University of Rhode Island

DigitalCommons@URI

Open Access Master's Theses

1982

Review and Evaluation of Two Models of Moral Development

Kevin Plummer

University of Rhode Island

Follow this and additional works at: <https://digitalcommons.uri.edu/theses>

Recommended Citation

Plummer, Kevin, "Review and Evaluation of Two Models of Moral Development" (1982). *Open Access Master's Theses*. Paper 1716.

<https://digitalcommons.uri.edu/theses/1716>

This Thesis is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Master's Theses by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu.

REVIEW AND EVALUATION OF
TWO MODELS OF
MORAL DEVELOPMENT

BY
KEVIN PLUMMER

A THESIS SUBMITTED IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF ARTS
IN
SCHOOL PSYCHOLOGY

UNIVERSITY OF RHODE ISLAND

1982

ABSTRACT

Moral development has been conceptualized in terms of simple and complex stage models. This study was aimed at comparing and evaluating these two models to determine which most adequately addresses the construct. Seventy college undergraduates were administered the Defining Issues Test (DIT) as a measure of moral reasoning. An analysis of their DIT protocols for response variation showed that a full range of reasoning was used to address each moral situation, supporting a complex stage orientation. This study was also designed to explore personal/situational factors which contribute to this variability in reasoning. Fifteen of the original 70 subjects were used for a follow-up interview. This revealed that severity of story consequences and personal relevance of the story theme were significant contributors to response variation. A complex stage orientation seems to be the most adequate way to conceptualize moral development, due to the degree of stage mixture noted in this study. This study revealed, however, that numerous methodological considerations need to be resolved before we can implement the complex stage model in the assessment of moral reasoning and design of moral education programs.

TABLE OF CONTENTS

<u>Chapter</u>	<u>Page</u>
I. Statement of the Problem.	1
II. Structural Theory	1
Structured whole	2
Invariant sequence	2
Hierarchical development	3
Theoretical conclusions.	4
III. Kohlberg's Structural Theory.	4
Six stages of reasoning.	5
The simple stage model	7
Theoretical conclusions.	10
IV. Theoretical Critique.	11
Naturalism	11
Invariant sequence	12
Structured whole	14
V. Methodological Considerations	15
VI. Theoretical Proposal.	17
Content based structuralism.	18
Complex stage model.	19
VII. The Purpose of the Study.	22
VIII. Moral Judgment and Abstract Reasoning	23
Identifying post conventional reasoners.	23
Moral reasoning and cognitive development.	24
IX. Research Questions.	28
X. Method.	29
Subjects	29
Measures	29
Similarities.	29
General Information	30
Defining Issues Test.	31
Structured Interview.	33
Procedure.	35
Inter-scorer Reliability.	35
XI. Results and Discussion.	37
Correlational Data	37
Zero Order Correlations	37
Partial and Semi-partial Correlations	38

Simple and Complex Stage Models.	43
Reasoning Across Situations	43
Reasoning Within Situations	46
Post Conventional Reasoning	48
Structured Interview, Results.	50
Structured Interview, General Discussion	59
Personal/Situational Factors.	59
Methodological Considerations	61
XII. Summary	66
APPENDICES	
A Similarities.	70
B General Information	71
C Defining Issues Test.	72
REFERENCES	86

LIST OF TABLES

<u>Table</u>		<u>Page</u>
1	Descriptive Statistics For The Reliability and Total Sample.	36
2	Zero Order Correlations Between Dependent Measures P% and T score and Independent Measures Similarities and General Information	39
3	Partial and Semi-Partial Correlations Between Dependent Measures P% and T score and Independent Measures Similarities and General Information	39
4	Stages Spanned When Reasoning Across Dilemmas	44
5	Stages Spanned When Reasoning Within Dilemmas, Considering All Four Choices.	46
6	Stages Spanned When Reasoning Within Dilemmas, Considering Only Two Choices.	47

LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
I	A Graphic Presentation Of The Simple Stage Model.	8
II	The Complex Stage Model. A Disjunctive Scale Formed By The Relative Usage Of Six Different Stages Across Development	20

Lawrence Kohlberg (1958) has proposed a structural developmental theory to account for moral development from childhood through adulthood. It is largely based on Piaget's model of moral development, however, Kohlberg claims that his theory is more adequate because (a) it accounts for life span development rather than stopping at early adolescence, and (b) he takes a stronger stance regarding the stage-like nature of his moral judgment sequence (Damon, 1980). Rest (1979) has developed a theory of moral development which is based largely on Kohlberg's paradigm. He extends Kohlberg's "simple stage" approach, however, to include more complex stage mixture.

The following review of the literature will examine Kohlberg's structural theory (i.e., his model of stage development). In addition, Kohlberg's "simple stage" model will be compared to Rest's "complex stage" orientation.

This study will document the degree of stage mixture across and within subjects. A large degree of stage mixture will indicate support for a complex stage model. In addition, this study will explore possible factors which could account for stage mixture. Also, evidence of post conventional reasoning will be documented in terms of a complex stage orientation, and methodological considerations in using Rest's method of assessment will be examined.

Structural Theory

Structuralism is a theoretical orientation which conceptualizes development in terms of an invariant sequence of discrete and

discontinuous stages or levels (Flavell, 1963). Each stage is internally consistent and qualitatively different from all other stages, and development is hierarchical as each successive stage in the sequence logically subsumes all lower stages (Rest, 1973). In addition, Gibbs (1977) suggests that successive stages of development are more adaptive with respect to human functioning, thus we should expect similar developmental patterns across cultures.

Structured Whole

Classifying development in terms of stages or levels is a means of organizing behavior according to discrete categories where each stage is internally consistent in terms of its unique logic. The logic within stages is internally consistent and qualitatively different from all other stages. This means that one's approach or response to various situations should reflect an orientation which is unified under a set of principles or rules for that stage. This response generalization, or unification of responses across diverse situations is referred to by Piaget as "structure d'ensemble" (Flavell, 1963). Liebert (1978) suggests that,

The concept of stages or levels of moral development implies that at any particular time in an individual's life one can find an organized system of feelings and beliefs that direct the individual's moral thought and action so as to produce similar responses to diverse situations (p. 8).

Invariant Sequence

Developmental progress, the move from one stage to the next, is marked by a consecutive, gradual, upward movement through the stage

3

sequence. Each major reorganization in the course of development represents a new stage. The sequence, or particular order of stage acquisition, is invariant as development proceeds one stage at a time, no stages are skipped, and there is resistance to extinction or regression (Gibbs, 1977).

Hierarchical Development

Each succeeding stage in the sequence represents an advance over previous stages as all the elements of the old stage are transformed and reorganized with new elements to form a more differentiated and integrated structure. The higher stage is more complex and adaptive as individuals are capable of using all lower stages as well as their present stage (Rest, 1973). The adaptive nature of development suggests that similar developmental patterns should be prevalent among members of the human species regardless of culture or other environmental factors (Gibbs, 1977). Developmental progress may be a function of culture or other environmental influences but the basic pattern of development should be consistent within the species (naturalism) Liebert (1978) suggests,

Just as well-watered and properly cared for plant will bear flowers and fruit sooner than one that is left to mature willy-nilly, so (according to structural theory) children may advance more or less quickly to mature thought depending on the degree to which their environments nurture cognitive and moral growth. But the tendency to blossom and to display distinctive characteristics in a fixed metamorphosis is to be found within the nature of the biological unit and cannot be said to arise from or be caused by the environment (p. 15).

Theoretical Conclusions

According to structural theory we should expect to discover similar developmental patterns among all members of the species, while specific developmental progress remains a function of cultural or other environmental factors. This developmental pattern is organized in terms of an invariant sequence of stages, where each stage is organized according to its own logic. Thus, individuals occupying the same stage should exhibit similar types of reasoning and respond consistently across a variety of situations. Each successive stage in this sequence logically presupposes all previous stages and represents a more advanced form of development.

Kohlberg has attempted to describe his model of moral development in terms of the structural approach reviewed above. The following discussion is a brief review of the major aspects of his structural theory.

Kohlberg's Structural Theory

Kohlberg's model covers the range of moral development from childhood through adulthood. His first level of development (stages one and two) is very similar to Piaget's description of early development, which includes children from ages 4 or 5 up to about 8 or 9. Kohlberg's later stages describe development beyond Piaget's highest stage, which extends only through early adolescence. Kohlberg believes that Piaget's stages are inadequately formulated, thus he has subdivided and reorganized them to be better representations of true stages, and he has extended them to include more advanced development

through adulthood (Damon, 1980). Thus, Kohlberg claims that his theory not only subsumes Piaget's stages of reasoning, but represents a reformulated and extended version to comply more rigidly with the structural developmental approach and cover life span development.

Six Stages Of Reasoning

Kohlberg initially formulated his six stage model based on a study of boys ages 10, 13, and 16 (Kohlberg, 1958). He included adults in later studies and conducted a series of longitudinal studies on his original group which has led to many theoretical revisions. The six stages are organized in terms of three levels; preconventional, conventional, and post conventional.

Individuals at the preconventional level of development do not understand or respond to the rules or expectations of society. Rather, preconventional reasoners interpret rules on a literal level in deference to an authority figure such as a parent. Right and wrong is strictly determined by whether one has obeyed or disobeyed these rules regardless of intent or the motives involved. Reasons for behavior at this level include self-interest, avoiding punishment, and deference to authority (Kohlberg, L., Colby, A., Gibbs, J., & Speicher-Dubin, B., 1976).

The preconventional level is subdivided into stages one and two. Stage one reasoners obey rules for the primary purpose of avoiding aversive consequences, and stage two reasoners operate for personal gain (Kohlberg, et al, 1976). Kohlberg (1976) claims that most children under age 9 are at the preconventional level, as well as some adolescents and adult criminal offenders.

Most adolescents and adults, however, reason according to conventional morality (Kohlberg, et al, 1976). Conventional level reasoners (stages three and four) define right in terms of the rules, roles and expectations of society or smaller groups such as religious and political organizations. The main distinction between stages at the conventional level is that stage three reasoners apply their conventional thinking to interpersonal situations and stage four individuals respond to the entire social order (Damon, 1980). Stage three reasoners, for example, conform to standards designed by others in pursuit of approval from these authority figures. Stage four reasoners, however, may conform to social norms in order to gain acceptance as good citizens. Damon (1980) adds, "Justice as stage four, thus, becomes establishing good citizenship, working hard, and maintaining the law of the land" (p. 42).

Post conventional reasoning (stages five and six) is the highest level in Kohlberg's paradigm. These individuals define right in terms of universal human rights, values and principles. They believe that it is usually right to uphold the law, however, violations of the law are justified when the law is not protecting human rights. In cases where principles come into conflict with society's rules the post conventional individual judges by principles rather than convention.

Stage five reasoners, for example, conceive of morality in terms of a social contract, "he or she conceives of moral responsibility as binding upon all those who claim the rights of society" (Damon, 1980, p. 42). Individuals feel obligated to obey the law because they have created a social contract to make and abide by laws for the good of all, to protect their own rights and the rights of others. According

to Kohlberg, et al, (1976) a social contract is equivalent to "... the notion that by living in society you have made a generalized commitment to respect and uphold the rights of others (and the laws this entails)" (p. 13).

Stage six reasoners are guided by self chosen ethical principles which are universal for all humanity.

Particular laws or social agreements are usually valid because they are based on such principles. When laws violate these principles one acts in accordance with the principle. Principles are universal principles of justice: The equality of human rights and respect for the dignity of human beings as individual persons (Kohlberg, et al, 1976, p. 20).

Kohlberg believes that his six stage model of moral development fits well within the structural developmental approach outlined earlier (Damon, 1980). His theoretical notions of development can best be understood in terms of the simple stage model (Rest, 1979).

The Simple Stage Model

Figure 1 (see page 8) is a graphic presentation of the simple stage model representing the major theoretical aspects of development according to Kohlberg (Rest, 1979).

Kohlberg suggests that individuals pass through alternate periods of transition and consolidation in the course of acquiring new modes of reasoning. Reasoners develop from a point of using one stage exclusively to using reasoning one stage above their dominant stage (+1) at increasing levels of frequency. The use of the +1 stage continues to increase until the former stage is dropped completely and again one stage of reasoning is used exclusively. "Higher stages displace the structures found at lower stages" (Kohlberg, 1967, p. 32).

FIGURE 1. A GRAPHIC PRESENTATION OF THE SIMPLE STAGE MODEL
(TAKEN FROM REST, 1979, P. 52.)

Thus, at times individuals utilize one stage exclusively after completely discarding the previous stage, and the +1 stage is still out of reach. This is a period of maximum consolidation and no further development of that stage occurs (Rest, 1979). Each stage peaks at 100% usage and has a turn at predominance in use over the other stages. In addition, the consolidation of successive stages is sequential.

