevant information on some of these important issues. (See Armon, 1995, for an example.)

Limitations of this Study

A major limitation of this study results from the size and nature of the sample, and its lack of generalizability to typical populations. While the sample was chosen purposefully to enhance the probability of observing adult development, it simultaneously limits how one can use the findings. Another cause for concern is the exclusive use of Western culture subjects, along with the corresponding reliance on Western philosophy for the formation of the Philosophical Orientations. The impact of this limitation cannot be tested without studying other populations. The Good Life Stage model, however, does not depend on the universalization of Western liberalism. On the contrary, one would expect if the study were to be carried out in the East, for example, other philosophies of the Good Life might be found (Zen Buddhist, for example). It has only been claimed here that there may be a finite number of Philosophical Orientations, not
that this study has discovered them all. Moreover, although it is expected that different cultures might provide alternative Philosophical Orientations, it is not expected that the structural aspects of value reasoning would be significantly different.

Another shortcoming of this study is that reasoning about the Good Life says little about real-life behavior. Construct validity, however, can be maintained by demonstrating that the interview and scoring scheme provide a valid assessment of value reasoning about the Good Life, or of value reasoning stage, rather than a valid assessment of individuals’ behavior or actual “good lives” (Colby et al., 1983). High correlations between stage attainment and Hollingshead occupational scores indicate that, on average, subjects at higher stages perform occupations with greater complexity and, possibly, social responsibility. To find out the extent to which they actually pursue the good lives they describe, however, would require in-depth case study methods.

In general, future research on structural development, particularly during adulthood, must maintain a rigorous theoretical and methodological approach so as not to fall prey to death by criticism. Moreover, the ethical dimension in social development research needs to be more clearly acknowledged. Without clearly specified constructs and a rigorous methodology to distinguish, for example, content from structure, or valuing from cognition, specific issues in human development will remain elusive. This is particularly true of the task of identifying the processes and causes of transition or development from one stage to another, which has eluded researchers for some time.

Implications for Education

The ways in which people decide what they value is of central importance to education. In the last decade, interest in value education has resurfaced. New emphasis is being placed on the role of the school and university in the formation of character, particularly since it is now recognized that all schools impact the character of its students, whether they do so planfully or not. In addition to issues of justice and fairness, character development includes reflection and action on issues of responsibility, general concern for others, and self respect. Such reasoning consists in part of value reasoning about the Good Life.

In particular, value reasoning about the Good Life in the moral domain identifies the moral good of responsibility, of persons, of the self, and even of justice (Rawls, 1971). For example, a principle of benevolence (Frankena, 1973) falls within the domain of the Good rather than the realm of the Right (justice, fairness, rights). In general, many real-life judgments concern neither physics (Piaget) nor justice (Kohlberg); rather, they are judgments of the Good Life—of values, aims, and ideals.

The developmental processes of value reasoning outlined here can inform a philosophy of education that goes beyond “values clarification” to identify reasonable and justifiable standards for values-guided education. That reasoning about the Good Life develops through a predictable sequence of stages allows educators to design curricula with developmentally appropriate methods and objectives. Moreover, our larger study, and others like it, make clear that young
students have their own concepts of the Good Life that are neither an incomplete notion of adult ideals nor simply a replication of their particular environment. These students need to be encouraged to articulate their ideals of the Good Life, to learn about those of others, and to question the meaning and outcomes of various good life concepts. These practices will serve to enhance students’ ethical thinking, encourage perspective-taking and mutual respect, and provide essential opportunities for reflection and critical analysis.

The findings of this study show that these issues should not be restricted to traditional elementary, high school, or college populations. It was demonstrated here not only that persons can develop their thinking about values throughout the lifespan, but also that postconventional stages of value reasoning do not emerge before adulthood. Increasing evidence supports these findings (see Kohlberg, 1986; Lam, 1994; also see reviews by Alexander et al., 1990; Richards et al., 1984). Moreover, it was shown here that development to those stages can be significantly influenced by educational experience. And, finally, today, more than one third of all university students are adults beyond the age of 25.

Unfortunately, typical adult-oriented education programs—whether they be pragmatic programs for continuing education, rehabilitation, or those of prestigious graduate schools—do not see values development as central to their endeavors. Rather, because it is assumed that the character of the adult student is already set, these programs are restricted to the accumulation of skills and knowledge and, occasionally, to the advancement of critical thinking. Findings of this and other studies, however, demonstrate the inadequacy of this view. Indeed, it appears that educational experience in adulthood can be extremely important in the development of autonomy and principled ethical reasoning. Philosophers of education, psychologists, and educators need to recognize that particular forms of educational experience are probably key factors in the creation of critically reflective, ethical adults who can participate effectively in our increasingly complex and often threatened civilization.

Although a general model of value reasoning about the Good Life will tell only a part of the story of human valuing, it nevertheless goes beyond and is to be preferred to the value relativism and subjectivism so prevalent in contemporary society, generally, and in psychology in particular. From homelessness to adolescent homicide, contemporary social problems are in part, a consequence of adult value reasoning. A stage model of value reasoning about the Good Life can inform our understanding as to some of the origins of such problems and contribute to education and intervention models that attempt to address them.

ACKNOWLEDGMENTS. 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 for sharing their thoughts with us over thirteen years. We are indebted to our interviewers Joyce Friedman, Larry Smiley, Lura Beth Illig, and Sandra Nahan who sometimes traveled thousands of miles to unexpected places to meet with participants. We thank Michael Commons and Bill Puka for valuable input on earlier versions of this paper. We also acknowledge Paul Holland and Michael Commons for their assistance with recent statistical analyses.
REFERENCES


moral development. Monograph of the society for research in child development, 48(4).

Adult development, Vol II: Models and methods in the study of adolescent and adult thought.
New York: Praeger.

Bauer, N. F. (1989). A multidomain study of adult development. In M. L. Commons, J. D. Sinnott,
F. A. Richards, & C. Armon (Eds.), Adult development, Vol. 1: Comparisons and applications of

Commons, M. L., & Richards, F. A. (1984). A general model of stage theory. In M. L. Commons,
F. A. Richards, & C. Armon (Eds.), Beyond formal operations: Late adolescent and adult cognitive
development (pp. 141–157). New York: Praeger.


34, 573–578.


DeVos, E. (1983). Socioeconomic influences on moral reasoning: A structural-developmental perspec-


developmental differences. Unpublished doctoral dissertation, University of California at
Berkeley, Berkeley, CA.


Ennis, (1978). Conceptualization of children’s logical competence: Piaget’s propositional logic and an
alternate proposal. In L. S. Seigel & C. J. Brainard (Eds.), Alternatives to Piaget: Critical essays on
the theory (pp. ). New York: Academic Press.

good, and the beautiful. The eighth annual proceedings of Piaget and the helping professions.
Los Angeles: University of Southern California Press.


Chicago: University of Chicago Press.


A Longitudinal Study of Adult Value Reasoning


