The Effects of Story Consequence and Story Theme on the Moral Reasoning of Seventh, Ninth and Eleventh Graders

Kevin Plummer

Follow this and additional works at: https://digitalcommons.uri.edu/oa_diss
THE EFFECTS OF STORY CONSEQUENCE AND STORY THEME ON THE MORAL REASONING OF SEVENTH, NINTH AND ELEVENTH GRADERS

BY
KEVIN PLUMMER

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE DEGREE OF DOCTOR OF PHILOSOPHY IN PSYCHOLOGY

UNIVERSITY OF RHODE ISLAND
1983
Abstract

Various stage models have been proposed as a means of organizing and understanding the construct of moral reasoning. Kohlberg's (1976) simple stage model is the most rigid application of structuralism as he suggests that moral reasoning is based almost exclusively on the structural capacities of the reasoner. Significant criticism has been levied against this model as it has not fared well under scientific scrutiny. Rest (1979) has extended Kohlberg's paradigm in his design of a complex stage model. Rest (1979) suggests that situational factors play a major role in reasoning about moral problems. This results in complex stage patterns as individuals use a variety of different stage responses to address moral situations. Rest's measure of moral maturity, though, does not systematically evaluate situational influence. In addition it does not fully address complex stage patterns because only the highest moral stage usage (stages 5 and 6) is evaluated in determining an individual's level of moral maturity. Carroll (1981) has refined the assessment of moral reasoning in complex stage terms by addressing lower stage usage. His Rejection Scale effectively distinguishes reasoners in terms of moral maturity by measuring the degree to which reasoners reject the lower stage answers (presented in multiple choice format) when asked to judge whether the answer is a good reason for making a moral decision.
The purpose of the present research was to further refine the complex stage model by systematically examining situational influence in terms of Carroll's notion of rejection. This resulted in an evaluation of the relative influence of story consequences, story themes and age effects on moral reasoning.

Moral reasoning was measured by adapting Carroll's assessment device to include the systematic presentation of mild, moderate and severe consequence conditions. The psychometric properties of this new Rejection Scale were established in a pilot study. The pilot data indicates that this new measure is reliable across test administrations: it correlates well with other measures of moral reasoning, and the junior high and high school populations interact appropriately with the test materials (e.g., comprehension of test items, understanding and following directions).

A total sample of 90 male and female students from 7th, 9th, and 11th grade were randomly selected to participate in the main study. An explanation of the study was provided and consent was secured from students, parents, and teachers. All students completed the Rejection Scale (see appendix B). The data were analyzed, using a three by three by six ANOVA with repeated measures, by grade, consequence condition, and story theme. In addition, Scalogram Analysis was used to determine the scalability of the data by consequence condition, story theme, and age.

The Scalogram Analysis demonstrated a very clear consequence effect as dilemmas were scalable by consequence condition with coefficients of reproducibility ranging from .85 to 1.0. This effect occurred for each age category.
The Analysis of Variance provided additional support for this consequence effect. Each age category showed a strong consequence effect and the mild, moderate and severe consequence conditions all differed significantly from each other in each dilemma.

An age effect was also demonstrated with Scalogram Analysis. The 7th graders had the lowest average stage rating, followed by the 9th graders, and the 11th graders achieved the highest average stage ratings.

This age effect was supported by the Analysis of Variance. There was a significant difference between all three ages at mild and moderate consequences, while 7th and 11th graders differed significantly at the severe consequence condition.

Finally, the dilemma effect was demonstrated with Scalogram Analysis and Analysis of Variance. The dilemmas differed in their average stage rating, and these differences held up across all three age groups and consequence conditions. In addition, the dilemmas differed in their degree of scalability. The Analysis of Variance showed significant differences between most of the dilemmas at mild and moderate consequences.

The results from this study demonstrate that story consequences and story themes are significant functional aspects of moral situations. In addition, the structural aspect of Kohlberg's stage model was also demonstrated as a developmental progression was documented between 7th, 9th, and 11th graders. These results are discussed in terms of a model of moral reasoning which incorporates both functional and structural aspects.
ACKNOWLEDGEMENT

I would like to thank my major professor, Dr. Robert Germain, for his support, encouragement and enthusiasm throughout this project and throughout my graduate training in general. Dr. Germain's sensitivity and concern for human development and professional growth is reflected in this project as well as in my training and growth as a professional psychologist.

I would also like to extend appreciation to my committee members, Dr. Albert Lott, Dr. Jerome Schaffran and Dr. William Vosburgh. My repeated requests for their input were always met with insightful feedback which has subsequently become an important part of this project.

Special thanks go to the Stonington School personnel, Mr. Norman Grabowski, Mr. Harold Hoyland, Mr. Philip House, Mr. Burt Stafford, and Ms. Jean Kenney for their cooperation in facilitating the collection of data.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>STATEMENT OF THE PROBLEM</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>STRUCTURAL STAGE THEORY</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Structured Whole</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Invariant Sequence</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Hierarchical Development</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Theoretical Conclusions</td>
<td>5</td>
</tr>
<tr>
<td>III</td>
<td>KOHLBERG'S STRUCTURAL THEORY: THE SIMPLE STAGE MODEL</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Six Stages of Reasoning</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>The Simple Stage Model</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Theoretical Conclusions</td>
<td>12</td>
</tr>
<tr>
<td>IV</td>
<td>THEORETICAL CRITIQUE</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Structured Whole</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Invariant Sequence</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Naturalism</td>
<td>16</td>
</tr>
<tr>
<td>V</td>
<td>REST'S STRUCTURAL THEORY: THE COMPLEX STAGE MODEL</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Complex Stage Model</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Theoretical Critique</td>
<td>20</td>
</tr>
<tr>
<td>VI</td>
<td>STRUCTURAL-FUNCTIONAL ASPECTS</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>The Functional Aspects</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>The Functional Aspect of Story Consequences</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Theoretical Conclusions</td>
<td>28</td>
</tr>
<tr>
<td>VII</td>
<td>RESEARCH QUESTIONS.</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Question One</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Question Two</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Question Three</td>
<td>30</td>
</tr>
<tr>
<td>VIII</td>
<td>METHOD.</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Pilot Study</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Subjects</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Measures</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Rejection Scale</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Structured Interview</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Defining Issues Test</td>
<td>35</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedure</td>
<td>38</td>
</tr>
<tr>
<td>Main Study</td>
<td>39</td>
</tr>
<tr>
<td>Subjects</td>
<td>39</td>
</tr>
<tr>
<td>Measures</td>
<td>40</td>
</tr>
<tr>
<td>Rejection Scale</td>
<td>40</td>
</tr>
<tr>
<td>Procedure</td>
<td>42</td>
</tr>
</tbody>
</table>

IX  HYPOTHESES AND PREDICTIONS | 43
- Hypothesis One | 43
  Predictions | 43
- Hypothesis Two | 43
  Predictions | 43
- Hypothesis Three | 44
  Predictions | 44

X  RESULTS | 45
- Pilot Study | 45
  Reliability | 45
  Validity | 46
  Concurrent Validity | 46
  Criterion Group Validity | 46
  Content Validity | 55
  Structured Interview | 55
- Pilot Summary | 57
- Main Study | 59
  Hypothesis One | 59
    Scalogram Analysis | 60
    Analysis of Variance | 60
    Hypothesis One Summary | 72
  Hypothesis Two | 77
    Scalogram Analysis | 77
    Analysis of Variance | 77
    Hypothesis Two Summary | 79
  Hypothesis Three | 79
    Scalogram Analysis | 80
    Analysis of Variance | 84
    Hypothesis Three Summary | 85
- Summary | 89

XI  DISCUSSION | 92
- Purpose of Study, Review | 92
- Pilot Study | 94
  Summary And Review | 94
  Implications And Limitations | 95
- Main Study | 98
  Situational Effects | 98
# TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary of Results</td>
<td>98</td>
</tr>
<tr>
<td>Implications And Limitations</td>
<td>99</td>
</tr>
<tr>
<td>Developmental Trend</td>
<td>103</td>
</tr>
<tr>
<td>Summary Of Results</td>
<td>104</td>
</tr>
<tr>
<td>Implications And Limitations</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>XII CONCLUSIONS</td>
<td>106</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>108</td>
</tr>
<tr>
<td>APPENDICES</td>
<td></td>
</tr>
<tr>
<td>A. Defining Issues Test</td>
<td>111</td>
</tr>
<tr>
<td>B. Rejection Scale</td>
<td>125</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>139</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>Retest Reliability Coefficients By Dilemma, Consequence and Total Score.</td>
</tr>
<tr>
<td>2</td>
<td>Three by Three by Six Analysis Of Variance On Pilot Data Rejection Scores By Age, Consequence, and Dilemma.</td>
</tr>
<tr>
<td>3</td>
<td>Simple Effects Tests On The Interaction Of Age At Each Consequence Condition (Pilot Study).</td>
</tr>
<tr>
<td>4</td>
<td>Simple Effects Tests On The Interaction Of Consequence At Each Age (Pilot Study).</td>
</tr>
<tr>
<td>5</td>
<td>Scalogram Analysis For Grade Seven By Dilemma.</td>
</tr>
<tr>
<td>6</td>
<td>Scalogram Analysis For Grade Nine By Dilemma</td>
</tr>
<tr>
<td>7</td>
<td>Scalogram Analysis For Grade Eleven By Dilemma</td>
</tr>
<tr>
<td>8</td>
<td>Scalogram Analysis Totals For All Three Grades And All Six Dilemmas</td>
</tr>
<tr>
<td>9</td>
<td>Three by Three by Six Analysis Of Variance On Rejection Scores From Main Study By Age, Consequence And Dilemma.</td>
</tr>
<tr>
<td>10</td>
<td>Three by Three by Six Analysis Of Variance On Rejection Scores Combined From The Pilot And Main Study By Age, Consequence, and Dilemma.</td>
</tr>
<tr>
<td>11</td>
<td>The Proportion Of Variance Accounted For By Variables In The Main Study And Main/Pilot Combined.</td>
</tr>
<tr>
<td>12</td>
<td>Simple Effects Tests On The Interaction Of Consequences At Each Age (Main Study And Combined Data).</td>
</tr>
<tr>
<td>13</td>
<td>Simple Effects Tests On The Interaction Of Age At Each Consequence Condition (Main Study And Combined Data).</td>
</tr>
<tr>
<td>14</td>
<td>Simple Effects Tests On The Interaction Of Consequence At Each Dilemma (Main Study And Combined Data).</td>
</tr>
<tr>
<td>15</td>
<td>Simple Effects Tests On The Interaction Of Dilemma At Each Consequence Condition (Main Study And Combined Data).</td>
</tr>
<tr>
<td>16</td>
<td>The Rating Of Consequence By Age And Dilemma.</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>A Graphic Presentation Of The Simple Stage Model</td>
</tr>
<tr>
<td>2</td>
<td>The Complex Stage Model. A Disjunctive Scale Formed By The Relative Usage Of Six Different Stages Across Development.</td>
</tr>
<tr>
<td>3</td>
<td>Design Sketch For Three-Way ANOVA With Repeated Measures, For Story Theme, Consequence Condition, And Grade Level.</td>
</tr>
<tr>
<td>4</td>
<td>The Difference Between Ages At Each Consequence Condition on Pilot Data Rejection Scores</td>
</tr>
<tr>
<td>5</td>
<td>The Difference Between Consequence Conditions At Each Age on Pilot Data Rejection Scores</td>
</tr>
<tr>
<td>6</td>
<td>The Difference Between Consequence Conditions At Each Age on Main Study Rejection Scores</td>
</tr>
<tr>
<td>7</td>
<td>The Difference Between Consequence Conditions At Each Age on Combined Data Rejection Scores</td>
</tr>
<tr>
<td>8</td>
<td>The Difference Between Ages At Each Consequence Condition on Main Study Rejection Scores</td>
</tr>
<tr>
<td>9</td>
<td>The Difference Between Ages At Each Consequence Condition on Combined Data Rejection Scores</td>
</tr>
<tr>
<td>10</td>
<td>The Difference Between Consequence Conditions At Each Dilemma on Main Study Rejection Scores</td>
</tr>
<tr>
<td>11</td>
<td>The Difference Between Consequence Conditions At Each Dilemma on Combined Data Rejection Scores</td>
</tr>
<tr>
<td>12</td>
<td>The Difference Between Dilemmas At Each Consequence Condition on Main Study Rejection Scores</td>
</tr>
<tr>
<td>13</td>
<td>The Difference Between Dilemmas At Each Consequence Condition on Combined Data Rejection Scores</td>
</tr>
</tbody>
</table>
CHAPTER I

STATEMENT OF THE PROBLEM

Various models have been proposed as a means of organizing and understanding moral development. The most noteworthy models are Kohlberg's simple stage model and Rest's complex stage approach. Kohlberg's simple stage model has not fared well under scientific scrutiny as researchers have failed to find support for his notion of structured whole, invariant sequence, and cultural universality. Rest has extended Kohlberg's six stage theory of moral development to account for situational influences on moral reasoning. His complex stage model is designed to measure moral reasoning by examining the variety of different stage responses which an individual uses to address each dilemma. The model is also designed to accommodate variation in stage responses across dilemmas. This is an advance over Kohlberg's simple stage model which assumes that individuals use one stage of reasoning, primarily, to address various moral situations. Rest's measure of moral maturity, though, does not systematically evaluate situational influence and it does not fully address stage pattern complexity because only the highest moral stage usage (stages 5 and 6) is evaluated in determining an individual's level of moral maturity.

Carroll has developed a more refined means of complex stage model measurement by looking at the degree to which individuals use lower stage reasoning. This has been shown to be more effective in distinguishing high and low level reasoners than measures which deal
strictly with preference for higher level thought. Carroll's device, however, also deals with situational variables unsystematically.

One remaining task, then, is to systematically examine situational factors in terms of the complex stage model, using Carroll's notion of moral maturity.

This situational investigation places this study within the bounds of Liebert's (1978) functional approach to moral reasoning. According to the functional approach individuals are not predisposed to think in certain ways due to their structural capacity. Rather, individuals are tuned into the payoff matrix of various situations and will respond to moral dilemmas in a way that will bring about the greatest payoff (Liebert, 1978). This functional approach to moral development for example, suggests that story theme and story consequences may be relevant functional aspects of moral situations. Studies conducted with children show that story consequence is a relevant situational factor (Armsby, 1971; Costanzo, 1973; Gutkin, 1972; Hewitt, 1975). Other studies show that story themes may have situational influence (Nucci, 1981).

The purpose of this study was to examine the effects of story consequences and story themes with individuals at various stages of moral maturity. This is a further refinement of the complex stage model. The goal was to create a hierarchical scale of development by showing an increase in use of Kohlberg's preconventional (Level I) reasoning under conditions of increasing severity of story consequence and across different story themes.

While Rest proposes a complex stage model and suggests that individuals use a variety of different kinds of reasoning in different moral situations, he does not systematically describe how
individuals are influenced by these different situational factors and his present model is only capable of gross stage descriptions. The model developed in this study outlines the specific effects of story consequences and story themes on the moral reasoning of individuals at various points along the continuum of moral maturity. Furthermore, it increases our descriptive capacities of the construct as one level of reasoning (i.e., the preconventional level) becomes subdivided in terms of a 7 step hierarchical sequence. This provides additional support for the continuous, as opposed to the discrete, model of development with respect to Kohlberg's six stage sequence. Finally, the results from this study represent one step in a potentially long line of research to build a structural-functional model which addresses a multitude of situational factors at each structural level.

CHAPTER II

STRUCTURAL STAGE THEORY

The structural-developmental approach, according to Kohlberg and other developmental researchers, is based on the key assumption that development is a process of restructuring or reorganizing internal thought patterns. Developmental change, then represents internal reorganization rather than a mere accumulation of new information. These internal structures account for the various stages in the development of moral reasoning (Damon, 1980). These stages are internally consistent and qualitatively different from each other, and development is invariantly sequenced and 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 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,
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). The rate of development may be a function of culture or other environmental influences but the basic pattern of development should be consistent within the species.

Theoretical Conclusions

According to the structural stage theory just reviewed, 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 their reasoning should be consistent 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.

CHAPTER III

KOHLBORG'S STRUCTURAL THEORY: THE SIMPLE STAGE MODEL

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 moral 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. The level of interest in this study is Level I, the preconventional level.

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, Colby, Gibbs, & Speicher-Dubin, 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 are 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 10) 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). 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 and hierarchical. Rest (1973) demonstrated that Kohlberg's six stages form an invariant sequence as subject's comprehension of stage prototype statements formed a Guttman scale.

The stages form a Guttman scale of comprehension as the attainment of a more advanced stage presupposes attainment of the simpler, less advanced stages. A new stage does not simply replace a previous stage,
FIGURE 1. A GRAPHIC PRESENTATION OF THE SIMPLE STAGE MODEL
(TAKEN FROM REST, 1979, P. 52.)
nor is it added to it, but rather the new stage is a transformation of elements of the old along with new elements into a new emergent structure... Each succeeding stage is said to be an advance over the preceding stage in being a more differentiated and integrated structure (Rest, 1973, p. 86).

Furthermore, Kohlberg and Kramer (1969) report that once a new more advanced stage has emerged, regression to a previous stage is rare.

In addition, Kohlberg et al, (1976) suggest,

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).