Stage theory holds that every single individual, studied longitudinally, should move only one step at a time through the stage sequence and always in the same order (p. 39). Stages imply distinct or qualitative differences in structure and the different structures form an invariant sequence (p. 32). (Kohlberg, et al, 1976)

The previous stage (-1) falls to 0% usage before the higher stage (+1) is ever used. "Individuals should be consistently at a stage unless they are in transition to the next stage" (Kohlberg, 1976, p. 47). This means that stage mixture is theoretically possible only between two adjacent stages (e.g., 2-3 and 3-4 is possible but 2-3-4 or 2-4 mixture is not possible). Individuals respond in a way characteristic of their stage, not in a manner more primitive or more mature (Bearison, 1974). "There is a disposition to prefer a solution of a problem at the highest level available to him" (Kohlberg, 1976, p. 32). Kohlberg (1973) also adds:

The stages form a clustered whole. There is a general factor of moral stage cross cutting all dilemmas, verbal or behavioral (p. 186).

Each of these different and sequential modes of thought forms a structured whole - a given stage response on a task does not just represent a specific response rather it represents an underlying thought organization (Kohlberg, et al, 1976, p. 32).

Thus, Kohlberg's description of development, in terms of transition and consolidation addresses the notion of invariant sequence, as well as structure d'ensemble.

Theoretical Conclusions

According to Kohlberg, stages of moral development are organized according to their own logic. They are internally consistent and qualitatively different from each other. Individuals use just one stage of moral reasoning across most situations unless they are in transition, which involves the use of +1 reasoning. Developmental progress consists of sequential consolidation of successive stages. This sequence is hierarchical and invariant, thus lower levels of moral reasoning are no longer used once higher levels have been attained. Kohlberg has also offered evidence of naturalism by citing examples of all his stages in other cultures. Furthermore, he claims that the same invariant sequence has been documented cross-culturally (Kohlberg, 1969).

Kohlberg's paradigm is a rigid application of structural developmental theory. Many researchers have attempted to test Kohlberg's model with a major emphasis on examining methodological and social psychological weaknesses. A brief review of this literature will follow. However, the most relevant issues for this study are the theoretical underpinnings of Kohlberg's model. Thus major theoretical notions such as naturalism, invariant sequence, and structured whole are examined at this point.

Theoretical Critique

Many researchers have reviewed the structural aspects of Kohlberg's theory (Edwards, 1978; Gibbs, 1977; Holstein, 1976; Kohlberg, 1968; Kohlberg, 1969; Kohlberg & Kramer, 1969; Kuhn (1976); Kurtines & Grief, 1974; McGeorge, 1974; Siegal, 1980). This has included an examination of naturalism, structured whole, and Kohlberg's invariant sequence of development.

Naturalism

Kohlberg (1968) cites cross-cultural evidence for his six stages of moral reasoning as well as his sequence of development. He claims to have discovered universal moral principles as his theory of development pertains to the human species in general. Kohlberg bases these claims largely on some unpublished work which he refers to in a 1968 popular magazine article. He collected data in America, Taiwan, Mexico, Turkey, and Yucatan which showed that 7% of the 16-year-olds in America and Mexico used stage six reasoning and 1% or less of a comparable Taiwan sample reasoned at this stage. None of the children in either Turkey or Yucatan were able to reach even stage five. Thus, stage five is missing in two of the five samples and stage six is absent in three of the samples. According to Kurtines and Grief (1974) "age trends in stage five and six are clearly present only in the United States sample -- the same group Kohlberg (1958) used to derive the stages" (p. 461). They conclude that there is no evidence to support Kohlberg's claim that the course of moral development is universal.

Siegal (1980) also reviewed cross-cultural research relevant to Kohlberg's paradigm. He concluded that stage five and six reasoning is generally attained only in western societies, and much of that post conventional reasoning can be attributed to scoring error.

The lack of evidence to support the development of post conventional reasoning in members of the same species across cultures renders Kohlberg's claim of universality somewhat questionable. At best, no evidence has been shown to document naturalism with respect to Kohlberg's theory of development. In addition, Siegal (1980) concludes that without a sufficiently large sample of persons who reason at stages five and six there can be no empirical support upon which to base an invariant six stage sequence.

Invariant Sequence

The strongest support for an invariant stage sequence is usually based on longitudinal research. Three major longitudinal studies have attempted to validate Kohlberg's sequence of development.

Kohlberg and Kramer (1969), in their only published longitudinal follow up, found that the subjects from Kohlberg's 1958 sample showed little systematic change in moral reasoning over time, and 20% regressed from stages four and five to stage two. Kohlberg explained this regression by inventing a new stage which he labeled 4B. This stage was designed to address the identity crisis which high school graduates encounter as they enter college.

Holstein (1976) conducted a three year longitudinal study, collecting data from parents, as well as their sons and daughters who

were 13-years-old at the first assessment. Holstein found no evidence of the stepwise progression described by Kohlberg's theory, and many of her subjects regressed from higher to lower stages across the three year period. Also, adults were just as likely to regress as adolescents. Even when Holstein used the new controversial stage 4B there was still 25% regression from stages four, five and six to stages one, two, and three.

Kuhn (1976) conducted a one year longitudinal study of 50 5-8 year olds with assessment at six month intervals. The first assessment indicated that equal numbers of subjects regressed and progressed slightly. The second assessment reported similar results. Overall results from the one year period showed that a total of 32 subjects progressed slightly and five subjects regressed. Although only five subjects showed regression overall, almost every subject showed both progression and regression at some point during the study, with amount progressed slightly more than amount regressed.

Kuhn (1976) and Kohlberg and Kramer (1969) claim that measurement error can account for the minor regressions noted in both of their studies. I suggest that if measurement error can account for slight regressions then it must also be credited with the slight progression (e.g., none of Kuhn's subjects progressed an entire stage).

Measurement error in Kuhn's study could have been a major problem due to her inappropriate use of global scoring. Global scoring is a gross estimate of the individual's stage of reasoning, insensitive to small changes in stage usage, thus it is most inappropriate for a one year longitudinal study. It is possible that considerably more regression occurred in Kuhn's study but it was not documented due to

the insensitivity of global scoring. The validity of Kuhn's study is questionable. At best, this study provides no evidence for Kohlberg's sequence of moral development.

Kohlberg's six stage sequence lacks empirical support. Regression was found in all the longitudinal research and no empirical evidence has been provided for stages five and six. Many researchers can document Kohlberg's early stages but cannot find support for a six stage sequence.

Regression not only represents a violation of Kohlberg's invariant sequence, it also violates his sense of structured whole because the same subjects are using reasoning from a wide variety of lower stages.

Structured Whole

Kohlberg claims that all of his stages are qualitatively different modes of thought and individuals are expected to respond in a way which is characteristic of their stage, not in some manner more primitive or more mature. However, Kohlberg reports that stage responses across all nine dilemmas correlate .31 to .75, with a mean correlation of .51. Sanstrock (1975) reports correlations ranging from .19 to .48 with a mean of .33. He found that subjects did not respond consistently across dilemmas, in terms of their stage usage, and concluded that there is more situational variability in moral judgment than Kohlberg assumes.

McGeorge (1974) noted significant variation between dilemmas in the responses of 40 12-year-old boys and 23 university students (range of correlation = .00 to .33). McGeorge suggests that the dilemmas are not pure measures of a single aspect of morality as Kohlberg believes.

The lack of internal consistency across dilemmas is another violation of Kohlberg's structural theory (structured whole). Other violations have also been noted. Regression in longitudinal research, for example, is an apparent violation of Kohlberg's invariant sequence. In addition, lack of empirical support for post conventional reasoning in cross-cultural studies is a violation of the six stage model as well as naturalism. In addition to these theoretical concerns, researchers have examined a variety of methodological considerations with respect to Kohlberg's paradigm.

Methodological Considerations

The present study consisted primarily of a theoretical examination, contrasting a simple stage versus complex stage approach to understanding and organizing moral development. Thus, the major emphasis is on a review of the theoretical literature. However, numerous studies have addressed the Kohlbergian methodology used to gain support for his theoretical notions. Therefore, a brief review of this literature follows.

Some weaknesses in methodology are evident in the process and conditions of test administration. When a dilemma is presented to a subject and the subject is having a difficult time making a clear judgment, the examiner asks probing questions (Kurtines & Grief, 1974). The fact that the examiner probes for more information implies an inadequate answer. The subject may alter or change his/her answer completely in an attempt to satisfy the examiner. Also, the same probing questions are not used in all cases. Since this process has

not been standardized, it is doubtful whether the data can be conclusively attributed to one specific factor such as moral reasoning.

It has also been discovered that testers often find Kohlberg's method time consuming (Kurtines & Grief, 1974). As a result these testers have not used all the dilemmas in their assessment. This creates problems because not all dilemmas are equally effective for assessing moral reasoning, and researchers do not always specify which dilemmas they use when presenting their data. Since each dilemma is designed to assess a specific portion of the overall stage development one cannot assume that the dilemmas are interchangeable (Kurtines & Grief, 1974). Failure to consistently follow standardized assessment means that research results based on Kohlberg's model may have little generalizability and cannot be considered supportive of his theory.

There are also many social psychological factors which seriously question the validity of research results based on the Kohlbergian paradigm. It has been hypothesized that children respond differently to young, old, male and female interviewers (Kurtines & Grief, 1974). A subject may respond to a dilemma according to what the subject feels the interviewer wants to hear. The interviewer needs to decide, then, whether a subject is responding only to the dilemma or partially to personal characteristics of the interviewer. In my view, the examiners are not in a position to decide this issue because they are not even considering it as a possibility.

Assuming that the issues concerning test administration and interpretation are resolved, there are still problems regarding the issue of content validity. The main characters in all the dilemmas are male. Kurtines & Grief (1974) have suggested that this creates a role

expectation bias. Any assumptions about the roles of males or their expected behavior could influence the subject's judgments.

A subject can be affected by a dilemma in ways which Kohlberg's method of assessment does not measure. For example, the subject may lose interest or get bored when responding to unrealistic dilemmas (Turiel, 1966). The lack of immediate relevancy can cause one to rely on expedient reasoning, and moral reasoning can be less mature under "deindividuating" circumstances (Arbuthnot & Andrasik, 1973).

It is obvious that there are many social psychological factors which contribute to an invalid assessment of moral reasoning. These factors, however, are not often considered by the evaluator when an assessment of moral reasoning is conducted.

Methodological weaknesses such as unstandardized test administration have been discussed. Social psychological problems have also been considered in terms of experimenter, subject and test bias. This evidence is important when examining the past research on moral development which has used Kohlberg's model.

Theoretical Proposal

Many researchers have failed to find support for Kohlberg's version of structural developmental theory. Specific theoretical problems include the lack of evidence for his notion of invariant sequence, structured whole, and naturalism. Pure methodological and social psychological problems have also been discovered. Rather than abandoning the theory completely, I suggest that we re-examine

Kohlberg's paradigm and adapt his approach to account for the theoretical and methodological violations described above.

Content Based Structuralism

Kohlberg has attempted to describe the structural nature of moral reasoning as it develops in all individuals. The evidence indicates that a rigid application of structural theory to the development of reasoning about moral situations is inappropriate. Rest (1979) believes that "no pure direct assessment of cognitive structure exists that is unaffected by the specific task, content, and response characteristics of the situation" (p. 64). For example, an individual's reasoning about interpersonal situations may be more advanced than his/her reasoning about labor strikes, due to the personal relevance or the individual's experience in this area.

The principles of structural developmental theory, then are more appropriate once they have been extended to include variation due to situational factors. Subjects may still pass through an invariant sequence of development with respect to each content, but developmental progress does not necessarily have to be at the same point with respect to all content areas. A structured whole can also be established for each individual with respect to various contents. An individual may respond with the same type of reasoning to dilemmas of similar content but we should not expect this consistency across different contents. In addition, substantial evidence of post conventional reasoning should be easier to document as modal use of stages five and six (which is Kohlberg's standard for judging a person as a post conventional reasoner) is not a necessary condition for verification of higher level thought.

Rather than assuming that fixed cognitive structures exist within individuals, allowing them to act on all moral situations in the same way, I suggest that situational factors such as the content of the moral dilemma interact with the cognitive structures to create different types of reasoning in a variety of situations. Thus, an individual who is capable of post conventional reasoning in some instances may use a variety of lower level responses in other moral situations as development with respect to some contents is more advanced than others. Rest (1979) adds,

A dilemma about mercy killing may evoke different organizing structures than a dilemma about distribution of wages or civil disobedience. We should recognize that assessment is content and method specific. A future goal is to identify the various attributions of test situations that affect the structural organizations of thinking manifested and the extent to which each of these attributes affect them (p. 68).

Thus we should expect a full range of stage responses across dilemmas, not just the dominant and +1 stage as Kohlberg suggests. In order to find results predicted by Kohlberg's structural theory we would need to control content and other situational variables for each subject, or pretend that these things do not matter and dismiss the regression and other theoretical inconsistencies as measurement error.

Complex Stage Model

The complex stage model (see figure 2) illustrates the extension of structural theory to account for the different types of reasoning across situations. Rest (1979) suggests that development can be assessed in terms of probability. Subjects begin by using a type of reasoning only in certain instances and move towards solidifying that

FIGURE 2. THE COMPLEX STAGE MODEL. A DISJUNCTIVE SCALE FORMED BY THE RELATIVE USAGE OF SIX DIFFERENT STAGES ACROSS DEVELOPMENT. (TAKEN FROM REST, 1979, P. 224)

reasoning and applying it to a wider variety of situations. The probability of observing a particular type of reasoning in an individual is a function of the degree of solidification. Thus the notion that a subject is on a particular stage is wrong. Rather than wondering whether a subject is on a stage or not we should be more concerned with the type of situations or conditions which are likely to induce individuals to organize their thinking in certain ways.

