THE RELATIONSHIP BETWEEN MORAL JUDGMENT AND CHRISTIAN BELIEFS WITH MODERATOR VARIABLES OF INTRINSIC RELIGIOUS ORIENTATION, SOCIOECONOMIC STATUS, YEAR IN SCHOOL, AND IQ

Dissertation Committee:
Dr. Theodore J. Kaul
Dr. Lyle D. Schmidt
Dr. Samuel H. Osipow

Approved by
Dissertation Committee:
Dr. Theodore J. Kaul
Dr. Lyle D. Schmidt
Dr. Samuel H. Osipow

Adviser
Department of Psychology
"Stop judging by mere appearances,
and make a right judgment."
John 7:24
To Elizabeth Hansen,
my soul mate, and faithful companion,
who supported me in every way throughout this Ph.D. endeavor
and to my parents,
who dedicated their lives to counseling and to serving God
ACKNOWLEDGMENTS

I truly could not have accomplished this task without assistance. And God has been faithful to supply me with many people who were willing to help and skilled in their respective areas. I am most grateful to Dr. Ted Kaul and to the other members of my committee, Dr. Lyle Schmidt and Dr. Sam Osipow, for their guidance, support, and encouragement throughout my graduate career at the Ohio State University, as well as their comments and suggestions on this Dissertation. I am also indebted to Dr. Richard Gorsuch, Dr. Ed Borgatta, Dr. James Rest, and Dr. James Hogge who graciously granted me permission to use the measures they created and answered my many technical questions in their respective areas of expertise. Dr. Bonnidell Clouse and Dr. Ray Paloutzian, both Christian psychologists interested in the scientific study of religion, have also encouraged me in undertaking this endeavor. Dr. John Gibbs and Dr. Bob Rodgers provided their perspectives on cognitive moral development and a sounding board in the early stages as I formulated this study. Thanks also to Tom Smith of the National Opinion Research Center for mailing me a complete listing of denominations coded in the General Social Survey.

Statistical consultants for the Psychology Department at the Ohio State University, especially Lee Fabrigar, helped me refine
methods to test my hypotheses, and provided a translation of statistical theory into computer programming language. Expert consultation in computing statistics was provided by Fred Ruland and his staff at the Ohio State University in Columbus, Barb House from the Ohio State University, and Kris Arhart from the University of Tennessee in Memphis.

Finally, my support network of friends and family sustained me physically, emotionally, and spiritually to do the mental work herein. I would like to thank the Ricketts and the Karls for hosting me during my pilgrimages from Wisconsin to Ohio; Rick Marrah for his encouraging prayers; and the Demien family who allowed me to use their phone and modem for endless hours. Last, but of course most importantly, I wish to acknowledge and thank my wife for her steadfast patience, her financial support, and her love, which she expressed in countless practical sacrifices over the last 4 years.
VITA

March 24, 1961 ................................ Born — Decorah, Iowa


1986–1987 ....................................... Assistant Hall Director, Residence and Dining Halls, The Ohio State University

1987–1988 ....................................... Learning Support Associate, Residence and Dining Halls, The Ohio State University

1988–1989 ....................................... Graduate Research Associate, Office for Disability Services, The Ohio State University

1989–1991 ....................................... Graduate Teaching Associate, Psychology Department, The Ohio State University

1990 ............................................... M. A., Counseling Psychology, Graduate School, The Ohio State University
1991-1993 . . . . . . . . . Career Counselor,
Colleges of Arts and Sciences,
The Ohio State University

1993-1994 . . . . . . . . . Psychology Intern,
University of Tennessee
Professional Psychology
Internship Consortium,
Memphis, Tennessee

FIELDS OF STUDY

Major Field: Psychology

Specialization: Counseling Psychology
# TABLE OF CONTENTS

DEDICATION ................................................................................................. ii

ACKNOWLEDGMENTS .................................................................................. iii

VITA .............................................................................................................. v

LIST OF TABLES ........................................................................................ viii

LIST OF FIGURES ...................................................................................... xvi

CHAPTER PAGE

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. LITERATURE REVIEW</td>
<td>7</td>
</tr>
<tr>
<td>III. METHOD</td>
<td>75</td>
</tr>
<tr>
<td>IV. RESULTS AND DISCUSSION</td>
<td>97</td>
</tr>
<tr>
<td>V. SUMMARY AND RECOMMENDATIONS</td>
<td>182</td>
</tr>
</tbody>
</table>

APPENDICES

<table>
<thead>
<tr>
<th>APPENDICES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Defining Issues Test and Directions</td>
<td>190</td>
</tr>
<tr>
<td>B. Examples of Principled Level Christian Teachings</td>
<td>196</td>
</tr>
<tr>
<td>C. Quick Word Test (First Five Items Only)</td>
<td>200</td>
</tr>
</tbody>
</table>

vii
D. Three Religious Scales and a Questionnaire including:
   Age Universal I/E-Revised Scales (items 1-14),
   Short Christian Orthodoxy Scale (items 15-20),
   Scriptural Literalism Scale (items 21-44)
   and Demographics Questionnaire (items 45-62).............202

E. Socioeconomic Index Coding Guidelines .........................207

F. Oral Instructions to Subjects ...........................................210

G. Debriefing ...................................................................214

H. Table 55: Frequency Table of Subject Denomination ..........216

I. Selected Raw Regression Equations and
   Standard Errors not Retrievable from
   Results Presented Above ................................................219

J. Regression Equations and Coordinates
   used to Graph Figures 3 and 5...........................................221

LIST OF REFERENCES................................................................224
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Relationship of the Three Levels to the Six Stages of Moral Judgement</td>
<td>12</td>
</tr>
<tr>
<td>2. Stages of Moral Judgment According to Rest</td>
<td>13</td>
</tr>
<tr>
<td>3. Total Number of DIT Queries Available at Each Stage</td>
<td>20</td>
</tr>
<tr>
<td>4. Descriptive Statistics for Study Variables</td>
<td>99</td>
</tr>
<tr>
<td>5. Denominational Classification Frequency Table based on Smith (1990)</td>
<td>104</td>
</tr>
<tr>
<td>6. Analysis of Variance Summary Table for Denominational Classification (Excluding Agnostics/Atheists) and Short Christian Orthodoxy Scale</td>
<td>106</td>
</tr>
<tr>
<td>7. Analysis of Variance Summary Table for Denominational Classification (Including Agnostics/Atheists) and Short Christian Orthodoxy Scale</td>
<td>106</td>
</tr>
<tr>
<td>8. Least Squares Means of Short Christian Orthodoxy Scale for Fundamentalists, Moderates, and Liberals Plus Agnostics/Atheists</td>
<td>107</td>
</tr>
<tr>
<td>9. Analysis of Variance Summary Table for Denominational Classification (Excluding Agnostics/Atheists) and Scriptural Literalism Scale</td>
<td>109</td>
</tr>
<tr>
<td>10. Least Squares Means of Scriptural Literalism Scale for Fundamentalists, Moderates, and Liberals</td>
<td>109</td>
</tr>
</tbody>
</table>
11. Analysis of Variance Summary Table for Denominational Classification (Including Agnostics/Atheists as Liberals) and Scriptural Literalism Scale .......................... 110

12. Least Squares Means of Scriptural Literalism Scale for Fundamentalists, Moderates, and Liberals Plus Agnostics/Atheists .............................................. 110

13. Analysis of Variance Summary Table for Denominational Classification (Excluding Agnostics and Atheists) and Quick Word Test ........................................................................ 112

14. Analysis of Variance Summary Table for Denominational Classification (Including Agnostics and Atheists as Liberals) and Quick Word Test .......................................................... 113

15. Analysis of Variance Summary Table for Denominational Classification (Excluding Agnostics and Atheists) and Socioeconomic Index .................................................. 114

16. Analysis of Variance Summary Table for Denominational Classification (Including Agnostics and Atheists as Liberals) and Socioeconomic Index .......................................................... 114

17. Least Squares Means of Socioeconomic Status for Fundamentalists, Moderates, and Liberals (Excluding Agnostics/Atheists) .............................................................. 115

18. Least Squares Means of Socioeconomic Status for Fundamentalists, Moderates, and Liberals Plus Agnostics/Atheists ................................................................................ 115

19. Zero-order Correlations between Defining Issues Test P score and IQ, Socioeconomic Status, and Year in School for Sample Including Agnostics/Atheists ........................................ 120
20. Zero-order Correlations between Defining Issues Test P score and IQ, Socioeconomic Status, and Year in School for Sample Excluding Agnostics/Atheists ............................................. 121

21. Zero-Order Correlations between Defining Issues Test P score and Short Christian Orthodoxy and Scriptural Literalism Scales .................................................. 122

22. Zero-Order Correlations among Defining Issues Test P score, Short Christian Orthodoxy Scale, Intrinsic Scale, and the Product of Short Christian Orthodoxy Scale, and Intrinsic Scale for Sample Excluding Agnostics/Atheists .......................... 132

23. Intermediate Regression Analysis Regressing Defining Issues Test P score on Short Christian Orthodoxy Scale, Intrinsic Scale, and the Interaction of Short Christian Orthodoxy Scale, and Intrinsic Scale for Sample Excluding Agnostics/Atheists .......................................................... 133

24. Simultaneous Regression Analysis Regressing Defining Issues Test P score on Short Christian Orthodoxy Scale and Intrinsic Scale for Sample Excluding Agnostics/Atheists .............................................................. 134

25. Zero-Order Correlations among Defining Issues Test P score, Short Christian Orthodoxy Scale, Intrinsic Scale, and the Product of Short Christian Orthodoxy Scale and Intrinsic Scale for Sample Including Agnostics .................................................................. 138

26. Intermediate Regression Analysis Regressing Defining Issues Test P score on Short Christian Orthodoxy Scale, Intrinsic Scale, and the Interaction of Short Christian Orthodoxy Scale and Intrinsic Scale for Sample Including Agnostics/Atheists ...................................................................... 139
27. Simultaneous Regression Analysis Regressing Defining Issues Test P score on Short Christian Orthodoxy Scale and Intrinsic Scale for Sample Including Agnostics/Atheists .......................................................... 140

28. Simultaneous Regression Analysis Regressing Defining Issues Test P score on Short Christian Orthodoxy Scale, Intrinsic Scale, and their Interaction for Sample Including Agnostics/Atheists .......................................................... 140

29. Coordinates for Graphing Three Lines demonstrating P Score Regression on Intrinsic Scale Moderated by Short Christian Orthodoxy Scale, with Interaction, for Sample Including Agnostics/Atheists .......................................................... 143

30. Coordinates for Graphing Three Lines demonstrating P Score Regression on Short Christian Orthodoxy Scale Moderated by Intrinsic Scale, with Interaction Term, for Sample Including Agnostics/Atheists .......................................................... 148

31. Zero-Order Correlations among Defining Issues Test P score, Scriptural Literalism Scale, Intrinsic Scale, and the Product of Scriptural Literalism Scale and Intrinsic Scale for Sample Excluding Agnostics/Atheists .......................................................... 151

32. Intermediate Regression Analysis Regressing Defining Issues Test P score on Scriptural Literalism Scale, Intrinsic Scale, and the Interaction of Scriptural Literalism Scale and Intrinsic Scale for Sample Excluding Agnostics .......................................................... 153

33. Simultaneous Regression Analysis Regressing Defining Issues Test P score on Scriptural Literalism Scale and Intrinsic Scale for Sample Excluding Agnostics/Atheists .......................................................... 153
<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>34. Zero-Order Correlations among Defining Issues Test P score, Scriptural Literalism Scale, Intrinsic Scale, and the Product of Scriptural Literalism Scale and Intrinsic Scale for Sample Including Agnostics/Atheists</td>
<td>154</td>
</tr>
<tr>
<td>35. Intermediate Regression Analysis Regressing Defining Issues Test P score on Scriptural Literalism Scale, Intrinsic Scale, and the Interaction of Scriptural Literalism Scale and Intrinsic Scale for Sample Including Agnostics</td>
<td>156</td>
</tr>
<tr>
<td>36. Simultaneous Regression Analysis Regressing Defining Issues Test P score on Scriptural Literalism Scale and Intrinsic Scale for Sample Including Agnostics/Atheists</td>
<td>156</td>
</tr>
<tr>
<td>37. Zero-order Correlations among the Defining Issues Test P Score, Socioeconomic Status, Year in School, IQ, Short Christian Orthodoxy Scale, and Intrinsic Scale for Sample Excluding Agnostics/Atheists</td>
<td>162</td>
</tr>
<tr>
<td>38. Intermediate Regression Analysis Regressing Defining Issues Test P score on Three Sets each Adding the Following Variables: 1) Socioeconomic Status, Year, IQ, 2) Short Christian Orthodoxy Scale, Intrinsic Scale, and 3) All Interactions between variables in Set 1 and Set 2 for Sample Excluding Agnostics/Atheists</td>
<td>163</td>
</tr>
<tr>
<td>39. Simultaneous Regression Analysis Regressing Defining Issues Test P score on Socioeconomic Status, Year, and IQ for Sample Excluding Agnostics/Atheists</td>
<td>164</td>
</tr>
<tr>
<td>40. Simultaneous Regression Analysis Regressing Defining Issues Test P score on Socioeconomic Status, Year, IQ, Short Christian Orthodoxy Scale, and Intrinsic Scale for Sample Excluding Agnostics/Atheists</td>
<td>164</td>
</tr>
</tbody>
</table>
41. Zero-order Correlations among Defining Issues Test P Score, Socioeconomic Status, Year in School, IQ, Short Christian Orthodoxy Scale, and Intrinsic Scale for Sample Including Agnostics/Atheists ...........................................166

42. Intermediate Regression Analysis Regressing Defining Issues Test P score on Three Sets each Adding the Following Variables: 1) Socioeconomic Status, Year, IQ, 2) Short Christian Orthodoxy Scale, Intrinsic Scale, and 3) All Interactions between variables in Set 1 and Set 2 for Sample Including Agnostics/Atheist ...........................................167

43. Simultaneous Regression Analysis Regressing Defining Issues Test P score on Socioeconomic Status, Year, and IQ for Sample Including Agnostics/Atheists ...........................................168

44. Simultaneous Regression Analysis Regressing Defining Issues Test P score on Socioeconomic Status, Year, IQ, Short Christian Orthodoxy Scale and Intrinsic Scale for Sample Including Agnostics/Atheists ...........................................168

45. Zero-order Correlations among the Defining Issues Test P Score, Socioeconomic Status, Year in School, IQ, Scriptural Literalism Scale, and Intrinsic Scale for Sample Excluding Agnostics/Atheists ...........................................171

46. Intermediate Regression Analysis Regressing Defining Issues Test P score on Three Sets each Adding the Following Variables: 1) Socioeconomic Status, Year, IQ, 2) Scriptural Literalism Scale, Intrinsic Scale, and 3) All Interactions between variables in Set 1 and Set 2 for Sample Excluding Agnostics/Atheists ...........................................172

47. Simultaneous Regression Analysis Regressing the Defining Issues Test P score on Socioeconomic Status, Year, IQ, Scriptural Literalism Scale, and Intrinsic Scale for Sample Excluding Agnostics/Atheists ...........................................174
### TABLE PAGE

<table>
<thead>
<tr>
<th>TABLE NUMBER</th>
<th>TABLE DESCRIPTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>48.</td>
<td>Zero-order Correlations among Defining Issues Test P Score, Socioeconomic Status, Year in School, IQ, Scriptural Literalism Scale, and Intrinsic Scale for Sample Including Agnostics/Atheists</td>
<td>176</td>
</tr>
<tr>
<td>49.</td>
<td>Intermediate Regression Analysis Regressing Defining Issues Test P score on Three Sets each Adding the Following Variables: 1) Socioeconomic Status, Year, IQ, 2) Scriptural Literalism Scale, Intrinsic Scale, and 3) All Interactions between variables in Set 1 and Set 2 for Sample Including Agnostics/Atheists</td>
<td>177</td>
</tr>
<tr>
<td>50.</td>
<td>Simultaneous Regression Analysis Regressing Defining Issues Test P score on Socioeconomic Status, Year, IQ, Scriptural Literalism Scale, and Intrinsic Scale for Sample Including Agnostics/Atheists</td>
<td>178</td>
</tr>
<tr>
<td>51.</td>
<td>T-Test Procedure of Defining Issues Test Stage 3 and Stage 4 Means for those with and those without a Currently Meaningful Christian Personal Religious Experience</td>
<td>179</td>
</tr>
<tr>
<td>52.</td>
<td>Zero-order Correlations among Stage 3, Stage 4, GPA, and Currently Meaningful Christian Personal Religious Experience</td>
<td>180</td>
</tr>
<tr>
<td>53.</td>
<td>Hierarchical Regression Analysis Regressing Defining Issues Test Stage 3 on GPA and Currently Meaningful Christian Personal Religious Experience</td>
<td>180</td>
</tr>
<tr>
<td>54.</td>
<td>Hierarchical Regression Analysis Regressing Defining Issues Test Stage 4 on GPA and Currently Meaningful Christian Personal Religious Experience</td>
<td>181</td>
</tr>
<tr>
<td>55.</td>
<td>Frequency Table of Subject Denomination</td>
<td>216</td>
</tr>
<tr>
<td>TABLE</td>
<td>PAGE</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>56.</td>
<td>Coordinates for Graphing Three Lines demonstrating P Score Regression on Intrinsic Scale Moderated by Short Christian Orthodoxy Scale, without Interaction, for Sample Including Agnostics/Atheists</td>
<td>222</td>
</tr>
<tr>
<td>57.</td>
<td>Coordinates for Graphing Three Lines demonstrating P Score Regression on Short Christian Orthodoxy Scale Moderated by Intrinsic Scale, without Interaction Term, for Sample Including Agnostics/Atheists</td>
<td>223</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Visual Analogy to the Interaction Expected Between Intrinsic Scale and Short Christian Orthodoxy Scale/Scriptural Literalism Scale for Sample Excluding Agnostics/Atheists</td>
<td>91</td>
</tr>
<tr>
<td>2. P Score Regression on Intrinsic Scale Moderated by Short Christian Orthodoxy Scale, with Interaction Term, for Sample Including Agnostics/Atheists</td>
<td>144</td>
</tr>
<tr>
<td>3. P Score Regression on Intrinsic Scale Moderated by Short Christian Orthodoxy Scale, without Interaction Term, for Sample Including Agnostics/Atheists</td>
<td>145</td>
</tr>
<tr>
<td>4. P Score Regression on Short Christian Orthodoxy Scale Moderated by Intrinsic Scale, with Interaction Term, for Sample Including Agnostics/Atheists</td>
<td>149</td>
</tr>
<tr>
<td>5. P Score Regression on Short Christian Orthodoxy Scale Moderated by Intrinsic Scale, without Interaction Term, for Sample Including Agnostics/Atheists</td>
<td>150</td>
</tr>
</tbody>
</table>
Chapter I
Problem Statement

Most humans live in community. In community one's behaviors may affect others for better or worse. History records both heinous, immoral actions such as the mass extermination of people in World War II Germany, and admirable, moral actions such as those of Ghandi and Martin Luther King, Jr. Morality may be defined as "a particular type of social value, that having to do with how humans cooperate and coordinate their activities in the service of furthering human welfare, and how they adjudicate conflicts among individual interests" (Rest, 1986, p. 3).

What factors go into "moral" behaviors? The answer to this question is not simple. Rest (1986) has theorized four components relevant to moral actions. These components are interactive, do not necessarily follow a sequence in time, and each involves aspects of cognition, affect and behavior. The first component involves *perceiving and comprehending* a situation as involving a moral issue. The second component involves *making a decision* about what is right, just, or fair, *ideally*. The third component involves the *actual priority* one gives to moral values against other values when deciding how to act. And the fourth component involves the *ability to follow*
through on one’s moral decision. The realm of cognitive moral development/moral judgment (I shall use these two phrases interchangeably in this paper) pertains to the second component of making a decision of what is the just thing to do in a situation. Rest (1986) has concluded from his review of the literature that moral judgment is pervasively and consistently associated with moral behavior to a moderate degree. Although moral judgment is only one component, it plays a significant role in the outcome of moral behavior. According to Clouse (1991), Kohlberg and others have also noted the relationship between moral judgment and moral behavior.