Developmental change, in Kohlberg's view, involves structural reorganization. "Development involves change in the general shape, pattern, or organization of responses rather than change in the frequency or intensity of emission of an already patterned response" (Kohlberg & Kramer, 1969, p. 98). Rest adds,

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 the individual will use this mode of thought all the time (p. 50).

Kohlberg (1976) concludes "individuals should be consistently at a stage unless they are in transition to the next stage" (p. 74). He believes that the stage forms a clustered whole as there is a general factor of moral stage cross-cutting all moral situations.

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 assumes that individuals will evidence the same stage of reasoning in response to different moral situations. In Kohlberg's simple stage view, the characteristics of structural theory do not
allow for stage mixture unless the individual is in transition to the next highest stage.

**Theoretical Conclusions**

According to Kohlberg, stages of moral development are organized according to a logic which is internally consistent and qualitatively different from the logic of every other stage. 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).

**CHAPTER IV**

**THEORETICAL CRITIQUE**

Kohlberg's paradigm is a rigid application of structural developmental theory. Many researchers have reviewed the structural aspects of Kohlberg's model (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 six stage sequence of development.
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.

McGeorge (1974) noted significant variation between dilemmas in the responses of forty 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 present author also documented significant stage mixture (Plummer, 1982). Seventy college sophomores each used a full range of reasoning (stages 2-5) to address different moral situations, rather than the dominant and +1 reasoning suggested by the simple stage model. Reasoning within each moral situation was also characterized by a full range of stage responses. Evidently subjects drew from a variety of different levels of thought in an attempt to address each dilemma most adequately. Assigning a single stage rating or expecting subjects to justify their answers with reasoning on just one stage, then, is unrealistic (Plummer, 1982).

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 of seventy-two 10, 13 and 16 year old boys 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 developing, post hoc, 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. That is, high school graduates shift from a societal rule orientation to an idealistic view of the world. They assume that everyone's ideas are equally valid, thus their judgments should be considered equal to others. As a result, this stage of development seemed to resemble the stage 2 self interest orientation.

Holstein (1976) conducted a three year longitudinal study, collecting data from 53 families, including both parents, as well as their sons and daughters who were 13-years-old at the first assessment and age 16 at the second 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 fifty 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 rather than admit that the theory is flawed. 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 progression and regression occurred in Kuhn's study but it was not documented due to the insensitivity of global scoring. At best, this study provides no evidence for Kohlberg's sequence of moral development.

Kohlberg's six stage invariant 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.

**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. He collected data in the United States, 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.

CHAPTER V

REST'S STRUCTURAL THEORY: THE COMPLEX STAGE MODEL

Many researchers have failed to find support for Kohlberg's version of structural developmental theory. Specific theoretical problems include lack of evidence for his notion of invariant sequence, structured whole, and naturalism. The evidence indicates that a rigid application of structural theory, such as the simple stage model, 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). Thus, Rest's complex stage model is an attempt to adapt Kohlberg's structural theory to consider the situational aspects along with the structural capacities of the reasoner.

Complex Stage Model

The complex stage model (see figure 2) illustrates the extension of structural theory to account for the different types of reasoning
FIGURE 2. THE COMPLEX STAGE MODEL. A DISSJUNCTIVE SCALE FORMED BY THE RELATIVE USAGE OF SIX DIFFERENT STAGES ACROSS DEVELOPMENT. (TAKEN FROM REST, 1979, P. 224)
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 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 in a particular stage is wrong. Rather than wondering whether a subject is in 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). 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 stage theory
which were violated in Kohlberg's structuralism are maintained by accounting for development in terms of the complex stage model.

Theoretical Critique

The complex stage model describes development as a continuous rather than discrete, step-like process. Development is characterized by increasing probability of higher stage usage and simultaneous decreasing probability of lower stage usage. This is because individuals who are capable of using post-conventional reasoning in some instances may use a variety of lower level responses in other situations as individuals respond to situational demands.

This model seems intuitively appealing as it links structural theory with a functional approach. Rest has not only designed a model which integrates these two approaches, he has also designed a method of assessment which yields a profile of stage usage across a variety of moral situations. However, this complex stage approach falls short of a thorough understanding of moral development. The significance of situational factors, for example, is the foundation of the complex approach, yet Rest has done little to explicate this facet. Rest does little more than suggest that situational factors must be relevant and individual stage patterns of moral reasoning are more complex than the simple stage model proposes. This leaves many questions unanswered:

1) What are the relevant situational factors?  
2) What is the precise effect of situational factors?  
3) Do these factors have a consistent influence on all reasoners at every step in the development sequence?

Given that situational factors play a major role in moral reasoning, we need to understand specifically what these factors are and the way they operate throughout the full range of moral development. We cannot assume, for example, that Rest's (1979) six
story moral preference test or Kohlberg's (1979) three story moral interview technique are systematically sampling the full range of relevant situational factors. In fact, the situational factor is dealt with quite haphazardly in these moral assessment devices as the total array of potentially influential factors varies unsystematically. We need to understand situational influence more fully by dealing with it more systematically in moral assessment.

Another difficulty with the complex stage model is the index of measurement ($P\%$). $P\%$ reflects the percentage of time an individual uses stages 5 or 6 while completing Rest's test of moral maturity. An increase in $P\%$ is considered developmental progression. This is based on the assumption that individuals decrease in lower stage thought as they increase in moral reasoning at stages 5 and 6. While $P\%$ is reported to be a reliable and valid index of moral maturity, it largely ignores development in terms of lower stage patterns. While an individual's $P\%$ may remain constant over time, development may indeed be occurring as lower stage patterns shift. We cannot be sure that $P\%$ will always reflect the shift in lower stage preference. In fact, the present author demonstrated that individuals with the same $P\%$ did not have similar lower stage patterns (Plummer, 1982).

Even though $P\%$ is reported to be a psychometrically sound index of moral maturity, it is deficient as a true measure of the construct, to the extent that it largely ignores the four lower stages in a six stage model without adequate justification. We need a more comprehensive way to look at the full spectrum of development.

Carroll and Rest (1981) have designed a method of moral assessment which considers development in terms of lower stage usage. Both Rest (1973) and Carroll and Rest (1981) report that the highest stages of
reasoning tend to be preferred by all reasoners regardless of their developmental level. Thus preference for higher stage responses makes no distinction between high and low level reasoners.

In each preference study, virtually all subjects preferred the highest stage statements. Hence preference of statements did not differentiate subjects. Differences were noted in the rating of lower stage items by age groups (Carroll & Rest, 1981, p. 538).

Thus, Carroll and Rest began looking at moral development in terms of rejection or "growing out" of the lower stage thought. This notion is based theoretically on the principle of hierarchy. The development of higher stage thought requires the incorporation of earlier, simpler stage reasoning. As individuals develop a more encompassing approach the lower stages are seen as simplistic or less adequate (Carroll & Rest, 1981).

Carroll and Rest (1981) discovered definite age trends in the rejection of lower stage reasoning. Older subjects (11th graders) rejected lower stage reasoning more frequently than younger subjects (7th graders). The groups did not differ on the rejection of higher stage thought. Carroll and Rest (1981) conclude, "when individuals give up thinking at stages lower than their model stage, this can be viewed as developmental advance in the sense of consolidation at the higher stage" (p. 43). Thus, an understanding of the development of moral reasoning can be made more comprehensive by considering rejection patterns of lower stage thought.

The "rejection score" approach addresses the concern outlined earlier, that is, to look more comprehensively at lower stage patterns. While rejection of lower stage reasoning may be a function of moral maturity, it may also vary differentially as a function of situational factors. Thus, we need to deal more systematically with
the situational factor within the framework of rejection of lower stage reasoning.

CHAPTER VI

STRUCTURAL-FUNCTIONAL ASPECTS

Various structural theories have been discussed, including Kohlberg's simple stage model and Rest's complex stage approach. Kohlberg's model is a rigid application of structuralism as he suggests that one's cognitive framework, in terms of a particular stage of reasoning, is the pervasive factor in the moral reasoning manifested in response to moral dilemmas. Rest has extended Kohlberg's model to conclude that one is not restricted to using just one stage of reasoning at a time. He believes that reasoners utilize a full range of stage responses to most adequately address each moral situation. He also suggests that situational factors, inherent in all moral dilemmas, play an important role in terms of influencing moral reasoning. His assessment of moral reasoning, though, a) does not systematically address situational influence and b) his index of measurement does not fully address stage pattern complexity because it largely ignores lower stage usage without sufficient justification. Carroll has extended Rest's work to look more closely at lower stage patterns. Carroll's measure yields a more comprehensive view of the complex stage patterns as it looks at moral maturity in terms of rejection of lower stage thought. Given Carroll's refinement of the structural stage model, the task remains to systematically examine situational influence. The functional approach to moral reasoning
which was outlined earlier has suggested some relevant situational factors. In addition, some studies have been conducted with young children which systematically examine story consequences as a functional factor. None of these studies has been conducted with adolescents or adults and none of them has incorporated Carroll's refinement of the structural model.

Functional Aspects

One feature of this functional approach is an emphasis on the importance of situational factors. It is assumed that situational contingencies influence all responses within the moral sphere, and that they also exert a considerable effect on the moral judgments and justifications which children and adults will advance in response to hypothetical moral dilemmas (Liebert, 1978, p. 45).

Functionalists do not assume that individuals are predisposed to think in certain ways due to their structural capacity. Rather, individuals are tuned into the payoff matrix of various situations and will respond to moral dilemmas in a way that will bring about the greatest payoff. The most important question in this functional approach, then, is: Which elements of the situation are functional? (Liebert, 1978).

Another important feature of this functional approach is the emphasis on the growth of knowledge and experience. The age-correlated changes, which are often referred to as development due to structural change, may be a reflection of changes in the information, knowledge, and experience that naturally comes with age (Liebert, 1978).

With increasing age and experience human beings acquire the ability to do many new things... In a very real way, for example, the average six-year-old child is barred from Kohlbergian stage 4 moral reasoning because of ignorance of the workings of the law (rather than limitations in structural capacity) (Liebert, 1978, p. 48).
One of the most salient functional variables in moral situations is the objectively presented or subjectively perceived or anticipated consequences. Individuals have the ability to infer consequences from moral situations so as to construct sophisticated payoff matrices.

A person's moral judgments, moral justifications, or personal conduct may be influenced by the cues the person is given before or while confronting a moral dilemma and these cues may be interpreted by the person as having implications about the actual payoff matrix that prevails in this situation (Liebert, 1978, p. 58).

Little has been done to systematically examine the functional aspects of moral situations using Kohlberg's, Rest's, or Carroll's paradigm. However, some researchers have investigated situational factors with children using Piaget's paradigm.

The Functional Aspect of Story Consequences

Considerable research has been conducted to examine the effects of story consequences on children's moral reasoning using Piaget's theoretical framework. The Piagetian assessment technique is to present children with story pairs and ask them which child (each story has one main child character) is naughtier. In one story the child causes severe damage while behaving according to good intentions. In the other story the child causes minimal damage while acting out of bad intentions. Children below seven to eight years judge according to damage level or consequences and older children consider intentions when making their moral judgments (Piaget, 1965).

Armsby (1971) noticed that regardless of intentions, the damage level in the two stories is not the same. In one instance considerable damage is caused and in the other story only minimal damage occurs. Since two variables, damage and intention, are being manipu-
lated at once it is difficult to determine how and why children change their focus. Armsby (1971) held intentions constant, at a "good" level, and varied the severity of damage. Results indicate that fewer children make intentionality judgments as the severity of consequences increases in the stories of good intention. As consequences become more severe intentionality loses its potency as the relevant variable. When children judge according to intention on Piaget's measure it does not mean the child considers consequences irrelevant. It may mean that consequences are being considered and put on low priority due to their level of severity.

Hewitt (1975) examined the effects of provocation, intentions and consequences on children's moral judgments. Serious injury to the victim (severe consequences) evoked more negative evaluations of the harm doers than did minor injury (mild consequences) when intentions and provocation were held constant. Hewitt noticed that even when subjects were capable of making judgements according to intentions they still had consideration for consequences.

Costanzo (1973) believes that a major problem with the Piagetian assessment device is that regardless of intentions the consequences are always negative. When Costanzo structured stories to include positive consequences it was discovered that the intentions of the story character were equally important for all age groups. Costanzo concluded that young children make their choices on the basis of consequences because it is the more salient and identifiable cue, not because intentions are irrelevant.

Gutkin (1972) examined the effect of systematic story changes on intentionality. This involved the systematic variation of two levels of intention (good and bad) and consequence (mild and severe).
Children who judged according to intention, using the standard Piagetian assessment, judged in terms of consequences when intentions were held constant. Based on these findings Gutkin proposed a four step developmental sequence to supplement Piaget's two stage model. Step one individuals always judge according to consequences regardless of intention. Step two individuals believe that intentions are important but only if level of damage (consequences) is equal. Step three reasoners focus on intentions but will shift to a consequence focus if intentions are equal. Step four subjects believe that damage is never a relevant factor.

Thus, it has been demonstrated that subjects can be influenced to shift their reasoning depending upon the severity of consequences and intentions. Most specifically, when one factor is held constant it is likely that the other variable plays a major role in the child's moral decision.

In addition, Nucci (1981) sorted moral dilemmas into categories of personal, social/conventional, and justice. It was discovered that these story themes had differential effects on moral reasoning in terms of perceived consequences. There was a significant tendency for subjects at college and high school age to judge an action as most wrong because they perceived the adverse consequences of the act as affecting many others. Dilemmas that were judged least wrong by this population were regarding rules or the prohibition of an act by an authority. Also, when the transgression creates disorder it is judged less severely than acts which resulted in adversely affecting many people.
Theoretical Conclusions: The Structural-Functional Model

The functional approach suggests that experience and situational factors, such as perceived consequences and story theme, are major factors influencing moral reasoning. The structural model has provided a framework for organizing different types of reasoning. This structural model is hierarchical and the sequence of development is invariant. The functional model has provided a number of situational factors which may impinge upon the reasoner in moral situations. The task remains to build a model which is capable of organizing the development of moral reasoning according to both structural and functional features.

Moral reasoning then, may best be understood through a combined structural-functional approach. This is an approach which illustrates how functional factors, such as story consequences and story themes impinge on the moral reasoning of individuals who differ in their structural capacity. Cognitive structures may manifest themselves in the form of a particular stage of reasoning, to the extent that an individual is consolidated at that stage and situational influence is minimal. Little is understood about this structural-functional interaction as the functional factors have not been systematically examined in the structural models.

Story consequences and story themes are two functional aspects examined in this study. These factors were examined in terms of their interaction with the structural-developmental period between 7th and 11th grade. Consequences varied in their severity and were hypothesized to influence moral reasoning by inducing subjects to reason at lower levels. Resistance to consequence influence, however, was expected to vary as a function of structural development.
(as depicted by the three grade levels). It was not clear exactly how story themes would influence reasoning, however, it was hypothesized that there would be differences in the reasoning used to address different moral stories which could not be accounted for by the severity of consequences.

CHAPTER VII

RESEARCH QUESTIONS

Question One

The first question addressed by this project is whether story consequences interact with an individual's structural capacity (the reasoner's stage or level of moral maturity) to form a cumulative or Guttman scale. That is, as story consequences increase in severity, will individuals be less likely to reject consequence level reasoning (Kohlberg's Level I, preconventional) as an appropriate answer to moral dilemmas? This first hypothesis is stated as follows:

Story consequence (mild, moderate, and severe) is a functional aspect of the moral dilemma which interacts with the reasoner's structural capacity to yield differential effects on the moral reasoning of junior high and high school students.

Question Two

The second question addressed by this project is whether the cumulative or Guttman scale described in question one can effectively distinguish the three grade levels of subjects. Is there a developmental trend with 7th, 9th, and 11th graders, marked by a lower
position on the Guttman scale of younger subjects, while older subjects occupy a higher step on this scale? Also, will each age group show the same story consequence effect? This second hypothesis is stated as follows:

The cumulative scale has criterion group validity as it represents a developmental continuum between younger and older students.

Question Three

The third question addressed by this project is whether story theme is a relevant functional aspect of moral reasoning. This third hypothesis is stated as follows:

Story theme is a relevant functional aspect of moral reasoning as different story themes have differential effects on the moral reasoning of junior high and high school students.

CHAPTER VIII

METHOD

Pilot Study

The primary measure used in this study is the Rejection Scale (see Appendix B and Measures section). The Rejection Scale used in this study, however, has been altered from the standardized version (Carroll, 1974) in order to test a new set of hypotheses. Carroll's scale presents four moral dilemmas with no control of functional factors such as story consequences. Carroll's instrument has been
altered to present three different consequence conditions for each of six moral dilemmas. Thus, the pilot study was used to establish reliability (test-retest) and validity (construct and criterion group) of a new Rejection Scale. Thus, the information gleaned from this pilot study was used to establish the psychometric properties of the new Rejection Scale, and facilitate the refinement of the instrument if necessary.

Subjects

The subjects in this pilot study included a total sample of 90 male and female junior high and high school students, 30 from each of the 7th, 9th, and 11th grades. Each group was comprised of approximately equal numbers of males and females. This sample was selected from a small rural school system in eastern Connecticut. The students were selected randomly from study halls which are part of every student's required program. The students participated through an arrangement made with parents, students, teachers, and school administration, after an explanation of the project was offered. Student confidentiality was maintained as no student names were used in the processing of results.