Since it is difficult to consider a subject at a particular stage, the complex stage model refers to development as the increasing probability of using higher stages of reasoning (Rest, 1979). As illustrated in figure 2, it is possible to advance in several organizations of thinking simultaneously (e.g., moving to advanced levels of stage three, moderate levels of stage four, the earliest levels of stage five, and decreasing use of stage two). Rest (1979) suggests that quality and quantity are important considerations for the complex stage model while the simple stage model is concerned only with quality.

The quality is critical in the simple stage model. The issue in question is what stage is being used, not how much. The quantitative aspect is irrelevant as it is assumed that the individual will use this mode of thought all the time (p. 50).

Kohlberg (1973b) adds,

quantitative considerations are antithetical to an interest in cognitive structures. The structural theory does not treat any change as a change in structural competence unless the change is evident in a qualitatively new pattern of responses (p. 181).

Rest (1979) argues that we need quantitative as well as qualitative descriptors.

We need qualitative descriptors to represent the different organizational patterns and we need quantitative descriptors to represent the degree to which a particular subject is manifesting one or another of those patterns (Rest, 1979, p. 54).

One implication of the complex stage model, then, is that when subjects become capable of higher level thought the lower reasoning is not totally abandoned (as proposed by the simple stage model). Subjects may prefer to use the higher stages as they become possible but still use lower stage reasoning in certain situations (Rest, 1979).

Development, as described by the complex stage model, is sequential as well as hierarchical. Also, the use of post conventional reasoning can be more widely found as it is not necessary to use this level exclusively before being credited with it, as Kohlberg's model assumes. Thus, all the general elements of structural developmental theory which were violated in Kohlberg's structuralism are maintained by accounting for development in terms of the complex stage model.

The Purpose of the Study

One purpose of this study was to find additional support for the complex stage model by noting the degree of stage mixture across and within situations. Another purpose of the study was to explore possible factors which could account for this variability in stage usage. In addition, I attempted to document substantial evidence of post conventional reasoning by using a complex stage orientation in the assessment of moral reasoning. I also attempted to gain evidence of post conventional thinking, as well as documentation of Kohlberg's six stage invariant sequence, by examining the correlation between moral

and abstract reasoning. Finally, this study was conducted to explore methodological considerations in using Rest's method of assessment. For example, Rest claimed to have developed a valid continuous index of moral reasoning (P%). This study was designed to examine that claim.

Moral Judgment and Abstract Reasoning

Many researchers have not been able to document Kohlberg's highest level of reasoning in their studies. Siegal (1980) concludes that there is no empirical support for stages five and six. Gibbs (1977) believes that development can only be documented through stage four. Kuhn (1976) and Holstein (1976) found longitudinal evidence of sequential development through stage three. Kurtines and Grief (1974) review Kohlberg's (1968) cross-cultural data and conclude that there is no evidence to support Kohlberg's claim that the course of moral development is universal, "age trends in stages five and six are clearly present only in the United States sample" (p. 461). While no one has been able to provide empirical support for post conventional reasoning, most researchers agree that there must be more advanced reasoning beyond the conventional level. Principles of justice and human rights are not reducible to stages three and four (Gibbs, 1977).

Identifying Post Conventional Reasoners

Rest (1979) describes some potential assessment problems created by using Kohlberg's interview technique, which may explain why this device identifies few post conventional reasoners.

The clinical interview may underestimate the structural competencies of the child. A child is credited with having a structure only if he can express it, explain it, justify it and sometimes argue against alternatives.... A person may be able to organize his or her actions and make discriminations using a structure without being able to talk about the structure (Rest, 1979, p. 60).

Development of a concept proceeds through a series of steps from preference to comprehension and finally verbal justification (Rest, 1979). Thus, individuals may be understanding and using post conventional reasoning before Kohlberg identifies this with his assessment device. Also, to be considered a post conventional reasoner, Kohlberg's assessment requires one to use this level predominantly throughout the assessment. Kohlberg has justified this in terms of the simple stage model (see figure 1). However, the complex stage model (see figure 2) describes development more adequately as most individuals do not use just one stage of reasoning (Rest, 1979). Thus Kohlberg's assessment device may be an inappropriate means of identifying post conventional reasoners. Therefore, we cannot assume that post conventional reasoning is as rare as Kohlberg's model suggests. "The post conventional level is reached by a minority of adults and is reached only after the age of 20-25" (Kohlberg, 1976, p. 48).

Moral Reasoning and Cognitive Development

Attempts have been made to demonstrate construct validity of Kohlberg's full six stage sequence by measuring the correlation between moral reasoning and other measures of cognitive development. Despite the assessment difficulties described above, a few researchers have

managed to show significant correlations between moral and general cognitive development.

Whiteman (1964) found that the ability to make more mature moral judgments increases with I.Q. across three age categories, 7-8, 9-10, 11-12. Each age category contained a full range of I.Q. scores (70-145) thus, Whiteman concluded that maturity of moral judgment is a function of increases in mental age rather than advances in chronological age.

Lee (1971) demonstrated that the structures involved in moral judgment developed concomitantly with cognitive structures in general, as measured by Piagetian tasks. Lee suggests that this is evidence for Piaget's assertion that changes in cognitive structure are essential for the development of moral judgment.

Tomlinson-Keasy (1974) examined the relationship between formal operations and principled moral reasoning, and concluded that there is a substantial predictable relationship between formal operations and principled level reasoning. More specifically, formal operational thought is a necessary, though not sufficient, condition for post conventional reasoning (Tomlinson-Keasy, 1974).

Kuhn, Langer, Kohlberg, & Haan, (1977) also studied the relationship between formal operational thought and post conventional reasoning. However, they also included I.Q. as a variable. Kuhn et al. concluded that formal operational thought is a necessary condition for the consolidation of conventional moral judgment. All post conventional reasoners had attained formal operational thought and the lower levels of moral judgment correlated with lower levels of performance on formal operational tasks. The full range of I.Q. scores occurred for all

groups, thus Kuhn et al. conclude that the relationship between levels of logical and moral development is not a reflection of their common relation to I.Q.

While there is some concentration of higher I.Q.'s at the higher logical and moral judgment stages there is wide variability in I.Q. at each logical-moral level. When the mental age variable is completely eliminated, in fact, the logical-moral correlation drops only slightly (Kuhn et al, 1977, p. 161).

Kohlberg (1979) adds "measures of moral judgment correlate with measures of moral attitudes, choices, and behavior to an extent not accounted for by I.Q. or other pure cognitive variables" (in Rest, 1979, p. xii). The relationship between I.Q. and moral judgment is moderate, since moral reasoning is one aspect of intellectual development, however, moral judgment is distinct from general intellectual development as correlations between logical and moral reasoning are higher than moral reasoning and I.Q. Kohlberg et al, (1976) comment further on this issue.

Since moral reasoning clearly is reasoning, advanced moral reasoning depends upon advanced logical reasoning. A person's logical stage puts a certain ceiling on the moral stage he can attain...logical development precedes moral development. Moral development depends upon intellectual development but intellectual development does not depend on moral development (pp. 5 and 6).

The correlations established between measures of cognitive development and moral development indicate that higher level moral reasoners are more advanced in terms of cognitive development. In fact, formal operational thought is a necessary condition for post conventional reasoning. Thus, higher level moral reasoners (stages 5 and 6) are distinct from lower level reasoners on a variety of measures other than moral reasoning. These distinctions suggest that post

conventional reasoning is a qualitatively different mode of thinking. Thus, the lack of empirical support for stages 5 and 6 is probably a function of the inability of Kohlberg's assessment device to identify this group, rather than evidence against its existence.

Correlational studies which attempt to document Kohlberg's sequence of development or distinguish post conventional reasoners in terms of other cognitive abilities have provided only limited evidence. This is primarily due to their use of Kohlberg's assessment device which is based on the simple stage model. The simple stage model limits the group of post conventional reasoners to individuals who use this level almost all the time. It also assumes that all post conventional reasoners are at the same developmental point. Thus, most researchers have justified comparisons between groups of reasoners (in the ANOVA sense) because they assume that their groups are homogeneous. However, the complex stage model assumes no specific stage assignments. Instead, it proposes a combination of stage usage which is not reducible to a single stage. Thus, according to the complex stage model, comparisons between groups of reasoners is inappropriate and less meaningful due to the inevitable variation within these groups.

Thus, another purpose of the present study was to find construct validity (convergence) for Kohlberg's full six stage sequence by measuring the correlation between abstract reasoning and moral judgment. This was an improvement over previous studies seeking the same goal as I used the complex stage model to assess moral reasoning, rather than previous simple stage approaches, and regression analysis was used rather than ANOVA.

Research Questions

A number of questions have been raised in this comparison and review of two models of moral development. This study was designed to evaluate the following questions.

1) Can we demonstrate construct validity for Rest's model of moral development? 1a) Can we show external validity for an invariant six stage developmental sequence? 1b) Can we find evidence of post conventional thinking by documenting higher levels of abstract reasoning in nonconventional reasoners?

2) Is there sufficient stage mixture within subjects to warrant a complex rather than simple stage orientation? 2a) Do subjects use a variety of stage responses to address different situations? 2b) Do subjects use a variety of stage responses to address the same situation?

3) Is post conventional reasoning more prevalent than Kohlberg assumes? Can we use the complex stage model to document post conventional thinking which would be missed with a simple stage approach?

4) What are some of the personal/situational factors which contribute to individual variability in reasoning?

5) What are some of the methodological concerns in implementing the complex stage model by using Rest's method of assessment? 5a) Is Rest's index of moral reasoning (P%) an adequate continuous measure?

Method

Subjects

All subjects included in this study were college undergraduate volunteers. These volunteers were solicited from introductory psychology classes at a rural state university in exchange for research course credit. Seventy subjects participated in the first phase of the study and 15 of these subjects were used to complete a second phase.

Measures

Similarities. A similarities subtest was administered to all subjects and used as a measure of abstract reasoning (see appendix A). This test consists of 15 items selected from the WAIS (1955) and WISC (1949) similarities subtests which are designed to measure abstract thinking ability (Kaufman, 1979; Matarazzo, 1979; Sattler, 1974). According to Matarazzo (1979),

This measures the individual's ability to perceive the common elements of the term he is asked to compare and, at a higher level, his ability to bring them under a single concept (p. 206).

Most correlational studies show that a well constructed similarities test is one of the most reliable measures of intellectual ability, however; there are instances where individuals do poorly on similarities and comparatively better in terms of I.Q. (Matarazzo, 1979). This demonstrates that abstract reasoning is only one part of I.Q. and is probably more directly related to other forms of reasoning such as moral judgment. "Subjects who perform poorly on the Similarities test may do so not because of intellectual lack but

because of an inner need for concretistic thinking" (Matarazzo, 1979, p. 490). Thus the similarities test sheds light on the nature of the subject's logical thinking process; the degree to which they can step outside the pure physical characteristics and describe functional, purposeful, or other abstract commonality. The WISC and WAIS Similarities subtests were combined because neither test contained enough difficult items to appropriately challenge the college subject population. Items were selected from both tests to increase the difficulty, allowing more variance in performance for more meaningful correlational results.

The similarities test was administered in group fashion and scored according to the standards described in the WISC (1949) and WAIS (1955) manuals. Good responses were given 2 points, fair responses were awarded 1 point and poor responses earned no point value. Points were summed across the 15 items to yield a total score. Standardized instructions for dealing with confusing responses include a request for clarification posed by the examiner, however, that was not part of the procedure in this study due to the limitations of group administration. Instead a method of interpolation which is commonly used in scoring standardized measures (e.g., Vineland Social Maturity Scale, 1965) was used. For example, when a confusing response (see appendix A) fell within a string of 0 point responses it was scored as 0. If it fell within a string of 1 or 2 point responses it was scored as 1. When it fell after a string of 1 or 2 point responses but was preceded by a 0 point response it was scored as 1/2 point.

General information. All subjects completed a test of general information (see appendix B). This test consists of items selected from

the WAIS (1955) and WISC (1949) information subtests which are designed to measure acquired general knowledge (Kaufman, 1979; Matarazzo, 1979; Sattler, 1974). It was administered in group fashion and scored according to the standards described in the WISC (1949) and WAIS (1955) manuals. Good responses were given 1 point and poor responses were awarded no point value. Partial credit of 1/2 point was given when a portion of the answer was correct. Points were summed across the 10 items to yield a total score.

Defining issues test (DIT). The DIT is a measure of moral preference designed by Rest (1979) (see appendix C). After subjects read through a moral situation they are expected to rate and rank a number of concerns and questions in terms of their relative importance in making a decision about what ought to be done. Item selection on the DIT is largely governed by two processes, the ability to comprehend an item and the sense of an item's conceptual adequacy (Lawrence, 1978).