The main focus of this study was the relationship between conservative Christian beliefs and moral judgment. Although Kohlberg repeatedly has suggested that moral judgment is independent of religion (Getz, 1984), researchers have reported results showing a fairly consistent pattern: conservatively religious subjects of high school age or younger obtain higher scores on measures of moral judgment and after high school show lower scores compared to more liberal religious persons and those who profess agnosticism, atheism, or no faith. These studies are generally cross sectional, so I do not mean to imply that the above reported results have followed the same individuals from junior high to post-high school. While inferences have been made concerning the relationship between conservative Christian beliefs and moral judgment, it is quite possible that the relationship is moderated by other variables and therefore not as open to simple inferences. For example, Rest (1986) suggested that in
studies where high school and younger conservative Christian subjects exhibited higher levels of moral judgment than more liberal, atheistic and agnostic subjects the conservative Christian subjects may have had a higher socioeconomic status (SES) than the liberal, atheistic, or agnostic subjects. The results of my study, which focused on college students, could provide information useful to interpreting the relationship between conservative Christian beliefs and moral judgment.

The relationship between religiosity and moral judgment may be moderated not only by socioeconomic status as Rest (1986) suggested, but also by intelligence. Both intelligence and socioeconomic status have been correlated with some measure of moral judgment. Intelligence has been reported as moderately positively correlated with all the major measures of moral judgment including the Sociomoral Reflection Measure-Short Form (Basinger, Gibbs, & Fuller, 1992), the Moral Judgment Interview (Colby, Kohlberg, Gibbs, & Lieberman, 1983), and the Defining Issues Test (Clouse, 1991; Rest, 1979a). Socioeconomic status has been reported as positively correlated with some of the major measures of cognitive moral development: the Sociomoral Reflection Measure-Short Form (Basinger, Gibbs, & Fuller, 1992) and the Moral Judgment Interview (Colby, Kohlberg, Gibbs, & Lieberman, 1983), with unclear and inconsistent findings on the Defining Issues Test (Rest, 1979a). Researchers on religion have noted a pattern in which more conservative Christian denominational affiliation has correlated with
lower socioeconomic status and lower intelligence (Argyle & Beit-Hallahmi, 1975). This relationship is not in conflict with a Scriptural description of Christians in the first century, and so it should not be surprising or offensive if it occurred in modern Christians. As the Apostle Paul wrote to the believers of the cosmopolitan city of Corinth:

Think of what you were when you were called. Not many of you were wise by human standards; not many were influential; not many were of noble birth" (1 Corinthians 1:26-29, New International Version).

One assumption implied by the above statements is that conservative affiliation would be correlated with conservative beliefs and therefore those with conservative beliefs may also have lower IQ and socioeconomic status than those with less conservative beliefs. Although conservative beliefs play a major role in determining whether a denomination is labelled conservative, some members of liberal denominations may hold conservative religious beliefs (Glock and Stark, 1965). Thus the correlation between religious beliefs and religious affiliation may not be as high as one might at first expect. This may serve to diminish the expected correlation of religious beliefs with IQ and socioeconomic status.

It is possible that conservative Christians may have lower intellectual ability and lower socioeconomic status, and that both intelligence and socioeconomic status may serve as moderator variables in the negative relationship between conservative Christianity and moral judgment that has been reported for post-high
school populations. While some studies have controlled for socioeconomic status, few, if any, studies have controlled for intelligence.

In summary, my goal was to study the relationships between conservative-liberal Christian beliefs and moral judgment in a manner similar to previous research with the added controls for intellectual ability and socioeconomic status. Perhaps the effects of conservative-liberal Christianity on moral judgment will be less (or greater) than previous research has shown. The results of this study may shed additional light on the mechanisms that may be mediating the reported relationship between Christian beliefs and moral judgment. Such information may prove useful to those who work with Christian populations.

Clouse (1985a) has noted that the cognitive moral developmental model is useful in crafting theology and educational programs so that those at each developmental level may readily comprehend material taught to them or be challenged to grow into higher developmental levels. Clouse adds that such information may also be useful in understanding and solving denominational schisms caused by persons of lower cognitive moral developmental levels misreading those at higher levels as giving up the basics of the faith. Finally, in addition to working with educational classes or groups with potential schisms, the cognitive moral developmental model may be useful in more individual encounters such as counseling and pastoral care. The counseling process may improve as the counselor correctly
assesses the client's cognitive moral developmental level and then speaks at the same level (to help the client understand the counselor's ideas) or perhaps one level higher (to promote the client's cognitive moral development; Hayes, 1991).
Chapter II
Literature Review

In her extensive review of the literature, Getz (1984) identified five aspects of religion that have been studied in relation to moral judgment: religious beliefs, intrinsic-extrinsic [I-E] motivation, affiliation, religious education, and religious knowledge. Religious beliefs showed the clearest relationship with moral judgment (at least in subjects who have graduated from high school), while intrinsic motivation served as an intensifier of one’s beliefs. Given these findings, religious beliefs and intrinsic religiousness were selected as the religious variables to concentrate on in this study, and the majority of the literature review will focus on studies using one or both of these two religious variables. After a brief overview I will describe the theory and major instruments used in the majority of the relevant studies, then review the studies themselves, including an overview of subjects from Christian higher educational settings. Finally, I will summarize the research suggesting the possibility of

---

1Intrinsic [I] religious motivation has been conceptualized as one of several potential theoretical components of intrinsic religious orientation. While religious orientation is not easy to define simply and precisely (Kirkpatrick & Hood, 1990, p. 443), Donahue (1985) has noted that "research indicates that I is a good, unidimensional, nondogmatic indicator of religious commitment, while E seems to measure the sort of religion that gives religion a bad name: prejudiced, dogmatic, fearful" (p. 422).
intelligence and socioeconomic status as moderators between religious beliefs and cognitive moral development.

Based on Getz (1984) there seems to be a trend that prior to high school graduation, subjects with greater religious knowledge and education score higher on measures of cognitive moral development, while after high school, conservatively religious subjects score lower than liberals, agnostics, or atheists when the measure of religion is beliefs. However, a number of disclaimers and qualifiers need to be made. Getz (1984) reported only one study of subjects prior to high school graduation that measured religious beliefs—Harris (1981). Harris (1981) reported that in a sample of 438 11th grade students religious beliefs were not significantly related to moral judgement. Because there is only one study of this type, one should be cautious in giving it too much weight. One study (Brown & Annis, 1978) has been interpreted by various authors in ways that are diametrically opposed. Another study (Clouse, 1991) reported a small, but statistically significant, positive correlation between conservative religious beliefs and moral judgment as measured by the P score (P-Principled) of the Defining Issues Test. The P score measures the highest two stages of moral judgment. These two stages (Stage 5 and Stage 6) added together form the principled level, thus P score. Finally, Holley (1989, 1991) failed to find a statistically significant relationship between P score and conservative Christian beliefs. Interestingly, in these last two studies, Clouse (1991) and Holley (1989, 1991), the samples were unusual. Holley (1989, 1991)
reported a mean P score for his sample that was significantly below the mean P score for college students. Similarly, Clouse (1991) reported a mean P score even lower than Holley's. This is further discussed in the Results and Discussion section below.

The evidence of a positive relationship between moral judgment as measured by the P score and aspects of religiosity prior to high school graduation occurs in the religious education variable and the religious knowledge variable (Getz, 1984). In most studies the higher levels of cognitive moral development among children attending religious schools as compared to those in public schools has been attributed to the environment created in those schools. However, Rest (1986) has suggested that the higher moral judgement scores obtained by students at religious schools were partly due to higher socioeconomic status of families who were able to send their children to private schools. It may also be possible that subjects with higher religious knowledge had a higher IQ in general which accounted for the higher moral judgement scores. Without controlling for socioeconomic status and IQ it is impossible to know the answer to these questions. As we shall see, both socioeconomic status and IQ should be controlled for when examining the relationship between religious variables and moral judgement.

Apart from the issue of IQ and socioeconomic status potentially moderating the relationship, why is it that samples composed of 12th graders and below tend to differ from post high school samples in the relationship between moral judgment and religiosity? Generally,
conservative religiosity in post high school subjects has been
negatively correlated with P scores on the Defining Issues Test (see
table of literature review in Getz, 1984), but as noted above this has
not been the case with 12th graders and below.

One possible explanation for this difference may be that after
high school some subjects are more likely than others to expose
themselves to teachings and experiences that facilitate their
preference for principled level items on the Defining Issues Test (e.g.,
through work and the college environment, Rest, 1988). Relevant to
this exposure to new teachings and experiences, an important but
often overlooked distinction exists between fundamentalist and non­
fundamentalist branches among the evangelical, conservative wing of
Christians. While I have not seen this distinction made in previous
studies, it should be kept in mind because the relationship of moral
judgment with conservative Christian religiosity may in fact apply
more to the fundamentalist components of conservative Christianity.
According to Ammerman (1982) several factors distinguishing
Fundamentalists from other evangelical Christians include the
following fundamentalist characteristics: a literal interpretation of the
Bible, a separatist attitude toward the world and other non­
fundamentalist Christians, and a belief in dispensational
premillennialism. Therefore, more fundamentalist Christians may not
be exposed to the same (cognitive moral development growth
inducing) influences, partially because they isolate themselves from
experiences that might challenge their world views (Ammerman, 1982).

Additionally, fundamentalist Christians may reject using their own intuitions of justice in favor of using explicitly Christian teachings found in the Bible (Lawrence, 1979; Rest, 1986), thus allowing those not holding to fundamentalist Christian beliefs (or perhaps this applies to mainline orthodox Christian beliefs as well) to outperform them on the P score. As we shall see below, it can be argued that some Defining Issues Test principled level items, especially at Stage 5, contain values counter to Biblical teachings. Thus, the Defining Issues Test P score, which is made of of Stage 5 and Stage 6, may be somewhat biased against conservative Christians and perhaps even more so against fundamentalist Christians (who are also conservative). These are only a few possibilities. There are many others (Hanson, 1991).

Cognitive Moral Development: Theory and Measurement

The most commonly used instrument for assessing cognitive moral development or moral judgment, according Gielen and Lei (1991), is the Defining Issues Test. The Defining Issues Test (DIT, Rest, 1990) is a multiple choice test which comes in a six-story form and a commonly used three-story form. The three-story form uses
the three stories that correlate most highly with the six-story form. Rest (1979a) based his Defining Issues Test on the six-stage model of cognitive moral development elaborated by Kohlberg and associates. Moral development, as defined by Rest, involves an increasingly sophisticated "understanding and application of principles of judging fairness" (McCrae, 1985, p. 439). According to Rest (1979a), "concepts of fairness and justice are essentially notions about the balancing of individual interests and the benefits of cooperation" (p. 20). Rest (1979a) stated that the Defining Issues Test is not appropriate for testing the hard stage model proposed by Kohlberg because Rest's stage model differs from Kohlberg's. Rather than assign an individual to one stage or a transition between stages as Kohlberg does, Rest theorized that individuals can operate at several stages at once (see Table 1). He defined "a stage [as] a logical organization of thinking assumed to underlie the manifestation of certain features of moral judgments" (p. 440). According to McCrae (1985), "defining 'stage' as an attribute of the reasoning rather than of the person is a major conceptual advance that preserves the qualitative framework so appealing to developmentalists while explaining the ubiquitous finding that individuals tend to straddle several stages" (p. 440). The six stages as modified by Rest (1979a) are presented below in Table 2.

---

2One set of three stories is not equivalent to the other set of three stories and so Rest (1986) advised against treating one set of three stories as an alternative form to the other set of three stories.
Table 1

*Relationship of the Three Levels to the Six Stages of Moral Judgement*

<table>
<thead>
<tr>
<th>Preconventional Level</th>
<th>Conventional Level</th>
<th>Postconventional or &quot;Principled&quot; Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stages 1 2</td>
<td>Stages 3 4</td>
<td>Stages 5 6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stage</th>
<th>Coordination of expectations about actions (how rules are known and shared)</th>
<th>Schemes of balancing interests (how equilibrium is achieved)</th>
<th>Central concept for determining moral rights and responsibilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>The caretaker makes known certain demands on the child's behavior</td>
<td>The child does not share in making rules, but understands that obedience will bring freedom from punishment</td>
<td>The morality of obedience: “Do what you're told.”</td>
</tr>
<tr>
<td>Stage 2</td>
<td>Although each person is understood to have his own interests, an exchange of favors might be mutually decided.</td>
<td>If each party sees something to gain in an exchange, then both want to reciprocate.</td>
<td>The morality of instrumental egoism and simple exchange: “Let's make a deal.”</td>
</tr>
<tr>
<td>Stage 3</td>
<td>Through reciprocal role taking, individuals attain a mutual understanding about each other and the on-going pattern of their interactions.</td>
<td>Friendship relationships establish a stabilized and enduring scheme of cooperation. Each party anticipates the feelings, needs, and wants of the other and acts in the other's welfare.</td>
<td>The morality of interpersonal concordance: &quot;Be considerate, nice, and kind, and you'll get along with people.&quot;</td>
</tr>
<tr>
<td>Stage 4</td>
<td>All members of society know what is expected of them through public institutionalized law.</td>
<td>Unless a society-wide system of cooperation is established and stabilized, no individual can really make plans. Each person should follow the law and do his particular job, anticipating that other people will also fulfill their responsibilities.</td>
<td>The morality of law and duty to the social order: &quot;Everyone in society is obligated and protected by the law.&quot;</td>
</tr>
</tbody>
</table>
Table 2 (continued)

<table>
<thead>
<tr>
<th>Stage 5</th>
<th>Formal procedures are institutionalized for making laws, which one anticipates rational people would accept.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Law-making procedures are devised so that they reflect the general will of the people, at the same time insuring certain basic rights to all. With each person having a say in the decision process, each will see that his interests are maximized while at the same time having a basis for making claims on other people.</td>
</tr>
<tr>
<td></td>
<td>The morality of societal consensus: &quot;You are obligated by whatever arrangements are agreed to by due process procedures.&quot;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stage 6</th>
<th>The logical requirements of non-arbitrary cooperation among rational, equal, and impartial people are taken as ideal criteria for social organization which one anticipates rational people would accept.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A scheme of cooperation that negates or neutralizes all arbitrary distribution of rights and responsibilities is the most equilibrated, for such system is maximizing the simultaneous benefit to each member so that any deviation from these rules would advantage some members at the expense of others.</td>
</tr>
<tr>
<td></td>
<td>The morality of non-arbitrary social cooperation: &quot;How rational and impartial people would organize cooperation is moral.&quot;</td>
</tr>
</tbody>
</table>

In the Defining Issues Test's stories (often referred to as dilemmas) the protagonists find themselves in situations where competing demands are placed on them and they must decide what to do. After reading the stories, subjects are asked to check one of three options: 1) The protagonist should do X, 2) The protagonist should not do X, or 3) The subject can't decide what the protagonist should do. For example, in story 1 "Heinz and the Drug" the protagonist's wife is dying of a special kind of cancer. A druggist has the only cure but is selling it for an exorbitant amount which Heinz can't afford no matter how he tries to raise the money. The druggist won't negotiate an alternate price or payment plan, so Heinz steals the drug. Subjects are asked "Should Heinz steal the drug?" then they must check either that Heinz "should steal it", "should not steal it", or that they "can't decide".

Following this decision, subjects are presented with twelve different phrases (statements and questions), which I will refer to as "queries" to be consistent with Holley's (1989, 1991) terminology. Some of the queries contain meaningless or irrelevant issues designed to detect if subjects are attracted to high sounding phrases or choosing irrelevant items rather than queries that are meaningfully related to the issues presented by the story. For example, in "Heinz and the Drug", query #4 "Whether Heinz is a professional wrestler, or has considerable influence with professional wrestlers" and query #7 "Whether the essence of living is more encompassing than the termination of dying, socially and individually" are categorized as "M" (i.e., meaningless, lofty and pretentious) and do not represent any of
the six stages. Another type of query, categorized as "A" is designed to attract those with purely anti-establishment leanings. The A score does not seem to be used. Each of the remaining queries represents one of the stages. For example, query #3 -- "Is Heinz willing to risk getting shot as a burglar or going to jail for the chance that stealing the drug might help" -- represents a stage 2 query, while query #8 -- "What values are going to be the basis for governing how people act towards each other" -- represents a stage 6 query. Subjects are asked to rate the importance of each of the twelve queries by checking one of the following: "Great", "Much", "Some", "Little", or "No". Irrelevant or meaningless queries are to be rated as having no importance. Next, subjects are asked to rank the top four of the twelve queries from first to fourth most important (see Appendix A).

A number of scores are available from the Defining Issues Test. Scoring is objective. The four highest ranked queries are given weights of 4, 3, 2, and 1, respectively, for each story. For each of the stages (2, 3, 4, 5A, 5B and 6) the weights across all stories are summed for each subject. (There are no queries for stage 1). Using the summed weights, percentages for each stage are calculated. Subjects' preferences for principled thinking are stipulated by the P score which is the sum of percentages from stages 5A, 5B and 6. The P score is the most frequently used indicator of moral judgement maturity. The D score is another overall measure of moral judgement maturity. It results from a complex formula combining and weighing the subject's ratings of all 72 arguments provided by the Defining
Issues Test and requires computer scoring. There are two checks on the reliability of subjects' scores: the M scores (meaningless queries) and the Consistency Check. If a subject was attracted by too many M queries (14% or more) or if there is sufficient inconsistency between the ratings and the rankings given to queries (such that one may doubt the seriousness with which the test was taken), then Rest (1990) suggested the subject's test be considered as possibly invalid and the data analyzed both with and without these subjects' scores. This suggestion was based on Rest's experience that results were more desirable when subjects who failed reliability checks were discarded. The standard in all of the studies I have read that relate religiosity to the Defining Issues Test is to remove subjects who fail on of the reliability checks and then analyze the data. I have done the same. Rest (1990) reported that it is typical to lose between 5 and 15% of a sample of volunteer subjects through reliability checks.

A number of reviews in the Mental Measurements Yearbook series have concluded that the Defining Issues Test has at least adequate validity and reliability. "It has good psychometric properties" (Sutton, 1992, p. 269). "The Defining Issues Test is a convenient and reliable objective measure of maturity in judging moral issues. It is backed by an impressive series of construct validity studies ..." (McCrae, 1985, p. 440). "The results of [numerous DIT] studies, while of course leaving room for improvement, suggest that the Defining Issues Test is at least as valid as other measures of moral development" (Moreland, 1985, p. 442).
More specifically, McCrae (1985) concluded that "many studies have been conducted showing that moral judgment as measured by the Defining Issues Test increases with age and education; that it is related to, but distinct from, general cognitive capacity; that it influences moral behavior — in short, that it conforms to the predictions of the theoretical model it is based on" (p. 440).

Martin, Shafto, and Vandeise (1977) have criticized the Defining Issues Test by noting that the number of queries available is not the same at each stage (see Table 3). They suggested it would make more sense, methodologically, if there were an equal number of queries available at each stage. Below is a table that I constructed based on Rest's (1990) Defining Issues Test Manual (Section 3.3). It indicates the number of Defining Issues Test queries representing each stage plus other special scores on both the six-story and the three-story versions.
### Table 3
**Total Number of DIT Queries Available at Each Stage**

<table>
<thead>
<tr>
<th>Stage or Reliability Check</th>
<th>6-story form number of queries</th>
<th>3-story form number of queries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>Stage 2</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Stage 3</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td>Stage 4</td>
<td>19</td>
<td>10</td>
</tr>
<tr>
<td>Stage 5A</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Stage 5B</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Stage 6</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>M</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>A</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>72</strong></td>
<td><strong>36</strong></td>
</tr>
</tbody>
</table>

A number of consequences may result from the unequal number of queries available at each stage. First, if subjects were to select queries at random one would expect their Stage 4 scores to be the highest. It would have been better if each stage had an equal chance of being chosen. Second, on some stories there are no queries representing a particular stage and therefore moral judgment representing that stage for that particular story can not be measured. For example, in the Prisoner story there are no Stage 2 options, and in
the Newspaper story there are no Stage 6 options. Finally, as explained to me by Clouse (personal communication, March, 1993), the reliability of Stage 2 is especially suspect and this makes testing hypotheses on Stage 2 questionable. She suspected politically and religiously liberal subjects would follow a bimodal distribution with a significant portion of them scoring at Stage 2, below the conservatives who tend to prefer Stages 3 and 4. With only two Stage 2 queries in the entire 3-story form of the Defining Issues Test, however, Clouse (1985) was concerned that any comparisons involving Stage 2 might be spurious and so Stage 2 comparisons were not included in her analyses.