This project was reviewed and approved by the superintendent, director of special education, and school principals. In addition, it was reviewed and approved by the Human Subjects Review Board at the University of Rhode Island.

Measures

Rejection Scale. The Rejection Scale is a measure of moral maturity based on the subject's rejection of lower stage reasoning.
Carroll and Rest (1981) demonstrated that rejection of lower stage reasoning increases with age and moral maturity. Carroll and Rest (1981) conclude, "when individuals give up thinking at stages lower than their modal stage, this can be viewed as developmental advance in the sense of consolidation at the higher stage" (p. 43). This notion of measuring moral maturity through rejection of reasoning led to the development of Carroll's objective instrument. Carroll's measure is based on Kohlberg's six stage model. His measure consists of four moral dilemmas each of which is followed by reasons for action, in the moral situation. Ten reasons, two at each of Kohlberg's stages 1-5, comprise the response choices. Subjects are expected to rate each reason in terms of one of four categories: 1) I accept the reason; 2) I tend to accept the reason; 3) I tend to reject the reason; or 4) I reject the reason. Scoring ranges from 1 point for every "I accept the reason" to 4 points for every "I reject the reason." In this manner a rejection rating is established. Theoretically the higher stage subjects should have higher rejection ratings.

Criterion group validity was established in Carroll's 1981 study as older subjects (11th graders) did indeed have higher rejection ratings than younger subjects (7th graders). Carroll (1974) also reports internal consistency for each stage (1 through 5), using Hoyt reliability coefficients, in the range of .48-.74.

The Rejection Scale used in this present study was altered from Carroll's version in order to test a new set of hypotheses (see appendix B for a copy of this new Rejection Scale).

This new Rejection Scale is comprised of five moral dilemmas which are currently being used in three other major moral assessment
devices, plus a dilemma designed by the present author for specific use in this study. Each of the six dilemmas has three forms so as to depict three consequence conditions (mild, moderate, and severe). The consequence levels were established by primarily two separate procedures. The first procedure included an analysis of the structured interview material from a previous study (Plummer, 1982). This analysis was used to determine how other subjects rated the consequences, presented unsystematically in the stories on the Defining Issues Test (Rest, 1979). Based on this information, and significant contributions from the committee supervising this project, the consequence levels were intuitively designed.

Following the intuitive derivation of mild, moderate, and severe consequences, the subjects in the pilot study were asked to rate all the consequences in the Rejection Scale, from 1 to 10 in terms of severity. The results depicted in Table 16 (page 88) show that all ages rated the mild consequence condition less severe than moderate, and the moderate was rated milder than the severe condition. This relationship was documented for each of the six stories. Thus, the intuitive design of consequence levels, based on structured interview data, was supported with empirical results by the same population that was used in the main study.

The main feature of this new Rejection Scale is the representation of three consequence conditions along with six different story themes (a total of 18 dilemmas). The test is designed to counter balance the order of story theme so that each story theme appears in each of the six positions an equal number of times over the entire sample.

A set of three response choices follows each story. These response choices are based on Kohlberg's three level system so that
each level is represented among the response choices. Levels are used to code responses, rather than stages, because the levels seem to be based on clearer structural and conceptual distinctions.

Subjects must rate each response choice in terms of the following categories: "G", this is a good reason; "U", I am uncertain about this reason; or "P", this is a poor reason. Scoring ranges from 0 for every "Good" response to 1 point for "Uncertain" responses, and 2 points for every "Poor" response. Scores are totaled for the Level I answers to yield a rejection score for each story.

Structured Interview. The structured interview was used to evaluate the Rejection Scale by addressing the following concerns.

1. Did the subjects understand the stories and response choices?

2. Why did subjects reject response choices?
   2a. Were they unable to understand them?
   2b. Did they understand response choices and place them on low priority?

3. Explore rejection of Level I thought.
   3a. Why is Level I thought rejected in some instances and accepted in others?
   3b. Is there consistency between stories on the same consequence condition? (in terms of rejection scores)

4. Do subjects consistently and correctly place consequences in the categories of mild, moderate, and severe?

5. Do subjects choose the "uncertain" response frequently enough to warrant this response category?

6. Can subjects think of a better response choice which was not listed?

The specific interview questions addressed the subject's particular pattern of responses to the Rejection Scale. These specific questions differed for each subject, since it was highly unlikely that subjects would generate identical response patterns. However, the purpose and format of the interview remained constant as it addressed the issues
and concerns listed earlier. The following five steps represent a model approach to the structured interview.

1. Subjects were given 5 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 this review period, 3 dilemmas were chosen to determine comprehension of stories and responses. The 3 dilemmas included 3 different story themes and 3 different consequence conditions. Comprehension was determined by requesting that subjects explain the stories and responses in their own words.

3. Each protocol contains 6 examples of each consequence condition. Each consequence condition was examined for consistency of rejection rating across stories. When inconsistency was discovered it was addressed in the manner of the following example.

"Why did you think the consequences for stealing the drug were good reasons for making a decision in the Heinz situation, while the consequences of hiring Mr. Lee were not good reasons in the Webster situation?"

4. The fourth step was comprised of a couple of general questions to further clarify the subject's response to the Rejection Scale. When subjects marked an item "uncertain" they were asked why the "accept" or "reject" choice was not suitable. In addition, subjects were asked if they could think of a better response choice which was not listed.

The structured interviews were conducted on an individual basis. They lasted approximately 15 minutes.

Defining Issues Test (DIT). The DIT is a measure of moral preference designed by Rest (1979) (see Appendix A). According to Kohlberg (1979) the DIT is an instrument which assesses a broad spectrum of moral reasoning: "The DIT assesses recognition, comprehension, and preference, and thus indirectly spontaneous production (of moral reasoning)" (Kohlberg, 1979, p. xv). Kohlberg (1979) also comments on the reliability and validity of the DIT.

The DIT can claim not only reliability but also construct validity as well, since results with it conform to expectations derived from the cognitive developmental theory and
cannot be accounted for by interpreting the test responses as other than cognitive developmental or other than moral (p. xiv).

In addition, Kohlberg (1979) has compared the DIT ("the Minnesota measure") with his measure of moral development ("the Harvard measure"). "From the point of view of the Harvard group, the moderate correlations between the DIT and our measure support the construct validity of the Harvard measure as well as the Minnesota measure" (p. xiv).

The DIT is an objective measure of moral preference. Subjects are required to read through six moral dilemmas and select a series of responses to best represent their moral preference. Response selections follow each dilemma. Subjects are required to rank this selection of responses in terms of relative importance. 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 response choices on the DIT were selected from the transcripts of subjects who were assessed with Kohlberg's interview technique. The issues and concerns which subjects typically raised in response to moral dilemmas from the Kohlberg interview were used as multiple choice answers on the DIT. All items on the DIT are matched by word length, syntactic complexity, and use of technical or specialized terminology (Rest, 1979).

DIT scoring yields a measure of how often various stages of reasoning are used to answer the dilemmas. The DIT score which is used most frequently in descriptions of moral reasoning is the P%. This 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 choice is weighted 4 points, second choice 3 points, third choice 2 points, and the last choice is weighted 1 point. 2) The weights assigned to each stage are summed and multiplied times that stage number. 3) These stage products are summed across the six moral dilemmas so that each stage has one grand numerical value. 4) These values are divided by the total points possible to obtain individual stage percentages. 5) The P% is derived by adding the percentage values assigned to stages 5 and 6.

Rest (1979) reports test-retest reliability, using the P% as the dependent measure, in the range of .70 to .80. A number of reliability studies have been conducted with retest intervals ranging from one week to five months. The most typical retest interval, though, has been a two to three month range. Test-retest reliability with individual stage scores is generally lower, in the range of .50 to .60. The standard error of measurement for the P% is 7.5%. Cronbach's alpha was used to establish a measure of internal consistency regarding the P%. Alpha was .77 for the P index.

Measures of convergent-divergent correlational validity have also been reported (Rest, 1979). The DIT correlates well with other measures of moral reasoning, such as measures of moral comprehension and Kohlberg's test. Correlations with moral comprehension range from .49-.65, and correlations with Kohlberg's measure of moral judgment range from .40-.70. Correlations between the DIT and measures of cognitive development (not specifically moral reasoning) range from .20-.50. Correlations with measures of general aptitude are .40, and tests of personality .25-.35. In general, the DIT
correlates best with other measures of moral development, moderately with measures of general cognitive development, and poorly with measures of personality.

The DIT protocols used in this study were hand scored by the principle investigator of this study according to the system briefly outlined above and detailed in Rest's (1979) scoring manual. This scoring system yields stage percentages as well as a P index.

Procedure

The pilot study sample of 90 subjects (7th, 9th, and 11th graders) completed the Defining Issues Test (DIT). This was group administered and took approximately 45 minutes to complete. Specific instructions for completing the instrument were included with the test, 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.

The instrument was hand scored according to the procedures described in the Measures section.

Following a one week interval the original sample of 90 subjects completed the Rejection Scale. This was administered in group fashion, preceded by the set of instructions used for the DIT. The Rejection Scale was hand scored according to the procedures described in the Measures section.

Following a two week interval, the original sample of 90 subjects completed another Rejection Scale. The process and conditions of test administration were identical to the first administration. This second administration of the Rejection Scale was followed by the
Structured Interview. At least one week after completing the second Rejection Scale, 5 subjects from each age category were selected at random to participate in the Structured Interview. This interview was conducted on an individual basis, following the guidelines outlined in the Measures section. The interviews were recorded in narrative form and subsequently summarized across subjects in terms of the concerns detailed in the Measures section. Interviews lasted approximately 15 minutes.

Main Study

Subjects

The subjects in this study included a total sample of 90 male and female junior high and high school students 30 from each grade of the 7th, 9th, and 11th grades. This sample was selected from a small rural school system in north east Connecticut. The same schools were used for both the pilot and main study. However, contact between subjects across studies was probably minimal because students' schedules were mutually exclusive. Each group contained 30 subjects and was comprised of approximately equal numbers of males and females. The students were selected randomly from study halls which are part of every student's required program. The students participated through an arrangement made with parents, students, teachers, and school administrators, after a complete explanation of the project was offered. Student confidentiality was maintained as no student names were used in the processing or reporting of results.

This project was reviewed and approved by the superintendent, director of special education, and school principals. In addition,
Measures

Rejection Scale. The Rejection Scale is a measure of moral maturity based on the subject's rejection of lower stage reasoning. Carroll and Rest (1981) demonstrated that rejection of lower stage reasoning increases with age and moral maturity. Carroll and Rest (1981) conclude, "when individuals give up thinking at stages lower than their modal stage, this can be viewed as developmental advance in the sense of consolidation at the higher stage" (p. 43). This notion of measuring moral maturity through rejection of reasoning led to the development of Carroll's objective instrument. Like the DIT, Carroll's measure is based on Kohlberg's six stage model. His measure consists of four moral dilemmas which are each followed by reasons for acting or not acting in the situation. Ten reasons, two at each of Kohlberg's stages 1-5, comprise the response choices. Subjects are expected to rate each reason in terms of one of four categories: 1) I accept the reason; 2) I tend to accept the reason; 3) I tend to reject the reason; or 4) I reject the reason. Scoring ranges from 1 point for every "I accept the reason" to 4 points for every "I reject the reason." In this manner a rejection rating is established. Theoretically the higher stage subjects should have higher rejection ratings.

Criterion group validity was established in Carroll's 1981 study as older subjects (11th graders) did indeed score higher rejection ratings than younger subjects (7th graders). Carroll (1974) also
reports internal consistency for each stage (1 through 5), using Hoyt reliability coefficients, in the range of .48-.74.

The Rejection Scale used in this study was altered from Carroll's version in order to test a new set of hypotheses (see appendix B for a copy of this new Rejection Scale). Reliability (test-retest) and validity (construct and criterion group) were established for this new Rejection Scale in a pilot study.

This new Rejection Scale consists of six moral dilemmas which are currently being used in three other major moral assessment devices. Each of the six dilemmas has three forms to depict three consequence conditions (mild, moderate, and severe). This yields a total of 18 dilemmas (six stories X three consequence conditions). The main feature of this new Rejection Scale is the representation of three consequence conditions along with six different story themes.

The test is designed to counterbalance the order of presentation of story theme so that each story theme appears in each of the six positions an equal number of times over the entire sample. Each consequence condition is paired with each story theme an equal number of times, and counterbalanced for order over the entire sample.

A set of three response choices follows each story. These response choices are based on Kohlberg's three level system so that each level is represented among the response choices. Levels are used to code responses, rather than stages, because the levels seem to be based on clearer structural and conceptual distinctions.

Subjects must rate each response choice in terms of the following categories: G, this is a good reason; U, I am uncertain about this reason; or P, this is a poor reason. Scoring ranges from 0 for every "G" response, 1 point for a "U" response and 2 points for a "P"
response. Scores are totaled for the level I answers to yield a rejection score for each story. This raw single story rejection score was used as one form of the dependent variable in the analysis to be described in the Results section of this proposal. A total rejection score can be obtained by summing all rejection scores of level I responses across all stories. This score was also used in the data analysis.

This instrument was group administered. It took approximately 25 minutes to complete.

Procedure

The Rejection Scale was administered in group fashion to a sample of ninety 7th, 9th, and 11th graders (n for each group = 30). This took approximately 25 minutes to complete. Specific instructions for completing the instrument were included with the test, however, subjects were given the same general instructions outlined in the Pilot Study.

The Rejection Scale was hand scored according to the procedures outlined earlier, to yield: a) an individual's 18 rejection scores, (six stories x three consequence conditions), b) a total rejection score, c) a rejection score for each consequence condition, and d) a rejection score for each story.

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

HYPOTHESES AND PREDICTIONS

Hypothesis One

Story consequence (from mild to severe) is a functional aspect of the moral dilemma which interacts with the reasoner's structural capacity to yield differential effects on the moral reasoning of junior high and high school students. This hypothesis leads to the following predictions about this structural-functional interaction.

Predictions.

1a. The three different consequence conditions will yield results, in the form of rejection scores, which form a cumulative or Guttman scale. That is, a subject who rejects Level I reasoning in one consequence condition of a dilemma will reject Level I reasoning in all milder consequence conditions on that dilemma (or story). In other words, subjects who accept Level I reasoning in one consequence condition will accept it in all more severe consequence conditions. More specifically, this Guttman scale will contain seven cumulative steps and yield a coefficient of reproducibility greater than or equal to .85. The Guttman scale will prevail for each grade level, 7th, 9th, and 11th, as well as across the entire sample.

1b. The total rejection scores across all dilemmas and ages will be greatest in the mild consequence condition, decrease in the moderate condition, and the rejection score total for the severe consequence condition should be the lowest. The consequence effect will prevail for each grade level, 7th, 9th, and 11th, as well as across the entire sample.

Hypothesis Two

The seven step rejection scale has criterion group validity as it depicts a developmental continuum between younger and older students.

Predictions.

2a. The seven step Guttman scale will show that 7th graders occupy a lower position on the scale than 9th graders, and both groups will be lower on this scale than 11th graders.
2b. Rejection scores will differ across grade level as the 7th graders will have the lowest rejection scores, 11th graders will have the highest rejection scores, and 9th graders will fall in between these two groups. This rejection score trend by age will hold in each consequence condition as well as across all consequence conditions. For example, given severe consequences, 7th graders will have the lowest rejection scores, followed by 9th graders, and 11th graders will have the highest rejection scores.

**Hypothesis Three**

The Story theme is a relevant functional aspect of moral reasoning as different story themes have differential effects on the moral reasoning of junior high and high school students.

**Predictions.**

3a. The six different story themes will yield different rejection scores across the entire sample. In addition, each age group, 7th, 9th, and 11th graders, will show a story theme effect.

3c. The functional importance of story theme will be demonstrated further as the stories will differ in their degree of scalability by consequence condition.
CHAPTER X

RESULTS

Pilot Study

The Rejection Scale used in this study was altered from the standardized version (Carroll, 1974) in order to examine the specific hypotheses described earlier. Thus, the pilot study was designed to establish the psychometric properties of this new instrument. The analysis of the pilot data will be presented in this section in terms of test-retest reliability and validity (content, criterion group and concurrent).

Reliability

A total sample of 90 subjects completed two Rejection Scales with a two week interval between administrations. A total rejection score was calculated for each student on each administration of the Rejection Scale. A rejection score for each dilemma and each consequence condition was also calculated by individual for each administration of the Rejection Scale. The test-retest data by dilemma, consequence condition and total score is summarized in Table 1.

The test-retest reliability of the total score (two week interval), .90, is well beyond the acceptable psychometric standard for retest reliability as suggested by Aiken (1974). In addition, this value is greater than the test-retest reliabilities reported by Carroll, Kohlberg, and Rest in reference to their own instruments.
The reliability coefficients for each dilemma and consequence condition are lower than the total score coefficient, which is consistent with general test theory (Aiken, 1974). However they account for 50 to 70 percent of the variance and exceed the general acceptable limit for retest reliability of individual test items (Aiken, 1974).

In general, the figures in Table 1 suggest that the Rejection Scale is a consistent measure from one administration to the next. Furthermore, the items which comprise the Rejection Scale are all similar in their high level of consistency across administrations.