The items on the DIT were selected from issues and concerns raised by subjects while they were being evaluated by Kohlberg's moral judgment interview. Typical stage responses given by subjects during interviews were extracted from Kohlberg's research transcripts. These stage responses were transcribed and placed as response choices on the DIT. All items are matched by word length, syntactic complexity, and use of technical or specialized terminology (Rest, 1979).

Internal consistency and reliability are reported on various DIT scoring methods. The P index refers to the percent of post conventional reasoning used by the subject throughout the DIT. The P index is derived in the following way. 1) Each of the subject's four choices to any one dilemma is differentially weighted. The first is weighted 4

points, second choice 3 points, third choice 2 points, and the last choice is weighted 1 point. 2) The weight is then multiplied times the stage rating of the response. 3) These products are summed by stage so that each stage has some numerical value. 4) These values are divided by the total possible point value, to obtain individual stage percentages. 5) To derive P%, the percentage values for stages 5 and 6 are summed.

The T score (also known as Kohlberg's Moral Maturity Score) was also used in this study, although no psychometric data exist on this measure. This score is derived in similar fashion to the P%. Following step 4, the stage percentages are multiplied times the stage number to yield a series of stage products. These products are summed across all stages to form the T score.

Rest (1979) reports DIT test - retest reliability in the range of .70-.80, and internal consistency of .77. Measures of convergent validity have also been reported (Rest, 1979). The DIT correlates with measures of moral comprehension .49-.65 and Kohlberg's measure of moral judgment .40-.70. Correlations between the DIT and I.Q. are approximately .36, general aptitude .40, and tests of personality .25-.35. In general, the DIT correlates best with other measures of moral development, moderately with measures of general intellectual development, and poorly with measures of personality. Since moral reasoning is presumed to be one aspect of general cognitive development we would expect moderate correlations with I.Q. However, the higher correlations with other measures of moral reasoning show that the DIT "is a distinct aspect of intellectual development and not simply the

application of general cognitive and linguistic skills to moral content" (Rest, 1979, p. 147).

The P% is the most common score reported by researchers using the DIT and its viability rests on the major assumptions of the complex stage model described earlier. Individual stage score percentages are also reported in Rest's DIT scoring system. The test-retest reliability of the P index ranges from .71-.82 and internal consistency is .77. The test-retest reliability of stage score percentages range from .50-.80 with an internal consistency of .28-.60.

The DIT protocols used in this study were scored according to the system described above and in Rest (1979). This system produces a P index as well as stage percentages.

Structured interview. The structured interview was designed to explore personal/situational factors which could account for the low internal consistency in stage responses across dilemmas as measured by the DIT. The questions differed for each subject, since it was unlikely that subjects would generate identical DIT response patterns. However, the purpose and format of the interview remained constant. The following four steps represent a model approach to the structured interview.

- 1) Subjects were given 5-10 minutes to review their protocol. This not only provided subjects with an opportunity to refresh their memory, it also gave them a chance to re-evaluate their choices.

- 2) Following the review period two or three dilemmas were chosen for further discussion. This choice was based on the subject's use of different stage responses across the different moral situations. The

purpose of step two was to generate discussion within each of the situations (comparisons across situations was the focus of step three). This step served as a means of clarifying the subject's reasoning as well as determining; a) if the subject really understood the dilemma, b) if their answer truly represents their reasoning, c) if they understood and rejected other stages of reasoning, and d) the situational factors which led to the rejection of higher and lower level reasoning. The following example is a model of this step.

In the Heinz story you said that a "husband's love for his wife" was more important than "upholding the community's laws". What are the key issues in your opinion? How important is it to consider whether the druggist deserves to be robbed for being so greedy and cruel? How does the law enter into a situation like this?

3) Following within situation exploration , a contrast across situations was conducted. This gave subjects an opportunity to justify their inconsistency in level of reasoning in terms of personal/-situational factors. The following is an example of this step.

Why is it important to uphold the law in the case of the escaped prisoner but acceptable to break the law and steal the drug in the Heinz situation?

4) This fourth step consisted of a series of additional questions which were designed to address situational factors. They were posed in a more general manner, asking subjects to comment on the experience of taking the DIT and comparing it to their own life experience. The following sample questions were characteristic of this step.

Considering all six situations, which ones did you feel most and least comfortable about answering? Why?

Which of your responses are you most and least satisfied with? Why?

Did any of these situations make you feel stressful or remind you of a real life experience which was stressful?

Which situations had issues which were most and least meaningful? Why?

Have you ever experienced a moral situation similar to the ones you have been asked to respond to?

The structured interview was conducted on an individual basis lasting about 45 minutes.

Procedure

Seventy subjects were administered the DIT, Similarities, and General Information in group fashion. This required approximately 50 minutes to complete. Specific instructions for completing this package are contained within the instrument, however, subjects were told:

This questionnaire is aimed at understanding how people think about social problems. Different people have different opinions about right and wrong thus these problems have no absolutely right answer. Please answer all questions as completely as possible and expect to complete everything in approximately 50 minutes.

Inter-scorer reliability. The measures of abstract reasoning (Similarities) and general knowledge (General Information) were scored by two psychology graduate students. A sample of 29 protocols was selected at random, from the total sample of 70, to establish reliability coefficients. The inter-scorer reliability on the Similarities measure was .971 and the reliability on the General Information measure was .944. This level of agreement between scorers suggests that the 29 protocols were scored accurately and it increases our confidence in the scores of the remaining sample. Table 1 displays

the means, standard deviations, and ranges of scores obtained by both scorers on each measure ($N = 29$), as well as summary statistics for the entire sample ($N = 70$). The similarity between the statistics in the reliability sample and the total sample indicates that the sample selected for establishing scorer reliability is truly representative of the total sample.

The DIT was scored according to the procedure described earlier. Fifteen subjects, who appeared to exhibit the greatest amount of inconsistency in their stage responses across the six moral dilemmas, were selected for the structured interview.

Table 1

Descriptive Statistics For Reliability And Total Sample

	Mean	Standard Deviation	Smallest Value	Largest Value
Similarities 1* ($N = 29$)	16.44	4.064	8.0	24.0
Similarities 2** ($N = 29$)	16.09	3.787	8.0	23.0
Similarities entire sample ($N = 70$)	16.02	4.21	6.0	27.0
General Information 1* ($N = 29$)	3.76	1.334	1.0	6.0
General Information 2** ($N = 29$)	3.72	1.327	1.0	6.5
General Information entire sample ($N = 70$)	4.01	1.72	1.0	6.5
P% ($N = 70$)	39.171	12.081	11.0	69.0
T score ($N = 70$)	376.343	42.308	248.0	466.0

Note. * scorer one
 ** scorer two

The structured interview was conducted on an individual basis lasting about 45 minutes. The information was recorded by the interviewer in the form of written transcripts and later summarized in terms of the situational factors which seemed to account for the disparity in reasoning.

Subjects were thanked for their participation and any questions were answered regarding the purpose of the study.

Results and Discussion

The following section will include an analysis and discussion of the correlational data (abstract and moral reasoning). Consistency in reasoning across and within different moral situations will also be explored and Rest's technique of indexing moral reasoning will be examined. In addition, the prevalence of post conventional reasoning will be evaluated to provide additional support for Kohlberg's full six stage model. Finally, the results of the structured interview will be presented to address situational variation in terms of personal/situational factors as well as methodological considerations.

Correlational Data

Zero order correlations. Can we demonstrate construct validity for Rest's model of moral development? Can we find external validity for an invariant six stage developmental sequence? Can we find evidence of post conventional thinking by documenting higher levels of abstract reasoning in nonconventional reasoners? I suggested that there would be a significant correlation between logical and moral reasoning.

Table 2 shows that this hypothesis was not supported as the correlation between similarities and P% was .194 ($F = 1.62$, $d.f. = 1/67$, $p > .05$, n.s.), and the correlation between Similarities and T score was .136 ($F = 1.26$, $d.f. = 1/67$, $p > .05$, n.s.). The General Information Score, however, correlated significantly with moral reasoning (P% general knowledge = .338, $F = 8.64$, $d.f. = 1/67$, $p < .01$, sig; T score, general knowledge = .273, $F = 5.39$, $d.f. = 1/67$, $p < .05$, sig; R^2 P% · Similarities, general knowledge = .126, $F = 4.84$, $d.f. = 2/67$, $p < .05$, sig.).

Partial and semi-partial correlations. Previous research has demonstrated a direct relationship between levels of logical and moral thought, not accounted for by I.Q. (Kohlberg, 1979; Kuhn, 1977). The limitation of previous research is the use of a simple stage model which leads to violations of homogeneity. In the present study I proposed that the correlation between abstract and moral reasoning would be significant following statistical control of general knowledge. Also, the unique contribution of general knowledge to the index of prediction would be nonsignificant. The figures in Table 3 indicate that neither hypothesis was supported (correlation between P% and Similarities with the General Information score partialled out = .116, $F = .93$, $d.f. = 1/68$, $p > .05$, ns; semi-partial correlation between P% and General Information = .297, $F = 6.785$, $d.f. = 1/67$, $p < .05$ sig.) In addition, we can see from the table that Similarities adds no significant unique information to our prediction, and the variance shared by moral reasoning and General Information remains significant after the contribution of logical reasoning is partialled out. Similar results are obtained using the T score as a dependent measure.

Table 2

Zero Order Correlations Between Dependent Measures P% and T score
and Independent Measures Similarities and General Information

	P%	T score	Similarities	General Information
P%	1.00	.805**	.194	.338**
T score		1.00	.136	.273*
Similarities			1.00	.263*
General Information				1.00

Note. * $p < .05$
 ** $p < .01$
 N = 70

Table 3

Partial and Semi-Partial Correlations Between Dependent Measures
P% and T score and Independent Measures Similarities and
General Information

	P%	T score
Similarities (1)	$r_{P1.2} = .116$ $r_{P(1.2)} = .108$	$r_{T1.2} = .069$ $r_{T(1.2)} = .067$
General Information (1)	$r_{P2.1} = .303^*$ $r_{P(2.1)} = .297$	$r_{T2.1} = .247^*$ $r_{T(2.1)} = .246^*$

Note. * $p < .05$
 N = 70

The poor correlation between measures of moral reasoning and Similarities (abstract logical reasoning) could be due to a variety of factors such as: poor measurement properties of the Similarities test, poor psychometric properties of the Defining Issues Test, subject response characteristics not measurable with the DIT format, or marked departure from the important assumption of linearity.

The Similarities test was constructed from items which comprise similarities subtests on the WAIS (1955) and WISC (1949). These subtests are proven effective measures of abstract logical thinking (Kaufman, 1979; Matarazzo, 1979; Sattler, 1974). However, since I did not use either subtest in its entirety it would have been useful to establish a correlation between my new Similarities measure and the standardized version. Otherwise I cannot assume that my measure has identical psychometric properties. It was necessary to alter the standardized measure to create a more difficult test which could generate a greater level of variability in the college population. The format of my Similarities test however, was identical to the standardized version and the few items which were substituted came unchanged from another standardized version. Therefore, even though we cannot be positive that the two measures are effectively evaluating the same thinging skills, we can be fairly confident that my measure has construct validity and the low correlations reported earlier are probably not due to poor psychometric properties of the experimental Similarities measure.

The low correlations may also be due to an ineffective dependent measure. The psychometric properties of the DIT reported earlier are fair at best. In addition, the present study revealed numerous

measurement problems ranging from content and response bias to inadequate indexing procedures. The results and discussion of this analysis will be presented in more detail later in this paper. Briefly stated, however, the P% does not hold up as a valid indexing procedure, not all dilemmas are capable of inducing thoughtful consideration of responses, the "meaningless" items do not distinguish high and low level reasoners as they were intended, the wording and presentation of items is confusing, and an individual's choice of response does not always represent his/her reasoning capacity. Since the DIT is psychometrically questionable at best we cannot expect meaningful correlational data when it is used as a dependent measure.

The poor correlations could be due to a particular response characteristic not appropriately assessed by the DIT. The structured interview revealed that subjects who were capable of understanding and using higher level thought did not always select the high level responses. They recognized, for example, that principles were important but found legitimate reasons for ranking consequences as more important. Thus, some individuals who were high level moral reasoners, and probably scored high on the measure of abstract thinking, were labeled by the DIT as lower level reasoners. Therefore, their high Similarities scores correlated poorly with their low DIT ratings. We cannot assume that a subject's multiple choice selection represents reasoning capacity in most instances, therefore the use of DIT information is limited in correlational research.

The low correlations may have been due to a marked departure from the important assumption of linearity. That is, the relationship between moral and abstract reasoning may not be linear. This means

that while moral and abstract reasoning may correlate moderately considering the full range of both constructs, the relationship may not hold up when we use just one portion of this range. This study, for example used a group of abstract reasoners which fall in the middle to high range of the construct. The relationship between moral and abstract reasoning may be different for this group than for the construct as a whole. To truly test the relationship between moral and abstract reasoning we should have included the full range of abstract and moral reasoners.

The most stable factor in the correlational aspect of the study was a weak, though significant, correlation between the General Information measure and DIT score. One possible explanation for this result could be that individuals who have more general knowledge have had a greater range of experience to develop this knowledge. This broader range of general experience may also include greater experience in processing moral situations.