A counter argument to these criticisms regarding an unequal number of queries at different stages and some stages totally unrepresented for particular stories is Rest's (1979a) statement that his multiple choice queries are based on the types of responses elicited during extensive interviews with subjects for his 1969 doctoral dissertation. He claimed that: 1) It would not be possible to create equal numbers of queries at each stage that are psychologically meaningful; and 2) If he were to leave out some of the more popular queries (such as those at Stage 4 for example) subjects who think in those terms for that particular content would be frustrated. This psychological pull of the content of certain stories for a greater or lesser number of queries across the six stages is why different stories have different numbers of same-stage queries.
In the 1990 manual Rest cited articles claiming that test-retest correlations for the major indices of P and D scores are "generally in the high .70's or .80's and Cronbach's alpha index of internal consistency is generally in the high .70's" (p. 51). For a sample of 33 Australian college students, however, the P and D score test-retest correlations were lower, .71 and .67, respectively, for the six-story form and .67 and .63, respectively for the three-story form. Rest noted that the test-retest reliabilities of the stage scores are even lower, in the .50's and .60's. He, therefore, recommended exercising "much caution" in using the stage scores. "I recommend using the stage scores only when the 6-story form has been used, and only when the information is presented in terms of group means or when the standard error of measurement has been taken into account" (Rest, 1990, 5.2).

Davidson (1979) elaborated this point:

The reliability data . . . indicate that using the shorter three-story DIT should have little effect in studies where the group means are the focus. In studies of group means, one need only have precise estimates of means; that is, estimates that have low standard errors. Regardless of a measure's reliability, mean estimates can be made as precise as desired by a sufficiently large sample size. In correlational studies, the drop in reliability associated with the shorter version cannot be overcome by an increase in sample size. According to classical reliability theory, the net effect should be a drop in the correlation between the DIT and any outside variable. The drop should be proportional to the square root of the three-story DIT reliability divided by the square root of the six-story reliability. For example, imagine that the six-story version correlates about .40 with some variable, and it has a reliability of .80 in the
population of interest, whereas the shorter version has a reliability of .70. The shorter version would be expected to have a correlation of $(\sqrt{.70} / \sqrt{.80}) \cdot (.40) = .37$. Using the shorter version would always reduce the observed correlation somewhat and may even reduce it to nonsignificance, but the reduction should usually be small [italics added]. (p. 245)

As noted above, Getz (1984) reviewed the religion/moral judgment empirical literature. Of the 24 studies reviewed by Getz, only three used Kohlberg dilemmas, the others used the Defining Issues Test. The most likely reason for most studies using the Defining Issues Test rather than Kohlberg's system is that Kohlberg's system is difficult to learn, expensive, and time consuming to administer and score. It requires extensive training of several weeks at a cost of hundreds of dollars. It also requires transcription of session tapes. Given these factors it seemed a rather impractical measure for the limited resources of subject time and experimenter money available for this study. Gielen and Lei (1991) wrote favorably of a third measure of cognitive moral development, the Sociomoral Reflection Measure (SRM). However, because the SRM was based on a revision of the theory of Kohlberg and does not acknowledge Stages 5 and 6, Gielen and Lei concluded that it "may not be the test of choice when the nature of post-conventional moral thinking is being explored" (p. 76). It is precisely at the post-conventional level where the differences are reported to occur between subjects with conservative versus liberal Christian beliefs. Therefore, it seems that the Sociomoral Reflection Measure would not be able to help at the
critical level of moral development. This leaves the Defining Issues Test. Although it has its problems, it is the most widely used measure and allowed me to measure the dependent variable in a comparable manner to the vast majority of previous studies.

Religious Orientation, Intrinsic-Extrinsic Scales

One of the more frequently used measures of religion is Allport and Ross's (1967) Religious Orientation Scale. Two of the studies reviewed below (Brown & Annis, 1978; Ernsberger & Manaster, 1981) used Allport and Ross's Religious Orientation Scale in slightly different ways. Therefore a brief overview of the uses of this instrument is in order. According to Allport and Ross there are four possible orientations to religion: extrinsic (E), intrinsic (I), indiscriminately proreligious (IPR) and indiscriminately antireligious (IAR). These constructs are defined as follows:

A person with an extrinsic religious orientation is using his religious views to provide security, comfort, status, or social support for himself—religion is not a value in its own right, it serves other needs, and it is a purely utilitarian formation. . . . Contrariwise, the intrinsic religious orientation is not an instrumental device. It is not a mere mode of conformity, nor a crutch, nor a tranquilizer, nor a bid for status. All needs are subordinated to an overarching religious commitment. In internalizing the total creed of his religion the individual necessarily internalizes its values of humility, compassion, and love of neighbor. (Allport & Ross, 1967, p. 441)
The indiscriminately proreligious exhibit "undifferentiated thinking" or excessive "category width." The indiscriminately proreligious respond positively to both extrinsic and intrinsic items, some of which don't really make sense together, according to Allport and Ross (1967). For example, the indiscriminately proreligious may respond positively to the following items: "My religious beliefs are what really lie behind my whole life" (intrinsic) and "Although I believe in my religion, I feel there are many more important things in my life" (extrinsic). "There seems to be one wide category—religion is OK. From the way in which the scale is constructed this undifferentiated endorsement can be the product of an agreement response set. . . . But if so, we are still dealing with undifferentiated cognitive disposition" (Allport & Ross, 1967, p. 441-442). Allport and Ross also noted that the indiscriminately proreligious had a lower level of formal education than the extrinsics and intrinsics, and that a lower level of formal education is relevant to forming and holding overwide categories.

Watson, Morris and Hood (1990) noted evidence suggesting that the indiscriminately proreligious (informally labeled "muddleheads" by Allport and Ross) are not necessarily muddleheaded. Kirkpatrick (1989) reported a factor analysis in which he found two dimensions for the Extrinsic Scale: a **Personal** extrinsic factor (exemplified by the item "the purpose of prayer is to secure a happy and peaceful life") and a **Social** extrinsic factor (exemplified by the item "the church is most important as a place to formulate good social relationships"). Of
the 11 extrinsic items, three loaded on the Personal factor, three on the Social factor and the other five were residuals. Among the five Extrinsic Scale residuals were items such as "although I am a religious person, I refuse to let religious considerations influence my everyday affairs" (essentially an anti-intrinsic item). Watson Morris, and Hood (1990) reported that among more orthodox subjects there is a stronger negative correlation between Intrinsic and Extrinsic Scales probably due to the fact that the Extrinsic Scale contains residual items which are essentially counter to an intrinsic commitment. However, one might score high on the Social and Personal extrinsic factors as well as the Intrinsic Scale without necessarily contradicting oneself, because these two extrinsic factors are orthogonally related to the Intrinsic Scale. With these complexities, Kirkpatrick (1989) recommended that researchers employing the I-E scale score the Social and Personal scales separately in addition to scoring E in the traditional way" (p. 28).

The indiscriminately antireligious or nonreligious category includes those who disagree with items on both the Extrinsic and Intrinsic Scales. This type of subject was not represented among Allport and Ross's (1967) churchgoing subjects. However, it was reported that among markedly liberal groups this type of religious orientation does exist "even among members of 'religious' organizations" (Allport & Ross, 1967, p. 438).

This brings up an interesting question: is it appropriate to interpret the results of Intrinsic-Extrinsic Scales when subjects are
Agnostics/Atheists? Because the Intrinsic-Extrinsic Scales were designed to measure the religious orientation of religious persons, many items explicitly assume such things as church membership/attendance, prayer, and belief (e.g., "One reason for my being a church member is . . .," "The prayers I say when I am alone . . .," "Although I believe in my religion . . ."). The same explicit assumptions of religiosity hold true for the Age Universal I/E-Revised Scales (Gorsuch & McPherson, 1989) which were used in this study. Therefore, because these items are worded to imply religiosity, it did not seem appropriate to include Agnostics/Atheists in the sample on which analyses of religious orientation were made. As Kirkpatrick (1989) put it, "It is not at all clear what a nonreligious respondent's agreement or disagreement with such statements would mean" (p. 6). Most (but not all) researchers in the study of religious orientation hold similar views (Kirkpatrick & Hood, 1990, p. 456), and Kirkpatrick (1989) recommended removing from secular university samples any respondents who claim to be nonreligious "provided that the range of the Intrinsic scale is not drastically restricted by doing so" (p. 7). Yet, I sense some theoretical ambivalence in Kirkpatrick's willingness to allow nonreligious respondents to be included in the sample given a "drastically restricted" range.

Kirkpatrick and Hood (1990) note an alternative theoretical view of intrinsic-extrinsic religious orientation that includes nonreligious respondents. However, they imply that this can not be
studied until the wording is revised to make the Intrinsic-Extrinsic Scales more meaningful and applicable to nonreligious respondents.

An alternative conceptualization of I-E is one in which each dimension is defined as a continuum from "not at all" to "very," where "not at all" on each dimension represents nonreligious individuals. From this perspective it would be entirely appropriate to use heterogeneous samples including both religious and nonreligious individuals; in fact such sampling would typically be desirable in order to examine the entire range of each dimension. It is only in light of this conceptualization that revised scales would be required for use with nonreligious respondents, as recommended by Donahue (1985b) [italics added]. If this stance is adopted, it makes little sense to restrict samples to religious respondents; this serves only to attenuate the magnitude of correlations between I and other variables. (p. 455)

Returning to the four-fold typology, most researchers have used the two subscales (I and E) independently without taking into account the four-fold typology. According to Kirkpatrick and Hood (1990) a good case has not been made to support using the four-fold typology:

In light of the statistical drawbacks involved in arbitrarily dichotomizing continuous variables, the long history of measuring I and E as continuous measures with multi-item scales, and the prevailing pro-dimensional (anti-type) sentiment in contemporary research psychology, we believe the burden is on those who wish to utilize types to demonstrate their greater utility over continuous scales. (p. 452)

One final point, as we shall see demonstrated below in Ernsberger and Manaster (1981), intrinsic religious orientation has
been found to act as a measure of commitment to one's beliefs (Getz, 1984; Donahue, 1985). Contrary to Allport and Ross's (1967) theory, this is true even when one's beliefs are racially prejudiced (Griffin, Gorsuch, & Davis, 1987). Therefore, I included the Intrinsic Scale of the Age Universal I/E-Revised Scales (Gorsuch & McPherson, 1989) as a measure of commitment to one's beliefs.

*Religious Beliefs (ideology)*

As mentioned above, after reviewing the cognitive moral development-religiosity literature in 1984 (Getz, 1984), Getz (1985) concluded that among several religious variables (religious belief, intrinsic-extrinsic religious orientation, religious affiliation, religious education, and religious knowledge) religious beliefs showed the clearest relationship with moral judgment, while intrinsic religious orientation (a measure of commitment) seemed to serve as an intensifier of religious beliefs. Religious beliefs typically have been measured along a conservative-liberal continuum and according to Gorsuch (personal communication, March, 1993) there are thousands of such measures. I am aware of 11 studies in which religious beliefs and moral judgment were examined and results reported (including Harris, 1981, summarized above).

Sanderson (1974) administered Kohlberg's then current revision of the Moral Judgement Interview along with a set of religious belief questions covering: God's existence, Jesus's deity, literal miracles in
the Bible, the afterlife, the Devil's existence, whether a member of another faith would be saved, the necessity of belief in Jesus for salvation, and whether those ignorant of Jesus would be saved. Subjects were 481 freshmen to senior Sociology students from the University of Nebraska, most of whom were freshmen and sophomores. Results indicated that the preconventional and conventional subjects were the most orthodox in belief followed by the postconventional (principled) subjects. The least orthodox group of subjects were those at stage 4 1/2 (the "ethical relativists") who were in transition from conventional to postconventional status. Brown and Annis (1978) used Hogge and Friedman's (1967) Scriptural Literalism Scale, which measures belief in the Bible as the literal word of God. The Scriptural Literalism Scale comes in three forms, two 12 item scales (Scriptural Literalism Scale-Form A and Form B) and a 16-item form which combines the best items from the 12 item forms (Scriptural Literalism Scale-Form C). Sample items include: "The scriptures contain religious truths," "Quotations appearing in the scriptures are accurate," "The scriptures are a collection of myths," and "The passage of time is accurately presented in the scriptures" (Hogge & Friedman, 1967, p. 276). Some items are reverse scored as appropriate. Items are scored on a 5 point Likert scale with higher scores indicating increasingly literal belief in the Bible. As evidence of construct validity for the Scriptural Literalism

---

3Interestingly, in an earlier version of the Moral Judgment Inventory Kohlberg had categorized this way of thinking (i.e., stage 4 1/2) as stage 2 relativism (Gielen & Lei, 1991).
Hogge and Friedman reported that a more conservative denomination (Baptist) scored higher than a very liberal denomination (Unitarian), with a moderate denomination (Methodist) scoring in between.

Along with an unspecified form of Hogge and Friedman's (1967) Scriptural Literalism Scale, Brown and Annis (1978) gave Allport and Ross's (1967) Religious Orientation Scale to 80 college students (40 males and 40 females), 20 from each class rank of freshman to senior, plus a measure of cognitive moral development—the Defining Issues Test (Rest, 1972). Since Brown and Annis did not discuss the four fold typology, it appears that they used the Religious Orientation Scale as it was first theoretically intended, a bipolar, unidimensional scale with intrinsic scores at the lower end and extrinsic scores at the higher end of the scale (see Allport & Ross, 1967, p. 435).

In their one page article, Brown and Annis (1978) did not report reliability checks on the Defining Issues Test so either 80 subjects were tested and 80 subjects retained, perhaps regardless of reliability problems, or 80 subjects were retained from a larger sample after reliability checks. Some of the results are quite

---

*I was unable to track down Dawn Brown, however, Lawrence Annis stated that he could "just about guarantee" they used the longest version possible (personal communication, April 26, 1993). This is still ambiguous; however, given that Brown and Annis (1978) cited the split half reliability of the SLS as .92 and Hogge and Friedman (1967) reported that the two 12 item forms correlated .92, it appears quite plausible that Brown and Annis gave both 12 item short forms (Scriptural Literalism Scale-Form A and Form B) as a single measure which they labelled the SLS.
reasonable, some very puzzling, and others have been interpreted by
various authors in diametrically opposed ways.

First the reasonable results. Pearson correlation between
intrinsic-extrinsic religious orientation and the Defining Issues Test (I
assume the P score\(^5\), although this was not explicitly stated) \((r = -.01, df = 79)\) was not significant (I assume at the commonly accepted .05
level, which was not stated). The correlation between the Scriptural
Literalism Scale and intrinsic-extrinsic religious orientation \((r = .64, df = 79, p < .01)\) was significant.

Now the puzzling results. "Mean scores were 105.8 for moral
However, neither the P score nor any of the Stage scores can possibly
go up that high. Typical P scores for a college sample range from 37 to
46; see Rest (1979b, p. 7.2). Finally, Brown and Annis reported that
the "correlation between subject's morality and literal scriptural belief
\((r = .44, df = 79, p < .01)\) was significant and accounted for 19% of the
total variance" (p. 1230). Based on the usual understanding of a
positive correlation, it seems that the higher the P score the higher the
score on Hogge and Friedman's (1967) Scriptural Literalism Scale
which is set so that higher scores indicate a more literal interpretation
of the Bible. However, as we shall see in the remainder of this
literature review, this seems to be a somewhat anomalous finding.

---

\(^5\)Getz (1984, 1985) and Rest (1986) reported that Brown and Annis (1978) used the
P score.
Except for three college samples with unusually small P scores, the general trend in the literature for a post high school population is for more liberal religious persons or groups to score higher than conservatives on the P score, not lower. Indeed, Getz (1984), in her review of the literature, interpreted Brown and Annis's findings as showing a "significant relationship between high P scores and low literal belief" (Getz, 1984, p. 109, Table 1) -- just the opposite of what Brown and Annis appear to have reported. Along with Getz (1984) and Rest (1986), I suspect that the correlation was negative, but Brown and Annis did not report this. Given the other puzzling aspects of their report, it does not seem too improbable that they left out a negative sign.

While I was not able to secure a copy of Volker (1979) or Cady (1982), Getz (1984) reported the following results from them:

In a study of 36 college students, Volker (1979) attempted to identify the kind of college experiences that may be associated with high levels of moral judgment. The DIT and an author-developed experience check-list were used in this

---

6In two studies no significant differences were reported (Clouse, 1985; Holley, 1989, 1991). In one study a quite small, but statistically significant positive correlation was reported between P score and conservative Christian beliefs. However, it vanished in a stepwise multiple regression (Clouse, 1991).

7I have been unable to contact Irene Getz. I wrote to her dissertation adviser, Dr. James R. Rest. Rest (1986) held the same interpretation as Getz (1984) regarding the Brown and Annis (1978) results, and Rest may have their original data since he requests data from each study using the DIT. Unfortunately, he did not respond to this portion of my letter.

8Ohio State University Interlibrary Loan service informed me that no library reported owning the unpublished manuscript Volker (1979), and the only library/libraries that owned Cady (1982) refused to loan or copy it.
as liberal or conservative and their participation in church activities, both on 7-point scales.

A positive correlation ($r = .22$) was found between principled reasoning and low religious activity, but it was not significant. A negative correlation ($r = -.31$) was found between principled reasoning and religious beliefs. The relationship was significant at the .05 level. Students who characterized their beliefs as conservative used a lower proportion of principled reasoning. A conservative ideology tended to be associated with lower $P$ scores. There were no significant findings in regard to religious experiences (included on Volker's checklist of college experiences) and principled reasoning. (Getz, 1984, pp. 105-106)

Cady (1982) assessed moral development, religious belief, dogmatism, and other variables among 57 clergy in Bloomington, Minnesota. His measure of religious belief consisted of 7-point rating scales from basically conservative to basically liberal choices for six beliefs. They were beliefs about the Bible, Jesus, the resurrection of the body, the rapture, the millennium, and the Ten Commandments. Cady used the Defining Issues Test to assess preference for principled reasoning.

Using the work of Schuller, Strommen, and Brekke (1980), Cady identified the denominations to which the clergy belonged as conservative or liberal. There were significant differences in preference for principled moral reasoning between the conservative and liberal clergy ($t(31) = 5.718$, $p < .0001$). Cady also found that those who had a flexible interpretation of the Bible showed a greater preference for principled reasoning ($r = .576$, $p < .05$). The correlation between moral judgment and dogmatism was nonsignificant.

---

9One third, or 2 of the 6 items of this scale would be likely to distinguish fundamentalists from other conservative believers: "the rapture" and "the millennium" both are necessary components to the fundamentalist doctrine of the premillennial return of Christ.