Validity

**Concurrent validity.** The scores on the rejection scale were compared to scores obtained on the Defining Issues Test (DIT) by using a Pearson Product Moment Correlation Coefficient. This correlation coefficient is .55. This value is high enough to indicate that the two measures are probably evaluating the same construct. In addition, the value is low enough to suggest that there may be a meaningful difference between the two instruments. Perhaps one value of this new Rejection Scale then, is its ability to measure something different about the construct of moral reasoning.

**Criterion group validity.** The results of the pilot study were also evaluated with a three way analysis of variance by consequence condition, age (grade level) and story theme (see figure 3 for a sketch of this design). The variable of primary interest in this measure of validity is age because moral reasoning is a developmental construct which should increase with age. Aiken (1974), and Rest (1979) report that the main purpose of criterion group validity is to
Table 1

Retest Reliability Coefficients By Dilemma, Consequence Condition and Total Score

<table>
<thead>
<tr>
<th>Dilemma, Consequence Condition and Total Score</th>
<th>Retest Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heinz</td>
<td>.81</td>
</tr>
<tr>
<td>Doctor</td>
<td>.70</td>
</tr>
<tr>
<td>Gail</td>
<td>.74</td>
</tr>
<tr>
<td>Joe</td>
<td>.82</td>
</tr>
<tr>
<td>Draft</td>
<td>.78</td>
</tr>
<tr>
<td>Webster</td>
<td>.70</td>
</tr>
<tr>
<td>Mild Consequences</td>
<td>.76</td>
</tr>
<tr>
<td>Moderate Consequences</td>
<td>.76</td>
</tr>
<tr>
<td>Severe Consequences</td>
<td>.82</td>
</tr>
<tr>
<td>Total Score</td>
<td>.90</td>
</tr>
</tbody>
</table>

*Note. N = 90*
<table>
<thead>
<tr>
<th><strong>STORY THEME</strong></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CONSEQUENCE</strong></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td><strong>GRADE LEVEL</strong></td>
<td>7th</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9th</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11th</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. * Six different story themes, 1-6.
** Three different consequence conditions.
    Mild = A, Moderate = B, Severe = C.

Figure 3. Design Sketch For Three-Way ANOVA With Repeated Measures, For Story Theme, Consequence Condition, and Grade Level.
demonstrate that subjects who ought to have different scores on a measure do in fact have different scores. Thus we expected older subjects to have higher scores, indicating a more advanced level of moral reasoning, than younger subjects.

As outlined in Table 2 the three way ANOVA reports a significant age effect \((F = 5.54, \text{df} = 2/87, p < .005)\). However, the interaction of age and consequence is also significant \((F = 5.35, \text{df} = 4/174, p < .0004)\) so the effects of age must be interpreted in terms of the consequence conditions. In addition, the issue of homogeneity of variance must be addressed.

The results of the F-Maximum Test for Homogeneity of Variance (Hays, 1973) indicate that the variances among the 54 cells were heterogeneous \((F_{\text{max}} = 11.25, K/\text{df} = 54/29, p < .01)\). However, since scores can only range from 0-2 we are dealing with a restricted range such that variances are considerably less than 1.0. When dealing with extremely small variances a significant \(F_{\text{max}}\) could result from trivial differences in variance (e.g., smallest variance = .20, largest variance = .80, \(F_{\text{max}} = 4.0\)). In addition, Hays (1973) firmly asserts that the effect of heterogeneity of variances on the F-distribution will most likely be negligible.

Ordinarily, other things being equal, this assumption of homogeneous variances can be violated without serious risk, provided that the number of cases in each sample is the same (Hays, 1973, p. 482).

A test for homogeneity of variance before the analysis of variance has rather limited practical utility, and modern opinion holds that the analysis of variance can and should be carried on without a preliminary test of variances, especially in situations where the number of cases in the various samples can be made equal (Hays, 1973, p. 484).
Table 2

Three by Three by Six Analysis of Variance on Pilot Data Rejection Scores by Age, Consequence, and Dilemma

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>D.F.</th>
<th>Mean Squares</th>
<th>F</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>28.24</td>
<td>2</td>
<td>14.12</td>
<td>5.54</td>
<td>&lt;.005</td>
</tr>
<tr>
<td>Dilemma</td>
<td>110.10</td>
<td>5</td>
<td>22.00</td>
<td>20.60</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Consequence</td>
<td>120.96</td>
<td>2</td>
<td>60.48</td>
<td>164.31</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Error (age)</td>
<td>221.58</td>
<td>87</td>
<td>2.55</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Age/Dilemma</td>
<td>8.20</td>
<td>10</td>
<td>.82</td>
<td>.77</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age/Consequence</td>
<td>7.88</td>
<td>4</td>
<td>1.97</td>
<td>5.35</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Dilemma/Consequence</td>
<td>22.26</td>
<td>10</td>
<td>2.23</td>
<td>7.47</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Error (Dilemma, Age/Consequence)</td>
<td>464.97</td>
<td>435</td>
<td>1.07</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Error (Consequence, Age/Consequence)</td>
<td>64.05</td>
<td>174</td>
<td>.37</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Age/Consequence/Dilemma</td>
<td>2.86</td>
<td>20</td>
<td>.143</td>
<td>.48</td>
<td>&lt; .97</td>
</tr>
<tr>
<td>Error (Age/Consequence/Dilemma)</td>
<td>259.33</td>
<td>870</td>
<td>.30</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

Note.  n per cell = 30  
N total sample = 90
Based on Hays' (1973) review of homogeneity of variance and the equal cell size in this study, the statistical analysis based on the three-way ANOVA was continued.

The interaction of age and consequence was followed up with simple effects tests (see Table 3, page 52) which revealed that there was an age effect at mild and moderate but not at severe consequences (Age/Mild, F = 11.35, d.f. = 2/174, p < .01; Age/Moderate, F = 4.91, d.f. = 2/174, p < .01; Age/Severe, F = 1.56, d.f. = 2/174, N.S.).

Simple effects tests of consequences at each age (see Table 4, page 52) revealed that there was a significant consequence effect at all three ages (Consequence/7th, F = 8.78, d.f. 2/174, p < .01; Consequence/9th, F = 20.3, d.f. = 2/174, p < .01; Consequence/11th, F = 30.14, d.f. = 2/174, p < .01).

The significant simple effects tests were followed by the Newman-Keuls Analysis to determine specific differences within ages and consequence conditions. The results depicted in Figure 4 (see page 53) show that all three ages are significantly different from each other under both mild and moderate consequence conditions. Ninth graders achieve higher scores than seventh graders, and eleventh graders achieve higher scores than ninth graders. Figure 5 (see page 54) shows that the differences between mild, moderate and severe consequence conditions are all significant at each age level. Mild consequences result in higher scores for all ages, compared to moderate and severe consequences, while moderate consequences result in higher scores than severe consequences.

Based on the ANOVA, simple effects tests, and Newman-Keuls it can be concluded that the Rejection Scale effectively distinguishes between all three age groups, with the significant differences
### Table 3
Simple Effects Tests On The Interaction of Age At Each Consequence Condition (Pilot Study)

<table>
<thead>
<tr>
<th>Age</th>
<th>Consequence Condition</th>
<th>F</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Across 7th, 9th and 11th Grade</td>
<td>Mild</td>
<td>11.35</td>
<td>&lt; .01</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>4.91</td>
<td>&lt; .01</td>
</tr>
<tr>
<td></td>
<td>Severe</td>
<td>1.56</td>
<td>N.S.</td>
</tr>
</tbody>
</table>

*Note. d.f. = 2,174  
n per cell = 30*

### Table 4
Simple Effects Tests On The Interaction of Consequence At Each Age (Pilot Study)

<table>
<thead>
<tr>
<th>Consequence Condition</th>
<th>Age</th>
<th>F</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Across Mild, Moderate and Severe Consequences</td>
<td>7th</td>
<td>8.78</td>
<td>&lt; .01</td>
</tr>
<tr>
<td></td>
<td>9th</td>
<td>20.30</td>
<td>&lt; .01</td>
</tr>
<tr>
<td></td>
<td>11th</td>
<td>30.14</td>
<td>&lt; .01</td>
</tr>
</tbody>
</table>

*Note. d.f. = 2,174  
n per cell = 30*
Figure 4. The Difference Between Ages At Each Consequence Condition on Pilot Data Rejection Scores

Note.  * $p \leq .01$
** $p \leq .001$

$N = 90$
Note. ** p < .001
N = 90

Figure 5. The Difference Between Consequence Conditions At Each Age on Pilot Data Rejection Scores
occurring at the mild and moderate consequence condition. Furthermore, the significant differences between consequence conditions hold for each age group. The rejection scale, then, has criterion group validity with reference to the seventh, ninth, and eleventh grade population sampled in this study.

**Content validity.** Aiken (1974) suggests that a test has content validity if it contains a representative sample of items from the universe of items commonly used to evaluate the construct. The Rejection Scale contains moral dilemmas selected from three major moral assessment instruments (Kohlberg's, Rest's and Carroll's). However, these researchers have not developed a categorization of moral content to determine if their sample dilemmas are in fact representative. The response format and response selections are modeled after two of these instruments (Rest's and Carroll's) which are based on the third (Kohlberg's).

**Structured interview.** Interviews were conducted with five subjects from each of the three age categories. The main purpose of the interviews was to evaluate whether subjects understood the stories and response choices and to determine some of the factors which subjects considered important in their evaluation of the moral dilemmas.

All subjects understood the stories, directions and pertinent responses well enough to explain them in their own words. In fact, without looking back at their test booklet, subjects were able to recall in great detail information such as the amount of money the drug cost in the Heinz story, the jail term in the Draft story, and the amount of money saved by Joe in the Joe story.
All subjects clearly understood the level one answers, which are used in scoring the Rejection Scale. However, approximately 50% of the 7th and 9th graders interviewed had a difficult time making the connection between the consequences in the Draft and Joe stories and the level one responses. Some students reported "I didn't know what you meant by too much trouble"... "I didn't know you meant the trouble Joe would get into"... "community service and paying money, I didn't take that to mean trouble". Even though some subjects did not immediately understand this connection, their scores fit the expected pattern according to the influence of consequences.

Subjects were also asked to explain why they thought Level I reasoning was acceptable in some instances and unacceptable in others. Subjects reported that the endings made a difference in how they evaluated the level one answers "doing a little work is nothing next to going to jail for life." In cases where the consequences did not appear to make a difference subjects were asked what it would take to make a difference to them. In most cases the students gave a response which was less severe than one of the originals. In response to the Gail story, for example, one ninth grader reported "I guess if I had to stay after school for a week I would tell the teacher." When it was pointed out how this answer was less severe than one of the originals the subject agreed and changed her mind regarding her original answer. A few other subjects had similar experiences. When they were required to review their protocol and justify their response choices some of them changed their answers. These few instances are part of test error.

In addition to specific consequences subjects also reported that other issues were relevant in their response choices. For example,
most subjects made a distinction between dilemmas such as "life and death is more important than somebody's business" (comparing the doctor story with the Webster story), and "the personal relationship makes it worth it (in the Heinz story) but here (in the Draft story) it's just himself".

In general, subjects seemed to understand the dilemmas, responses, and directions for completing the instrument. The seventh graders, however, needed a concrete example in their directions before they fully understood. In addition, seventh graders frequently asked for clarification throughout the task. Seventh graders completed the instrument in 35 minutes while ninth and eleventh graders required only 20 minutes. Some students had difficulty making the connections between consequences in the story and reference to those consequences in the response choices. This did not seem to influence the results significantly as the scores still fit the expected pattern based on the consequence effect. Finally, subjects at all ages reported that they noticed the differences in consequences across stories and this affected their response choices. The results from the ANOVA reported earlier strongly support this. Other relevant factors reported by subjects include theme issues such as personal relationship and life and death.

Pilot Summary

The procedures chosen to evaluate reliability and validity of the Rejection Scale indicate that this new instrument has strong psychometric properties. For example, test-retest reliability is .90 while the coefficients for individual test items ranges from .70-.82. In addition, the Rejection Scale correlates .55 with the Defining
Issues Test. This indicates that the two measures are probably evaluating the same construct and correlation coefficient is low enough to suggest that the two instruments may be different in some meaningful way. Thus, this new Rejection Scale may be measuring something different about the construct of moral reasoning.

The consequence effect which we expected from the Rejection Scale is significant at the \( p < .001 \) level for all ages sampled. Furthermore, the Rejection Scale effectively distinguishes the three ages from each other. Follow-up analysis shows that the mild and moderate consequence conditions are responsible for making this distinction. We expect an increase in moral reasoning between seventh, ninth, and eleventh grade, which the Rejection Scale measures quite effectively.

Finally, the structured interview showed that all three ages understood and interacted appropriately with the test materials. However, seventh graders needed more clarification and concrete examples to understand the directions. Subjects reported that they noticed differences in story consequences which made a difference in their evaluation of level one responses. Furthermore, they identified themes or moral issues, such as life and death and personal relationships, which also made a difference in the evaluation of level one responses.

In conclusion, the Rejection Scale appears to be psychometrically sound based on the analyses described above. Research with different populations and other measures of reliability and validity would further clarify this issue.
Main Study

The Rejection Scale reviewed above was used to address three major questions: 1) as story consequences increase in severity, will individuals be less likely to reject consequence level reasoning (Kohlberg's Level I, preconventional) as an appropriate answer to moral dilemmas? 2) Is there a developmental trend between 7th, 9th, and 11th graders regarding this consequence effect? 3) Do story themes have differential effects on the moral reasoning of junior high and high school students?

The data from the Rejection Scale was evaluated with Scalogram Analysis (Edwards, 1957) by consequence condition, age, and story theme. In addition, the data was evaluated by analysis of variance with repeated measures by grade level, consequence condition and story theme (see Figure 3 for a sketch of this design). The results of this main study will be presented in this section to address the three questions outlined above. Each question will be presented in terms of its hypothesis and predictions, followed by a summary of the appropriate statistical analysis.

Hypothesis One

Story consequence (mild, moderate, and severe) is a functional aspect of the moral dilemma which interacts with the reasoner's structural capacity to yield differential effects on the moral reasoning of junior high and high school students. Based on this hypothesis it was predicted that the consequences would form an ordered scale, based on their severity, which would conform to the characteristics of a Guttmann scale. In addition, an analysis of
variance, including follow-up tests, should reveal a significant consequence effect which should hold for each grade level as well as the entire sample.

Scalogram analysis. Scalogram Analysis is the appropriate statistical technique for evaluating data to determine whether it meets the conditions of a cumulative or Guttman scale (Edwards, 1957). Tables 5, 6, 7, and 8 summarize the Scalogram Analysis by grade and dilemma (see pages 61-64). The coefficient of reproducibility by age ranges from .92 to .96, well beyond the .85 mark necessary to consider the data scalable (Edwards, 1957). In addition, the coefficient of reproducibility by dilemma ranges from .90 to .95 and the dilemma within age coefficients range from .88 to 1.0. The results from all three grades and all six dilemmas indicate that consequences have differential effects on moral reasoning. The Scalogram Analysis shows that consequence level reasoning is more difficult to reject when consequences are severe, as opposed to moderate and mild, and moderate consequences make it more difficult to reject level 1 reasoning than mild ones. Thus a seven step Guttman scale can be formed (see Tables 5-8) where each successive step represents a higher stage as it becomes increasingly more difficult to reject level 1 reasoning. The data clearly indicates that subjects often accept level 1 reasoning under severe consequence conditions while rejecting it under the mild condition within the same dilemma. Clearly, consequences have a significant effect on moral reasoning according to the Scalogram Analysis.

Analysis of variance. The data from the main study and the combined data (main and pilot study) were both evaluated with a three by three by six ANOVA by consequence condition, age (grade level) and
### Table 5
Scalogram Analysis For Grade Seven By Dilemma

<table>
<thead>
<tr>
<th>Consequence* Condition</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>X</th>
<th>Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0 0</td>
<td>55</td>
<td>10</td>
<td>42</td>
<td>55</td>
<td>30</td>
<td>75</td>
<td>45</td>
<td>1</td>
</tr>
<tr>
<td>1 0 0</td>
<td>8</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>2 2 1</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>2 2 2</td>
<td>13</td>
<td>20</td>
<td>18</td>
<td>13</td>
<td>17</td>
<td>2</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>Errors</td>
<td>5</td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>12</td>
<td>6</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

**Average Stage Rating**

| 2.46 | 4.0 | 3.12 | 2.7 | 3.45 | 1.47 | 2.87 |

**Coefficient of Reproducibility**

| .95 | .93 | .88 | .93 | .88 | .93 | .92 |

**Note.** Response frequency is in percentage units.

* $C_1 =$ Mild, $C_2 =$ Moderate, $C_3 =$ Severe

** $1 =$ Heinz, $2 =$ Doctor, $3 =$ Gail, $4 =$ Joe, $5 =$ Draft, $6 =$ Webster

$N = 60$ (repeated across dilemmas)
Table 6
Scalogram Analysis For Grade Nine By Dilemma

<table>
<thead>
<tr>
<th>Condition*</th>
<th>Dilemma**</th>
<th>Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1 C2 C3</td>
<td>1 2 3 4 5 6 X</td>
<td></td>
</tr>
<tr>
<td>0 0 0</td>
<td>32 10 27 32 25 68 32</td>
<td>1</td>
</tr>
<tr>
<td>1 0 0</td>
<td>7 5 8 5 2 10 6</td>
<td>2</td>
</tr>
<tr>
<td>2 0 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 1 0</td>
<td>27 20 17 8 22 8 17</td>
<td>3</td>
</tr>
<tr>
<td>2 1 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 1 1</td>
<td>8 8 7 13 5 5 8</td>
<td>4</td>
</tr>
<tr>
<td>2 2 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 1 1</td>
<td>2 27 7 12 17 0 11</td>
<td>5</td>
</tr>
<tr>
<td>2 2 1</td>
<td>2 8 5 5 5 0 4</td>
<td>6</td>
</tr>
<tr>
<td>2 2 2</td>
<td>8 13 22 18 18 3 14</td>
<td>7</td>
</tr>
</tbody>
</table>

Errors: 15 8 8 7 7 5 8

Average Stage Rating: 2.7 4.2 3.7 3.6 3.8 1.6 3.3

Coefficient of Reproducibility: .85 .92 .92 .93 .93 .95 .92

Note. Response frequency is in percentage units.