Due to the psychometric properties of the instruments involved, however, we must assume that the constructs proposed for correlation were measured only to a limited degree. Their relationship remains largely unknown. Before these relationships can be validly established a number of issues need to be addressed. When using measures of moral reasoning one needs to recognize the discrepancy between capacity and performance and decide upon which aspect it is most appropriate to focus. We need a more valid measure of the construct before we use moral reasoning in any meaningful research. Also, no one has ever used an established standardized measure of abstract reasoning in their correlational research. When referring to a construct such as abstract

logical reasoning we need to become more operational in terms of description as well as measurement before designing proposals and making generalizations.

Simple and Complex Stage Models

Is there sufficient stage mixture within subjects to warrant a complex rather than simple stage orientation? There are a variety of difficulties in defining the construct of moral reasoning. One issue is the design of an appropriate model of organization. Within structural theory alone, for example, there is little agreement regarding the process of moral development. Some conceptualize development as discrete steps while others believe that it is a continuous progression. Some theorists focus strictly on qualitative change while others look for quantitative changes. Furthermore, some believe that individuals can be adequately described with a single stage label while others suggest that some rating of stage mixture is most appropriate. These major theoretical issues have been contrasted as the simple versus complex stage models.

Reasoning across situations. Do subjects use a variety of stage responses to address different situations? It was proposed that individuals would use different types of reasoning in response to different situations. In keeping with the complex stage model I expected stage mixture across situations to cover the full range for each individual, rather than only adjacent stage usage as the simple stage model suggests. To illustrate this point most dramatically only the subject's first choice responses to each of the six stories were considered. Table 4 shows that no one responded with the same stage across all

stories and only four subjects (6%, $N = 70$) responded within the same level (two adjacent stages). Many subjects spanned all five stages (39%) with their six first choice responses, and 77% of the subjects spanned four or five stages. The mean number of stages spanned was 4.2 out of 5. This suggests that individuals use a full range of reasoning to respond to different situations, thus it is inappropriate to label them in terms of any one stage.

Table 4
Stages Spanned When Reasoning Across Dilemmas

Stages Spanned Across Six Stories (using only first choice responses)	Number Of Subjects	Percentage Of Subjects
1	0	0
2	4	6%
3	11	16%
4	27	38%
5	28	39%

Note. $\bar{X}$ stages spanned = 4.2
 $N = 70$

As Rest (1979) suggests, the notion that a subject is on a stage is wrong. Individuals continue to use lower stage reasoning long after they are capable of higher level thought.

Stage mixture is one explanation for the regression noted previously in longitudinal research. Rather than interpret regression as evidence contrary to the invariant developmental sequence we can think of it as selective use of alternative forms of reasoning. Since individuals do not discontinue their use of lower level reasoning when they become capable of higher level thought we should expect them to utilize seemingly "regressed" reasoning occasionally. Rather than

taking just two measures (before and after) to document longitudinal change it seems more appropriate to take several measures before and after to determine changes in terms of probability of stage usage. Development could then be considered the increasing probability of using higher level thought. This conclusion would not be altered drastically by a few cases of lower stage usage following the longitudinal period, which is evidently what has happened in previous studies employing the simple stage model.

Another implication of stage mixture concerns moral education programs. Turiel (1966) has done extensive research to determine the most productive means of facilitating moral growth. His general conclusions suggest that; 1) cognitive moral conflict is a necessary condition for moral growth and 2) moral discussions where subjects are presented with reasoning which is $1/3$ to 1 stage higher than their own stage is a most effective means of inducing cognitive moral conflict. This "+1" approach has been worked into educational curriculums. However, given the complex stage model and the evidence presented in this study, we would be hard pressed to identify anyone in terms of a particular stage. Thus, which stage do we use for the +1 reference when an individual uses stages 2 through 5? The +1 approach to promoting moral growth, then, not only becomes impractical but also theoretically impossible.

The information on stage mixture suggests that assessing moral reasoning is not as simple a task as once assumed. When a single stage rating is assigned to individuals based on their answers to a few moral stories, we are discarding valuable information in the interest of simplicity and losing sight of the true purpose of our investigation.

Reasoning within situations. Do subjects use a variety of stage responses to address each situation? The complex stage model predicts that individuals will use a full range of reasoning when responding to moral dilemmas, as different situations are capable of inducing different levels of thought, and personal characteristics of the reasoner interact with specific features of the dilemma. Subjects may use a variety of stage responses regardless of the personal/situation factor. The current study provides an opportunity to examine this question. To obtain this comparison we treated each dilemma as a unit and established a measure of consistency within the dilemma. Since 70 subjects responded to six dilemmas we had a sample of 420 dilemmas to examine. We measured consistency by considering the number of stages spanned by the four responses within each dilemma. Table 5 shows that only one dilemma (.4%) was answered with four responses at the same stage. Most dilemmas contained responses which spanned four or five stages (65%).

Table 5

Stages Spanned When Reasoning Within Dilemmas,
Considering All Four Choices

Stages Spanned Within Dilemma	Number Of Subjects	Percentage Of Subjects
1	1	.4%
2	33	8%
3	111	26%
4	164	39%
5	111	26%

Note. $\bar{X}$ stages spanned = 3.8
N = 420

Since the Defining Issues Test is a forced choice multiple selection survey, some subjects may have chosen responses unrepresentative of their reasoning because they were required to make a choice even though there were not enough reasonable answers provided at their stage. Thus, the stages spanned within dilemma was calculated using only the first two responses chosen for each situation. Table 6 shows that subjects are more consistent in their first two choices, however, there is still a large amount of variability. The mean stages spanned was 2.6 with 22% of the stories involving a span of four to five stages.

Table 6
Stages Spanned When Reasoning Within Dilemmas,
Considering Only Two Choices

Stages Spanned Within Dilemma	Number Of Subjects	Percentage Of Subjects
1	85	20%
2	140	33%
3	103	24%
4	57	14%
5	35	8%

Note. $\bar{X}$ stages spanned = 2.6
N = 420

Again, this makes it difficult to assign a single stage rating even when using just one dilemma. Assigning a single stage rating or expecting subjects to justify their answers with reasoning on just one stage is unrealistic. An individual's decision about whether or not to steal a drug to save a life may be based on principles, as well as

laws, and consequences or personal gain. The individual who can consider and integrate all three levels of thought may be more successful than someone who narrowly focuses on only one level. Kohlberg's simple stage approach forces the data into artificial simplicity.

The data summarized within and across situations suggests that reasoning is not consistent within the one or two stage framework proposed by the simple stage model. The across situation variability indicates that subjects respond differently to different situations. The within situation variability suggests that although subjects are more consistent within situations, the situational factor cannot account for all the variability. Evidently subjects draw from a variety of different levels of thought in an attempt to address each dilemma most adequately.

Post conventional reasoning. Is post conventional reasoning more prevalent than Kohlberg assumes? Can we use the complex stage orientation to document post conventional thinking which would be missed with a simple stage approach? The fact that subjects draw from a variety of different levels of thought as they address moral dilemmas implies considerable use of post conventional reasoning. Many researchers have questioned Kohlberg's six stage model because evidence is lacking for stages 5 and 6. Another purpose of this study was to substantiate Kohlberg's full six stage model by documenting evidence of post conventional reasoning in terms of the complex stage model.

According to the simple stage model and Kohlberg's scoring methods, individuals are required to use post conventional reasoning predominantly throughout the assessment before they can be considered

post conventional reasoners. This creates a problem as only a rare sample of people fit these restricted qualifications. It is not necessary to find pure post conventional reasoners before we can report that post conventional reasoning is being used, especially in view of the stage mixture evidence presented earlier. In fact, we should not expect examples of any pure stage reasoners (individuals using the same stage of reasoning 100% of the time).

Evidence of post conventional reasoning was found in 100% of the sample. In fact, individuals used stages 5 and 6 an average of 39.17% of the time. However, none of the subjects in the study would have been considered post conventional reasoners by Kohlberg's standards (i.e., global or moral maturity score). The complex stage model allows documentation of the six stage developmental sequence, in terms of increasing probability of higher stage usage, long before an individual reaches a point of exclusive stage 6 usage.

Empirical results have shown that moral reasoning does not exist in pure stage form. Thus we can accept that post conventional reasoning is being used even though no one would be labeled as a post conventional reasoner according to Kohlberg's standards.

The implications of this conclusion are especially important for cross-cultural studies. Researchers have failed to find evidence of post conventional reasoning in other cultures. This may be due primarily to their dependence on the simple stage model. Individuals may understand and be capable of post conventional thinking but choose not to practice this level of reasoning because it has no relevance in their society. Also, post conventional reasoning may be used in part to address moral issues but it does not exist in pure form so people

are not credited with this level of sophistication. Studies have not addressed the issue of preference versus capacity with respect to moral reasoning, and cross cultural studies have not been conducted based on the complex stage model. We need to direct ourselves to these issues before we make conclusions about the reasoning capacity of people in other cultures.

The complex stage model seems to be the most adequate way of organizing and understanding moral development because of its flexibility in considering personal and situational factors. However, Rest and others do little more than propose the model and mention that their data seem to fit a complex pattern. This does not explain why data patterns are complex.

Structured Interview, Results

What are some of the personal/situational factors which contribute to individual variability in reasoning? The structured interview was designed to explore possible explanations for situational variation in reasoning. In addition, it also revealed many methodological considerations regarding the validity of the DIT.

The results presented in this section will address two major issues: 1) does the DIT accurately evaluate an individual's reasoning? If not, what are some of the methodological considerations? 2) what are some of the personal/situational factors leading to variation in reasoning across dilemmas? The results of the structured interview will be presented on a case by case basis, using the subject's initials to identify each case.

L.M. Accuracy. In the "Heinz" story L.M. chose a stage 2 answer and defended it with stage 2 and 4 reasoning "I think of myself before anyone else...is it really worth the risk of getting shot just to save someone else?...you have to put yourself first in those situations." When pressed further, L.M. responded "under no circumstances would I steal...stealing is wrong."

In the "Prisoner" story L.M. chose an anarchistic response which was scored as stage 2. However, her justification of this answer was on a principled level "I guess I would approach him first (before deciding to turn him in)...he would have to demonstrate to me that he could make a worthwhile contribution to society again."

In the "Doctor" story L.M. chose a stage six response and defended it with stage six and two reasoning "People should be allowed to make their own decision about living and dying...there should not be a law about it because no one has a right to tell someone when they can live and die...you have got to think of yourself."

The three dilemmas chosen for follow up discussion were largely accurate indicators of L.M.'s true reasoning. However, L.M. remarked that "most of the questions seemed ambiguous and a little confusing." This is another source of measurement error.

Personal/situational factors. The varying level of consequence across stories influenced L.M. to attend to them in some instances and overlook them in others. The severity of the consequences also affected the difficulty of the dilemma "the doctor dilemma was the hardest for me because the stakes were high...the Heinz story was easiest because it seemed more clear cut. I've always been brought up to believe that stealing is wrong." Thus, personal experience and

severity of consequences were two factors contributing to variation in reasoning.

J.D. Accuracy. In the "Heinz" story J.D. chose a "meaningless" response and justified it with principled level reasoning "life is more important than laws and getting in trouble...no matter who is dying."

In the "Newspaper" story J.D. chose a stage 3 response and justified it with principles and purposes "student's rights are important...if they had good intentions and some important issues to write about then the paper should be allowed to continue."

In the "Doctor" story J.D. chose a stage 5 response and justified it with stages 4 and 2 reasoning "I would gain some kind of legal permission...there is a pretty heavy penalty for killing someone."

None of the dilemmas chosen for further discussion was an accurate measure of J.D.'s reasoning. In addition, the "meaningless" response did not serve its purpose as J.D. attributed a meaning to this answer which would have been scored as stage 6.

Personal/situational factors. The varying level of consequences across stories was a factor influencing J.D.'s reasoning. J.D. remarked "the consequences for stealing are nothing next to what would probably happen if you murdered someone." J.D. also mentioned that the dilemmas varied in their ability to engage the reader "some stories seemed more realistic, more vivid, and easier to imagine or get involved with." Thus, severity of consequence, interest level, and realistic nature of the story were all factors contributing to response variation.

L.W. Accuracy. In the "Webster" story L.W. chose a stage 4 answer and defended it with stage 5 and 6 reasoning "he should hire the best

mechanic and laws or races should have nothing to do with it...he should hire the individual for his abilities...laws are designed to guarantee freedom to all, if this is not working what is the point of the law?"

In the "Student take-over" story L.W. chose a stage 4 answer and defended it with stage 5 and 6 reasoning "I don't believe that anyone has the right to exercise their rights when they infringe on others... in the long run this approach will not benefit society as a whole, just individual small groups. If it meant saving lives I would probably break the law."

In the "Newspaper" story L.W. chose a stage 6 response and justified it with stage 5 and 6 reasoning "I think freedom of speech is critical. This public discussion is where a lot of learning about real issues takes place."

L.W. was consistent in her reasoning across situations but the DIT did not evaluate her this way. In most cases L.W. had some higher level rationale for choosing a lower level response. Evidently some DIT responses are not understood in the same way by all subjects. The DIT, however, makes no distinction between the more sophisticated and less capable person's version of the same response.