10"Flexible interpretation" seems to be the opposite of a "literal interpretation," again, another case where a marker of fundamentalism (Scriptural literalism) -- which is not a marker of all conservative Christians -- correlated negatively with $P$ score.
Getz (1985) tested a sample of 249 subjects, 77 (30%) were dropped due to reliability problems on the Defining Issues Test or response sets. Of the remaining 172 subjects, 52 were from a Baptist congregation (conservative), 53 from a United Church of Christ congregation (liberal), and 67 were university students. Getz administered Brown and Lowe's Inventory of Religious Belief, Hoge's (1972) Intrinsic Religious Motivation Scale, an author-devised Attitudes toward Human Rights Inventory, a self-rating political ideology scale ranging from 1 (conservative) to 5 (liberal) and a demographics questionnaire that included use of the Duncan Socioeconomic Index (see method section for more description of this index). Relevant results included a significant negative correlation ($r = -.47, p < .002$) reported between P score and conservative religious beliefs and a positive correlation between stage 4 and conservative religious beliefs ($r = .53, p < .001$). However, in a multiple regression religious beliefs did not contribute a significant amount to the variance of the P score. An analysis of variance resulted in the church sample showing a significant negative relationship between intrinsic motivation and P score ($F = 5.95, p < .01$). The relationship was in the same direction but not significant in the student sample ($F = .41, p > .05$). With the higher N, the total sample showed a significant negative relationship between intrinsic motivation and P score ($F = 9.65, p < .01$). Education level resulted in a significant positive relationship
with P score in the church sample \(F = 12.96, p < .001\) but not in the student sample \(F = 1.43, p > .05\) which again was in the same direction as the church. This difference in significance may have been due to the restricted range in the student sample (all undergraduates) as compared to the church sample (which had the possibility of grade school to Ph.D. level). Again, the total sample, with its higher N, showed a significant positive relationship between P score and educational level \(F = 8.60, p < .001\).

Clouse (1985) gave her own measure of political and religious attitudes (to be described below) along with the 3-story Defining Issues Test to 371 college students at a midwestern university who were taking an education course and volunteered for the study. Forty-nine students' were discarded (13%) due to failing the reliability checks leaving 322 subjects.

The Clouse Politics-Religion Attitude Scale (Clouse, 1985) contains 38 items. Religious attitudes are measured by 10 items regarding liberal ("agnostic" or "nonrestricted") beliefs versus conservative religious beliefs on such issues as 'belief in the inspiration of Scripture, the deity of Jesus Christ, life after death, miracles recorded in the Bible, and the importance of religious experience" (p. 193). Subjects respond on a five point Likert scale ranging from strongly agree to strongly disagree. Both conservative and liberal items were included to lower response set. Liberal statements were given reverse weights from conservative statements such that lower scores are more conservative, higher scores are more
liberal. Scores could range from 10-50, with the 10-29 range categorized conservative and the 30-50 range considered liberal. Spearman Brown reliability corrected for attenuation equalled .86. Validity was assumed based on the scale's careful method of construction:

The religion statements were chosen from an original list of 33 statements sent to the 79 ministers of churches listed in the local telephone directory. The 33 statements were constructed after noting specific beliefs linked with the Christian faith in numerous books, journal articles, and public opinion polls. Sixty-three (80 percent) of the ministers responded to the questionnaire in terms of whether a religious conservative would agree with the statement, whether a religious liberal ("liberal" was defined for the ministers as "nonrestricted" or "agnostic") would agree with the statement, whether both would agree with the statement, or whether neither would agree. The 10 statements chosen were those for which there was 70 percent or more agreement by the ministers that a person conservative in religious belief would agree with the statement or that a person liberal in religious belief would agree with the statement. (Clouse, 1985, p. 193)

Religious conservatives scored significantly lower on the Defining Issues Test's P score than did religious liberals, $F(1, 314) = 5.13, p = .023$. Pearson zero-order correlations between P score (or any of the stage scores) and Clouse's Politics-Religion Attitude Scale were not reported. The scores of religious conservatives did not significantly differ from religious liberals at Stage 4 or Stage 3 of the Defining Issues Test. Stage 2 was not analyzed because there are too few items and the results may have been spurious. There were no
significant differences between men and women on the P score or Stage 4, however, males made significantly more choices at Stage 3 than did females, $F(1, 314) = 12.36, \rho = .001$.

In a more recent study, Clouse (1991) gave two measures of religiosity (religious beliefs and personal religious experience) along with the Defining Issues Test and collected information on GPA, class rank, and gender. The Clouse Politics-Religion Attitude Scale's 10 religion questions (already described above) were given along with the following set of questions regarding one's religious experience:

"Mark which of the following statements comes closest to your...religious experience: (a) I have never had a personal religious experience. (b) I have had a personal religious experience but it doesn't mean much to me now. (c) I have had a personal religious experience which is very important to me now." Students who marked either "b" or "c" then responded to one of the following statements: (a) My personal religious experience is with Christianity, or (b) My personal religious experience is with a religion other than Christianity. (p. 341)

Subjects who marked that their personal religious experience was with a religion other than Christianity (about 4%) were not used in the study.

There was a significant zero-order correlation in Clouse's (1991) sample between P score and religious belief ($r = -.12, N=335 \rho = .017$); however, since her scale puts conservatives at the low end of the Likert scale they actually scored higher on P score than liberals. This is the only college sample in which this has been reported (depending
on how one interprets Brown & Annis, 1978). Although the effect size is relatively small and only accounts for 1.4% of the variance in P, the sample size was large enough to make it significant.

Clouse (1991) drew the following four conclusions:

1) "GPA was a better predictor of stage of moral reasoning than gender, year in college, religious experience or religious belief" (p. 346) especially at the P level where it accounted for three times the combined variance of the four other predictors. In her 1985 discussion, Clouse recommended that rather than assuming that "conservatives and liberals do not differ in intellectual ability ... future research should be conducted in such a way that mental acumen is kept constant between groups or the extent of its influence is reported" (p. 197). This advice is supported by McCrae (1985) who recommended that researchers be aware of "the possible influence of third variables [such as intelligence and education] on results" (p. 440). In her 1991 discussion Clouse noted again that the diverse conclusions of past research relating Christian belief to cognitive moral development may be due in part to the different cognitive abilities among the subjects. "Unless some measure of intelligence or correlate of intelligence is included, this would not be known" (p. 346).

2) Using stepwise multiple regression analysis, Christian belief was not significant at any stage of moral development. However, a currently meaningful personal Christian experience was significantly negatively related to Stage 3 and positively related to Stage 4, after
GPA had been selected and entered through the stepwise multiple regression.

3) When using stepwise multiple regression there were no significant differences between men and women on any Defining Issues Test score. This is consistent with the literature review of Walker (1991) who concluded that empirical evidence has shown a pattern of nonsignificant differences between women and men on measures of cognitive moral development. Interestingly, however, when using a Pearson correlation matrix, Clouse reported that, contrary to Gilligan (1982), it was men (not women) who were more likely to score higher on Stage 3, while women tended to score higher at the P level.

4) The overall preference in this sample was for Stage 4. However, as explained above under *Cognitive Moral Development: Theory and Measurement*, this reported finding could be influenced by the greater number of Stage 4 queries compared to queries at other stages.

In another study, Ernsberger and Manaster (1981) gave the Defining Issues Test and Allport and Ross's (1967) Religious Orientation Scale to 360 church members. A stratified random sample was used to select members of four churches in Minneapolis.

---

11 While Ernsberger and Manaster (1981) did not test the individual beliefs of each subject, they did carefully assess the beliefs taught in each local congregation, and so, with some hesitation, I have reported their findings here under the category of "religious beliefs" as Getz (1984) and others have done in their literature reviews. As noted in the introduction chapter, however, there is no guarantee that those who attend a particular congregation will hold the same beliefs, especially in the liberal congregations (Glock & Stark, 1965).
Minnesota (60 members from each church in one stratum and 30 members from each church who fit the category of "teachers and leader-participants" in the other stratum). The return rate was slightly less than 50% and reasonably balanced between congregations (49 Methodists, 41 Unitarian-Universalists, 38 Missouri Synod Lutherans, and 41 Conservative Baptists). "Each sample was fairly evenly divided between male and female [subjects]" (p. 33), 80 males and 89 females. Socioeconomic status was measured using a combination of occupational and educational levels (see Manaster & Havighurst, 1972) but fell within a very narrow range of 9.12 to 10.63 (possible range was 5-30). Two churches (Unitarian-Universalist and United Methodist) were considered "principled" and two other churches (one affiliated with Missouri Synod Lutherans and the other affiliated with Conservative Baptist Association of America) were considered "conventional" according to the following criteria described in Ernsberger and Manaster (1981):

The principled congregations are of denominations which recognize ethical principles, beyond those literally permitted or endorsed in the Bible, on such moral issues as abortion, just and unjust war theory, and women's political, economic, and ecclesiastical rights. The conventional congregations belong to denominations which tend to regard implicit principles as unwarranted human interpretations of Biblical [sic] data, and which regard only the literal moral commands contained in the Bible as authoritative for moral judgment. Information was gathered from the research headquarters of the American Lutheran Church, from executives of the local urban council of churches, from on-site inspection of the churches, and from interviews with the pastors, to determine that these congregations were not doctrinally atypical of the
denominations to which they belonged. In each congregation, denominationally produced or approved curricula and program resources materials were frequently used with children, youth, and adults. (p.29)

Although Ernsberger and Manaster (1981) refer to these congregations as principled and conventional, I will refer to them as liberal and conservative, respectively, for ease of communication and comparison across studies. According to Glock and Stark's (1965) categorization of denominations based on theological orthodoxy, Methodists are categorized as liberals, American Baptists are categorized as conservative and Missouri Synod Lutherans are categorized as fundamentalists. Unitarians are not categorized and as Ernsberger and Manaster pointed out "a majority of Unitarian-Universalists... would not characterize or identify themselves as 'Protestants'" (p. 35). Although I am referring to the "principled" categories as liberal, the reader should keep in mind that while members of liberal or principled congregations are all likely to be exposed to sermons containing extra-Biblical ethical positions as described above, a minority of the members may still hold to orthodox theological beliefs. Glock and Stark (1965) noted that within more liberal congregations there is an orthodox minority who tend to socialize together and may be unaware of the less orthodox views of the majority who tend to socialize with those outside the church. Sermons in liberal churches tend to cover topical and ethical issues rather than theological issues, yet the creeds and rituals remain traditional (i.e. orthodox). Thus liberals may recite orthodox creeds
but agree less with these creeds than do the conservatives, but the
two groups never dialogue on the issue. These factors tend to
decrease the likelihood of any manifest conflict arising in liberal
churches over the issue of doctrinal orthodoxy (see Glock & Stark,
1965, p.119).

Ernsberger and Manaster (1981) reported that conservative
church members scored lower on the P score than liberal church
members, however, no significance tests were reported. Following the
pattern of Allport and Ross (1967) the indiscriminately proreligious
were defined as those who scored at least 12 points less on the
intrinsic scale than on the extrinsic scale, and were removed from the
sample. However, unlike Allport and Ross who used theoretical scale
mid-points, Ernsberger and Manaster chose four different median
splits based on how each congregation scored on the Religious
Orientation Scale. They concluded that intrinsic conservative church
members scored significantly lower on the P score while intrinsic
United Methodist (a liberal church)\(^1\) members scored significantly
higher even after socioeconomic status was controlled\(^2\). Conservative
church members clearly preferred Stage 4 \(\rho < .001\) while liberal
church members more often chose Stage 5A \(\rho < .001\) or 5B \(\rho < .01\).

\(^1\)Ernsberger and Manaster (1981) reported that the Unitarians evidently had
trouble with the theistic items of the Religious Orientation Scale and so their
scores did not parallel those of the United Methodists, i.e. did not reach
significance.

\(^2\)However, one should keep in mind that the SES range in the sample of
churches was small. This would lower the likelihood of finding significant
effects related to SES and make SES an irrelevant variable for this particular
sample.
On Stage 6 the differences were not statistically significant. Thus we might be safe to assume that the Conservatives as a whole scored statistically lower on the P score than the Liberals as a whole.

Finally, the more active members (those who volunteered as teachers and leaders for social action committees) in the liberal churches (especially the teachers) had significantly higher P scores than the members who were chosen at random, while in the conservative churches the teachers had lower P scores than the randomly chosen members. However, no statistical tests were reported for the conservative church comparisons. Getz (1984) has pointed out that Ernsberger and Manaster's findings support the notion that "intrinsic religious orientation may be an intensifier variable" (p. 107). This is consistent with Donahue (1985) who concluded in his review of the literature that "I is a good, unidimensional, nondoctrinal indicant of religious commitment" (p. 422). But, by I (Intrinsic) Donahue was referring only to the nine items that make up the Intrinsic Scale, while Getz (1984) and Ernsberger and Manaster (1981) were referring to the Intrinsic group in the four fold typology described above. This four fold typology required using both the Intrinsic and Extrinsic Scales, not just the Intrinsic Scale.

As we have seen, post-high school religious conservatives tend to score lower on the Defining Issues Test P score than religious liberals. One reason for this might be that conservative Christians who are intrinsically committed to their faith use the teachings of their
faith to make moral decisions in moral matters rather than using their own ideas about what should or should not be done. This was reportedly the case with Lawrence’s (1979) sample of fundamentalist seminarians. They scored below 9th graders and graduate students of Philosophy on the Defining Issues Test P score. These seminarians had the highest percentage of Stage 4 scores of any reported sample in Rest’s files (Rest, 1986).14

Following the Defining Issues Test, Lawrence (1979) interviewed all subjects by asking them to think out loud as they again completed the rating and ranking tasks on two Defining Issues Test stories: Escaped Prisoner and Doctor’s dilemma. Next, Lawrence asked each subject to complete card sorts on how easy it was to understand selected high stage and low stage queries and how likely each subject was to use these queries in evaluating what should be done by the character in each of the two dilemmas. Finally, Lawrence had the subjects write out paraphrases of certain queries from the Doctor’s Dilemma and the Escaped Prisoner stories (reproducibility tests).

For the card sorts and the reproducibility tests, Lawrence chose to analyze 62 of her 75 subjects. This sample included graduate philosophy students with D scores higher than 26, ninth graders with D scores lower than 24, and all 12 of the fundamentalist seminary students. Although the fundamentalist seminarians reported

---

14Rest asks users of the Defining Issues Test to send him their results and reports.
significantly greater understanding of the principled concepts of justice than did the 9th graders, they did not score significantly higher on the reproducibility tests. Typically, subjects score from one to one and a half stages higher on the Defining Issues Test than on Kohlberg's stages, most likely because it is easier to recognize than it is to produce higher levels of moral reasoning (Gielen, 1991). Therefore subjects tend to prefer some higher level queries which they can understand but not necessarily reproduce. The seminarians rejected these higher level queries which they claimed to understand in favor of stage 4 responses. As Rest (1986) put it:

The seminarians explained their thinking process in this way: they understood that they were being asked to make moral judgments. A tenet of their religious faith is that moral values come from divine revelation and that mortals should not depend on their own intuitions for making decisions about such matters. Therefore they deliberately suppressed their own personal notions about what was fair or just, and rated the Defining Issues Test items in terms of whether the item was consonant with some biblical passage, or some part of church doctrine that they could remember. And so in solving the Defining Issues Test dilemmas, they self-consciously discounted their own concepts of justice (their own intuitions of fairness), and deliberately derived a judgment of what was right by substituting religious ideology. The high Stage 4 scores comes (sic) from advocating external authority for the solution of moral dilemmas. The Lawrence study gives clear demonstration that people may have certain concepts of justice, yet they may not use them [italics added]. In the case of the fundamentalist seminarians, it is conservative religious ideology that overrides their own intuitions of justice. (p. 127)
It seems that in the case of subjects with a strong commitment to very conservative beliefs, such as these fundamentalist seminarians, the results on the Defining Issues Test might easily be interpreted in a way that is biased against their ability to understand principled levels of justice concepts\(^{15}\). This criticism might apply equally to Kohlberg's measure of moral development, which, according to Lawrence (1979), subsumes any type of religious material "under stage 4 Law and Order Orientation. In [Kohlberg's] scheme religious criteria are not differentiated from law and order, legalistic rule-applying religious orientation and reasoned or principled religious thinking are both categorized with other types of legalism" (Lawrence, 1979, p. 147-148).

Holley (1989, 1991) noted that the Defining Issues Test's multiple choice format, which employs fixed choices with a set content of responses, mixes the structural *form* of the six stages of moral development with specific *content*. This mixture of form and content did not trouble its developer, Dr. Rest, although he did acknowledge the possibility that a subject may respond differently to the content of one moral dilemma than another in terms of the stages (structural form) that the subject prefers (Rest, 1979a). But Holley (1991) relayed the argument that "the content of many psychological tests

\(^{15}\)While the typical seminarian in Lawrence (1979) may not have been able to paraphrase higher level moral reasoning, they did claim to understand it. It is very conceivable that a person with fundamentalist Christian beliefs could both understand and reproduce higher level moral reasoning and yet prefer Stage 4 queries on Rest's Defining Issues Test due to some Stage 5A and perhaps Stage 5B queries (see below under Shaver, 1987) being antithetical to their beliefs."
has an anti-religious bias" (p. 325) and went on to apply this
statement to the content of the Defining Issues Test:

Some have contended that there is a poor fit between the
present Stage 5 and 6 queries and conservative religiousness
which arises from the "non-Christian view of life" represented
by the DIT, in which love of this life, one's self, and one's society
is emphasized. On the other hand, the "Christian" orientation is
directed toward love of God as an ultimate goal. This difference
in orientation could explain why religious conservatives might
shun Stage 5 and 6 queries and thus, lower their DIT scores.
For instance, a person who believes that "... God is no respecter
of persons" [i.e., doesn't discriminate between people based on
wealth, sex, nationality, etc.] (Acts 10:34) might well consider
the DIT query "Do a majority of people in Mr. Webster's society
feel like his customers or are a majority against prejudice"
(Webster, #7 [Stage 5A] ) an irrelevant and perhaps even
senseless, argument. (p. 325)

In an insightful and creative attempt to rectify this problem in
content, Holley (1989) altered the queries of the Defining Issues Test
to include references to God. For example, instead of citing obedience
to the the "community's laws" as a reason for not stealing, Holley
substituted "God's law." However, because Holley was attempting to
make the standard Defining Issues Test and his "Christian content
version" comparable, his method may have failed to achieve his goals.
By substituting in "God" for "society," Holley may have created queries
that only superficially appear to meet the Christian criterion of being
"directed toward love of God as the ultimate goal" (p. 325). It seems
that at least some of these principled level queries still retain some
basic underlying humanistic philosophies and world views that do not necessarily agree with conservative Christian teachings.

For example, compare the Defining Issues Test version of Heinz #12 (a Stage 5A query), "Would stealing in such a case bring about more total good for the whole society or not," with Holley's "Christian version," "Would stealing in such a case further God's efforts [italics added] to bring about more total good for the whole society or not?" (Holley, 1989, p. 123). Both versions contain the underlying principle of utilitarianism, a philosophy aimed at achieving the greatest good for the greatest number of people. The only difference is that in Holley's Christian version God is purported to have the utilitarian goals. But the conservative Christian understanding of God is not necessarily compatible with utilitarianism. For example, a conservative Christian might reason that if God were a utilitarian then the greatest good for the greatest number of people might require God to force every one into believing that "Jesus died for the sins of the world" so that all could go to heaven. But conservative Christian doctrine teaches that God appears to value individual free will above forcing people to do what they do not wish to do, even if that means the majority choose to be in Hell rather than Heaven. The words of Jesus found in Matthew 7:13 (New International Version) support this doctrine: "Enter through the narrow gate. For wide is the gate and broad is the road that leads to destruction, and many enter through it. But small is the gate and narrow the road that leads to life, and only a few find it." Given this one example, it seems that at least some of Holley's
Christian version queries have retained their bias and still may inhibit conservative Christians from choosing some principled level queries. In addition to changing the wording of the queries, Holley (1989, 1991) also made changes in the dilemmas such that the characters involved in the dilemmas (both the originator of the action and the recipient of the action) were both Christians, or in another case, only the originator of the action was a Christian. This was done under the assumption that the context of the dilemma would affect the type of moral reasoning used by conservative Christians. Holley argued that if the characters involved in the dilemmas were identified as Christians, then conservative Christian subjects may be more likely to use "other-worldly" principles (e.g., principles based on a "salvation" orientation, Holley, 1989, p. 30) rather than confining themselves to principles typically based only on the existence of this physical world. Holley predicted that a negative relationship would be found between the original Defining Issues Test P score and conservative Christian religiosity, consistent with past research. Furthermore, he predicted that this negative relationship would be decreased in his Christian versions both by labelling the characters in the dilemmas Christian and by rephrasing the queries to include references to God.