* C1 = Mild, C2 = Moderate, C3 = Severe
** 1 = Heinz, 2 = Doctor, 3 = Bill, 4 = Joe, 5 = Draft, 6 = Webster
N = 60 (repeated across dilemmas)
Table 7
Scalogram Analysis For Grade Eleven By Dilemma

<table>
<thead>
<tr>
<th>Consequence* Condition</th>
<th>Dilemma**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>0 0 0</td>
<td>22</td>
</tr>
<tr>
<td>1 0 0</td>
<td>3</td>
</tr>
<tr>
<td>2 0 0</td>
<td></td>
</tr>
<tr>
<td>1 1 0</td>
<td>25</td>
</tr>
<tr>
<td>2 1 0</td>
<td></td>
</tr>
<tr>
<td>1 1 1</td>
<td>10</td>
</tr>
<tr>
<td>2 2 0</td>
<td>7</td>
</tr>
<tr>
<td>2 2 1</td>
<td>3</td>
</tr>
<tr>
<td>2 2 2</td>
<td>20</td>
</tr>
<tr>
<td>Errors</td>
<td>10</td>
</tr>
<tr>
<td>Average Stage Rating</td>
<td>3.7</td>
</tr>
<tr>
<td>Coefficient of</td>
<td>.90</td>
</tr>
</tbody>
</table>

Note. Response frequency is in percentage units.

* \( C_1 = \text{Mild}, C_2 = \text{Moderate}, C_3 = \text{Severe} \)
** \( 1^* = \text{Heinz}, 2^* = \text{Doctor}, 3 = \text{Gail}, 4 = \text{Joe}, 5 = \text{Draft}, 6 = \text{Webster} \)
N = 60 (repeated across dilemmas)
Table 8  
Scalogram Analysis Totals For All Three Grades and all Six Dilemmas

<table>
<thead>
<tr>
<th>Consequence* Condition</th>
<th>Grade</th>
<th>Dilemma**</th>
<th>Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Response frequency is in percentage units.

* C1 = Mild, C2 = Moderate, C3 = Severe
** 1 = Heinz, 2 = Doctor, 3 = Cail, 4 = Joe, 5 = Draft, 6 = Webster
N = 60 (repeated across dilemmas)
story theme (see Figure 3). The variable of interest with respect to hypothesis one is consequence.

As outlined in Table 9 (page 66), a significant consequence effect was found ($F = 107.73, d.f. = 2/174, p \leq .001$). However, the interactions of consequence and age as well as consequence and dilemma are also significant ($F = 2.59, d.f. = 4/174, p \leq .04$ and $F = 5.95, d.f. = 10/870, p \leq .001$). Therefore the effects of consequence must be interpreted in terms of age and dilemma. In addition, the issue of homogeneity of variance will be addressed.

While the interactions may be statistically significant, the issue of practical significance should also be addressed. Table 11 reports the proportion of variance accounted for by each of the variables and interactions for the main study and the combined data from the pilot and main study (Hays, 1973). Each interaction term accounts for 1 to 5 percent of the variance in the main study, while the main effect for consequence accounts for 24 percent of the variance. The significant interaction effects, then, could be considered trivial. Therefore, even though the interactions will be explored in more detail with simple effects tests and Newman-Keuls, the main effect of consequence is fairly straightforward.

The results of the F-Maximum Test for Homogeneity of Variance (Hays, 1973) indicate that the variances among the 54 cells were heterogeneous ($F_{\text{max}} = 10.15, K/d.f. = 54/29, p \leq .01$). As reported earlier, Hays (1973) claims the assumption of homogeneous variances can be violated without serious risk, provided that the number of cases in each sample is the same (p. 482)... modern opinion holds that the analysis of variance can and should be carried on without a preliminary test of variances, especially in situations where the number of cases in the various samples can be made equal (484).
Table 9

Three by Three by Six Analysis of Variance on Rejection Scores
From Main Study By Age, Consequence, and Dilemma

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>D.F.</th>
<th>Mean Squares</th>
<th>F</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>30.97</td>
<td>2</td>
<td>15.48</td>
<td>5.03</td>
<td>≤ .009</td>
</tr>
<tr>
<td>Dilemma</td>
<td>110.35</td>
<td>5</td>
<td>22.07</td>
<td>20.40</td>
<td>≤ .001</td>
</tr>
<tr>
<td>Consequence</td>
<td>90.71</td>
<td>2</td>
<td>45.36</td>
<td>107.73</td>
<td>≤ .001</td>
</tr>
<tr>
<td>Error (age)</td>
<td>267.76</td>
<td>87</td>
<td>3.08</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Age/Dilemma</td>
<td>9.12</td>
<td>10</td>
<td>.91</td>
<td>.84</td>
<td>≤ .59</td>
</tr>
<tr>
<td>Age/Consequence</td>
<td>4.36</td>
<td>4</td>
<td>1.09</td>
<td>2.59</td>
<td>≤ .04</td>
</tr>
<tr>
<td>Dilemma/Consequence</td>
<td>16.04</td>
<td>10</td>
<td>1.60</td>
<td>5.95</td>
<td>≤ .001</td>
</tr>
<tr>
<td>Error (Dilemma, Age/Dilemma)</td>
<td>470.70</td>
<td>435</td>
<td>1.08</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Error (Consequence, Age/Consequence)</td>
<td>73.26</td>
<td>174</td>
<td>.42</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Age/Consequence/Dilemma</td>
<td>2.97</td>
<td>20</td>
<td>.15</td>
<td>.55</td>
<td>≤ .95</td>
</tr>
<tr>
<td>Error (Age/Consequence/Dilemma)</td>
<td>234.66</td>
<td>870</td>
<td>.27</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

Note. n per cell = 30
N total sample = 90
Table 10

Three by Three by Six Analysis of Variance on Combined Rejection Scores From The Pilot and Main Study By Age, Consequence, and Dilemma

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>D.F.</th>
<th>Mean Squares</th>
<th>F</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>57.50</td>
<td>2</td>
<td>28.75</td>
<td>10.56</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Dilemma</td>
<td>212.37</td>
<td>5</td>
<td>42.47</td>
<td>39.97</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Consequence</td>
<td>209.86</td>
<td>2</td>
<td>104.93</td>
<td>267.90</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Error (age)</td>
<td>481.99</td>
<td>177</td>
<td>2.72</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Age/Dilemma</td>
<td>9.44</td>
<td>10</td>
<td>.94</td>
<td>.89</td>
<td>&lt; .54</td>
</tr>
<tr>
<td>Age/Consequence</td>
<td>11.71</td>
<td>4</td>
<td>2.93</td>
<td>7.48</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Dilemma/Consequence</td>
<td>36.09</td>
<td>10</td>
<td>3.61</td>
<td>12.38</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Error (Dilemma, Age/Dilemma)</td>
<td>940.41</td>
<td>885</td>
<td>1.06</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Error (Consequence, Age/Consequence)</td>
<td>138.65</td>
<td>354</td>
<td>.39</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Age/Consequence/Dilemma</td>
<td>3.01</td>
<td>20</td>
<td>.15</td>
<td>.52</td>
<td>&lt; .96</td>
</tr>
<tr>
<td>Error (Age/Consequence/Dilemma)</td>
<td>516.00</td>
<td>1770</td>
<td>.29</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

Note. n per cell = 60
N total sample = 180
Table 11
The Proportion of Variance Accounted For by Variables in The Main Study and Main/Pilot Combined

<table>
<thead>
<tr>
<th>Variable</th>
<th>Proportion Variance Accounted For</th>
<th>Main Study*</th>
<th>Main/Pilot Combined**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1%</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>Dilemma</td>
<td>10%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Consequence</td>
<td>24%</td>
<td>15%</td>
<td></td>
</tr>
<tr>
<td>Age/Consequence</td>
<td>1%</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>Dilemma/Consequence</td>
<td>5%</td>
<td>3%</td>
<td></td>
</tr>
</tbody>
</table>

Note.  * N = 90  ** N = 180
Based on Hays' (1973) review of this issue and the equal cell size in this study, the statistical analysis based on the three-way ANOVA was continued.

The interaction of consequence and age was followed up with simple effects tests (see Table 12, page 70) which revealed that there was a consequence effect at each age (Consequence/7th grade, $F = 19.06$, d.f. = 2/174, $p < .01$; Consequence/9th grade, $F = 40.68$, d.f. = 2/174, $p < .01$; Consequence/11th grade, $F = 29.70$, d.f. = 2/174, $p < .01$). Similar results were obtained from the combined analysis of pilot and main study (see Table 12).

The significant simple effects were followed by the Newman-Keuls analysis to determine specific differences within consequence conditions and ages. The results depicted in Figure 6 (see page 73) show that consequence conditions are significantly different from each other at all three ages with the largest differences occurring for ninth and eleventh graders. Similar results are found with the combined data (see Figure 7). At all three ages mild consequences result in higher scores than moderate and severe consequences while moderate consequences result in higher scores than severe.

Figure 8 (see page 75) shows that for the main study sample, there are significant differences between all three ages at the mild consequence condition while the moderate condition distinguishes 9th from 11th and 11th from 7th graders, but not 7th from 9th graders. The severe condition makes the same distinction but at a lower level of significance. The combined data shows a similar pattern of significance. Mild consequences distinguish all three ages, and moderate consequences make the same distinctions but the level of significance is higher between 9th and 11th graders than it is
### Table 12

Simple Effects Tests On The Interaction Of Consequences At Each Age (Main Study and Combined Data)

<table>
<thead>
<tr>
<th>Consequence Condition</th>
<th>Age (Grade)</th>
<th>F Ratio</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Main</td>
<td>Combined</td>
</tr>
<tr>
<td>Across Mild,</td>
<td>7th</td>
<td>19.06</td>
<td>45.12</td>
</tr>
<tr>
<td>Moderate, and</td>
<td>9th</td>
<td>40.68</td>
<td>100.03</td>
</tr>
<tr>
<td>Severe Consequences</td>
<td>11th</td>
<td>29.70</td>
<td>140.91</td>
</tr>
</tbody>
</table>

*Note. d.f. Main Study = 2,174
d.f. Combined Study = 2,354
n per cell Main Study = 30
n per cell Combined Study = 60*

### Table 13

Simple Effects Tests On The Interaction Of Age At Each Consequence Condition (Main Study and Combined Data)

<table>
<thead>
<tr>
<th>Consequence Condition</th>
<th>Age (Grade)</th>
<th>F Ratio</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Main</td>
<td>Combined</td>
</tr>
<tr>
<td>Across 7th,</td>
<td>Mild</td>
<td>22.18</td>
<td>53.02</td>
</tr>
<tr>
<td>and 9th and 11th grade</td>
<td>Moderate</td>
<td>14.52</td>
<td>30.36</td>
</tr>
<tr>
<td></td>
<td>Severe</td>
<td>3.66</td>
<td>4.49</td>
</tr>
</tbody>
</table>

*Note. d.f. Main Study = 2,174
d.f. Combined Study = 2,354
n per cell Main Study = 30
n per cell Combined Study = 60*
between 7th and 9th graders. Severe consequences only distinguish between 11th and 7th graders.

The interaction of consequence and dilemma was followed up by simple effects tests (see Table 14, page 79). This revealed that there was a consequence effect at each dilemma (consequence/Heinz, $F = 15.53$, d.f. = $2/870$, $p < .01$; Doctor, $F = 75.32$, d.f. = $2/870$, $p < .01$; Gail, $F = 21.82$, d.f. = $2/870$, $p < .01$; Joe, $F = 20.18$, d.f. = $2/870$, $p < .01$; Draft, $F = 57.11$, d.f. = $2/870$, $p < .01$; Webster, $F = 7.75$, d.f. = $2/870$, $p < .01$). Similar effects were obtained from the combined analysis of pilot data and main study (see Table 15). In addition, simple effects tests of dilemma at each consequence condition revealed that there was a significant dilemma effect at all three consequence conditions (dilemma/mild consequences, $F = 46.95$, d.f. = $5/870$, $p < .01$; moderate consequences, $F = 35.90$, d.f. = $5/870$, $p < .01$; severe consequences, $F = 11.48$, d.f. = $5/870$, $p < .01$). The combined data yielded similar results.

The significant simple effects were followed by a Newman-Kuels analysis to determine specific differences within consequence conditions and dilemmas. The results depicted in Figure 10 (see page 82) show that all consequence conditions are significantly different from each other in three of the six dilemmas (Doctor, Gail, and Draft). In two other dilemmas there is a significant difference between severe and moderate consequences as well as the severe and mild consequence condition (Heinz and Joe). Finally, in the Webster dilemma there is a significant difference only between severe and mild consequences. The results from the combined data show a stronger effect as all dilemmas have significant differences between all three consequence conditions. Thus, the consequence effect is clearly
demonstrated in most dilemmas, and every dilemma has at least some significant consequence effects.

**Hypothesis one summary.** The consequence effect was significant in all forms of analysis. The Scalogram Analysis demonstrated that the moral dilemmas could be arranged in the form of a cumulative or Guttman Scale based on the severity of consequences. For example, when subjects rejected level 1 reasoning in the severe consequence condition they also rejected it in all milder consequence conditions. This graded scale of difficulty by consequence condition held up for each age group, each moral dilemma, each dilemma and age combination, as well as when the data was totaled across all ages and dilemmas.

The ANOVA demonstrated that there was a significant consequence effect at each age with the strongest effects occurring for 9th and 11th graders. Furthermore, mild consequences were most effective at distinguishing the three ages, followed by moderate then severe. While mild consequences distinguished between all ages the moderate condition distinguished effectively between 9th and 11th and 11th and 7th graders. The severe condition distinguished between 11th and 7th graders and 11th and 9th graders in the main study only.

The consequence effect is clearly demonstrated in most dilemmas. Three dilemmas show significant differences between all three consequence conditions. Two dilemmas show significant differences between severe and moderate and between severe and mild consequences. One dilemma has significant differences between mild and severe consequence conditions. Thus, every dilemma has at least some significant consequence effects. In general, the Scalogram Analysis and ANOVA show a very strong consequence effect, which is further
Figure 6. The Difference Between Consequence Conditions At Each Age on Main Study Rejection Scores

Note.  * p < .01  
** p < .001  
N = 90
Figure 7. The Difference Between Consequence Conditions at Each Age on Combined Data Rejection Scores

Note. ** p < .001
N = 180
Figure 8. The Difference Between Ages At Each Consequence Condition on Main Study Rejection Scores

Note.  * p ≤ .05  
      ** p ≤ .01  
      *** p ≤ .001  
N = 90
Figure 9. The Difference Between Ages At Each Consequence Condition on Combined Data Rejection Scores

Note.  * p < .01
** p < .001
N = 180
supported by the data presented in Table 11 that indicate the consequence variable accounts for 24% of the total variance.

**Hypothesis Two**

The Rejection Scale has criterion group validity as it depicts a developmental continuum between younger and older students. Based on this hypothesis it was predicted that 7th graders would occupy a lower position on the Guttman Scale than 9th graders, and both groups would be lower on this scale than 11th graders. It was also predicted that this relationship would hold for each dilemma. In addition, it was predicted that an ANOVA by age would show significance between the three groups at each consequence condition as well as a general effect across consequence conditions.

**Scalogram Analysis.** The Scalogram analysis reviewed earlier (see Tables 5, 6, 7 and 8) was one of the statistical techniques used to determine age effects. As outlined in Table 8, the 7th graders have a lower average stage rating than the 9th graders, and they are both lower than 11th graders (7th = 2.87, 9th = 3.3, 11th = 3.9). In addition, when individual dilemmas across ages are compared (see Tables 5, 6 and 7) all six dilemmas show the same developmental trend. In each case the lowest scores are obtained by 7th graders, middle scores are obtained by 9th graders, and 11th graders achieve the highest scores on each dilemma. The results from Scalogram Analysis, then, show a consistent age effect.

**Analysis of Variance.** The data from the main study and the combined data were both evaluated with a three by three by six ANOVA by age (grade level), consequence condition, and story theme (see
Figure 3 for a sketch of this design). The variable of interest with respect to hypothesis two is age (grade level).

As outlined in Table 9, a significant age effect was found ($F = 5.03$, d.f. = $2/87$, $p < .009$). However the interaction of age and consequence is also significant ($F = 2.59$, d.f. = $4/174$, $p < .04$). Similar results were found with the combined data (see Table 10). Therefore, the effects of age must be interpreted in terms of consequences.