Personal/situational factors. While there was actually little variation in reasoning across situations, L.W. mentioned that some dilemmas were more difficult to answer than others "the social decisions were easier to answer than the personal issues. Making decisions for large groups seems less personal and easier to detach from."

A.A. Accuracy. In the "Heinz" story A.A. chose a stage 2 response and justified it with stage 3 reasoning "I guess I would end up taking the drug...I'd probably do just about anything for my family." She also had some meaning for the meaningless response which matched her own level of reasoning. When the people involved in the story are not direct family members A.A. falls back to a consequence, personal gain orientation "it just isn't worth the risk...you've got to think of yourself you know." A.A. had a difficult time imagining some situations because of the lack of personal concrete experience "its to save his wife so no one will look down on him...he should consider the consequences of losing his wife versus going to jail."

In the "Student take-over" story D.B. chose a stage 4 response and justified it with stage 2 and 3 reasoning "the president has been around, he's experienced...he is the authority."

In the "Doctor" story D.B. selected a stage 6 response and justified it with principled level reasoning "her life is her responsibility, it is her own decision...the doctor is there to save lives and make people more comfortable but ultimately it is her decision."

Personal/situational factors. The "Student take-over" was more meaningful for D.B. because of personal experience. D.B. has already thought a great deal about these issues thus his answers seemed more practical than his responses to other stories.

A.V. Accuracy. In the "Heinz" story A.V. chose a stage 2 answer and justified it with stage 2 and 3 reasoning "even if the law isn't right you need to abide by it...if you get caught you must pay the consequences."

In the "Prisoner" story A.V. chose a stage 4 answer and justified it with stage 5 reasoning "the severity of the crime makes a difference ...the particular crime is a more important factor than the law." This person was able to consider individual circumstances and the purpose behind the law in a way not measured by her selection of a stage 4 response.

In the "Doctor" story A.V. chose a stage 4 response and justified it with stage 5 reasoning "its her decision, she should know best, its her body and human life is more important than laws...law doesn't take into consideration special circumstances."

While the DIT was a little more accurate in evaluating A.V. the overall DIT rating was still an underestimate of A.V.'s actual level of reasoning. Some of this may be due to structural properties of DIT. A.V. commented that the answers were not phrased clearly and it was not clear whose point of view one should be answering from. She also discovered that occasionally her preferred response to a story was not listed among the options.

Personal/situational factors. A.V. noticed that "in the "Heinz" story only two people are affected but in the "Prisoner" story many people are." When thinking in terms of society the consequences were not as important for A.V. as when dealing with a personal situation. A.V. also mentioned that the most difficult dilemmas to decide were the ones dealing with life and death. Evidently the various stories induced different kinds of thinking for A.V., depending on the personal nature of the story and the extent to which it dealt with life or death.

N.M. Accuracy. In the "Heinz" story N.M. chose a "meaningless" response and justified it with stage 5 and 6 reasoning. "Laws should

not get in the way of principles...laws are not immutable...there are circumstances under which it is o.k. to break the law." Evidently this "meaningless" answer had considerable meaning for N.M. Again, these meaningless items are not serving their purpose of separating high and low level reasoners. They are not identifying "fakers", which was their original purpose.

In the "Newspaper" story N.M. chose another "meaningless" response and justified it with principled level reasoning "the points of view of everyone should be considered in making the most equitable choice."

In the "Student take-over" story N.M. chose a stage 4 response and justified it with principled level reasoning "the ROTC should not be disbanded based on the needs of just a few students, you should really consider the rights of all."

The DIT was particularly inaccurate in this case due to the ineffectiveness of "meaningless" responses. Even though they are supposed to be identifying low level reasoners who fake a high answer, there seem to be just as many high level reasoners choosing these answers.

Personal/situational factors. Personal experience played a role in the decision making "I have no real life experiences to help me answer the Heinz and Protest stories, but my high-school experience working on a newspaper helped me decide about that story."

M.M. Accuracy. In the "Heinz" story M.M. chose a stage 2 response and justified it with stage 2 reasoning. An important aspect of M.M.'s answer was the recognition that laws and personal obligations were important considerations but "the most important thing to consider is whether I live or die" (consequences for getting caught). M.M. said

that she would steal the drug only if she knew she was not going to be caught.

In the "Doctor" story M.M. chose a stage 3 answer and justified it with stage 3 reasoning "only God should decide."

In the "Webster" story M.M. chose a stage 4 answer and justified it with stage 5 reasoning "I would hire him as his talent as a mechanic is the most important consideration...he deserves a job, if his customers don't like it they can go elsewhere."

With the exception of the last story the DIT was accurate in assessing this individual. However, this subject got little credit for laws and personal obligations even though they were relatively important for her. This information was discovered during the follow up discussion but was not indicated by the DIT because thought process is not scored.

Personal/situational factors. M.M. mentioned that consequences were more or less important depending on their severity "running a red light may be worth the risk but getting shot is not." Also, previous experience in a philosophy class helped her sort out the issues in the "Doctor" dilemma. Thus, previous experience and varying severity of consequence were two factors contributing to the variability in reasoning across situations.

A.R. Accuracy. In the "Heinz" story A.R. chose a "meaningless" response and justified it with a law and order rationale, mixed with personal gain (stages 2 and 3) "stealing is wrong...it isn't going to help him."

In the "Doctor" story A.R. chose a stage 3 answer and justified it with stage 3 reasoning "Its God's responsibility not the family's or the doctor's."

In the "Newspaper" story A.R. chose a stage 6 answer and justified it with a principled level response. In general the DIT was fairly accurate in assessing A.R. Unfortunately no interview information was provided to explain her variability in reasoning across situations (consequences in some instances and principles in other cases.)

V.S. Accuracy. In the "Student take-over" story V.S. chose a stage 2 answer and justified it with a principled level response "the principle behind the movement is more important than the law."

In the "Newspaper" story V.S. chose a stage 3 answer and justified it with a principled level response "the general principle of freedom of speech is most important."

In the "Webster" story V.S. chose a stage 5 response and justified it with a principled approach "everyone is equal...laws are irrelevant to the issue."

In most cases V.S. understood the lower level responses in higher level terms. The DIT has no provision for this kind of thinking. Rather, the DIT assumes that everyone will understand an answer for the meaning it was intended to convey, given the individual's capacity to reason on that level.

Personal/situational factors. V.S. has encountered many of the same issues of protest and equal rights in his own personal experience, thus these issues have the most meaning for him. This made a difference in how he was able to get involved in the story and generate meaningful answers.

S.M. Accuracy. In the "Student take-over" story S.M. chose a stage 2 answer and justified it with a stage 5 response "personal beliefs are the motivating factor...they are more important than laws or consequences." S.M. also mentioned that she had not given the stories much thought, they were not able to stimulate her involvement.

In the "Prisoner" story S.M. chose a stage 3 answer and justified it with a stage 5 response "he should be evaluated as a person and a judgment should be made depending on the situation...only then can the best decision for him and society be made."

In the "Doctor" story S.M. chose a stage 5 answer and justified it with a stage 3 response "God has the ultimate authority."

In all instances S.M. changed her point of view during the interview from her original feelings at the time she took the DIT. This was mainly due to her more thorough examination of the situations. The DIT, in its original format, was not capable of stimulating an intense level of thought in the case of S.M.

Personal/situational factors. The major situational factor was thought provoking ability of the dilemma. As stories were able to induce more intense thinking, S.M.'s reasoning expanded to include more encompassing principles. Also her uncle recently became terminally ill so she has very strong feelings on this issue.

Structured Interview, General Discussion

Personal/situational factors. Personal experience and severity of consequence were the two most common factors contributing to variation in response level across situations.

Most subjects noticed that the severity of consequences were different across stories. They also reported that this influenced their choice of response "the consequences for stealing are nothing next to what would happen if you killed someone." Most individuals admitted that consequences played some role in their decision, regardless of their overall level of reasoning. When consequences were less severe they could more easily be placed on low priority. Thus, an individual's choice of consequence level reasoning is not purely a function of reasoning capacity, but also situational demands. We could probably create situations which would influence the highest level reasoner to focus on consequences or level I responses. This is not as much an indication of low level reasoning as much as it may be a measure of adaption. An individual who is able to understand principles, laws, and consequences, and can select the most appropriate levels to address the specific situation, is showing signs of adaption not measured by the DIT. There is no provision to distinguish this person, when they select level I reasoning, from the individual who can only reason on a consequence level.

Subjects also reported that their own personal experience affected their response to the DIT "I've been through that with my grandfather, that made a lasting impression on me.... It made it much easier to put myself into the situation...it makes more sense to me." The personal relevance factor made the dilemma clearer and easier to understand. In addition, having dealt thoroughly with these issues in the past, subject's answers were more organized and well thought out. Also, as dilemmas were more personally relevant subjects showed more interest and were better able to remember story details.

Memory and interest level also affected reasoning in dilemmas that were not personally relevant. Subjects reported that some stories were more thought provoking than others. As dilemmas seemed uninteresting or difficult to imagine subjects became expedient in their approach to the task. They changed their answers during the follow up interview and admitted that they had previously not given the story much thought. Thus, reasoning is variable across situations as these situations vary in their ability to induce deeper level thought. Subjects also reported varying difficulty imagining themselves as the main characters in the dilemmas.

The results from the structured interview suggest that individual variation in reasoning across situations was due partially to severity of consequence, personal relevance, thought provoking ability of the dilemma, and egocentrism in unrealistic situations. Thus measuring moral reasoning is even more complex than the complex stage model assumes. Since an assessment procedure has not been designed to accommodate these personal/situational factors, we are without a measure of moral reasoning capacity that can accurately describe an individual's thought processes.

Methodological considerations. What are some of the methodological concerns in using Rest's method of assessment? In addition to personal/situational factors, the present study showed that numerous methodological factors also contribute to response variation. One of these factors is P%, Rest's index of moral reasoning.

Rest (1979) suggests that the variability in stage usage can be thought of in terms of overlapping probability curves, each curve representing a different stage. Rest's complex stage model gives us a

continuous measure of moral reasoning (P%) which is supposed to describe individuals in terms of these complex stage patterns. This means that we should be able to describe individuals in terms of their lower stage usage based on their P%.

Is Rest's index of moral reasoning an adequate continuous measure? I attempted to validate the P% by examining the relationship between P% and lower stage patterns. This was accomplished by using the T score (also known as Kohlberg's moral maturity score). This is a total measure of stage usage, principled and lower. The T score is a sum of stage products. A product consists of the stage used times the percentage it was used. If stage patterns were similar for individuals with the same P%, then the correlation between T score and P% would have been nearly perfect. However, $r_{P\%T \text{ score}} = .80$, thus the scores share only 64% common variance. The two scores are calculated from some of the same information as they both use percentages of stage five and six responses. Since the average P% is 39.2%, the P% and T score share about 40% of the variance just by virtue of calculation method. Also, an average of 8% of the subject's responses were meaningless answers which have no influence on either score. Thus, 48% of the variance is either controlled or removed. The 64% total common variance, then, is less than impressive. Thus, P% is not a good continuous measure since individuals with the same P% do not have similar lower stage patterns. We cannot accurately describe moral reasoning in terms of P% alone. Without an adequate continuous measure we cannot expect meaningful correlational data. The theory proposed by Rest and supported in this study indicates that moral development is a continuous rather than discrete construct. Due to Kohlberg's simple

stage orientation he does not adequately address stage mixture. Rest claimed to have developed an index to measure this stage mixture. Evidently his measure is also inadequate. Kohlberg and Rest both discard large amounts of information without sufficient justification. Neither theorist has an adequate measure of the construct moral reasoning.

Without an adequate measure of moral reasoning the implications for research and practical interventions are distressing. For example, when organizing groups for different experimental treatments we cannot assume homogeneity based on P%. Also, we cannot accurately measure change over time with P% because many changes in lower stage usage could occur and go unnoticed. Finally, it is difficult to prescribe individual educational interventions when individuals cannot be adequately described in terms of the trait of interest. For example, Turiel's (1966) +1 approach of facilitating moral growth is useless without an adequate means of measuring an individual's baseline level of reasoning.

The stage mixture documented in this study makes it difficult to measure and describe individuals in terms of moral reasoning. Even though an adequate continuous measure has not been designed, I am confident that it is most appropriate to organize the construct in continuous rather than discrete fashion. For example, this helps up explain the regression in reasoning which is found in longitudinal research based on the discrete model. Thus, the notion of invariant sequential development is preserved with the continuous model.

Another source of methodological difficulty involves the "meaningless" responses on the DIT. These items are syntactically

complex and designed to be totally meaningless. Their purpose is to identify those individuals who do not understand higher level reasoning but choose complex answers in an attempt to fake high.

The most common response choice on the "Heinz" story was one of the "meaningless" items. Fifty-six percent of the subjects chose a "meaningless" response as one of their first or second answers to this dilemma. However, equal numbers of high and low level reasoners answered with this response. The mean P% in the college population was 39.2. Nineteen subjects who chose a meaningless response to the "Heinz" story had a P% greater than 39.2, and the remaining twenty meaningless responders had a P% below 39.2. Furthermore, most subjects had some sensible meaning for this response which was consistent with their other answers.