The various "versions" of the Defining Issues Test dilemmas were distributed to 163 Introduction to Psychology students from a conservative Midwestern university who had volunteered by signing up on a bulletin board. These tests were appropriately counterbalanced and all subjects took both the standard and the
Christian version of the Defining Issues Test queries. Subjects were also given five religious measures: Batson's Doctrinal Orthodoxy Scale, Clouse's Religious Attitude Scale (described above in Clouse, 1985), Hoge's Intrinsic Motivational Scale, the Indiscriminate Proreligiousness Scale, and The Religious Problem-Solving Scales (Self-Directive, Collaborative, and Deferring). Holley (1991) reported that "none of the religious measures used in this study correlated significantly with the P-score" (p.330). Holley (1991) went on to conclude:

no strong support was obtained for any of the hypotheses. However, the pattern of the correlations between Stages 4 and 5A of the DIT suggests that Stage 4 moral reasoning tends to be positively associated with conservative religiosity, whereas, Stage 5A moral reasoning tends to be inversely associated with it. This pattern of relationships is consistent with Rest's interpretation of past research in this area, but inconsistent with the hypotheses advanced in the introduction. (p. 331)

In his abstract Holley (1991) suggested three possible reasons for his hypotheses being rejected. First, the sample had an overall low P score compared with the the average sample for college students reported by Rest (1979b, p. 7.2). Apparently, neither the conservative Christians nor those who are not conservative Christians were scoring high on the P score compared to Rest's sample of over 2,000 college subjects from various parts of the United States. Also test retest reliability on a number of measures was poor. Low reliability will serve to diminish potential correlations among variables. Second, "the alterations [on the Defining Issues Test] may have been too obscure to
influence the results" (Holley, 1991, p.323). Third, "religiously oriented moral development might evolve in a manner not yet described in the literature" (p. 323). For example, moral judgment in conservative Christians may better be measured by dilemmas with God than by dilemmas among members of society. Although Holley suspected bias in the Defining Issues Test, his sample responded in a way that made it impossible to test his hypotheses adequately.

Moral Judgment Development at Christian Colleges

Whereas it may be true that some of the Defining Issues Test P score queries are biased against subjects with conservative Christian beliefs, there is evidence of mean P scores for Christian liberal arts college student samples in the average range for general college samples (see McNeel, 1991, p. 317, Table 2 compared with Rest, 1979b, p. 7.2).

Again, self-selection related to the moderator variables of intelligence and socioeconomic status must be considered. For example, in Shaver (1987) subjects who chose to go to a Bible college had significantly lower average entering P scores (M-33.45, SD-12.73) than those who chose to go to a Christian liberal arts college (M-41.51, SD-17.22). However, the Christian liberal arts college subjects were likely to have an overall higher intelligence level given that "on the basis of entering freshmen's scores on the American College Testing program (ACT), the liberal arts college would be considered highly
selective and the Bible college would be considered of low selectivity" (Shaver, 1987, p. 212). However, this particular Christian liberal arts college (Wheaton College) has higher selection standards than many Christian liberal arts colleges. Wheaton College "is a fairly selective institution (incoming freshman SAT averages: V-543, M-483 compared to V-463, M-495 in the Christian College Consortium as a whole; data from Cass & Birnbaum, 1985)," according to McNeel (1991, p. 314).

Many subjects from the Christian College Consortium are also likely to be of fairly high socioeconomic status based on a frequently used indicator: parental education. Buier, Butman, Burwell and Van Wicklin (1989) reported that, overall, a sample from three of the colleges in the Christian College Consortium had parents who were well-educated. Percentage of subjects' parents who had a college or graduate degree was 70% for the sample of 99 freshmen and 63% for the sample of 93 seniors. Both samples were obtained through a computer-generated random list or from class volunteers, thus samples are not entirely random. This level of parental education is likely to be quite higher than that of the average Bible college subject's parents and also likely to be higher than that of the average state university student's parents.

According to McNeel (1991), subjects in several Christian liberal arts colleges made significant increases in their mean P scores over the course of four years in longitudinal studies. And in cross sectional studies seniors had significantly higher P scores than freshmen. Only
McNeel's design provided for collecting longitudinal data from one sample both in their freshmen and senior years, and then in the seniors' last year simultaneously collecting cross sectional data randomly on another sample of seniors and freshmen (i.e., individual senior subjects were not in both longitudinal and cross-sectional samples). McNeel (1991) stated:

This is an important design enhancement because it makes it easier to draw firm conclusions about the meaning of freshman-senior differences. The longitudinal sample minimizes differential selection and dropout as simple explanations for any freshman-senior differences, while the cross-sectional sample minimizes such threats as history and testing (Campbell & Stanley, 1963). If similar results are obtained in each sample, then the above four threats become less likely as rival explanations to the college impact hypothesis. (p. 314)

Additionally, among individual subjects in the longitudinal sample, McNeel (1991) reported that 67.9% showed increases greater than the estimated standard error of 7.5 and only 7.1% showed decreases greater than the estimated standard error. The other 25% did not exhibit changes larger than the standard error and thus their changes were not counted because they could be due to measurement error in the six-story P score. McNeel added, "these percentages are very similar to the four-year change percentages reported by Rest (1979[a], p. 131) for college students of ages 17-21 (63% and 8% respectively)" (p. 318). Whereas there was good evidence supporting the hypothesis that four years at a Christian liberal arts college generally improves P
score. Shaver (1987) reported that the mean P score for the Bible college sample stayed the same over four years.

An important distinction among evangelicals who call themselves fundamentalists and those who don't may be relevant to what type of Christian higher education conservative evangelical Christian students choose. Ammerman (1982) and others have noted that

Fundamentalism and evangelicalism are not the same phenomenon; they cannot be hyphenated. Fundamentalists subscribe to inerrancy; not all evangelicals do. . . . One of the most important of the differences, in fact, is fundamentalism's insistence on being different. The issue of "separation from the world" results in distinctive fundamentalist lifestyles and in organizational isolation from the evangelical mainstream. In addition, fundamentalists are different from the rest of their evangelical brothers and sisters in their belief in dispensational pre-millennialism. This is no insignificant doctrine. . . . Their eschatology, along with their literalism and separatism, identifies fundamentalism as a distinct group within evangelicalism. (p. 170-171)

Perhaps the Bible college was more of a fundamentalist institution than the Christian liberal arts colleges, and it may have attracted students who were more fundamentalist than those attracted to the Christian liberal arts colleges.

Shaver (1987) reported that the failure of the P score to change did not tell the whole story. On average, Bible college subjects made statistically significant increases in Stage 4 of +2.81 ($p < .05$), Stage 6 of +3.11 ($p < .001$) and a significant decrease in Stage 2 of -1.73
(\(p < .05\)), Stage 3 of -1.67 (\(p < .05\)), and Stage 5 of -3.70 (\(p < .05\)) with a mean \(P\) of 33.21, SD=11.09. Since Stage 5 went down about the same as Stage 6 went up, there was no change in \(P\) score. Shaver then interpreted this result:

The significant increase in preference for Universal-Ethical-Principled Reasoning indicates that the Bible college students are organizing some system of principles and values in their lives and are attempting to live consistently by these values. Rest (1979) stated that the universal-ethical-principled reasoner comprehends not only the procedure for making just laws (Stage 5) but also understands the substantive aspects behind the procedure. The Stage 6 reasoner anticipates the principles by which members of a totally just society would choose to live. The results indicated that the Bible college students seem to have increased in understanding the substance of the principles of a just society but may have difficulty in their ability to participate in the procedure required to implement them in society. The significant increase in Stage 4 reasoning further supports the concept that the Bible college students may have difficulty participating in an environment that requires negotiation and compromise.

The liberal arts students followed the predicted sequential stage development. The Bible college students, however, displayed a unique form of development. The pattern of increases at Stages 4 and 6 and no change at Stage 5 is contrary to the step-by-step sequential development described by Kohlberg (1976). It seems that the entering Bible college students may have selected an institution that, after 4 years, reinforces their initial pattern of reasoning.

Other questions to be addressed include (a) Are there specific values, reflected in the Stage 5 responses offered in the DIT dilemmas, to which Bible college students respond negatively? (b) If so, could Stage 5 responses be written in value terms that would be acceptable to these students? (c) Did the Bible college students use less Stage 5 reasoning on all the
dilemmas or just on dilemmas dealing with clear Biblical mandates (e.g., stealing)? . . . (p.216)

Shaver concluded that if both Stage 5 and Stage 6 scores are not reported, significant changes may be masked by P score.

Shaver (1987) proceeded to call for replication at other institutions and this is a good suggestion. I would add a cautionary note to his interpretation of the increase in Stage 6. In his study he used the three-story form of the Defining Issues Test, which has only two Stage 6 queries (see Table 3). Test-retest reliability is rather low for individual stage scores on the three-story form. Stability coefficients for Stage 6 range from .26 (for ninth graders) to .52 (for college students) with the combined sample's stability coefficient at .47. Another sample, which included subjects from junior high, high school, college, and graduate school, showed a Stage 6 stability coefficient of .50 (Rest, 1990, p.5.2). Of the all the stage scores, Stage 6 has the greatest percentage increase in standard error of measurement from the six-story form to the three-story form (Rest, 1990, p. 5.3). Therefore I question the reliability of the two-query Stage 6 score on the three-story form. Rest (1990, p. 5.2) recommends interpreting individual stages only for the six-story form. I would hesitate to put much weight in Shaver's (1987) Stage 6 increases until they are replicated on a sample using the six-story form of the Defining Issues Test (and even then there may be another interpretation to the increase in preference for Stage 6 queries).
Specifically, I am concerned that, due to their ambiguous wording, the two Stage 6 queries of the three-story form are open to being reinterpreted by a subject in a Stage 4 manner. (This may be why its standard error of measurement is relatively high.) Shaver (1987) helps me make my point when he writes: "Are there specific values, reflected in the Stage 5 responses offered in the DIT dilemmas, to which Bible college students respond negatively?" (p.217). Yes, there probably are values in Stage 5 that offend Stage 4 thinkers... "liberal values" such as: "encouraging the break down of society by saying things are relative, so it's OK to steal sometimes and break the Ten Commandments; also, letting criminals off scot-free when they deserve to be punished and serve their time." Try to imagine how a Bible college student with this Stage 4 mind-set would make meaning of the Heinz story, Stage 6 query #8 that reads: "What values are going to be the basis for governing how people act towards each other?" What values? "Traditional Values!" "This society is going to Hell in a handbasket and we've got to get back to our traditional values of God and Country. Whatever happened to 'Thou Shalt not Steal'? We can't allow stealing, for any reason. Sure, I'm all for getting back to values; this has got to be one of the more important items for me." Or consider the other three-story, Stage 6 query #5 (from Escaped Prisoner): "Would society be failing what Mr. Thompson should fairly expect?" "Yeah, he should expect to serve his time. Not break out of jail and think it's going to be OK. Criminals need to be punished to set an example. And if she failed to turn him
in, it wouldn't be fair. Although this Stage 6 query is much less likely to be misinterpreted at a Stage 4 level than the Stage 6 query in the Heinz story, it's possible to do so. Thus, with enough Stage 4-thinking Bible college subjects reinterpreting either one or both of these two queries, one could find a significant increase in mean Stage 6 scores after the first year, especially when the tendency toward Stage 4 thinking is increasing over the years as in this sample. This reinterpretation theory could be examined by following Rest's (1986) suggestion of measuring subjects' comprehension of moral judgment queries, as well as their Defining Issues Test scores.

Finally, Shaver (1987) asks the question of whether Stage 5 queries are written in value terms to which Bible college students respond negatively. This question has been asked more recently by others, including Holly (1991). As noted above, Lawrence (1979) provided evidence that for at least two stories, Escaped Prisoner and Doctor's Dilemma, fundamentalist seminarians found a number of queries causing them negative reactions. Generally these queries were above Stage 4 and were disagreeable to seminarians because they failed to match their understanding of Biblical teaching related to the issues in the story.

Of the published studies I was able to obtain, two broke down their P score results into separate stages. Ernsberger & Manaster (1981) reported that on the six-story form, Defining Issues Test Stage 6 mean scores failed to show significant differences between conservative and liberal congregations, however, liberal congregations
showed significantly higher levels of Stage 5B, and even higher levels of 5A, than conservative congregations, while conservative congregations showed significantly and dramatically higher levels of Stage 4 than liberal congregations. Holly (1991) reported that a combination of all religious variables correlated significantly and positively with the six-story form Defining Issues Test Stage 4 and significantly and negatively with Stage 5A, with no mention of results for Stages 5B and 6, so I assume correlations for these stages were not significant. Thus, in these two studies, there is a consistent pattern of nonsignificant differences in Stage 6 and significant differences in 5A, with one study noting significant differences in 5B on a conservative-liberal continuum of religiosity.

This pattern could be interpreted as bias against fundamentalist Christians operating in Stage 5A and possibly also in Stage 5B, but not in Stage 6. Or it could be consistent with the Shaver's (1987) first interpretation that the type of person attracted to Bible college skips over Stage 5 but does develop in Stage 6. Or it could be consistent with my theory that Stage 6 queries are being reinterpreted as Stage 4 queries by subjects who are predominantly Stage 4-thinkers.

Two of the three remaining Stage 6 queries on the six-story form seem vulnerable to such reinterpretation. Subjects are to choose for or against the action stated in the story. Then they are to choose among queries that relay important criteria to solving the problem. Given that when subjects rank a particular query high they may still apply it to either choice, for or against a particular action, it is
perfectly legitimate for subjects to respond negatively to a query suggesting a choice that was *opposite* the one they felt was moral. Such a query would be very important in helping them select what "should" be done, and would likely be ranked as one of the most important queries. Following are two more examples of how Stage 6 queries may be reinterpreted as Stage 4 by Stage 4-thinkers. (1) Stage 6 query #7 (from Student Take-over): "Is taking over a building consistent with principles of justice?" "Heck NO! Not if it's against the law, if it's against the law then it's never *consistent with the principles of justice*. They have no right to be there. This is a good item, I'm going to rank it fairly high."

(2) Stage 6 query #8 (from The Doctor's Dilemma): "Is helping to end another's life ever a responsible act of cooperation." "NO, NEVER! Gee, this one is almost as good as 'Whether only God should decide when a person's life should end.' But then there's that one about it being against the law for the doctor to kill someone, and there is also the one about not letting people kill themselves whenever they want to. Hmmm, maybe I'll have to rank it fourth, no maybe third."

As you can see, it is very possible to reinterpret four of the five Stage 6 queries into Stage 4-thinking. For some queries it seems more likely than for than others. But the point remains: until we can be sure that subjects are interpreting Stage 6 queries at the Stage 6 level and not at the Stage 4 level, we can not be certain of the meaning of

---

16 Bible college students and fundamentalist seminarians tend not to say "Hell", at least not as an expletive.
longitudinal leaps over Stage 5 for improvements in Stage 6, or the
meaning of nonsignificant differences in Stage 6 between groups
differing on some measure of religiosity. Again, comprehension
measures as recommended by Rest (1986) would be useful for the
purpose of ruling out the reinterpretation hypothesis.

While Rest's Defining Issues Test and some versions of
Kohlberg's Moral Judgement Inventory may be open to accusations of
bias against respondents with conservative Christian beliefs, Gibbs'
system (Basinger, Gibbs, & Fuller, 1992; Gibbs, Basinger, & Fuller,
1992), on the other hand, may not be biased against conservative
Christian subjects because his system views stage 4 as the highest
level of moral development and dispenses with stages 5 and 6. Stages
1 and 2 are considered immature and stages 3 and 4 are considered
mature. According to Basinger, Gibbs, and Fuller (1992):

'the rarity of not only stage 6 but also stage 5 suggests a datadriven need for further modifications in Kohlberg's
developmental typology. We suggest that the discarding of the
preconventional—conventional—postconventional trichotomy
allows the emergence of a more valid moral judgment maturity
construct, that is, one less dependent on philosophical
sophistication.' (p. 11)

More specifically, stage 5 descriptions are based upon the
Western philosophies of Kant, Rawls and others. Gibbs has repeatedly
argued that a cognitive moral development that is said to be naturally
occurring across all cultures should not be dependent on Western
philosophical sophistication (Basinger, Gibbs, and Fuller, 1992).
It could be that post high school subjects are exposed more to the Western philosophy exhibited by Stages 5 and 6 than are high school subjects. Those post-high school subjects who are more committed to conservative Christian beliefs prefer to use their beliefs to guide their moral choices, while the less committed and the less conservative are open to using new "principled" (philosophically sophisticated) levels of moral judgment. This could explain why it is that not until after high school do subjects with conservative Christian beliefs tend to score lower on the P score than subjects with more liberal beliefs. However, this simple explanation does not seem to take into account that subjects from Christian liberal arts colleges also make progress in moral judgement levels as measured by the P score (McNeel, 1991). On the other hand, this explanation could remain plausible if the conservative beliefs that were negatively associated with the P score were mainly fundamentalist beliefs, and if the subjects from the Christian liberal arts colleges were not Fundamentalists.

Perhaps the difference is that conservative Christian students at secular state universities are so threatened by the non-Christian ideologies that they try to avoid these alternative perspectives, shut down, and refuse to accommodate the information. McNeel (1991) who was "both a student and faculty member at Christian liberal arts

---

17 This line of reasoning raises the question "are there principled level beliefs and teachings in the Christian faith?" It seems there are. For the interested reader, Holley (1989, 1990) has convincing evidence supporting the existence of principled level thinking in the Christian tradition. Other examples of principled level thinking based on Christian teachings (including the quote found on the frontispiece of this document) may be found in Appendix B.
colleges (Westmont and Bethel) as well as state universities (University of California and Southern Illinois University)” (p.322), supports this theory from his own experience:

During seven years of teaching at Southern Illinois University, several times I had conservative Christian students ask me which professors in my department they, as committed Christians, should avoid. In their minds there was a clear need to protect themselves from certain kinds of information and from the influence of certain kinds of people. (Note 1, p. 320)

At a Christian liberal arts college, on the other hand, students come in expecting the values, ideas, and perspectives to be Christian and thus "safe." Thus they are more open to these ideas and more likely to accommodate them. Unfortunately, not all Christian colleges will use this greater openness as an opportunity to challenge students' world views and encourage growth in moral judgement, but some will:

While a kind of ideology which might inhibit growth in principled moral reasoning undoubtedly is present in some conservative Christian higher education settings, the ideology of many Christian liberal arts colleges is different. For example, when discussing the mission of the Christian liberal arts college, the Christian College Consortium states:

The Christian liberal arts college, as a community of learners, attempt to bring its Christian commitment into critical and constructive interaction with the arts and sciences in honesty, self-reflection and intellectual rigor . . . The personal development of students in this learning environment includes their response to Christian values as they learn to exercise self-reflective, prophetic leadership and service in all communities of which they
are members, from church to the larger society . . . Any permeation of life and thought by Christian commitment must transcend defensive apologetics, devotional applications or moralizing . . . The Christian [liberal arts] college becomes the locus for an internal criticism of Christianity which is rare or absent from those Christian institutions existing primarily for evangelistic, devotional or polemically apologetic purposes. (Christian College Consortium, 1979, pp. 20, 22, and 26)

Thus, remaining committed to conservative theology, these colleges are also strongly committed to an open-minded examination of their own and a serious consideration of alternative perspectives . . . As Kohlberg (1976) has shown, such an openness is one of the basic cognitive prerequisites of principled reasoning. Thus, it might be expected that [an open-minded, hermeneutically humble] Christian liberal arts education [based on an ontological commitment to a loving personal God] would lead to growth in principled reasoning. (McNeel, 1991, pp. 313, 314)

It is also possible that some subjects who score higher on the Defining Issues Test don't really understand the meaning of the principled queries at all. Rest (1986) noted that:

some subjects may give high ratings to Stage 6 statements because they understand the distinctive Stage 6 way of thinking and appreciate its greater adequacy, but other subjects may give high ratings to Stage 6 because they project their own thinking into the statement and like its fancier style [emphasis added]. The problem with rating and ranking data is that one cannot always be certain of the basis on which the subject is judging the statement. (p. 81)

In order to circumvent these problems, Rest (1986) recommended a time-consuming process in which subjects'
comprehension of moral judgment queries, as well as their Defining Issues Test scores, are measured. Unfortunately, such methods are beyond the constraints of this study. But fortunately, the Defining Issues Test has M (meaningless) queries, which will be helpful to some extent in weeding out subjects who are ranking queries mainly based on fancy language. If subjects are merely attracted to queries with a fancier style then they are likely to score high on M queries and hence will be removed from the study. Rest (1986) also cited research showing that when subjects attempted to "fake good" on the Defining Issues Test by pretending to take the test to show "the highest principles of justice" they did no better than a control group. "These findings suggest that under the usual conditions, subjects are giving their best notions of the highest principles of justice, and that the test-taking set of 'faking good' does not appreciably increase scores" (p. 5.6).