The interaction of age and consequence was followed up with simple effects tests (see Tables 12 and 13, page 70) which revealed that there was an age effect at each consequence condition (age/mild consequences, $F = 22.18$, d.f. = $2/174$, $p < .01$; age/moderate consequences, $F = 14.52$, d.f. = $2/174$, $p < .01$; age/severe consequences, $F = 3.66$, d.f. = $2/174$, $p < .05$). Similar results were obtained from the combined analysis of pilot and main study (see Table 13).

The significant simple effects were followed by a Newman Keuls analysis to determine specific differences within ages and consequence conditions. The results depicted in Figure 8 (see page 75) show that there is a significant difference between all ages at the mild consequence condition while the moderate consequence condition distinguishes 11th and 9th graders and 11th and 7th graders, but not 7th and 9th graders. There is a significant difference between 7th and 11th graders at the severe consequence condition. The combined data reveal similar findings except 9th and 7th graders are distinguished at the moderate condition in addition to the other distinctions shown by the main study.
Figure 6 shows that there are significant differences between all consequence conditions at each age. This effect is most significant for the 9th and 11th graders. The combined data show a similar pattern of significance.

**Hypothesis two summary.** The Rejection Scale has criterion group validity as it effectively distinguishes between 7th, 9th, and 11th graders. This was demonstrated with Scalogram Analysis as the 7th graders had a lower average stage rating than the 9th graders and they were both lower than the 11th graders. This trend was matched without exception by all six moral dilemmas.

The developmental trend was also demonstrated with Analysis of Variance. There was a significant age effect at each consequence condition with the strongest effect occurring at the mild consequence condition. The mild condition distinguishes all three ages while the moderate condition distinguished effectively between 9th and 11th, and 11th and 7th graders. The severe condition distinguished between 11th and 7th graders and, in the main study, 11th and 9th graders. In general the Scalogram Analysis and ANOVA show a strong developmental effect across the three age groups when the data is analyzed by consequence condition, dilemma, and total scores.

**Hypothesis Three**

Story theme is a relevant functional aspect of moral reasoning as different story themes have differential effects on the moral reasoning of junior high and high school students. Based on this hypothesis it was predicted that the six story themes would yield different average stage ratings, the coefficient of reproducibility would vary across stories and the stories would differ according to
their total rejection scores. Scalogram Analysis was used to evaluate the story themes in terms of their average stage ratings and coefficients of reproducibility. Analysis of variance was used to evaluate the story themes in terms of differences in rejection scores.

**Scalogram analysis.** Tables 5, 6, 7 and 8 (see pages 61-64) summarize the Scalogram Analysis by grade and dilemma. The dilemma effect can be shown by comparing the average stage ratings across dilemmas. This is done by age as well as across the entire sample. Each age group shows the same relationship between dilemmas in terms of the average stage rating (see Tables 5, 6 and 7). The average stage ratings ranged from 1.47 to 4.0 for 7th graders, 1.6 to 4.2 for 9th graders, and 2.4 to 4.8 for 11th graders. The rank ordering of these dilemmas by average stage rating is the same for all three grade levels. The average stage rating across the entire sample ranges from 1.8 to 4.3 (see Table 8). Thus, there are some fairly suggestive differences among the six dilemmas, in terms of their average stage ratings, whether the analysis is conducted by age or across the entire sample.

The coefficients of reproducibility (C.O.R.) show a similar pattern of results. The C.O.R. across dilemmas ranged from .88 to .95 for 7th graders, .85 to .95 for 9th graders, and .90 to 1.0 for 11th graders (see Tables 5, 6 and 7). The rank ordering of dilemmas by C.O.R., however, was different for each age. Thus, some fairly suggestive differences among the six dilemmas were evidenced in terms of C.O.R., but the dilemma effect interacts with age as some dilemmas are more scalable than others depending on the age of the subject.

According to the Scalogram Analysis, then, the dilemma effect is very clear. The dilemmas are responded to with different levels of
Table 14
Simple Effects Tests On The Interaction of Consequence At Each Dilemma (Main Study and Combined Data)

<table>
<thead>
<tr>
<th>Consequence Condition</th>
<th>Dilemma</th>
<th>F RATIO</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Main</td>
<td>Combined</td>
</tr>
<tr>
<td>Across Mild, Moderate and Severe Consequences</td>
<td>Heinz</td>
<td>15.53</td>
<td>202.93</td>
</tr>
<tr>
<td></td>
<td>Doctor</td>
<td>75.32</td>
<td>173.05</td>
</tr>
<tr>
<td></td>
<td>Gail</td>
<td>21.82</td>
<td>56.78</td>
</tr>
<tr>
<td></td>
<td>Joe</td>
<td>20.18</td>
<td>37.19</td>
</tr>
<tr>
<td></td>
<td>Draft</td>
<td>57.11</td>
<td>107.72</td>
</tr>
<tr>
<td></td>
<td>Webster</td>
<td>7.75</td>
<td>16.02</td>
</tr>
</tbody>
</table>

Note. d.f. Main Study = 2/870
d.f. Combined Data = 2/1770
n per cell Main Study = 30
n per cell Combined Study = 60

Table 15
Simple Effects Tests On The Interaction of Dilemma At Each Consequence Condition (Main Study and Combined Data)

<table>
<thead>
<tr>
<th>Consequence Condition</th>
<th>Dilemma</th>
<th>F RATIO</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Main</td>
<td>Combined</td>
</tr>
<tr>
<td>Across All Six Dilemmas</td>
<td>Mild</td>
<td>46.95</td>
<td>94.55</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>35.90</td>
<td>58.04</td>
</tr>
<tr>
<td></td>
<td>Severe</td>
<td>11.48</td>
<td>19.12</td>
</tr>
</tbody>
</table>

Note. d.f. Main Study = 5/870
d.f. Combined Data = 5/1770
n per cell Main Study = 30
n per cell Combined Study = 60
Note.  * \( p \leq 0.05 \)
** \( p \leq 0.01 \)
*** \( p \leq 0.001 \)

\( N = 90 \)

Figure 10. The Difference Between Consequence Condition At Each Dilemma on Main Study Rejection Scores
Figure 11. The Difference Between Consequence Condition At Each Dilemma on Combined Data Rejection Scores

Note.  * p ≤ .05  
** p ≤ .01  
*** p ≤ .001  
N = 180
reasoning, as shown by the comparison of average stage ratings, and they are also different in terms of their scalability as measured by the C.O.R.

**Analysis of variance.** The data from the main study, along with the combined data (main and pilot study) were both evaluated with a three by three by six ANOVA by consequence condition, age, and story theme (see Figure 3 for a sketch of the design, page 48). The variable of interest with respect to hypothesis three is dilemma.

As outlined in Table 9 (see page 66), a significant dilemma effect was found ($F = 20.4$, d.f. $= 5/435$, $p < .001$). However, the interaction of dilemma and consequence was also significant ($F = 5.95$, d.f. $= 10/870$, $p < .001$). Therefore, the dilemma effect must be interpreted in terms of consequence.

The interaction of dilemma and consequence was followed up by simple effects tests (see Tables 14 and 15, page 81). This revealed that there was a dilemma effect at each consequence condition (dilemma/mild consequences, $F = 46.95$, d.f. $= 5/870$, $p < .01$; dilemma/moderate consequences, $F = 35.90$, d.f. $= 5/870$, $p < .01$; dilemma/severe consequences, $F = 11.48$, d.f. $= 5/870$, $p < .01$). Similar effects were obtained from the combined analysis of pilot data and main study (see Table 15).

The significant simple effects were followed by a Newman-Kuels analysis to determine specific differences between dilemmas at each consequence condition and between consequence conditions at each dilemma. The results depicted in Figure 12 (see page 86) show that within mild consequences all except three possible pairwise comparisons of the six dilemmas are significant. Two of the dilemmas are significantly different from all of the other five dilemmas while the
remaining four dilemmas differ significantly from four out of five dilemmas. Under moderate consequence conditions one dilemma differs significantly from the five remaining dilemmas while the other five dilemmas differ significantly from four out of five dilemmas. Under severe conditions only three dilemmas show significant differences. Thus significant differences were found between most dilemmas under mild and moderate consequence conditions and severe consequences generated similar scores across dilemmas. The results from the combined analysis (see Figure 13) were similar.

Thus, when consequences were held constant, meaningful differences emerged between dilemmas. Since mild consequences in one story may not equal mild consequences in another story, in terms of severity, subjects were asked to rank order all the consequence conditions in terms of severity. These results (see Table 16, page 88) show that the rank ordering of dilemmas by severity of consequences did not always match the rank ordering of dilemmas by rejection score (see Figures 12 and 13, pages 86 and 87). Thus, differences among dilemmas emerged due to more than just perceived story consequences.

Hypothesis three summary. According to the Scalogram Analysis, the dilemma effect is fairly suggestive. The average stage rating shows that the dilemmas evoke different levels of reasoning. This relationship holds within each age and across ages. In addition, the coefficients of reproducibility (C.O.R.) vary across dilemmas within age groups to show that the dilemmas differ in their rate of scalability.

The dilemma effect is also supported by Analysis of Variance. There are several significant differences between dilemmas at the mild and moderate consequence conditions. In addition, each dilemma shows a strong consequence effect. Thus, there are some very meaningful
Figure 12. The Difference Between Dilemmas At Each Consequence Condition on Main Study Rejection Scores
Figure 13. The Difference Between Dilemmas At Each Consequence Condition on Combined Data Rejection Scores

Note. * p < .05  
** p < .01  
*** p < .001  
N = 180
Table 16
The Rating of Consequence By Age and Dilemma***

<table>
<thead>
<tr>
<th>Consequence Condition</th>
<th>Grade Level</th>
<th>7th</th>
<th>9th</th>
<th>11th</th>
<th>X</th>
<th>7th</th>
<th>9th</th>
<th>11th</th>
<th>X</th>
<th>7th</th>
<th>9th</th>
<th>11th</th>
<th>X</th>
<th>X</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEINZ</td>
<td></td>
<td>4.3</td>
<td>4.8</td>
<td>3.8</td>
<td>4.3</td>
<td>7.3</td>
<td>7.4</td>
<td>7.3</td>
<td>7.3</td>
<td>8.9</td>
<td>9.8</td>
<td>9.6</td>
<td>9.4</td>
<td>9.4</td>
<td>7.0</td>
</tr>
<tr>
<td>DOCTOR</td>
<td></td>
<td>3.7</td>
<td>3.3</td>
<td>3.5</td>
<td>3.5</td>
<td>5.6</td>
<td>7.1</td>
<td>7.5</td>
<td>6.7</td>
<td>9.9</td>
<td>10.0</td>
<td>9.9</td>
<td>9.9</td>
<td>9.9</td>
<td>6.7</td>
</tr>
<tr>
<td>JOE</td>
<td></td>
<td>2.3</td>
<td>2.9</td>
<td>2.3</td>
<td>2.5</td>
<td>4.1</td>
<td>5.0</td>
<td>4.8</td>
<td>4.6</td>
<td>6.1</td>
<td>6.6</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>4.5</td>
</tr>
<tr>
<td>GAIL</td>
<td></td>
<td>3.4</td>
<td>3.5</td>
<td>2.5</td>
<td>3.2</td>
<td>5.1</td>
<td>5.1</td>
<td>4.4</td>
<td>4.9</td>
<td>6.3</td>
<td>7.2</td>
<td>6.0</td>
<td>6.5</td>
<td>6.0</td>
<td>4.9</td>
</tr>
<tr>
<td>WEBSTER</td>
<td></td>
<td>4.5</td>
<td>4.1</td>
<td>3.5</td>
<td>4.0</td>
<td>5.6</td>
<td>5.3</td>
<td>6.0</td>
<td>6.0</td>
<td>7.5</td>
<td>7.7</td>
<td>7.8</td>
<td>7.7</td>
<td>7.7</td>
<td>5.9</td>
</tr>
<tr>
<td>DRAFT</td>
<td></td>
<td>3.9</td>
<td>4.6</td>
<td>4.3</td>
<td>4.3</td>
<td>7.2</td>
<td>7.6</td>
<td>7.7</td>
<td>7.5</td>
<td>5.7</td>
<td>8.8</td>
<td>9.6</td>
<td>8.7</td>
<td>8.7</td>
<td>5.3</td>
</tr>
</tbody>
</table>

Note. * X = Average score by consequence condition across age.
** X = Grand average across all consequence conditions and ages.
*** Consequence rating can range from 1-10, 1 = mild, 10 = severe.
N = 90
differences between dilemmas when consequence and age are controlled, as shown by Scalogram Analysis and ANOVA.

Summary

A pilot study was conducted to evaluate the psychometric properties of the new Rejection Scale. This new Rejection Scale was designed to test three main hypotheses concerned with the effects of consequence and dilemma on moral reasoning of junior high and high school students.

The reliability measures indicate that the Rejection Scale is a consistent measure from one administration to the next (test-retest reliability = .90). In addition, the individual test items (i.e., dilemmas and consequence conditions) were also consistent across administrations as test-retest reliability ranged from .70 to .82.

Concurrent validity measures show that the Rejection Scale is measuring the same construct as another well known measure of moral reasoning (correlation with Defining Issues Test = .55). The correlation is high enough to indicate that the two measures are probably evaluating the same construct and it is low enough to suggest that the two instruments may be different in some meaningful way.

Criterion group validity measures show that the Rejection Scale effectively distinguishes the three age groups from each other. Seventh graders achieve lower scores than ninth graders, who achieve lower scores than eleventh graders. Follow-up tests show that this occurs at mild and moderate consequences but not at the severe consequence condition. In addition, each age group showed the same significant consequence effect.
The structured interview showed that all three ages understood and interacted appropriately with the test materials. This adds a final vote of support that the new Rejection Scale is a psychometrically sound instrument. Research with different populations and other measures of reliability and validity would further clarify this issue.

It was hypothesized that story consequences would influence the reasoning level of 7th, 9th, and 11th grade. This was confirmed with two forms of analysis. Scalogram Analysis demonstrated that the moral dilemmas could be arranged in the form of a cumulative or Guttman scale based on the severity of consequences. This graded scale of difficulty by consequence condition held for each age group, each moral dilemma, each dilemma and age combination, as well as when the data was totalled across all ages and dilemmas.

The Analysis of Variance demonstrated that there was a significant consequence effect at each age with the strongest effects occurring for 9th and 11th graders. Furthermore, mild consequences were most effective at distinguishing the three age groups, followed by moderate then severe consequences.

The consequence effect was also demonstrated in most dilemmas. Most dilemmas distinguished between all three consequence conditions while some distinguished between only two.

Thus, the Scalogram Analysis and Analysis of Variance show a very clear consequence effect.

It was also hypothesized that the Rejection Scale would identify a developmental trend among 7th, 9th, and 11th graders. This was confirmed using two forms of analysis.
Scalogram Analysis demonstrated that 7th graders had a lower stage rating than 9th graders, who had a lower stage rating than 11th graders. This trend occurred without exception in all six dilemmas.

The developmental trend was also demonstrated with Analysis of Variance. There was a significant age effect at each consequence condition with the strongest effect occurring at the mild consequence condition.

In general, the Scalogram Analysis and ANOVA show a strong developmental effect across the three age groups when the data is analyzed by consequence condition, dilemma, and total scores.

It was also hypothesized that story themes (dilemmas) would vary in terms of the level of reasoning they evoked. This was also confirmed using two forms of analysis.

Scalogram Analysis demonstrates a fairly suggestive dilemma effect. The average stage rating shows that the dilemmas evoke different stages of reasoning. This effect holds within each age and across ages. In addition, the coefficients of reproducibility vary across dilemmas within age groups to show that the dilemmas differ in their rate of scalability.

The dilemma effect is also supported by Analysis of Variance. There are several significant differences between dilemmas at the mild and moderate consequence conditions. In addition, each dilemma shows a strong consequence effect. Thus, there are some very meaningful differences between dilemmas when consequence and age are controlled, as shown by the Scalogram Analysis and Analysis of Variance.
CHAPTER XI

DISCUSSION

Purpose of Study, Review

Model theories have been used to understand the reasoning. Kohlberg's simple stage model is the application.

that individuals will evidence the same stage of to different moral situations. In Kohlberg's thought organization, the cognitive structure, is variable in the measurement of moral judgement. The outcome of the moral situation is irrelevant.

Many researchers have failed to find support for Kohlberg's model. The theoretical problem which is most relevant to this study is the lack of evidence for Kohlberg's notion of structured whole.

Rest (1979) has addressed this concern with his complex stage model. He 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). Thus, Rest has adapted Kohlberg's model to accommodate situational aspects along with structural capacity. The significance of these situational factors then, is the foundation of the complex stage model. However, Rest has done little to explicate this facet. The situational factor is dealt with quite haphazardly in Rest's moral assessment device as the total array of potentially influential factors varies unsystematically.
One purpose of the present study was to systematically examine situational influence. This situational concern places this study within the bounds of a functional approach. The functional approach reviewed in this paper suggests that individuals are tuned into the payoff matrix of various situations and will respond to moral dilemmas in a way that will bring about the greatest payoff (Liebert, 1978). One of the most salient functional variables in moral situations is the objectively presented or subjectively perceived or anticipated consequences. As reported earlier, little has been done to examine this functional aspect in Kohlberg's or Rest's paradigm.