Considering meaningless responses across all dilemmas, 83% of the subjects selected at least one "meaningless" response. These subjects were also divided equally in terms of high and low level reasoners.

Thus, these meaningless items serve no useful function. They work poorly as an identification of subjects faking high, introducing more error to a psychometrically questionable instrument. Also, since they have no point value they significantly reduce the person's total score.

Subjects also reported that response choices were not phrased clearly. The accuracy results reported earlier indicate that subjects justified their response choice with a different stage of reasoning 25 out of 34 times (73%). In addition, there was no systematic or consistent direction of error in terms of always justifying with a higher or lower stage response. Much of this inaccuracy may be due to the poor phrasing or ambiguous nature of the response choices.

Another critical measurement issue is the subject's varying interpretation of the same response choice. There were many instances of subjects choosing the same response and justifying it with very different reasoning. A stage 4 response choice, then, could be understood two different ways by individuals using stage 2 and stage 6. However, once the response has been chosen, the individual is labeled by the stage rating it represents regardless of the reasoning involved. This is another explanation of the poor correlational data reported earlier. As subjects are incorrectly labeled we cannot expect meaningful correlations.

Subjects were also incorrectly labeled as low level reasoners when they understood yet placed principles on low priority. The structured interview revealed that some reasoners who understood principles, and used them in other situations occasionally, chose lower level responses to address some situations more adequately. These reasoners are given a low rating as if they are not capable of principled level thought. The DIT cannot distinguish between this individual and one who can only reason at a lower level.

The structured interview revealed that the psychometric properties of the DIT are questionable at best. Meaningless answers do not serve their purpose, there is no control for interest level, and many items are difficult to understand due to awkward wording. Expediency is promoted in some individuals due to the length and unrealistic nature of the dilemmas. Forced choice responding creates errors as there are not enough adequate answers at each stage, and the reasoning behind the choice does not always match the reasoning intended by the authors of the DIT. Severity of consequence and personal relevance create

variation in response choice across situations. Memory for story details is also influenced by personal relevance, and the dilemmas vary in their ability to stimulate thoughtful consideration for different subjects. In general, variability in reasoning is to be expected due to numerous measurement issues as well as personal/situational factors.

Summary

This review and comparison of a simple and complex stage model has addressed a variety of questions and concerns. For example, I attempted to show a significant positive correlation between moral and abstract reasoning using a complex stage orientation. However, the correlation between moral and abstract reasoning was not significant. This could have been due to poor measurement properties of my test of abstract reasoning. However, it is more likely that the poor psychometric properties of the DIT were a major factor. Before we can expect meaningful research results we need to design a more valid measure of moral reasoning.

This study also contrasted the simple and complex stage model by noting stage mixture across and within situations. The degree of stage mixture documented in this study indicates that a simple stage approach is an inappropriate way to classify or organize moral reasoning. College age individuals use a full range of reasoning to respond to different situations, rather than the dominant and +1 reasoning suggested by the simple stage model. Since it is inappropriate to label anyone in terms of a single stage, much of the research and many of the educational programs based on this model must be questioned.

Another purpose of this study was to document Kohlberg's full six stage sequence, particularly stages 5 and 6. Can we use the complex stage model to document post conventional thinking which would be missed with a simple stage approach? I found evidence of post conventional reasoning in 100% of the sample. The complex stage model permitted documentation of the six stage developmental sequence, in terms of increasing probabilities of higher stage usage, long before simple stage criteria could have been met. Thus, post conventional reasoning is not as rare as Kohlberg suggests.

This study was also designed to explore personal/situational factors which contribute to inconsistency in moral reasoning. Two personal/situational factors were discovered to be common to most subjects examined in this study: personal experience and severity of consequence influenced the response choice of many subjects.

Since not all dilemmas are equal with respect to severity of consequences, according to complex stage model, variation in reasoning is inevitable on the Kohlberg or Rest measure of moral reasoning. Subjects reported that consequences played a role in most of their decisions, regardless of their overall level of reasoning, and they were able to disregard them to the extent that their severity was minimal. None of the major assessment devices are designed to accommodate this situational factor.

Subjects also reported that their personal experience affected their response to the dilemmas. They were able to organize their answers, attend more to the details, and remember these "personal" dilemmas with greater accuracy. Since assessment devices are not prepared to consider the role of personal/situational factors we cannot

be sure that these instruments are effectively measuring moral reasoning capacity.

I proposed that variation in reasoning would be due largely to personal/situational factors. The across situation variability supports this hypothesis, however, the situational factor does not account for all the variability. Subjects were more consistent within each situation than across situation, though their within-dilemma reasoning still included a great variety of stage responses. Evidently subjects draw from a variety of different levels of thought in order to most adequately address each moral situation.

Finally, this study was concerned with methodological considerations in using Rest's method of assessment. I discovered that methodological concerns further cloud the issue of assessment.

Subjects justified their response choice (DIT) with a different stage of reasoning (structured interview) 73% of the time. This inaccuracy could have been due to a variety of factors. For example, the "meaningless" items on the DIT were largely a source of error. Response choices were not phrased clearly. Also, different subjects often interpreted the same response choice in terms of a different stage of reasoning. In other words, the selection of a stage 4 answer by a particular subject tells us very little about that person's level of moral reasoning.

Other methodological problems were also documented. Subjects responded expediently to dilemmas of low interest, hence their answers did not represent their true reasoning capacity. Also, there were not enough adequate answers at each stage so some subjects were forced to make a choice which was not representative of their true thinking.

In general, the psychometric properties of the DIT are questionable at best. This is not only important in terms of future research in the area of moral development but we must also question past research which has been based on this assessment device.

One of the most problematic issues in moral development research is finding an adequate index of measurement. One of the major methodological questions addressed in this study was: Is Rest's index of moral reasoning (P%) an adequate continuous measure? The complex stage model is appealing because it does not label subjects in terms of a single stage. This study revealed, though, that the index of measurement used by the complex stage model (P%) is inadequate. Without a good measure of moral reasoning it is difficult to conduct meaningful research or design appropriate educational programs.

The construct of moral reasoning is more complex than the DIT and other assessment devices are prepared to handle. We have yet to discover an appropriate means to measure the construct. The results presented in this study indicate that a complex stage orientation is the most appropriate way to understand moral development. However, we need to achieve the same level of sophistication with respect to assessment that has been achieved in terms of theory. The problem remains to create a method of assessment which can fully accommodate all the implications of a complex stage model.

Appendix A:

SIMILARITIES

Directions: Please answer the following items by explaining how they are alike or similar.

Example A: Cat - Mouse

best response: both are animals
 fair response: both have tails and fur
 poor response: they chase each other
 confusing response: they eat the same food

Example B: Orange - Banana

best response: both are fruit
 fair response: both have peels
 poor response: both are good for you
 confusing response: they taste the same

Provide your best response in the space below the word pair. Remember to explain how the two items are similar, not how they are different.

1. Piano - Violin

2. North - West

3. Air - Water

4. Pound - Yard

5. Mountain - Lake

6. Scissors - Copper Pan

7. Egg - Seed

8. First - Last

9. Poem - Statue

10. Wood - Alcohol

11. Praise - Punishment

12. Liberty - Justice

13. Fly - Tree

14. Salt - Water

15. The Numbers 49 and 121

GENERAL INFORMATION

Directions:

Please answer these questions of general information in the space provided next to each item.

1. How far is it from
Paris to New York?

2. Who discovered the
South Pole?

3. What is the capitol of
Greece?

4. What does C.O.D. mean?

5. From what source do we
get turpentine?

6. How does yeast cause
dough to rise?

7. What is the population
of the United States?

8. How many senators are there
in the United States Senate?

9. What is the main theme of
the Book of Genesis?

10. What is ethnology?

OPINIONS ABOUT SOCIAL PROBLEMS

This questionnaire is aimed at understanding how people think about social problems. Different people often have different opinions about questions of right and wrong. There are no "right" answers in the way that there are right answers to math problems. We would like you to tell us what you think about several problem stories. The papers will be fed to a computer to find the average for the whole group, and no one will see your individual answers.

Please give us the following information:

Name _____ female
Age _____ Class and period _____ male
School _____

* * * * *

In this questionnaire you will be asked to give your opinions about several stories. Here is a story as an example. Read it, then turn to the next page.

Frank Jones has been thinking about buying a car. He is married, has two small children and earns an average income. The car he buys will be his family's only car. It will be used mostly to get to work and drive around town, but sometimes for vacation trips also. In trying to decide what car to buy, Frank Jones realized that there were a lot of questions to consider. On the next page there is a list of some of these questions.

If you were Frank Jones, how important would each of these questions be in deciding what car to buy?

On the left hand side of the page check one of the spaces by each question that could be considered.

GREAT importance
MUCH importance
SOME importance
LITTLE importance
NO importance

- ____ 1. Whether the car dealer was in the same
block as where Frank lives.
- ____ 2. Would a used car be more economical in the
long run than a new car.
- ____ 3. Whether the color was green, Frank's
favorite color.
- ____ 4. Whether the cubic inch displacement was at
least 200.
- ____ 5. Would a large, roomy car be better than a
compact car.
- ____ 6. Whether the front connibilies were
differential.

PART B. (Sample)

From the list of questions above, select the most important one of the whole group. Put the number of the most important question on the top line below. Do likewise for your 2nd, 3rd, and 4th most important choices.

Most important	<u>5</u>
Second most important	<u>2</u>
Third most important	<u>3</u>
Fourth most important	<u>1</u>

HEINZ AND THE DRUG

In Europe a woman was near death from a special kind of cancer. There was one drug that the doctors thought might save her. It was a form of radium that a druggist in the same town had recently discovered. The drug was expensive to make, but the druggist was charging ten times what the drug cost to make. He paid \$200 for the radium and charged \$2000 for a small dose of the drug. The sick woman's husband, Heinz, went to everyone he knew to borrow the money, but he could only get together about \$1000, which is half of what it cost. He told the druggist that his wife was dying, and asked him to sell it cheaper or let him pay later. But the druggist said, "No, I discovered the drug and I'm going to make money from it." So Heinz got desperate and began to think about breaking into the man's store to steal the drug for his wife.

Should Heinz steal the drug? (Check one)

☐ Should steal it

☐ Can't decide

☐ Should not steal it

HEINZ STORY

On the left hand side of the page
check one of the spaces by each
question to indicate its importance.

GREAT importance
MUCH importance
SOME importance
LITTLE importance
NO importance

— — — — —	1. Whether a community's laws are going to be upheld.
— — — — —	2. Isn't it only natural for a loving husband to care so much for his wife that he'd steal?
— — — — —	3. Is Heinz willing to risk getting shot as a burglar or going to jail for the chance that stealing the drug might help?
— — — — —	4. Whether Heinz is a professional wrestler, or has considerable influence with professional wrestlers.
— — — — —	5. Whether Heinz is stealing for himself or doing this solely to help someone else.
— — — — —	6. Whether the druggist's rights to his invention have to be respected.
— — — — —	7. Whether the essence of living is more encompassing than the termination of dying, socially and individually.
— — — — —	8. What values are going to be the basis for governing how people act towards each other.
— — — — —	9. Whether the druggist is going to be allowed to hide behind a worthless law which only protects the rich anyhow.
— — — — —	10. Whether the law in this case is getting in the way of the most basic claim of any member of society.
— — — — —	11. Whether the druggist deserves to be robbed for being so greedy and cruel.
— — — — —	12. Would stealing in such a case bring about more total good for the whole society or not.

From the list of questions above, select the four most important:

Most important	—
Second most important	—
Third most important	—
Fourth most important	—

STUDENT TAKE-OVER

At Harvard University a group of students, called the Students for a Democratic Society (SDS), believe that the University should not have an army ROTC program. SDS students are against the war in Viet Nam, and the army training program helps send men to fight in Viet Nam. The SDS students demanded that Harvard end the army ROTC training program as a university course. This would mean that Harvard students could not get army training as part of their regular course work and not get credit for it towards their degrees.

Agreeing with the SDS students, the Harvard professors voted to end the ROTC program as a university course. But the President of the University stated that he wanted to keep the army program on campus as a course. The SDS students felt that the President was not going to pay attention to the faculty vote or to their demands.

So, one day last April, two hundred SDS students walked into the university's administration building, and told everyone else to get out. They said they were doing this to force Harvard to get rid of the army training program as a course.

Should the students have taken over the administration building? (Check one)

☐ Yes, they should take it over

☐ Can't decide

☐ No, they should not take it over

GREAT importance
 MUCH importance
 SOME importance
 LITTLE importance
 NO importance

STUDENT TAKE-OVER

— — — — —	1. Are the students doing this to really help other people or are they doing it for kicks.
— — — — —	2. Do the students have any right to take over property that doesn't belong to them.
— — — — —	3. Do the students realize that they might be arrested and fined, and even expelled from school.
— — — — —	4. Would taking over the building in the long run benefit more people to a greater extent.
— — — — —	5. Whether the president stayed within the limits of his authority in ignoring the faculty vote.
— — — — —	6. Will the takeover anger the public and give all students a bad name.
— — — — —	7. Is taking over a building consistent with principles of justice.
— — — — —	8. Would allowing one student take-over encourage many other student take-overs.
— — — — —	9. Did the president bring this misunderstanding on himself by being so unreasonable and uncooperative.
— — — — —	10. Whether running the university ought to be in the hands of a few administrators or in the hands of all the people.
— — — — —	11. Are the students following principles which they believe are above the law.
— — — — —	12. Whether or not university decisions ought to be respected by students.