In addition to fundamentalist Christian subjects comprehending but refusing to choose Stage 5 (and possibly State 6 queries), along with the possibility of reinterpreting some Stage 6 queries as Stage 4 queries, many other variables and factors may moderate the relationship between conservative Christian beliefs and P scores. (Some examples are interpersonal negotiation, perspective-taking, and cognitive disequilibration, see Hanson, 1991). My study, however, focused on the potential moderators of intelligence and socioeconomic status, which I will outline next following a brief summary of this section.
To summarize the studies reviewed above, in most cases, among samples who have graduated high school, higher levels of principled moral judgment were significantly associated with lower levels of conservative or fundamentalist beliefs. There were two clear exceptions to the above pattern. Clouse (1991) showed a quite small, but statistically significant, positive correlation of .12 between P score and conservative Christian beliefs and Holley (1991) reported no significant correlations between P score and any of his measures of religiosity. Thus, while the tendency for moral judgment to be negatively associated with conservative Christian beliefs may appear quite strong, a few of the more recent studies have failed to support this trend. The results of one study (Brown and Annis, 1978) seemed somewhat unreliable and untrustworthy. In accordance with the majority of the studies reviewed above, I hypothesized a significant negative correlation between the P score and my measure of conservative Christian beliefs (Short Christian Orthodoxy Scale) and a significant negative correlation between P score and my measure of one aspect of fundamentalist beliefs (Scriptural Literalism Scale).

**Intelligence**

Clouse (1985, 1991) has noted that intellectual ability should be controlled for in studies relating religion to cognitive moral

---

18 The studies were not designed to make distinctions between conservative and fundamentalist beliefs.
development. I have yet to see intelligence controlled for in a post high school sample in any literature review or in any of the studies I have located through PsychLit. In her 1991 study, Clouse controlled for GPA; but GPA is not purely a measure of intellectual ability but also a measure of perseverance and study habits. Kahoe (1974) noted that scores on Allport and Ross's (1968) Extrinsic Religious Orientation subscale correlated negatively with ACT scores ($r = -.19, N=188, p<.01$), while scores on the Intrinsic Religious Orientation subscale were not significantly correlated with ACT scores. Thus some religious subjects may indeed have lower aptitude for obtaining good grades. Extrinsic religious orientation was negatively correlated with college first year GPA ($r = -.23, p<.01$), however, this relationship vanished when aptitude (ACT score) was controlled for. Intrinsic religious orientation was positively correlated with first year college GPA ($r = .25, p<.01$) with or without statistical controls for college aptitude. Thus it seems that intrinsics try harder, so GPA may be an overestimate of the effects of intellectual ability on intrinsics' Defining Issues Test scores. This interpretation is not inconsistent with Zern (1987) who stated that self-reported religiosity was negatively related to ability as measured by SAT scores and not related to achievement, as measured by GPA. Zern concluded:

"In all cases, the highest percentage of students achieving more than expected on the basis of their SAT scores came from the most religiously committed, and the lowest percentage came from the least religiously committed, with the moderately involved always having an intermediate percentage of above
Argyle and Beit-Hallahmi (1975) have noted that various studies show denominational differences in intelligence which appear to associate more conservative Christian denominations with lower intelligence and more liberal denominations with higher intelligence:

Jews and Episcopalians come out high on average, Baptists and Catholics low. The most likely explanation of these denominational differences is in terms of the social class differences between denominations (p. 166f). These in turn are partly due to the fact that particular sets of beliefs and practices have more appeal to one class than another.... A further factor is family size: Catholics, Mormons and others have larger families (p. 159f), and large family size may be a source of lower intelligence, since the children receive less attention from their parents....

Studies comparing IQ scores in religious groups have become rare.... Rhodes and Nam (1970) arranged categories of religious identification according to their degree of fundamentalism and anti-intellectualism, with Baptists highest and Jews lowest. When the religious groups were ranked according to their distance from fundamentalism, there was a positive significant correlation of 0.17 with IQ. Belonging to the Jewish or major Protestant group was associated with a higher IQ, compared with Catholics, small Protestant groups, and Baptists. (pp. 93-94)

Rest (1979a) has noted significant moderate correlations between the Defining Issues Test and measures of intelligence, aptitude and achievement. The relationship occurs from junior high school to college and on into adult samples. For example, in an
unpublished doctoral dissertation by Coder (1975) the Quick Word Test (QWT) correlated .42 with the Defining Issues Test on an adult sample. This relationship between the Defining Issues Test and measures of intellectual aptitude and achievement occurs both with verbal and nonverbal measures and does not consistently favor verbal over nonverbal. According to Rest (1979):

> whenever subtests are given, the verbal subtest is not consistently more highly correlated with the DIT than the nonverbal subtest (in only three out of seven comparisons verbal is more highly correlated than nonverbal) and that the composite indices (verbal plus nonverbal) are consistently more highly correlated than just the verbal subtests alone. This suggests that the linkage between intellectual aptitude and the DIT is not just verbal fluency or knowledge of special vocabulary, but the linkage is in a common general factor of intelligence. (p. 147)

Given the moderate positive correlations between the Defining Issues Test and intelligence and the possible association between conservative religious affiliation and lower intelligence it seemed wise to follow Clouse's (1985, 1991) advice and control for intelligence, which has seldom, if ever, been done for a post high school sample with Christian beliefs as a predictor variable.

**Socioeconomic Status**

As alluded to by Argyle and Beit-Hallahmi (1975) above, socioeconomic status also tends to vary by denomination. Again, the
more conservative denominations tend to have lower socioeconomic status than the more liberal denominations. According to Spilka, Hood, and Gorsuch (1985) the "sociology of religion tells us that American religious groups are not socioeconomically equivalent. Episcopals, Presbyterians, and Jews tend to be rather high on the class ladder. Catholics, Pentecostal sects, and Baptist bodies group toward the lower end of the economic distribution" (p. 307).

Although Rest (1979a) stated that "SES has not been a consistent or powerful correlate of the DIT" (p. 120), he added that the socioeconomic status data has only been used as a covariate and no studies have been reported which used the full range of socioeconomic status while controlling for other variables. So it seemed possible that the Defining Issues Test would correlate with socioeconomic status and was worth including in this study given the positive correlation between liberal denominations and higher socioeconomic status and the generally positive correlation between liberal beliefs and the P score of the Defining Issues Test.

Conclusion

The central aim in this study was to explore the relationship between religious beliefs and moral judgment/cognitive moral development. Given that denominations having higher socioeconomic status and intelligence levels tend to teach more liberal doctrines, and more liberal beliefs have been associated (e.g., positive correlations)...
with principled moral reasoning, one might wonder what proportion of
the variance in principled moral reasoning is left to differences in
religious beliefs once the variance associated with intelligence and
socioeconomic status is removed.

Past reports of the tendency toward a negative association
between moral judgment as measured by the P score on the Defining
Issues Test and conservative Christian beliefs led me to expect a
significant negative correlation between my two measures of religious
beliefs (the Scriptural Literalism Scale and the Short Christian
Orthodoxy Scale) and the P score of the Defining Issues Test. This
negative correlation was expected to be greater for those religious
subjects who were more intrinsically motivated by their beliefs
(scored higher on the Intrinsic Scale) than for those who were less
intrinsically motivated by their beliefs (scored lower on the Intrinsic
Scale). Thus, I hypothesized an interaction between my measures of
belief and the Intrinsic Scale in predicting principled moral judgement
as measured by the P score.

Intelligence and socioeconomic status were expected to
moderate these correlations. There is some evidence to suggest that
those with more conservative denominational affiliation have lower
cognitive ability than those with less conservative denominational
There is also evidence to suggest that those with more conservative
denominational affiliation have lower socioeconomic status than those
with more liberal denominational affiliation (Argyle, 1975; Chalfant,
Beckley, & Palmer, 1981; Spilka, Hood, and Gorsuch, 1985). Since intelligence has been positively correlated with the P score on the Defining Issues Test (Rest, 1986), and socioeconomic status theoretically should covary with measures of moral judgment (as it did with Kohlberg’s measure, Colby, Kohlberg, Gibbs, & Lieberman, 1983), I expected the combined factors of intelligence and socioeconomic status to significantly correlate with the P score. Once intelligence and socioeconomic status were partialled out from the P score relationship, given the potential correlation between intelligence/socioeconomic status and conservative religious beliefs, I expected the correlation between conservative religious beliefs and P score to drop dramatically (perhaps cut in half). However, I still expected there to remain a significant negative correlation between conservative Christian beliefs (on both the Scriptural Literalism Scale and the Short Christian Orthodoxy Scale) and P score. The main purpose of this study, then, was to determine whether intelligence and socioeconomic status would account for a significant proportion of variance in the expected negative correlation between conservative Christian beliefs and moral judgment as measured by the P score of the Defining Issues Test. I expected they would.

One concern I had regarding the reliability of my measure of socioeconomic status made a change in analysis a possibility. Although I had refined my instructions to ask for specific information on parents’ career, it was likely that the some of the information I received would be somewhat vague and thus vary the accuracy
slightly in matching the parent's actual career with the appropriate score on the Duncan Index. Additionally, there was the unresolvable issue of using only one parent's career even when both were working (thus creating a higher actual socioeconomic status than my index was designed to measure). I wanted to be prepared to consider dropping the socioeconomic variable if there was good evidence for its unreliability. Fortunately, inter-rater reliability was quite high ($r = .93, N=222, p<.01$)\textsuperscript{19}. 

I did not expect class rank to account for substantial or statistically significant variance in my sample because of the restricted range (mostly first and second year students). However, because amount of education has been shown to significantly correlate with moral judgment (Rest, 1990), I included class rank to reduce my error variance in case I had some more highly educated subjects in my sample.

\textsuperscript{19}This fairly high inter-rater reliability was partially due to a set of guidelines I developed for interpreting and categorizing somewhat ambiguous types of potential responses (see Appendix E).
Chapter III

Method

Subjects

Subjects were 268 Introductory Psychology students at a large midwestern university who signed up for this study during Spring Quarter, 1993. Sign up sheets contained the following restriction: "To participate in this study you must have learned English as your first language and been born in the U.S." Of the 268 subjects, 50 (18.7%) were removed from the analysis because of the potential unreliability of their Defining Issues Tests, as recommended by Rest (1983, 1990). Of the 218 subjects remaining, two failed to complete significant portions of the religious or demographic questionnaires making their protocols unusable. Of the remaining 216 subjects, 10 subjects were removed because they did not provide enough information to classify their general religious affiliation or they reported belief in a religion other than Christianity, leaving 206 subjects. An additional seven subjects were dropped because they reported a personally meaningful

---

20Requiring English as a first language made the English vocabulary based IQ test scores more evenly comparable. Restricting sample to those born in the US was done to limit the number of subjects who were raised in a culture that taught a religion other than Christianity, and thus to limit subject loss.
religious experience with a religion other than Christianity. This left 199 subjects on which the majority of the analyses were performed.

However, among these 199 subjects, particular variables showed missing data for a variety of reasons. Twenty-three of the 199 subjects did not provide enough information to determine their family's socioeconomic status, leaving 176 subjects on which the majority of the multiple regressions were performed. Six subjects reported religious denominations that were not classifiable under Smith's (1990) system, leaving 193 subjects for hypotheses using denominational classification. Additionally, a number of other variables have varying amounts of missing data. Therefore, the number of subjects used to analyze results varied based on the number with available data.

Based on the 199 subjects the following demographics were calculated. Male subjects numbered 119 (59.8%), female 80 (40.2%). The great majority of subjects were white—180 (90.9%). The remaining subjects reported their race as follows: African Americans—10 (5.1%), Hispanics—4 (2.0%), Asians—2 (1.0%), American Indians—0, and 2 subjects who checked the “other” category. The non-white race categories were combined due to the small number of subjects in each specific category. First year students numbered 101 (50.8%), second year—52 (26.1%), third year—28 (14.1%), fourth year—18 (9.0%), and no one beyond the fifth year or greater. Subject ages
ranged from 18 to 51 with the average age 20.38 years old (SD = 3.94).

Measures

1) Defining Issues Test (DIT, short form) See Appendix A. This group-administered multiple choice test is currently the most frequently used test of moral judgment (Gielen & Lei, 1991). For each of three stories, twelve arguments or queries (reflecting different moral stages) were provided to solve the story-dilemma. Subjects were asked to rate the importance of each query. Furthermore, subjects were required to select and rank the four most important queries.

Rest (1990) cited articles claiming that test-retest correlations for the major indices of P and D scores were "generally in the high .70's or .80's and Cronbach's alpha index of internal consistency is generally in the high .70's" (p. 5.1). For a sample of 33 Australian college students, however, the P and D score test-retest correlations were lower, .71 and .67, respectively, for the six-story form and .67 and .63, respectively, for the three-story form. Rest noted that the test-retest reliabilities of the stage scores were even lower, in the .50's and .60's. He, therefore, recommended exercising much caution in using the stage scores. "I recommend using the stage scores only when the 6-story form has been used, and only when the information is presented in terms of group means..." (Rest, 1990), 5.2). McCrae (1985) attested to the validity of the Defining Issues Test as
conforming to the predictions of its theoretical model and being "at least as valid as other measures of moral development" (p. 442).

2. The Quick Word Test (QWT) See Appendix C21. The Quick Word Test (Borgatta & Corsini, 1960a) is a multiple choice vocabulary test that was used to assess intelligence. The old Quick Word Test correlated with the verbal score on the Wechsler Adult Intelligence Scale (WAIS) in the .80's (Borgatta & Corsini, 1960b) and the total score correlations range from .79 to .84 across various forms. The old Quick Word Test also correlated .42 with the Defining Issues Test (Coder, 1975 as cited in Rest, 1979a). The revised Quick Word Test has two 80 item forms (A and B) which combine the best items of the former four forms and remove items that contained culturally and sexually biased terms. It is expected to have increased reliability and validity over the old form according to E. Borgatta (personal communication, March 8, 1993).22 E. Borgatta stated Cronbach's alpha for Form A was .92 among 12th graders. He added that the raw scores will follow a normal distribution and do not need to be standardized for my purposes (E. Borgatta, personal communication, April 27, 1993). Given the somewhat restricted range of my sample (college subjects taking an Introduction to Psychology course) he felt Cronbach's alpha would decrease slightly to .88 or .89 (E. Borgatta, personal communication, November 11, 1994).

---

21 Borgatta gave permission to print only the first block of five items.

22 Borgatta wanted me to use both forms for the sake of his reliability studies...
3) *Age Universal I/E-Revised Scales (I/E-R)* See Appendix D, items 1-14. Religious orientation was measured using this fourteen item scale scored on a five point Likert continuum from (1) strongly disagree to (5) strongly agree. Higher scores indicate a higher degree of the particular orientation being measured. This scoring is reverse that normally used; but this reverse format matched the other religious scales in this study and so was less likely to confuse subjects and result in accidental reversal by them. The Intrinsic (I) Scale has 8 items (range of 8 to 40) and the Extrinsic (E) Scale has 6 items (range of 6 to 30) with three items each for Personal and Social factors.

Unlike Allport-Ross's Religious Orientation Scale, the Age Universal I/E-Revised Scales' Extrinsic Scale contains no *residual items* (i.e., items that are essentially counter to an intrinsic commitment). With the removal of the residual items, subjects who score high on both Intrinsic and Extrinsic scales can no longer be considered "muddleheaded" or classified as "indiscriminately pro-religious." Therefore, unlike subjects in Ernsberger and Manaster (1981), subjects who strongly agree with both Intrinsic and Extrinsic scales were not removed from this study. Although these scales have fewer items than the original Allport-Ross, they have reliabilities equal to or better than the original version, plus a revised Intrinsic Scale that is partially counterbalanced for acquiescence (Gorsuch & McPherson, 1989). For a sample of 771 college students at both secular and religious colleges the internal consistencies were reported to be I (.82) and E (.66). Finally, the Age Universal I/E-Revised Scales items are
more simply worded than the original Allport-Ross version and applicable across a larger range of the population. Although the Age Universal I/E-Revised Scales have not been used with the Defining Issues Test before and I did not anticipate problems of item difficulty with my sample, using this scale created the opportunity for future research results on subjects with lower reading abilities to be compared with these results.

4) Short Christian Orthodoxy Scale (SCO). See Appendix D, items 15-20. Orthodox Christian beliefs were measured using this six item scale. The scale was designed to be scored on a seven point Likert continuum from (-3) strongly disagree to (+3) strongly agree, with an option to write 0 if subject felt "exactly and precisely neutral about an item" (Hunsberger, 1989, p. 361). However, for the purposes of clarity for the subjects and ease of administration and scoring for the researcher, the Likert continuum using negative numbers was altered so that the Short Christian Orthodoxy Scale was anchored in a similar fashion to all other religious instruments in the study: from (1) strongly disagree to (7) strongly agree. Thus scores can range from 6 to 42, with higher scores indicating more orthodox beliefs.

The Short Christian Orthodoxy Scale was designed to use the best items of the 24-item Christian Orthodoxy (CO) Scale (Fullerton & Hunsberger, 1982) so as to obtain "the best combination of strong psychometric properties without compromising the CO scale's breadth and attention to different facets of religious orthodoxy. The items finally selected addressed issues of the divinity of Christ, inspiration
of the Bible, the concept of God as superstition, forgiveness of sin, God's awareness of human actions and the resurrection" (Hunsberger, 1989, p. 361). The Christian Orthodoxy Scale was based on the Apostles' and Nicene Creeds, and "beliefs which seem to be universally endorsed by orthodox Christian groups even though they are not mentioned in the Creeds: ... the Divine inspiration of the Bible: ... miracles; and ... the efficacy of prayer" (Fullerton & Hunsberger, 1982, p. 318). As in the Christian Orthodoxy Scale, half the items on the Short Christian Orthodoxy Scale were reverse scored which reduced its vulnerability to acquiescence. The short form correlated with the long form .98.

Evidence of validity for the Christian Orthodoxy Scale, reported by Fullerton and Hunsberger (1982), included significant correlations (ranging from .57, N=720 to .77, N=647) between the Christian Orthodoxy Scale and measures of the frequencies of religious service attendance, prayer, Scriptural-devotional reading and the extent of trust in religious guidance from the Bible and the Church. Further evidence of validity was obtained by comparing the mean scores of "apostates" (those who were raised in a Christian religion but no longer identified with any religious denomination) and "switchers" (those who were raised in a particular Christian denomination, but subsequently switched to another Christian denomination) with their matched controls (those of the same sex, year in school, and approximate age who were raised in the same faith as the particular apostate or switcher but who remained in that denomination). "A
highly significant difference emerged between the mean Christian Orthodoxy scores of the apostates (61.8) and their matched controls (115.4), \( t(153) = -10.13, \ p < .001 \). This is to be expected since apostates have by definition severed their formal relationship with Christian religious denominations, while their matched controls have maintained their associations with the same denominations..." (Fullerton & Hunsberger, 1982, p. 324). The mean scores of switchers were somewhat higher than those of their matched controls, however, Fullerton and Hunsberger did not mention significance tests, so I assume the differences were not statistically significant.