Some researchers, however, have examined this variable in children's studies on moral development. It has been demonstrated several times, using Piaget's paradigm with young children, that subjects can be influenced to shift their reasoning depending on the severity of story consequences (Armsby, 1971; Costanzo, 1973; Gutkin, 1972; Hewitt, 1975). In addition, story themes had differential effects on moral reasoning (Nucci, 1981).

Moral reasoning, then, may best be understood through a combined structural-functional approach. This is an approach which illustrates how functional factors such as story consequences and story themes, impinge on the moral reasoning of individuals who differ in their structural capacity. A combined structural-functional model was explored in this research.
Pilot Study

Summary and Review

One problem with Rest's moral assessment device, as outlined earlier, is that it deals haphazardly with the situational factor. Another problem is that it evaluates moral reasoning by reviewing an individual's preference for higher stage reasoning (stages 5 and 6), largely ignoring lower stage reasoning without sufficient justification.

Carroll (1981) addressed this issue by designing an objective measure of moral reasoning based on rejection of lower stage reasoning. This technique is more comprehensive than Rest's because it examines moral reasoning in terms of lower stage patterns. However, the situational factors in Carroll's instrument are still allowed to vary unsystematically. The present study was designed to systematically examine some of the situational factors by bringing them under experimental control. Carroll's instrument was adapted to measure situational factors.

Since this new Rejection Scale was a revision of Carroll's standardized version, its psychometric properties were unknown. The pilot study was designed to establish the psychometric properties of this new instrument.

The results from the pilot study indicate that the instrument is a reliable measure across a retest interval of two weeks. Furthermore, the items which comprise the Rejection Scale are all similar in their high level of consistency across administrations. In addition, its correlation with another well established moral assessment device suggests that the two measures are probably
evaluating the same construct while there may be some meaningful difference between them. The Rejection Scale, then, may be measuring something different about the construct of moral reasoning. The Rejection Scale also has criterion group validity with reference to the seventh, ninth, and eleventh grade population sampled in this study. These three age groups achieved significantly different scores which were consistent with the expected developmental trend.

Implications and Limitations

One of the important features of this new Rejection Scale is that the test-retest reliability coefficients for test items and the total score test-retest reliability were much higher for the Rejection Scale than any of the other major moral assessment devices reviewed in this paper. This may be due partially to the control of story consequences as a situational variant. Since the most current research by Carroll and Rest (1981) and Rest (1979) suggest that situational variables influence moral reasoning, perhaps control of some of these variables will bring about more refined measurement of the construct. The higher test-retest reliability coefficients confirms this. Further research is needed to operationalize and control other relevant situational variables, such as story themes. Based on the present study we can assume that this future research may lead to even greater psychometric accomplishments in the measurement of moral reasoning.

In addition, research should be conducted with other populations and different age groups, using various test-retest intervals. This type of research is necessary for generalization of the results
reported in this study. Since there has been only one study on this new Rejection Scale its psychometric properties should be accepted with some caution.

The concurrent validity statistics reported on the new Rejection Scale suggest that the Rejection Scale is an accurate and unique measure of moral reasoning. One important feature of this concurrent validity is that the Rejection Scale may be measuring something unique or different than the Defining Issues Test (Rest's measure). The Rejection Scale was designed to be sensitive to features of moral reasoning that the Defining Issues Test (DIT) does not measure. More specifically, the Rejection Scale controls story consequence as a situational variant. Perhaps this additional situational control has distinguished the Rejection Scale from the DIT. The increase in reliability coefficients of the Rejection Scale over the DIT also support this. Additional research is needed to clarify and support this assumption. A thorough study of construct validity, correlating the Rejection Scale with other measures of moral development, cognitive development, achievement, intelligence, personality, etc. is necessary at this point.

The criterion group validity results reported earlier show clear and consistent age trends on the new Rejection Scale. This age effect occurs for each dilemma and each consequence condition. Thus, the Rejection Scale seems to be sensitive to the age-developmental trend, with respect to moral reasoning, which has been well documented in the past. In addition, the structured interviews revealed that the seventh, ninth and eleventh graders all interacted appropriately with the test materials. There was some concern at the beginning of this project that the younger subjects would not be able
to understand the task and would not interact appropriately with the test materials. In addition, concern was expressed about whether the story endings could communicate significant differences in consequences. All three age levels reported that the consequences seemed to vary across stories. In addition, even the youngest subjects comprehended and recalled story details and specifics about consequences with remarkable accuracy.

Additional research should be conducted to see if the developmental trend can be documented among younger and older subjects. Some revisions would have to be made for the Rejection Scale to be used with elementary school age children, especially in terms of the reading level of the instrument and the complexity of some stories. In addition, it is unclear how large a role age plays in more advanced levels of moral development. Once subjects have reached eleventh grade, for example, they should have the cognitive capacity to function on the highest level of moral development. Thus, if the upper end of the developmental continuum is to be documented with the Rejection Scale other factors besides age should be considered. Some studies, for example, have used graduate philosophy students to document the upper end of the scale.

An interesting area of research, then, would be the development of a Rejection Scale that could be used with all age groups. Currently, Piaget's measure of moral reasoning is most appropriate for ages 4 to 11, Kohlberg's instrument is appropriate for ages 12 to adult, as is Rest's instrument.

In conclusion, the psychometric properties of this new Rejection Scale need to be established with different populations to include the full range of demographic characteristics. In addition, a
thorough study of construct validity is needed. Finally, situational variables in addition to story consequences should be operationalized and controlled. The evidence from the present study suggests that this additional situational control could improve the reliability of the measure. In general, though, the data generated by this pilot study indicate a trend which suggests that the instrument is psychometrically sound. Hopefully this will encourage others to explore this area in more detail according to some of the suggestions I have detailed above.

Main Study

Situational Effects

A structural-functional approach was explored in this study by examining two situational factors across a four year developmental span. It was hypothesized that story themes and story consequences would interact with structural capacity to influence the moral reasoning of 7th, 9th, and 11th graders. The situational effects of story theme and story consequences will be reviewed in this section.

Summary of results. The consequence effect was significant in all forms of analysis. For example, the Scalogram Analysis demonstrated that the moral dilemmas could be arranged in the form of a cumulative or Guttman Scale based on the severity of consequences. This occurred for each age group and across the entire sample. The ANOVA demonstrated that there was a significant consequence effect at each age, with the strongest effects occurring for the 9th and 11th graders.
The results also show a strong dilemma effect. According to the Scalogram Analysis the average stage ratings show that the different dilemmas evoke different levels of reasoning. This relationship holds for each age group and across ages. The dilemma effect is also supported by Analysis of Variance as there were several significant differences between dilemmas at the mild and moderate consequence conditions. In addition, each dilemma shows a strong consequence effect.

Implications and limitations. The scalability of dilemmas by consequence conditions, along with the significant differences between consequence conditions shown by the ANOVA, suggests that moral dilemmas can be arranged in a cumulative order of difficulty based on consequences. Difficulty is defined in terms of level of reasoning. Situations which are most difficult are those which stimulate the lowest level of reasoning. Also, both forms of analysis also show that there are meaningful differences among dilemmas in addition to consequence differences. Subjects rated all the dilemmas on a scale of one (mild consequences) to ten (severe consequences) in terms of severity of consequences (see Table 16, page 88). These results showed that the rank ordering of dilemmas by severity of consequences did not always match the rank ordering of dilemmas by rejection score (see Figures 10 and 11). Thus, the difficulty of dilemmas is due to more than just perceived story consequences.

Perhaps story themes should be operationalized and controlled, as consequences were in this study, to more systematically examine their effects. One relevant dimension upon which story themes could be ordered is the protagonist-antagonist relationship. Subjects who
were interviewed in this study made distinctions between dilemmas based on the nature of this relationship, as it involved personal relationships in some cases (e.g., Heinz, Doctor), social/conventional relationships in other instances (Webster) and one dilemma involved a relationship with the world or human race in general (Draft). This relationship dimension fits well within Piaget's cognitive theory and Kohlberg's theory of moral development. According to Piaget (Flavell, 1963) we would expect development to proceed from a focus on personal, immediate, concrete experiences to a rule orientation that overrides the immediate and concrete, to an abstract level of functioning based on principles that are not necessarily defined by rules or convention. Kohlberg suggests a similar scheme of development with respect to the moral sphere.

Thus, based on the information from the structured interviews, and the models of cognitive and moral development proposed by Piaget and Kohlberg, I suggest that moral themes can be ordered along a dimension of story character relationship. This issue should be addressed empirically along with the exploration of other dimensions along which story themes could be ordered.

The story consequences and other features of the moral dilemmas are relevant situational factors that contribute to the difficulty of these moral dilemmas. As dilemmas vary in their level of difficulty we must expect variation within individuals in terms of the type of reasoning used across the different stories.

Bem and Allen (1974) suggest that in our search for cross situational consistencies in behavior we must consider the relative difficulty of various situations. Thus, we cannot expect to find
consistency based solely on our knowledge of person variables such as developmental level. Consider the example of children learning math. Some math problems are passed and others failed depending on the interaction of difficulty and the child's level of development. Bem and Allen report that as long as the pattern of passes and failures matches our preconceived notion of difficulty then we can regard behavior as consistent even though not all items are passed or all items failed. For example, when a young child passes an addition problem yet fails a calculus problem he/she is regarded as consistent. This same analysis can be applied to the area of moral reasoning. Just as there are features of math problems which make them more or less difficult there are also features of moral problems (such as story consequences) that allow these dilemmas to be scaled along a dimension of difficulty. Thus we should expect subjects to pass certain moral situations (reject level 1 reasoning) and not others because of difficulty (situational or functional influence) and developmental level (structural influence). Variability in reasoning, then, should not be associated with inconsistency. In fact, cross-situational variability may well be the mark of a highly refined discriminative facility - the ability to respond appropriately to subtle changes in situational contingencies.

Individuals, then, are not entirely predictable based on a trait or structure of moral reasoning and they are not predictable by the situational contingencies alone. Perhaps the greatest degree of predictability comes from knowledge about the person and the situation and how these two variables interact.

The trait approach, espoused by rigid structuralists such as Kohlberg, leaves us short of a complete understanding. Individuals
will appear to be inconsistent in their reasoning across situations if we attend only to their structural capacity. The situational approach, espoused by the pure functionalists, also leaves us short of a thorough understanding as the same situations bring about very different response patterns in individuals who are hypothesized to have structural differences. An integrated approach yields the greatest level of understanding as individuals with more advanced structural capacity pass the more functionally difficult moral items. We must identify the relevant features of moral situations that make them more or less difficult to better understand how the functional interacts with the structural to generate moral responses.

Bowers (1973) reports that both the trait and situation approach explain little of the change in experimental studies compared to the interaction approach:

An interactionist or biocognitive view denies the primacy of either traits or situations in the determination of behavior; instead it fully recognizes that whatever main effects do emerge will depend entirely upon the particular sample of settings and individuals under consideration...interactionism argues that situations are as much a function of the person as the person's behavior is a function of the situation (p. 327).

Bowers (1973) claims that individuals or behaviors differ in their vulnerability to situational influence. The age and consequence interaction reported in the results demonstrates this clearly. Under mild consequences definite age trends appear as the structural capacity of the older subjects leads to higher levels of moral reasoning. Under moderate consequence conditions there are no longer any significant differences between 7th and 9th graders, while 11th graders still show a higher level of reasoning. Finally, under the severe condition the age groups converge, reducing the significant
differences. Vulnerability to situational influence, then, is a function of structural capacity and the strength of the functional aspect. Again, knowledge of the situation or person variables alone have limited value in terms of predictability next to the integrated structural-functional model.

Consequences are clearly relevant functional aspects of moral situations, and dilemmas can be ordered in terms of difficulty to form a Guttman or cumulative scale based on these consequences. A look at story themes leads to similar conclusions. There are meaningful differences among the dilemmas in addition to the consequence effect. Future research should focus on discovering other relevant situational variables and documenting their interaction with person variables to further refine the structural-functional model. Based on the results of this study we can assume that greater understanding or predictability will be achieved in this area as more relevant situational variables are documented and a system of assessment is designed to accommodate this. Perhaps this level of understanding would help explain how human atrocities such as the Holocaust and the Mai Lai Massacre could occur in societies comprised of individuals capable of reasoning on the highest moral levels.

Developmental Trend

It was hypothesized that the Rejection Scale would have criterion group validity, thus it would be sensitive to the developmental trend in moral reasoning between 7th and 11th grade. The developmental trend in moral reasoning, as measured by the Rejection Scale, was documented across three age groups while situational variables were
controlled. This is the structural portion of the structural-functional model. Regardless of situational influence, differences will occur across individuals, hypothesized to be due to the structural or developmental factor.

Summary of Results. Scalogram Analysis showed that 7th graders had a lower average stage rating than 9th graders and they were both lower than the 11th graders. This trend was matched without exception by all six moral dilemmas.

The developmental trend was also documented with Analysis of Variance. There was a significant age effect at each consequence condition with the strongest effect occurring at the mild consequence condition. Moderate consequences distinguished 11th from 7th and 9th graders, as did the severe condition at a lower level of significance.

Implications and limitations. Perhaps the most interesting age effect is the interaction with consequences. As Bowers (1973) suggests, behaviors or personality traits differ in their vulnerability to situational influence. This point is clearly illustrated in the age/consequence interaction. Figures 8 and 9 show that there are significant differences between all three age groups under the mild consequence condition. However, moderate consequences have a greater relative effect on 9th graders than they do with the 7th and 11th graders. Thus the difference between 7th and 9th graders decreases under moderate consequences while the difference between 9th and 11th graders increases. Finally, the 11th graders are more affected by severe consequences than the other two age groups. Thus individuals differ in their vulnerability to situational influence, illustrating the unique structural-functional interaction.
While 11th graders may have the cognitive capacity to achieve the highest structural levels of reasoning, age does not guarantee development. A developmental trend was documented across 7th, 9th and 11th grade, however, these subjects occupied the middle stages of this new scale. It would be interesting to explore the person-situation interaction at both ends of the developmental scale. The current trend suggests that the more developmentally advanced individuals will manifest a smaller rate of vulnerability to situational influence. However, just as with the 11th graders in this study, potent situational influences could probably be created to influence the reasoning of individuals at the upper end of the scale.

The application of this finding to "real world" problems is somewhat unsettling. For example, we may select individuals of the highest moral character to govern our countries and participate in sophisticated scientific ventures such as space exploration, disease control, and nuclear weaponry. The potency of these situations, in terms of influence on level of reasoning, is extremely high. Individuals with the most sophisticated reasoning capacity, then, may find themselves in situations that have an overwhelming influence on their level of reasoning. A decrement in reasoning level when dealing with issues such as the prospect of nuclear war can be devastating.

Perhaps there are situations that are so potent that they cause even the most sophisticated reasoners to regress. Perhaps these are the very situations where the highest levels of reasoning are necessary to make critical decisions which affect the lives of thousands of people. The relationship of situational influence and
regression in reasoning is unsettling because regression is always in the direction of egocentricity or self interest, and most of the potent situations which would bring about regression involve decisions about the lives of many or the general welfare of the human species.

Obviously we need to understand more about the effects of situational influence on higher level reasoning. The trend documented in this study is suggestive but it only deals with the middle range of reasoning.

CHAPTER XII

CONCLUSIONS

The purpose of this project was to integrate the structural and functional approach of moral reasoning. This was accomplished by measuring functional aspects such as story consequences and story themes at different structural-developmental levels (7th, 9th and 11th graders). The result was a structural-functional model which showed unique effects of story theme, consequence, and age, along with some very meaningful interactions of these variables. For example, as consequences increased in severity, Rejection Scale scores decreased. This trend was consistent for each age group, however, the age groups differed in their vulnerability to this situational influence. Under mild consequence conditions the three age groups were significantly different. Moderate consequences affected the 9th graders more than the other two age groups and severe consequences had the greatest effect on 11th graders.
In general, this project generated support for both a structural and functional model. However, the greatest level of understanding is achieved from the integrated approach which considers unique structural-functional interactions.

Further research is needed to identify additional functional aspects of moral situations which impact on the reasoner, and document other unique structural-functional interactions. The conclusions from this study were based on 7th, 9th and 11th graders who occupied the middle of the developmental scale. Additional research is needed to build a structural-functional model that spans the full range of development.

Finally, to facilitate the external validity of this new structural-functional model, several studies should be conducted with populations that vary in their demographic characteristics. In addition, the laboratory approach should be adapted for field study so that moral reasoning and behavior can be studied under more realistic conditions. At this point the results of this study are highly suggestive. Field studies are necessary to confirm these results.
REFERENCES


Kuhn, D. Short term longitudinal evidence for the sequentiality of Kohlberg's early stages of moral judgment. Developmental Psychology, 1976, 12, 162-166.


Appendix A:

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?
PART A. (SAMPLE)

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

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 5
Second most important 2
Third most important 3
Fourth most important 1
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.