From the list of questions above, select the four most important:

Most important	_____
Second most important	_____
Third most important	_____
Fourth most important	_____

ESCAPED PRISONER

A man had been sentenced to prison for 10 years. After one year, however, he escaped from prison, moved to a new area of the country, and took on the name of Thompson. For 8 years he worked hard, and gradually he saved enough money to buy his own business. He was fair to his customers, gave his employees top wages, and gave most of his own profits to charity. Then one day Mrs. Jones, an old neighbor, recognized him as the man who had escaped from prison 8 years before, and whom the police had been looking for.

Should Mrs. Jones report Mr. Thompson to the police and have him sent back to prison? (Check one)

- ☐ Should report him
- ☐ Can't decide
- ☐ Should not report him

ESCAPED PRISONER

GREAT importance
 MUCH importance
 SOME importance
 LITTLE importance
 NO importance

— — — — —	1. Hasn't Mr. Thompson been good enough for such a long time to prove he isn't a bad person?
— — — — —	2. Everytime someone escapes punishment for a crime, doesn't that just encourage more crime?
— — — — —	3. Wouldn't we be better off without prisons and the oppression of our legal system?
— — — — —	4. Has Mr. Thompson really paid his debt to society?
— — — — —	5. Would society be failing what Mr. Thompson sould fairly expect?
— — — — —	6. What benefits would prisons be apart from society, especially for a charitable man?
— — — — —	7. How could anyone be so cruel and heartless as to send Mr. Thompson to prison?
— — — — —	8. Would it be fair to all the prisoners who had to serve out their full sentences if Mr. Thompson was let off?
— — — — —	9. Was Mrs. Jones a good friend of Mr. Thompson?
— — — — —	10. Wouldn't it be a citizen's duty to report an escaped criminal, regardless of the circumstances?
— — — — —	11. How would the will of the people and the public good best be served?
— — — — —	12. Would going to prison do any good for Mr. Thompson or protect anybody?

From the list of questions above, select the four most important:

Most important	_____
Second most important	_____
Third most important	_____
Fourth most important	_____

NEWSPAPER

Fred, a senior in high school, wanted to publish a mimeographed newspaper for students so that he could express many of his opinions. He wanted to speak out against the war in Viet Nam and to speak out against some of the school's rules, like the rule forbidding boys to wear long hair.

When Fred started his newspaper, he asked his principal for permission. The principal said it would be all right if before every publication Fred would turn in all his articles for the principal's approval. Fred agreed and turned in several articles for approval. The principal approved all of them and Fred published two issues of the paper in the next two weeks.

But the principal had not expected that Fred's newspaper would receive so much attention. Students were so excited by the paper that they began to organize protests against the hair regulation and other school rules. Angry parents objected to Fred's opinions. They phoned the principal telling him that the newspaper was unpatriotic and should not be published. As a result of the rising excitement, the principal ordered Fred to stop publishing. He gave a reason that Fred's activities were disruptive to the operation of the school.

Should the principal stop the newspaper? (Check one)

_____ Should stop it

_____ Can't decide

_____ Should not stop it

GREAT importance

MUCH importance

SOME importance

LITTLE importance

NO importance

NEWSPAPER

1. Is the principal more responsible to students or to parents?
2. Did the principal give his word that the newspaper could be published for a long time, or did he just promise to approve the newspaper one issue at a time?
3. Would the students start protesting even more if the principal stopped the newspaper?
4. When the welfare of the school is threatened, does the principal have the right to give orders to students?
5. Does the principal have the freedom of speech to say "no" in this case?
6. If the principal stopped the newspaper would he be preventing full discussion of important problems?
7. Whether the principal's order would make Fred lose faith in the principal?
8. Whether Fred was really loyal to his school and patriotic to his country.
9. What effect would stopping the paper have on the student's education in critical thinking and judgment?
10. Whether Fred was in any way violating the rights of others in publishing his own opinions.
11. Whether the principal should be influenced by some angry parents when it is the principal that knows best what is going on in the school.
12. Whether Fred was using the newspaper to stir up hatred and discontent.

From the list of questions above, select the four most important:

Most important _____

Second most important _____

Third most important _____

Fourth most important _____

WEBSTER

Mr. Webster was the owner and manager of a gas station. He wanted to hire another mechanic to help him, but good mechanics were hard to find. The only person he found who seemed to be a good mechanic was Mr. Lee, but he was Chinese. While Mr. Webster himself didn't have anything against orientals, he was afraid to hire Mr. Lee because many of his customers didn't like orientals. His customers might take their business elsewhere if Mr. Lee was working in the gas station.

When Mr. Lee asked Mr. Webster if he could have the job, Mr. Webster said that he had already hired somebody else. But Mr. Webster really had not hired anybody, because he could not find anybody who was a good mechanic besides Mr. Lee.

What should Mr. Webster have done? (Check one)

☐ Should have hired Mr. Lee

☐ Can't decide

☐ Should not have hired him

WEBSTER

GREAT importance
 MUCH importance
 SOME importance
 LITTLE importance
 NO importance

- ____ 1. Does the owner of a business have the right to make his own business decisions or not?
- ____ 2. Whether there is a law that forbids racial discrimination in hiring for jobs.
- ____ 3. Whether Mr. Webster is prejudiced against orientals himself or whether he means nothing personal in refusing the job.
- ____ 4. Whether hiring a good mechanic or paying attention to his customers' wishes would be best for his business.
- ____ 5. What individual differences ought to be relevant in deciding how society's roles are filled?
- ____ 6. Whether the greedy and competitive capitalistic system ought to be completely abandoned.
- ____ 7. Do a majority of people in Mr. Webster's society feel like his customers or are a majority against prejudice?
- ____ 8. Whether hiring capable men like Mr. Lee would use talents that would otherwise be lost to society.
- ____ 9. Would refusing the job to Mr. Lee be consistent with Mr. Webster's own moral beliefs?
- ____ 10. Could Mr. Webster be so hard-hearted as to refuse the job, knowing how much it means to Mr. Lee?
- ____ 11. Whether the Christian commandment to love your fellow man applies to this case.
- ____ 12. If someone's in need, shouldn't he be helped regardless of what you get back from him?

From the list of questions above, select the four most important:

Most important _____

Second most important _____

Third most important _____

Fourth most important _____

THE DOCTOR'S DILEMMA

A lady was dying of cancer which could not be cured and she had only about six months to live. She was in terrible pain, but she was so weak that a good dose of pain-killer like morphine would make her die sooner. She was delirious and almost crazy with pain, and in her calm periods, she would ask the doctor to give her enough morphine to kill her. She said she couldn't stand the pain and that she was going to die in a few months anyway.

What should the doctor do? (Check one)

- ☐ He should give the lady an overdose that will make her die
- ☐ Can't decide
- ☐ Should not give her an overdose

GREAT importance
 MUCH importance
 SOME importance
 LITTLE importance
 NO importance

DOCTOR

1. Whether the woman's family is in favor of giving her the overdose or not.
2. Is the doctor obligated by the same laws as everybody else if giving an overdose would be the same as killing her.
3. Whether people would be much better off without society regimenting their lives and even their deaths.
4. Whether the doctor could make it appear like an accident.
5. Does the state have the right to force continued existence on those who don't want to live.
6. What is the value of death prior to society's perspective on personal values.
7. Whether the doctor has sympathy for the woman's suffering or cares more about what society might think.
8. Is helping to end another's life ever a responsible act of cooperation.
9. Whether only God should decide when a person's life should end.
10. What values the doctor has set for himself in his own personal code of behavior.
11. Can society afford to let anybody end their lives when they want to.
12. Can society allow suicides or mercy killing and still protect the lives of individuals who want to live.

From the list of questions above, select the four most important.

Most important _____

Second most important _____

Third most important _____

Fourth most important _____

References

- Arbuthnot, J. Modification of moral judgment through role playing. Developmental Psychology, 1975, 11, 319-324.
- Bearison, D.J. The construct of regression: A Piagetian approach. Merrill-Palmer Quarterly, 1974, 20, 21-30.
- Damon, W. Structural-developmental theory and the study of moral development. In M. Windmiller, N. Lambert and E. Turiel (Eds.), Moral Development and Socialization. Boston: Allyn and Bacon, Inc., 1980.
- Doll, E.A., Vineland Social Maturity Scale. Minnesota: American Guidance Service, Inc., 1965.
- Edwards, J.B. A developmental study of the acquisition of some moral concepts in children aged 7 to 15. Educational Research, 1974, 16, 83-93.
- Flavell, J. The Developmental Psychology of Jean Piaget. Princeton: Van Nostrand, 1963.
- Gibbs, J.C. Kohlberg's stages of moral judgment: a constructive critique. Harvard Educational Review, 1977, 47, 43-61.
- Holstein, C.B. Irreversible, stepwise sequence in the development of moral judgment: a longitudinal study of males and females. Child Development, 1976, 47, 51-61.
- Kaufman, A.S. Intelligent Testing with the WISC-R. New York: John Wiley & Sons, 1979.
- Kohlberg, L. The development of modes of moral thinking and choice in the years 10 to 16. Unpublished doctoral dissertation, University of Chicago, 1958.
- Kohlberg, L. Moral and religious education and the public schools: a developmental view. In T. Sizer (Ed.), Religion and Public Education. Boston: Houghton-Mifflin, 1967.
- Kohlberg, L. The child as a moral philosopher. Psychology Today, 1968, 2, 25-30.
- Kohlberg, L. Stage and sequence: the cognitive-developmental approach to socialization. In D. Goslin (Ed.), Handbook of Socialization Theory and Research. Chicago: Rand McNally, 1969, 347-480.
- Kohlberg, L. Continuities in childhood and adult moral development revisited. In P.B. Baltes & K.W. Schaie (Eds.), Life-Span Developmental Psychology: Personality and Socialization. New York: Academic Press, 1973.

- Kohlberg, L. Moral stages and moralization: the cognitive-developmental approach. In T. Lickona (Ed.), Moral Development and Behavior: Theory, Research, and Social Issues. New York: Holt, Rinehart, and Winston, 1976.
- Kohlberg, L. Introductory note in J. Rest, Development in Judging Moral Issues. Minneapolis: University of Minnesota Press, 1979, pp vii-xvi.
- Kohlberg, L., Colby, A., Gibbs, J. & Speicher-Dubin, B. Moral Stage Scoring Manual. Cambridge: Center for Moral Education, Harvard University, 1976.
- Kohlberg, L. & Kramer, R. Continuities and discontinuities in childhood moral development. Human Development, 1969, 12, 93-120.
- Kuhn, D. Short term longitudinal evidence for the sequentiality of Kohlberg's early stages of moral judgment. Developmental Psychology, 1976, 12, 162-166.
- Kuhn, D., Langer, J., Kohlberg, L., & Hoan, N. The development of formal operations in logical and moral judgment. Genetic Psychology Monographs, 1977, 95, 97-188.
- Kurtines, W. & Grief, E. The development of moral thought: Review and evaluation of Kohlberg's approach. Psychological Bulletin, 1974, 81, 453-470.
- Lawrence, J.A. The component procedures of moral judgment making. Unpublished Doctoral Dissertation, University of Minnesota, 1978.
- Lee, L.C. The concomitant development of cognitive and moral modes of thought: A text of selected deductions from Piaget's theory. Genetic Psychology Monographs, 1971, 83, 93-146.
- Liebert, R. Moral development: a theoretical and empirical analysis. Unpublished manuscript, S.U.N.Y. at Stony Brook, 1978.
- Matarazzo, J.D. Wechsler's Measurement and Appraisal of Adult Intelligence. New York: Oxford University Press, 1979.
- McGeorge, C.M. Situational variation in level of moral judgment. British Journal of Educational Psychology, 1974, 44, 116-122.
- Rest, J. The hierarchical nature of moral judgment. Journal of Personality, 1973, 41, 86-109.
- Rest, J. Development in Judging Moral Issues. Minneapolis: University of Minnesota Press, 1979.
- Sattler, J. Assessment of Children's Intelligence. Philadelphia: W.B. Saunders Company, 1974.

- Siegal, M. Kohlberg versus Piaget: To what extent has one theory eclipsed the other? Merrill-Palmer Quarterly, 1980, 26, 285-297.
- Tomlinson-Keasy, C. The mediating role of cognitive development in moral judgment. Child Development, 1974, 45, 291-298.
- Turiel, E. An experimental test of the sequentiality of developmental stages in the child's moral judgments. Journal of Personality and Social Psychology, 1966, 3, 611-618.
- Wechsler, D. Manual for the Wechsler Intelligence Scale for Children. New York: Psychological Corporation, 1949.
- Wechsler, D. Manual for the Wechsler Adult Intelligence Scale. New York: Psychological Corporation, 1955.
- Whiteman, P.H. Development of children's moralistic judgments: age, sex, I.Q. and certain personal-experiential variables. Child Development, 1964, 35, 843-850.