According to Hunsberger (1989):

This short version of the Christian Orthodoxy scale retains psychometric properties very close to the original. As expected, the mean inter-item correlation of the six items for all participants (669 successfully completed at least four of the six items, the minimum criterion for inclusion) was .72, and Cronbach’s alpha was .94. All items correlated at more than .76 with the total scale score, and a PA2 factor analysis (SPSSX, 1983) revealed a single large factor with an eigenvalue of 4.61 accounting for 76.9% of the variance, with all items loading higher than .78 on this factor. This next largest factor had an eigenvalue of just .39, so there was no justification for rotation. These properties were also evident for various subgroups. For example, Cronbach’s alpha was .93 for 395 females (M = 33.4), .94 for 262 males (M = 30.2), .93 for 333 respondents for Protestant backgrounds (M = 33.1), and .93 for 191 participants from Catholic backgrounds (M = 34.5).

... scores on the Short Christian Orthodoxy scale correlated with a variety of variables and scales virtually as strongly and significantly as did the original Christian Orthodoxy scale, adding to the impression that the Short...
Christian Orthodoxy scale is valid. As one would expect, religiously orthodox people (as measured by the [Short Christian Orthodoxy scale] also tend to report high levels of authoritarianism (see Altemeyer, 1988), less doubt about religious teachings, more interest in religion, more emphasis on religion in the family background, more agreement with parental religious teachings, higher frequency of church attendance, and stronger religious socialization influences in their childhood. (Hunsberger, 1989, p. 362)

5) Scriptural Literalism Scale (SLS) See Appendix D, items 21-44. This 24 item scale, scored on a five point Likert continuum from (1) strongly disagree to (5) strongly agree, was used to measure "the degree to which an individual believes in a literal, God-related interpretation of scripture" (Hogge & Friedman, 1967, p. 275). Lower scores indicate that an individual "believes that scriptural writings should be interpreted freely, and as ordinary literature not related to a deity" (p. 275). Scores may range from 24 to 120. As evidence of construct validity for the Scriptural Literalism Scale, Hogge and Friedman reported that a more conservative denomination (Baptist) scored higher than a liberal denomination (Unitarian), with a moderate denomination (Methodist) scoring in between. Hogge and Friedman reported the split-half reliability (in this case the correlation between the two 12 item forms) to be .92.

6) A Questionnaire See Appendix D, items 45-62. The questionnaire included subject's age, sex, race, years of education and current GPA, years of education for parents, denominational affiliation (if any) of parents and subject, and subject's recent pattern of
attendance at religious activities, parents' occupations, and parents' educational level. Socioeconomic status was determined using the Socioeconomic Index recommended by Duncan: "MSEI2" (Men's Socioeconomic Index, version 2), a revised measure which is based on educational and income characteristics of the U.S. labor force and coded for the occupation of the head of the household (Stevens & Featherman, 1981). See Appendix E for guidelines used by raters in this study. Inter-rater reliability for Duncan's Socioeconomic Index in this study was quite high ($r = .93$, df = 221, $p < .0001$). This Pearson correlation coefficient was based on 222 subjects because 46 of the original 268 subjects did not provide enough information to determine Socioeconomic Index. Of the final subject sample of 199, we were not able to determine Socioeconomic indices for 23 subjects, leaving 176 subjects. Clouse's (1991) question regarding a personal religious experience was asked. Finally, Subjects were asked if English was their native language and where they were born.

Subject's denominational affiliation was classified under Smith's (1990) system as being Fundamentalist, Moderate or Liberal. While Smith conceded it was difficult to rigorously define these groups, he outlined the main issues separating fundamentalist from liberal theologies and his methods of classification as follows:

At one end [of the Fundamentalist-Liberal continuum] we find the Fundamentalists, a movement of conservative or traditionalist Protestant denominations that grew largely out of the Holiness and Pentecostal movements (and later denominations) of the nineteenth century. The movement was
formed in the early 20th century as a reaction to what was seen as the secularization and modernization of religious beliefs and practices within many mainstream and established Protestant denominations. Its keys beliefs were first articulated in a series of pamphlets called *The Fundamentals* (1909). In addition to their opposition to the growth of secular influence in society, the Fundamentalists are distinguished by belief in 1) the inerrancy of the Bible (or more technically in the verbal, plenary inspiration of the Bible), 2) personal salvation by accepting Christ as their Saviour in what is often called the born-again experience, 3) the personal premillennial imminent return of Christ, 4) an evangelical or revivalist desire to reach out to save and convert others, and 5) acceptance of most traditional Protestant beliefs such as in Trinity, the Virgin birth, and the existence of angels and devils. The position of liberal denominations is perhaps less clear than that of the Fundamentalists, but tends to 1) emphasize concerns about the nature and operation of this world more than salvation in the next which lends to support for social action and progressive reform, 2) accept secular change and science as probably worthwhile and at least not as anti-religious, 3) have little faith in the literal message of the Bible and particularly in Biblical miracles which are seen as either questionable as historical or metaphorical in nature, and 4) be nonadventist. The large group of moderates between the two poles tends of course to reflect varying elements of both the polar groups. They tend for example to reject the extreme inerrancy and anti-science leanings of the Fundamentalists while sharing with them many other traditional beliefs. Likewise, they tend to share liberal acceptance of modernization and some of their leanings toward humanitarian reform, but share less of the deism or even agnosticism that pervades some liberal faiths. . . .

To categorize denominations along Fundamentalist-Liberal continuum, we used five different techniques: 1) utilization of prior classification schemes, 2) membership in theologically oriented ecumenical associations, 3) surveys of denominational members, 4) surveys of denominational clergy, and 5) theological beliefs of denominations. (p. 226)
The self-admitted Agnostics/Atheists in this sample were not "liberal Christians" under Smith's (1990) system, yet in past research on this topic they have been indirectly equated with liberal Christians. For example, when Clouse formed the The Clouse Politics-Religion Attitude Scale (Clouse, 1985) she asked pastors of various churches to state whether a religious conservative, religious liberal, both or neither would agree to each particular item of her scale. For the purposes of developing Clouse's (1985) scale "'liberal' was defined for the ministers as 'nonrestricted' or 'agnostic'" (p. 341). Clouse’s scale has been used in a number of studies reviewed above (Clouse, 1985, 1991; Holley, 1991). In these studies agnostics were not removed from the samples. Since the beliefs of agnostics have been equated with those of liberal Christians on a conservative to liberal continuum, and Agnostics/Atheists have not been removed from samples in the studies reviewed above, it seemed inappropriate to remove them from this sample in which a continuum of Christian beliefs was one of the main independent variables. On the other hand, since Agnostics/Atheists are not Christians and were not considered as part of the sample of liberal Christians on whom the conclusions regarding socioeconomic status and intelligence (two important covariates of this study) were based, it did not seem appropriate to leave them in this sample. What to do? I analyzed the sample both ways, with and without Agnostics/Atheists, and reported results both ways.
Procedure

Subjects completed materials in the following order: 1) Defining Issues Test, 2) Quick Word Test-Form A, 3) Age Universal I/E-Revised Scales (I/E-R), 4) Short Christian Orthodoxy Scale, 5) Scriptural Literalism Scale, 6) Questionnaire, and 7) Quick Word Test-Form B. Packets (face down) and pencils were arranged at tables around the room. Subjects seated themselves. Four seats nearest the door were reserved for those who arrived late. At the scheduled time for the experiment to begin I closed the door to the room and read the "Oral Instructions to Subjects" (Appendix F) while the subjects followed along with the "Directions for Opinions About Social Problems." See Appendix A, first page. Subjects who arrived late were instructed by a sign on the door to enter quietly and pick up a packet. After reading the oral instructions, I asked subjects who arrived on time to begin completing the packets as directed. Those who arrived late (but during the oral instructions) were taken into another room, where I read them the "Oral Instructions to Subjects" from the beginning. I ceased reading at the point where the last subject who had entered the room recognized what was being read. Just before I had lead the late arriving subjects from the experiment room, I placed a "Do Not Enter--Experiment in Progress" sign on the door so that any additional late arriving subjects were turned away. When subjects completed the packets they turned in their packets and pencils and were reminded to pick up a debriefing sheet (see Appendix G). Time to complete the packets ranged from 30 minutes to 60 minutes with
the average subject spending 40-45 minutes. This was less time than expected based on estimates gathered from earlier studies.

**Hypotheses and Data Analysis**

1. Denominational classification will be related to Christian beliefs as measured by the Short Christian Orthodoxy Scale and the Scriptural Literalism Scale. Fundamentalists will have the highest scores, Liberals the lowest (one group of liberals included Agnostics/Atheists and the other group excluded them), and Moderates in between scores. Hypothesis 1 was tested using analysis of variance as described below.

2. Both intelligence, as measured by the Quick Word Test, and socioeconomic status, as measured by the Socioeconomic Index recommended by Duncan: "MSEI2" (Men’s Socioeconomic Index, version 2), will be related to denominational classification with Fundamentalists having the lowest intelligence and socioeconomic status, Liberals having the highest intelligence and socioeconomic status, and Moderates in between. Hypothesis 2 was tested using analysis of variance as described below.

Because the number of subjects in each denominational classification was unequal the data were unbalanced. Hypotheses 1 and 2 were tested using SAS’s PROC GLM analysis of variance procedure which is appropriate for unbalanced data. If the overall $F$ was significant then all pairwise differences were tested using the
LSMEANS statement which uses least-squares means. "Least-squares estimates of marginal means (LSMs) are to unbalanced designs as class ... arithmetic means are to balanced designs. LSMs are simply estimators of the class ... marginal means that would be expected had the design been balanced" (SAS Institute Inc., 1985, p. 444). If the overall $F$ was not significant then no additional tests were conducted.

3. Intelligence, as measured by the Quick Word Test, socioeconomic status, as measured by the Socioeconomic Index recommended by Duncan: "MSEI2" (Men's Socioeconomic Index, version 2), and year in school will be significantly positively correlated with moral judgment as measured by the Defining Issues Test P score. Hypothesis 3 was examined using Zero-order Pearson Correlation.

4. Higher scores (more conservative) on the Short Christian Orthodoxy Scale will correlate negatively with P score on the Defining Issues Test. Hypothesis 4 was examined using Zero-order Pearson Correlation.

5. Higher scores (more conservative) on the Scriptural Literalism Scale will correlate negatively with P score on the Defining Issues Test. Hypothesis 5 was examined using Zero-order Pearson Correlation.

6. There will be a significant interaction between Intrinsic Scale scores and the Short Christian Orthodoxy Scale scores in predicting P score. As intrinsic religiousness *increases*, those who score higher on the Short Christian Orthodoxy Scale will score lower on P score, while
those who score lower on the Short Christian Orthodoxy Scale will score higher on the P score. However, this will not apply to those who do not consider themselves to be of the Christian faith (i.e. agnostics, atheists, Hindus, Muslims, Jews, etc.) As intrinsic religiousness decreases, the differences noted on the Short Christian Orthodoxy Scale as a predictor of P score will decline. For clarity see Figure 1. Hypothesis 6 was examined using a regression analysis described below.

7. There will be a significant interaction between Intrinsic Scale scores and the Scriptural Literalism Scale scores in predicting P score. As intrinsic religiousness increases, those who score higher on the Scriptural Literalism Scale will score lower on P score, while those who score lower on the Scriptural Literalism Scale will score higher on the P score. However, this will not apply to those who do not consider themselves to be of the Christian faith (i.e. agnostics, atheists, Hindus, Muslims, Jews, etc.) As intrinsic religiousness decreases, the differences noted on the Scriptural Literalism Scale as a predictor of P score will decline. For clarity see Figure 1. Hypothesis 7 was examined using a regression analysis described below.
Although Figure 1 divides the Intrinsic into two groups (High Intrinsics and Low Intrinsics) this is only done for clarity. To actually dichotomize the many points of the Intrinsic Scale into two points so that the traditional Analysis of Variance with its interaction component could be performed would be to unnecessarily throw away data and decrease the power of the statistical tests. As Cohen and Cohen (1983) put it:

"It is intuitively obvious that when one reduces a graduated many-valued scale to a two-point scale, one is willfully throwing away information. This has immediate negative consequences to the amount of variance such a crippled variable can account for and concomitantly to the
power of the statistical test of its contribution. Now, the MRC [Multiple Regression/Correlation] gives the same results as the (exact) AV [Analysis of Variance], once the IVs have been dichotomized, and there is the rub! It would be dismal enough a prospect if it were necessary to dichotomize in order to analyze the data, but of course no such necessity exists. . . . the general model for interaction holds that for quantitative (graduated) variables, as indeed for variables or variable sets of any kind, the interaction $u \times v$ represented by $uvu,v$, and its effects are readily determined by the MRC model." (pp.309-310)

Hypotheses 6 and 7 were analyzed by an "intermediate" regression analysis as recommended by Cohen and Cohen (1983, pp. 148-149). Intermediate analysis is neither fully simultaneous, nor fully hierarchical, but uses the best aspects of both in a flexible manner. For hypothesis 6 the set of variables containing the Short Christian Orthodoxy Scale ($u$) and the Intrinsic subscale ($v$) were simultaneously analyzed followed by an analysis of this same ($u,v$) set plus the appropriate interaction variable ($u \times v$), which was also simultaneously analyzed. The increase in $R^2$, also called the squared semipartial correlation coefficient ($sr^2$) for $uvu,v$ is the variance accounted for by the interaction once the main effects have been partialled out. Hypothesis 7 was analyzed in the same way as hypothesis 6, except substituting in the Scriptural Literalism Scale for the Short Christian Orthodoxy Scale. For all regressions Type II sums of squares and Type II partial correlations were obtained. Thus each effect (e.g., variable, interaction) in the model has been adjusted for all other effects at that stage of the intermediate regression. This is the
same as if each effect were entered last in the model and Type I sums of squares and partial correlations were obtained for each effect in the model at that stage.

As with the analyses of variance above, in carrying out all the analyses reported in this study, I followed the general protected $t$ test procedure recommended by Cohen and Cohen (1983 pp. 172-176, 390) for its combined "simplicity and effectiveness in balancing Type I and Type II errors in inference." Thus, additional analyses within a set of variables were conducted only if the overall $F$ test for the set was significant.

8. Once socioeconomic status, year in school and intelligence are partialled out, higher scores (more conservative) on the Short Christian Orthodoxy Scale will still correlate negatively with $P$ score on the Defining Issues Test, although to a lesser degree. Hypothesis 8 was tested with a regression analysis described below.

9. Once socioeconomic status, year in school and intelligence are partialled out, higher scores (more conservative) on the Scriptural Literalism Scale will still correlate negatively with $P$ score on the Defining Issues Test, although to a lesser degree. Hypothesis 9 was tested with a regression analysis described below.

Hypotheses 8 and 9 were tested using an "intermediate" regression analysis with Type II sums of squares and partial correlations (as defined above under hypotheses 6 and 7) with the covariates of socioeconomic status, year in school, and intelligence
entered first as a set, followed by the relevant religious belief variable.

10. There will be a significant interaction between Intrinsic Scale scores and the Short Christian Orthodoxy Scale scores in predicting P score. As intrinsic religiousness increases, those who score higher on the Short Christian Orthodoxy Scale will score lower on P score, while those who score lower on the Short Christian Orthodoxy Scale will score higher on the P score. However this will not apply to those who do not consider themselves to be of the Christian faith (i.e. agnostics, atheists, Hindus, Muslims, Jews, etc.) As intrinsic religiousness decreases, the differences noted on Short Christian Orthodoxy Scale as a predictor of P score will decline. See Figure 1 above for clarity. Once socioeconomic status, year in school and intelligence are partialled out, the above relationships will still hold, although to a lesser degree. Hypothesis 10 was tested with a regression analysis described below.

11. There will be a significant interaction between Intrinsic Scale scores and the Scriptural Literalism Scale scores in predicting P score. As intrinsic religiousness increases, those who score higher on the the Scriptural Literalism Scale will score lower on P score, while those who score lower on the Scriptural Literalism Scale will score higher on the P score. However this will not apply to those who do not consider themselves to be of the Christian faith (i.e. agnostics, atheists, Hindus, Muslims, Jews, etc.) As intrinsic religiousness decreases, the differences noted on the Scriptural Literalism Scale as a
predictor of P score will decline. See Figure 1 above for clarity. Once socioeconomic status, year in school and intelligence are partialled out, the above relationships will still hold, although to a lesser degree. Hypothesis 11 was tested with a regression analysis described below.

Hypotheses 10 and 11 were tested using an "intermediate" regression analysis with the covariates of socioeconomic status, year in school, intelligence, and the main effects of the Intrinsic Scale and the relevant religious belief scale entered first as a set, followed by the interaction variable of the Intrinsic Scale with the relevant religious belief scale (as in hypotheses 6 and 7).

12. A meaningful Christian personal religious experience - as measured by Clouse's (1991) technique - will be significantly negatively associated with Stage 3 and positively associated with Stage 4 on the Defining Issues Test. Hypothesis 12 was tested in two ways. First, according to Rest's suggested method of finding the mean stage percentages (Stages 3 & 4) for those who have had vs. those who have not had a Christian religious experience that is currently meaningful to them. Mean differences were tested for significance using t tests. Second, in order to exactly replicate Clouse's (1991) reported result that in a stepwise regression GPA was selected first and a meaningful Christian personal religious experience second. A multiple regression using Type I sums of squares and partial correlations was performed with GPA entered first, followed by meaningful Christian personal religious experience, one regression with Stage 3 as the dependent variable and one with Stage 4 as the
dependent variable. Based on Clouse (1991) Stage 3 was predicted to correlate negatively with GPA and negatively with religious experience, while Stage 4 was predicted to correlate negatively with GPA and positively with Stage 4.

Additional post-hoc analyses were performed and will be reported in chapter 4, Results and Discussion.
Chapter IV
Results and Discussion

Results and discussion are combined in this chapter. First descriptive statistics are presented and discussed. Then hypotheses are presented and results discussed in numerical order. Relevant post hoc analyses are presented in conjunction with their related hypotheses. Interpretations, theoretical musings and suggestions for future research are incorporated.

Descriptive statistics are found in Table 4. Please note the low Cronbach's coefficient alpha (.40) for the P score. This was well below the .76 Cronbach's coefficient alpha reported by Rest (1990) for the three-story Defining Issues Test. This relatively low reliability seriously limits P score's ability to correlate with other variables. "Generally, the maximum possible correlation between the test and some external criterion is equal to the square root of the reliability coefficient" (Walsh & Betz, 1990, p.57). The square root of .40 is .63. Thus, the maximum correlation possible with another variable is .63, down from the expected maximum possible correlation of .87 (the square root of .76). This may explain a portion of the lower than expected correlations between P score and other variables.
Cronbach's coefficient alpha for the P score was calculated by following Rest's (1990) method of "finding a stage score for each story, then looking at the consistency across all stories on that score" (p. 5.3). The three stage scores (5A, 5B, and 6) that go into the P score were added together so that the P score for each of the three stories (Heinz, Prisoner and Newspaper) could be correlated using Cronbach's coefficient alpha in the same manner as for three individual items on any other three item test. No coefficient alphas were reported in Rest (1990) for any of the stages. Indeed, for many of the stages I don't know how one could determine a coefficient alpha because some of the stories do not have any items for a given stage. If the number of items were merely different, one could take the average instead of the sum of the scores of the items. Fortunately, this was not a problem on P score because it had an equal number of items on every story.