<table>
<thead>
<tr>
<th>GREAT importance</th>
<th>MUCH importance</th>
<th>SOME importance</th>
<th>LITTLE importance</th>
<th>NO importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Whether a community's laws are going to be upheld.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Isn't it only natural for a loving husband to care so much for his wife that he'd steal?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Is Heinz willing to risk getting shot as a burglar or going to jail for the chance that stealing the drug might help?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Whether Heinz is a professional wrestler, or has considerable influence with professional wrestlers.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Whether Heinz is stealing for himself or doing this solely to help someone else.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Whether the druggist's rights to his invention have to be respected.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Whether the essence of living is more encompassing than the termination of dying, socially and individually.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. What values are going to be the basis for governing how people act towards each other.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Whether the druggist is going to be allowed to hide behind a worthless law which only protects the rich anyhow.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Whether the law in this case is getting in the way of the most basic claim of any member of society.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Whether the druggist deserves to be robbed for being so greedy and cruel.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Would stealing in such a case bring about more total good for the whole society or not.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 Students for a Democratic Society (SDS), believe 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
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:

<table>
<thead>
<tr>
<th>Most important</th>
<th>Second most important</th>
<th>Third most important</th>
<th>Fourth most important</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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

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 could 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
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
<table>
<thead>
<tr>
<th>Importance</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Great</td>
<td>1. Is the principal more responsible to students or to parents?</td>
</tr>
<tr>
<td>Much</td>
<td>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?</td>
</tr>
<tr>
<td>Some</td>
<td>3. Would the students start protesting even more if the principal stopped the newspaper?</td>
</tr>
<tr>
<td>Little</td>
<td>4. When the welfare of the school is threatened, does the principal have the right to give orders to students?</td>
</tr>
<tr>
<td>None</td>
<td>5. Does the principal have the freedom of speech to say &quot;no&quot; in this case?</td>
</tr>
<tr>
<td></td>
<td>6. If the principal stopped the newspaper would he be preventing full discussion of important problems?</td>
</tr>
<tr>
<td></td>
<td>7. Whether the principal's order would make Fred lose faith in the principal?</td>
</tr>
<tr>
<td></td>
<td>8. Whether Fred was really loyal to his school and patriotic to his country.</td>
</tr>
<tr>
<td></td>
<td>9. What effect would stopping the paper have on the student's education in critical thinking and judgment?</td>
</tr>
<tr>
<td></td>
<td>10. Whether Fred was in any way violating the rights of others in publishing his own opinions.</td>
</tr>
<tr>
<td></td>
<td>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.</td>
</tr>
<tr>
<td></td>
<td>12. Whether Fred was using the newspaper to stir up hatred and discontent.</td>
</tr>
</tbody>
</table>

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
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
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
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. This project is not part of the school program and no one in school will see any of these completed papers. Also, your work will be kept strictly confidential as your names will not appear in any report of this project.

Please give us the following information:

Name ________________________________ female
Age ______ School ________________________ male

______________________________
In this questionnaire you will be asked to give your opinions about several stories. Here is a story as an example.

Frank Jones

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 also. He has the chance to buy a 1970 station wagon from a friend. In trying to decide whether to buy this car, Frank Jones realized that there were many reasons for making important decisions like this.

When deciding about what Frank Jones should do, consider the reasons listed below. First choose yes or no about whether Mr. Jones should buy the car, then answer the questions under the choice you made. Answer these questions by circling G for a good answer (it is a good reason for making the decision in this case); P for a poor answer (it is not a good reason for making the decision in this case); or U for uncertain (you are not sure if it is a good reason for making the decision in this case).

Should Frank Jones buy this car?

<table>
<thead>
<tr>
<th>YES</th>
<th>Uncertain</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
</tbody>
</table>

A large room car is better than a compact car.

The color of the car is the same as the color of Frank's house.

The price of the car is low.

This car may use too much gas.

People should not buy cars from friends.

The price of the car is too high.

Note: The person in the sample decided that Frank should buy the car, so only the answers under the YES choice were circled and none of the NO responses were circled.

Now read through the next stories and answer yes or no about what the person should do. Then circle G (good reason), U (uncertain about this reason), or P (poor reason) next to each reason. You must answer yes or no for each story. Both choices may seem reasonable but choose just one anyway.

Each main story will be repeated three times with a different ending each time. Read the main story and all three different endings first, then go back and answer the questions under each ending. Once you have done this for one main story, move on to the next main story and answer it in the same manner (read all three endings first).
THE DOCTOR'S DILEMMA

A lady in China 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 that she couldn't stand the pain and that she was going to die in a few months anyway.

Ending 1

The penalty in China, for doctors who kill out of mercy, is that they must give at least half of their salary to charity every year for the rest of their life. The doctor was not sure if anyone would realize that the death was caused by him.

Should the doctor give the lady the overdose that will cause her death?

<table>
<thead>
<tr>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>YES</th>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>It will relieve the doctor of some responsibility.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>The doctor would get into too much trouble.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>If, and only if, there is a law allowing the doctor to give the overdose.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>If, and only if, it is against the law.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>The lady, as an individual, has the right to choose to die.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>No individual has the right to interfere with life's course.</td>
</tr>
</tbody>
</table>
There is no penalty in China, for doctors who end life out of mercy, if the judge decides the doctor made the best decision. If the judge decides that the doctor did not make the right decision the doctor must give free medical treatment to the family of the victim.

Should the doctor give the lady the overdose that will cause her death?

<table>
<thead>
<tr>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>YES</th>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>It will relieve the doctor of some responsibility.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>The doctor would get into too much trouble.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>If, and only if, there is a law allowing the doctor to give the overdose.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>If, and only if, it is against the law.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>The lady, as an individual, has the right to choose to die.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>No individual has the right to interfere with life's course.</td>
</tr>
</tbody>
</table>

The penalty in China for killing someone, by accident or on purpose, is the death penalty. The doctor knew that the people in charge of the hospital were watching him closely to see what he was going to do.

Should the doctor give the lady the overdose that will cause her death?

<table>
<thead>
<tr>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>YES</th>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>It will relieve the doctor of some responsibility.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>The doctor would get into too much trouble.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>If, and only if, there is a law allowing the doctor to give the overdose.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>If, and only if, it is against the law.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>The lady, as an individual, has the right to choose to die.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>No individual has the right to interfere with life's course.</td>
</tr>
</tbody>
</table>
Ann took the school lunch money which had been collected from all the students for a full week. When she saw Gail, her best friend, she told her that she had something very serious and personal that she wanted to talk about. But, before she told Gail anything about the incident, she asked for her promise not to tell anyone what she said. Gail agreed to keep everything she heard a secret. When Ann told Gail she took the money, Gail was in a difficult position.

**Ending 1**

She had given her word not to tell anyone but she knew that until the thief was caught there would be no passes or special privileges for any students, and the whole class would have to stay after school every day until the truth was discovered. As long as Gail stayed after school she would be unable to try out for the tennis team, even though she was supposed to be the captain this year.

**Should Gail tell the teacher that Ann took the money?**

<table>
<thead>
<tr>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>YES</th>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Because it creates too much trouble if she does not turn her in.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Because Ann might take revenge and hurt Gail.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Because Ann broke a school rule and should be reported.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Because it could ruin her friendship with Ann.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Because Ann's behavior goes against principles of right and wrong. Even if it were not against the rules, it should be reported.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Because trust between friends is the most important principle.</td>
</tr>
</tbody>
</table>
Ending 2

She had given her word not to tell anyone but she knew that until the thief was caught there would be no passes or special privileges for any students. Also, students would have to use a different procedure for buying lunches, which would create an extra hassle for them.

Should Gail tell the teacher that Ann took the money?

<table>
<thead>
<tr>
<th>YES</th>
<th>GOOD</th>
<th>UNCERTAIN</th>
<th>POOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>P</td>
<td>G</td>
<td>U</td>
</tr>
<tr>
<td>P</td>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
</tbody>
</table>

Because it creates too much trouble if she does not turn her in.

Because Ann broke a school rule and should be reported.

Because Ann's behavior goes against principles of right and wrong. Even if it were not against the rules, it should be reported.

<table>
<thead>
<tr>
<th>NO</th>
<th>GOOD</th>
<th>UNCERTAIN</th>
<th>POOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>G</td>
<td>P</td>
</tr>
<tr>
<td>U</td>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
</tbody>
</table>

Because Ann might take revenge and hurt Gail.

Because it could ruin her friendship with Ann.

Because trust between friends is the most important principle.

Ending 3

She had given her word not to tell anyone but she knew that until the thief was caught students would have to use a different procedure for buying lunches, which would create an extra hassle for them.

Should Gail tell the teacher that Ann took the money?

<table>
<thead>
<tr>
<th>YES</th>
<th>GOOD</th>
<th>UNCERTAIN</th>
<th>POOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>P</td>
<td>G</td>
<td>U</td>
</tr>
<tr>
<td>P</td>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
</tbody>
</table>

Because it creates too much trouble if she does not turn her in.

Because Ann broke a school rule and should be reported.

Because Ann's behavior goes against principles of right and wrong. Even if it were not against the rules, it should be reported.

<table>
<thead>
<tr>
<th>NO</th>
<th>GOOD</th>
<th>UNCERTAIN</th>
<th>POOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>G</td>
<td>P</td>
</tr>
<tr>
<td>U</td>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
</tbody>
</table>

Because Ann might take revenge and hurt Gail.

Because it could ruin her friendship with Ann.

Because trust between friends is the most important principle.
In Europe a woman was suffering unbearable pain from a kidney infection. There was one drug that the doctors thought might cure this illness. It was a rare form of penicillin 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 him to make. He paid $200 for the penicillin and was charging $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 $1,000 which is half of what it cost. He told the druggist that his wife was extremely ill and in severe pain, 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.

**Ending 1**

Since the store is well guarded, it is likely that Heinz would be caught and arrested. The penalty for theft over $1,000 in this country is life in prison with hard labor.

Should Heinz steal the drug?

<table>
<thead>
<tr>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>YES</th>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Heinz might be miserable throughout his life if his wife stays sick.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Heinz would get in too much trouble.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Heinz has a responsibility to take care of his wife.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>It is wrong even if he does not get caught because stealing is against the law.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Saving a life is more important than the druggist's legal right of property.</td>
<td>U</td>
<td>P</td>
<td>The druggist's right to property should be respected.</td>
<td></td>
</tr>
</tbody>
</table>
Ending 2

The store had a burglar alarm but since it was located in a secluded section of town Heinz could probably take the drug without getting caught. The penalty for theft in this country involves returning the stolen item and one week of hard work for the victim.

Should Heinz steal the drug?

<table>
<thead>
<tr>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>YES</th>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Heinz might be miserable throughout his life if his wife stays sick.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Heinz would get in too much trouble.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Heinz has a responsibility to take care of his wife.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>It is wrong even if he does not get caught because stealing is against the law.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Saving a life is more important than the druggist's legal right of property.</td>
<td>U</td>
<td>P</td>
<td></td>
<td>The druggist's right to property should be respected.</td>
</tr>
</tbody>
</table>

Ending 3

The store is part of the security rounds of the local police so Heinz will probably get caught soon after he delivers the drug to his wife. The penalty for theft in this country is a two year jail sentence.

Should Heinz steal the drug?

<table>
<thead>
<tr>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>YES</th>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Heinz might be miserable throughout his life if his wife stays sick.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Heinz would get in too much trouble.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Heinz has a responsibility to take care of his wife.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>It is wrong even if he does not get caught because stealing is against the law.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Saving a life is more important than the druggist's legal right of property.</td>
<td>U</td>
<td>P</td>
<td></td>
<td>The druggist's right to property should be respected.</td>
</tr>
</tbody>
</table>
Joe is a fourteen year old boy who wanted to buy a radio for his room. His father promised him he could buy the radio if he saved the money for it himself. So Joe did some extra chores around the house and odd jobs for neighbors. So, he saved the $50.00 he needed for the radio and a little more besides. But just before he purchased the radio his father changed his mind and wanted to spend Joe's money on a fishing trip with his friends. So he asked Joe to give him the money he had saved.

**Ending 1**

While Joe knew that he would not get in any trouble for refusing to give the money to his father, he realized that his father might get a little angry.

Should Joe give his father the money?

<table>
<thead>
<tr>
<th>YES</th>
<th>Uncertain</th>
<th>Poor</th>
<th>NO</th>
<th>Uncertain</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>Because otherwise he will get in too much trouble.</td>
<td></td>
<td>Because promises are always supposed to be kept.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td>A good son should do what his father wants.</td>
<td>Joe has the right to keep the money he earned regardless of his duty as a son.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td>Joe should trust that his father would be doing what is best for everyone.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
He explained that if Joe did not give him the money he would lose his bicycle and not be allowed to spend his summer vacation at camp. Riding the bicycle is Joe's favorite activity and he looks forward to summer camp all year.

Should Joe give his father the money?

<table>
<thead>
<tr>
<th></th>
<th>YES</th>
<th></th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Good</td>
<td>Uncertain</td>
<td>Poor</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Because otherwise he will get in too much trouble.

Because he wants the radio.

A good son should do what his father wants.

Because promises are always supposed to be kept.

Joe should trust that his father would be doing what is best for everyone.

Joe has the right to keep the money he earned regardless of his duty as a son.

He explained that if Joe did not give him the money he would not be allowed to watch television or see his friends that week.

Should Joe give his father the money?

<table>
<thead>
<tr>
<th></th>
<th>YES</th>
<th></th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Good</td>
<td>Uncertain</td>
<td>Poor</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Because otherwise he will get in too much trouble.

Because he wants the radio.

A good son should do what his father wants.

Because promises are always supposed to be kept.

Joe should trust that his father would be doing what is best for everyone.

Joe has the right to keep the money he earned regardless of his duty as a son.
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.

Ending 1

Most of his customers not only promised to take their business elsewhere, but also threatened to destroy Mr. Webster's property.

Should Mr. Webster hire Mr. Lee?

<table>
<thead>
<tr>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>YES</th>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>A good mechanic could help business.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>It might hurt Mr. Webster's business too much.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Because racial discrimination is against the law.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>If most other people in the community would agree with not hiring Mr. Lee.</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Hiring should be based on ability not ethnic background.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Fair hiring procedures will not do anyone any good if the business fails.</td>
</tr>
</tbody>
</table>
Ending 2

Some of his customers didn't like Orientals and would probably take their business elsewhere if Mr. Lee was working in the gas station.

Should Mr. Webster hire Mr. Lee?

<table>
<thead>
<tr>
<th></th>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>YES</th>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>A good mechanic could help business.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>It might hurt Mr. Webster's business too much.</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Because racial discrimination is against the law.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>If most other people in the community would agree with not hiring Mr. Lee.</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Hiring should be based on ability not ethnic background.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Fair hiring procedures will not do anyone any good if the business fails.</td>
<td></td>
</tr>
</tbody>
</table>

Ending 3

Some of his customers didn't like Orientals and they might get angry.

Should Mr. Webster hire Mr. Lee?

<table>
<thead>
<tr>
<th></th>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>YES</th>
<th>Good</th>
<th>Uncertain</th>
<th>Poor</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>A good mechanic could help business.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>It might hurt Mr. Webster's business too much.</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Because racial discrimination is against the law.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>If most other people in the community would agree with not hiring Mr. Lee.</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Hiring should be based on ability not ethnic background.</td>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Fair hiring procedures will not do anyone any good if the business fails.</td>
<td></td>
</tr>
</tbody>
</table>
DRAFT DECISION

Bill has reached the age for draft registration for the armed forces. He is wondering whether he should register or not. Where Bill lives people who have not registered for the draft have been asked to do two years of community service, such as working in a hospital.

Should Bill register for the draft?

<table>
<thead>
<tr>
<th>YES</th>
<th>Uncertain</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>Because otherwise Bill may get in too much trouble.</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>Because it is a citizen's duty and a national law.</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>The leaders of the country are better able than any individual to make decisions concerning the well-being of the entire nation.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NO</th>
<th>Uncertain</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>Because Bill could get hurt in a war.</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>If a majority of Bill's family and friends are against the draft.</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>Individuals should have the right to decide whether or not to risk their lives.</td>
<td></td>
</tr>
</tbody>
</table>
Ending 2

sentenced to jail for ten years with no chance for parole.

Should Bill register for the draft?

<table>
<thead>
<tr>
<th>Uncertain</th>
<th>Poor</th>
<th>YES</th>
<th>Uncertain</th>
<th>Poor</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Because otherwise Bill may get in too much trouble.</td>
<td>G</td>
<td>U</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Because it is a citizen's duty and a national law.</td>
<td>G</td>
<td>U</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>The leaders of the country are better able than any individual to make decisions concerning the well-being of the entire nation.</td>
<td>G</td>
<td>U</td>
</tr>
</tbody>
</table>

Ending 3

required to donate one half of their salary to the government for ten years.

Should Bill register for the draft?

<table>
<thead>
<tr>
<th>Uncertain</th>
<th>Poor</th>
<th>YES</th>
<th>Uncertain</th>
<th>Poor</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Because otherwise Bill may get in too much trouble.</td>
<td>G</td>
<td>U</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>Because it is a citizen's duty and a national law.</td>
<td>G</td>
<td>U</td>
</tr>
<tr>
<td>G</td>
<td>U</td>
<td>P</td>
<td>The leaders of the country are better able than any individual to make decisions concerning the well-being of the entire nation.</td>
<td>G</td>
<td>U</td>
</tr>
</tbody>
</table>


Kuhn, D. Short term longitudinal evidence for the sequentiality of Kohlberg's early stages of moral judgment. Developmental Psychology, 1976, 12, 162-166.