Denominational classification frequency data are summarized in Table 5 with specific denominations' frequency data in Appendix H. Other relevant subject frequency data may be found above in the Method section under Subjects.
Table 4
Descriptive Statistics for Study Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defining Issues Test</td>
<td></td>
<td></td>
<td></td>
<td>Alpha</td>
</tr>
<tr>
<td>Stage 2 with Agnostics/Ath.</td>
<td>199</td>
<td>7.91</td>
<td>7.05</td>
<td></td>
</tr>
<tr>
<td>Stage 2 without Agnostics/Ath.</td>
<td>159</td>
<td>7.82</td>
<td>6.97</td>
<td></td>
</tr>
<tr>
<td>Stage 3 with Agnostics/Ath.</td>
<td>199</td>
<td>20.00</td>
<td>11.95</td>
<td></td>
</tr>
<tr>
<td>Stage 3 without Agnostics/Ath.</td>
<td>159</td>
<td>19.54</td>
<td>11.45</td>
<td></td>
</tr>
<tr>
<td>Stage 4 with Agnostics/Ath.</td>
<td>199</td>
<td>34.20</td>
<td>14.63</td>
<td></td>
</tr>
<tr>
<td>Stage 4 without Agnostics/Ath.</td>
<td>159</td>
<td>34.34</td>
<td>14.28</td>
<td></td>
</tr>
<tr>
<td>Stage 5A with Agnostics/Ath.</td>
<td>199</td>
<td>24.84</td>
<td>14.63</td>
<td></td>
</tr>
<tr>
<td>Stage 5A without Agn./Ath.</td>
<td>159</td>
<td>24.42</td>
<td>15.05</td>
<td></td>
</tr>
<tr>
<td>Stage 5B with Agnostics/Ath.</td>
<td>199</td>
<td>2.86</td>
<td>4.83</td>
<td></td>
</tr>
<tr>
<td>Stage 5B without Agnostics/Ath.</td>
<td>159</td>
<td>3.04</td>
<td>4.82</td>
<td></td>
</tr>
<tr>
<td>Stage 6 with Agnostics/Ath.</td>
<td>199</td>
<td>3.32</td>
<td>4.73</td>
<td></td>
</tr>
<tr>
<td>Stage 6 without Agnostics/Ath.</td>
<td>159</td>
<td>3.33</td>
<td>4.82</td>
<td></td>
</tr>
<tr>
<td>P score with Agnostics/Ath.</td>
<td>199</td>
<td>30.98</td>
<td>16.63</td>
<td>see note b</td>
</tr>
<tr>
<td>P score without Agn./Ath.</td>
<td>159</td>
<td>30.79</td>
<td>17.24</td>
<td>see note b</td>
</tr>
<tr>
<td>P score by groups (see note c)</td>
<td></td>
<td></td>
<td></td>
<td>Alpha</td>
</tr>
<tr>
<td>Fundamentalists</td>
<td>26</td>
<td>27.69</td>
<td>16.81</td>
<td></td>
</tr>
<tr>
<td>Moderates</td>
<td>90</td>
<td>32.52</td>
<td>17.73</td>
<td></td>
</tr>
<tr>
<td>Liberals</td>
<td>37</td>
<td>30.61</td>
<td>15.70</td>
<td></td>
</tr>
<tr>
<td>Liberals Plus Agn./Ath.</td>
<td>77</td>
<td>31.20</td>
<td>14.80</td>
<td></td>
</tr>
</tbody>
</table>
Table 4 (continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQ (based on Quick Word Test scores) with Agn./Ath.</td>
<td>199</td>
<td>53.80</td>
<td>11.13</td>
<td>see note d</td>
</tr>
<tr>
<td>IQ without Agnostics/Ath</td>
<td>159</td>
<td>52.83</td>
<td>10.77</td>
<td>see note d</td>
</tr>
<tr>
<td>IQ by groups:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fundamentalists</td>
<td>26</td>
<td>52.92</td>
<td>10.99</td>
<td></td>
</tr>
<tr>
<td>Moderates</td>
<td>90</td>
<td>53.47</td>
<td>10.83</td>
<td></td>
</tr>
<tr>
<td>Liberals</td>
<td>37</td>
<td>52.70</td>
<td>10.13</td>
<td></td>
</tr>
<tr>
<td>Liberals Plus Agn./Ath</td>
<td>77</td>
<td>55.26</td>
<td>11.24</td>
<td></td>
</tr>
<tr>
<td>Intrinsic/Extrinsic-Revised (IER)e with Agn./Ath.</td>
<td>199</td>
<td>2.88</td>
<td>0.60</td>
<td>.81</td>
</tr>
<tr>
<td>Intrinsic/Extrinsic-Revised (IER)e without Agn./Ath.</td>
<td>159</td>
<td>3.04</td>
<td>0.51</td>
<td>.76</td>
</tr>
<tr>
<td>Intrinsic (I)e with Agn./Ath</td>
<td>199</td>
<td>3.06</td>
<td>0.72</td>
<td>.79</td>
</tr>
<tr>
<td>Intrinsic (I)e without Agn/Ath</td>
<td>159</td>
<td>3.24</td>
<td>0.67</td>
<td>.80</td>
</tr>
<tr>
<td>Extrinsic (E)e with Agn./Ath</td>
<td>199</td>
<td>2.55</td>
<td>0.72</td>
<td>.71</td>
</tr>
<tr>
<td>Extrinsic (E)e without Agn/Ath</td>
<td>159</td>
<td>2.70</td>
<td>0.63</td>
<td>.62</td>
</tr>
<tr>
<td>Extrinsic Social Subscale (ES)e with Agn./Ath</td>
<td>199</td>
<td>2.15</td>
<td>0.83</td>
<td>.79</td>
</tr>
<tr>
<td>Extrinsic Social Subscale (ES)e without Agn./Ath</td>
<td>159</td>
<td>2.25</td>
<td>.82</td>
<td>.77</td>
</tr>
</tbody>
</table>
Table 4 (continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extrinsic Personal Subscale</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(EP) with Agn./Ath.</td>
<td>199</td>
<td>3.16</td>
<td>1.00</td>
<td>.68</td>
</tr>
<tr>
<td>Extrinsic Personal Subscale</td>
<td>159</td>
<td>3.39</td>
<td>1.06</td>
<td>.50</td>
</tr>
<tr>
<td>Short Christian Orthodoxy Scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(SCO) with Agn./Ath.</td>
<td>199</td>
<td>5.67</td>
<td>1.44</td>
<td>.91</td>
</tr>
<tr>
<td>Short Christian Orthodoxy Scale</td>
<td>159</td>
<td>6.10</td>
<td>1.06</td>
<td>.86</td>
</tr>
<tr>
<td>SCO by groups:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fundamentalists</td>
<td>26</td>
<td>6.44</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>Moderates</td>
<td>90</td>
<td>6.09</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>Liberals</td>
<td>37</td>
<td>5.90</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>Liberals Plus Agn./Ath.</td>
<td>77</td>
<td>4.89</td>
<td>1.65</td>
<td></td>
</tr>
<tr>
<td>Scriptural Literalism Scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(SLS) with Agn./Ath.</td>
<td>199</td>
<td>3.28</td>
<td>0.77</td>
<td>.96</td>
</tr>
<tr>
<td>Scriptural Literalism Scale</td>
<td>159</td>
<td>3.48</td>
<td>0.67</td>
<td>.95</td>
</tr>
<tr>
<td>SLS by groups:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fundamentalists</td>
<td>26</td>
<td>3.93</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>Moderates</td>
<td>90</td>
<td>3.42</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>Liberals</td>
<td>37</td>
<td>3.21</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Liberals Plus Agn./Ath.</td>
<td>77</td>
<td>2.86</td>
<td>0.75</td>
<td></td>
</tr>
</tbody>
</table>
Table 4 (continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Coefficient</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year in School (YEAR) with Agn./Ath.</td>
<td>199</td>
<td>1.81</td>
<td>0.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year in School (YEAR) without Agn./Ath.</td>
<td>159</td>
<td>1.72</td>
<td>0.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socioeconomic Status (SES) with Agn./Ath.</td>
<td>176</td>
<td>50.14</td>
<td>20.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socioeconomic Status (SES) without Agn./Ath.</td>
<td>142</td>
<td>49.73</td>
<td>20.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES by groups:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fundamentalists</td>
<td>25</td>
<td>37.60</td>
<td>20.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderates</td>
<td>79</td>
<td>52.45</td>
<td>19.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liberals</td>
<td>32</td>
<td>53.55</td>
<td>19.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liberals Plus Agn./Ath.</td>
<td>66</td>
<td>52.68</td>
<td>19.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPA with Agn./Ath.</td>
<td>196</td>
<td>2.73</td>
<td>0.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPA without Agn./Ath.</td>
<td>156</td>
<td>2.74</td>
<td>0.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age with Agn./Ath.</td>
<td>199</td>
<td>20.38</td>
<td>3.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age without Agn./Ath.</td>
<td>159</td>
<td>20.09</td>
<td>3.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father’s Education (EDDAD) with Agn./Ath.</td>
<td>199</td>
<td>3.79</td>
<td>0.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father’s Education (EDDAD) without Agn./Ath.</td>
<td>159</td>
<td>3.84</td>
<td>0.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>N</td>
<td>M</td>
<td>SD</td>
<td>Alpha</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>-----</td>
<td>-------</td>
<td>-----</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>Mother’s Education (EDMOM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with Agn./Ath.</td>
<td>199</td>
<td>3.52</td>
<td>0.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother’s Education (EDMOM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>without Agn./Ath.</td>
<td>159</td>
<td>3.57</td>
<td>0.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yearly Attendance at Religious Activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with Agn./Ath.</td>
<td>196</td>
<td>26.92</td>
<td>38</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>Yearly Attendance at Religious Activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>without Agn./Ath.</td>
<td>156</td>
<td>33.31</td>
<td>41</td>
<td>78</td>
<td></td>
</tr>
</tbody>
</table>

*a* Cronbach’s coefficient alphas were based on raw variables.

*b* Cronbach’s coefficient alpha for P score was .41 based on N=200 subjects for which computer scoring program was able to obtain scores. Subjects who did not pass Rest’s reliability tests had already been removed. Additionally, some protocols that needed to be hand scored were not included due to missing data which the computer program could not handle properly.

*c* Mean P score differences among denominational classification groups were not statistically significant.

*d* It was impractical to obtain Cronbach’s coefficient alpha for the Quick Word Test.

*e* In this study the Age Universal I/E-Revised Scales were reverse scored compared to the usual scoring of these scales. High scores are more intrinsic.
Table 5
Denominational Classification Frequency Table based on Smith (1990)

<table>
<thead>
<tr>
<th>Denominational Classification</th>
<th>Freq</th>
<th>Percent</th>
<th>Cum. Freq</th>
<th>Cum. Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fundamentalists</td>
<td>26</td>
<td>13.1</td>
<td>26</td>
<td>13.1</td>
</tr>
<tr>
<td>Moderates</td>
<td>90</td>
<td>45.2</td>
<td>116</td>
<td>58.3</td>
</tr>
<tr>
<td>Liberals</td>
<td>37</td>
<td>18.6</td>
<td>153</td>
<td>76.9</td>
</tr>
<tr>
<td>Agnostics/Atheists</td>
<td>40</td>
<td>20.1</td>
<td>193</td>
<td>97.0</td>
</tr>
<tr>
<td>Not classified in Smith (1990)</td>
<td>6</td>
<td>3.0</td>
<td>199</td>
<td>100.0</td>
</tr>
</tbody>
</table>

**Fundamentalists:**
- Apostolic Christian (2)
- Baptist, Don't know which (12)
- Bible Teaching, Any Christian Church (1)
- Church of Christ (4)
- Evangelical, Evangelist (1)
- Lutheran Church, Missouri Synod (1)
- Mennonite (1)
- Pentecostal (1)
- Southern Baptist Convention (1)
- The Way Ministry (1)
- Wesleyan (1)

**Moderates:**
- African Methodist Episcopal Church (1)
- Catholic (65)
- Evangelical Lutheran (1)
- Lutheran, Don't know which (11)
- No Denomination/Non-Denominational (12)

**Liberals:**
- Episcopal (4)
- Presbyterian, Don't know which (8)
- Quaker (1)
- United Church of Christ (6)
- United Methodist Church (16)

**Not classified under Smith's (1990) system:**
- Coptic Orthodox (2)
- Greek Orthodox (1)
- Protestant (3)
Hypothesis 1

Denominational classification will be related to Christian beliefs as measured by the Short Christian Orthodoxy Scale and the Scriptural Literalism Scale. Fundamentalists will have the highest scores, Liberals the lowest (one group of Liberals included Agnostics/Atheists and the other group excluded them), and Moderates in between scores.

Hypothesis 1 received support for the Short Christian Orthodoxy Scale only when Agnostics/Atheists were included in the Liberal denominational classification. When the Liberal group did not include the Agnostics/Atheists, the overall $F$ test (Table 6) showed that there were no significant differences among the means of the various denominational classifications on the Short Christian Orthodoxy Scale, $F(2, 150)=2.06$, $p=.13$. As seen in Table 7, when the denominational classification of Liberals included Agnostics/Atheists there was a significant difference, $F(2, 190)=23.24$, $p<.001$, among the mean scores on the Short Christian Orthodoxy Scale. As evidenced in Table 8, the only group that differed significantly from the others was the Liberal denominational classification that included Agnostics/Atheists.
Table 6
Analysis of Variance Summary Table for Denominational Classification (Excluding Agnostics/Atheists) and Short Christian Orthodoxy Scale

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denominational Classification</td>
<td>4.52</td>
<td>2</td>
<td>2.26</td>
<td>2.06</td>
</tr>
<tr>
<td>Error</td>
<td>164.31</td>
<td>150</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>168.83</td>
<td>152</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7
Analysis of Variance Summary Table for Denominational Classification (Including Agnostics/Atheists) and Short Christian Orthodoxy Scale

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denominational Classification</td>
<td>78.42</td>
<td>2</td>
<td>39.21</td>
<td>23.24***</td>
</tr>
<tr>
<td>Error</td>
<td>320.59</td>
<td>190</td>
<td>1.69</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>399.01</td>
<td>192</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

***$p < .001$
Table 8

Least Squares Means of Short Christian Orthodoxy Scale for
Fundamentalists, Moderates, and Liberals Plus Agnostics/Atheists

| Denominational Classification | SCO LSMEAN | Std Err LSMEANa | Pr > |T| H0: LSMEAN(i) - LSMEAN(j) |
|-----------------------------|------------|----------------|-----------------|-----------------|
|                             | (1)        | (2)            | (3)             |                 |
| (1) Fundamentalists         | 6.44       | 0.25           | --              |                 |
| (2) Moderates               | 6.09       | 0.14           | .2232           | --              |
| (3) Liberals and Agnostics/Atheists | 4.89   | 0.15           | .0001           | .0001           | --              |

SCO = Short Christian Orthodoxy Scale

aStandard Errors are estimated based on a pooled error term and the sample size of the group contributing to each mean. See Table 4 for actual standard deviations.

This failure of the Short Christian Orthodoxy Scale to distinguish between Fundamentalists, Moderates, and Liberals could be occurring for a number of reasons. There is a good degree of variability among particular churches within some denominations. Perhaps a number of subjects belonged to churches that did not fit the general pattern of their respective denomination as classified in Smith (1990). Or, the anomaly could be with the individual subjects in this study. As noted above, Glock and Stark (1965) reported a pattern in which a minority of members of liberal congregations hold very conservative beliefs.

While the Short Christian Orthodoxy Scale mean scores did not differ significantly across denominational classifications of Christians, the Scriptural Literalism Scale mean scores did differ significantly, in partial support of Hypothesis 1; see Table 9, $F(2, 150)=8.81, p<.001$. 
As seen in Table 10, the mean scores of the Moderates and Liberals did not differ significantly, but the mean score of the Fundamentalists was significantly higher than both Moderates and Liberals. Given the fact that one of the distinguishing characteristics of Fundamentalists is their literal view of the Bible (Ammerman, 1982; Kellstedt & Smidt, 1991; Smith, 1990), it makes sense that they would score significantly higher than the Moderates and Liberals on this particular measure.

When Agnostics/Atheists were added to the Liberal group the overall $F$ test continued to be significant, Table 11, $F(2, 190)=28.34$, $p<.001$, with significant differences among all three groups (see Table 12). Thus, in a pattern consistent with their lower scores on the Short Christian Orthodoxy Scale, the Liberal group that included the Agnostics/Atheists also scored significantly lower on the Scriptural Literalism Scale.
Table 9
Analysis of Variance Summary Table for Denominational Classification (Excluding Agnostics/Atheists) and Scriptural Literalism Scale

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denominational</td>
<td>7.09</td>
<td>2</td>
<td>3.54</td>
<td>8.81***</td>
</tr>
<tr>
<td>Classification</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>60.31</td>
<td>150</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>67.39</td>
<td>152</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

***p < .001

Table 10
Least Squares Means of Scriptural Literalism Scale for Fundamentalists, Moderates, and Liberals

<table>
<thead>
<tr>
<th>Denominational Classification</th>
<th>SLS</th>
<th>Std Err</th>
<th>Pr &gt;</th>
<th>H0: LSMEAN(i) = LSMEAN(j)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSMEAN</td>
<td>LSMEANa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i/j</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>(1) Fundamentalists</td>
<td>3.93</td>
<td>0.12</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>(2) Moderates</td>
<td>3.42</td>
<td>0.07</td>
<td>.0004</td>
<td>--</td>
</tr>
<tr>
<td>(3) Liberals</td>
<td>3.29</td>
<td>0.10</td>
<td>.0001</td>
<td>.2915</td>
</tr>
</tbody>
</table>

SLS = Scriptural Literalism Scale

*aStandard Errors are estimated based on a pooled error term and the sample size of the group contributing to each mean. See Table 4 for standard deviations.
### Table 11

**Analysis of Variance Summary Table for Denominational Classification (Including Agnostics/Atheists as Liberals) and Scriptural Literalism Scale**

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denominational Classification</td>
<td>26.26</td>
<td>2</td>
<td>13.19</td>
<td>28.34***</td>
</tr>
<tr>
<td>Error</td>
<td>88.00</td>
<td>190</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>114.26</td>
<td>192</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

***p < .001

### Table 12

**Least Squares Means of Scriptural Literalism Scale for Fundamentalists, Moderates, and Liberals Plus Agnostics/Atheists**

<table>
<thead>
<tr>
<th>Denominational Classification</th>
<th>SLS Std Err</th>
<th>Pr &gt;</th>
<th>HO: LSMEAN(i) - LSMEAN(j)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSMEAN</td>
<td>[IT]</td>
<td></td>
</tr>
<tr>
<td>i/j</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) Fundamentalists</td>
<td>3.93</td>
<td>0.13</td>
<td>--</td>
</tr>
<tr>
<td>(2) Moderates</td>
<td>3.42</td>
<td>0.07</td>
<td>.0008</td>
</tr>
<tr>
<td>(3) Liberals and Agnostics/Atheists</td>
<td>2.86</td>
<td>0.08</td>
<td>.0001 .0001 --</td>
</tr>
</tbody>
</table>

SLS = Scriptural Literalism Scale;
aStandard Errors are estimated based on a pooled error term and the sample size of the group contributing to each mean. See Table 4 for standard deviations.
Hypothesis 2

Both intelligence, as measured by the Quick Word Test, and socioeconomic status, as measured by the Socioeconomic Index recommended by Duncan: "MSEI2" (Men's Socioeconomic Index, version 2), will be related to denominational classification with Fundamentalists having the lowest intelligence and socioeconomic status, Liberals having the highest intelligence and socioeconomic status, and Moderates in between.

Hypothesis 2 was not supported for differences in IQ among denominational classifications. There were no significant differences among denominational classifications on the Quick Word Test mean scores either when the Liberals were considered excluding Agnostics/Atheists, Table 13, \( F(2, 150) = 0.08, p = .93 \), or when Agnostics/Atheists were combined with Liberals, Table 14, \( F(2, 190) = 0.72, p = .49 \).

This failure to replicate past findings of a positive correlation between intelligence and liberal denominational affiliation may be due to the restricted nature of this sample of college students who were probably disproportionately higher in IQ than the general population of Fundamentalists, Moderates and Liberals. While Rhodes and Nam (1970) reported a significant .17 correlation \((p < .01)\) between subject's IQ and liberal denominational affiliation of subject's mother, their sample did not include college students but rather white children ages 14-19 enrolled in elementary or secondary public or private schools. Their sample's data base was gathered through the 1965 U.S. Census Bureau's Current Population Survey and so is broader than the
college sample in this study. Similarly, Argyle and Beit-Hallahmi (1975) made no mention of samples being limited to college students in the studies they alluded to in support of their conclusion that higher intelligence is related to more liberal denominational affiliation. Had my sample included a non-college group of subjects with a broader range of IQs perhaps it would have replicated past studies that indicated a positive correlation between intelligence and liberal denominational affiliation.

Table 13

*Analysis of Variance Summary Table for Denominational Classification (Excluding Agnostics and Atheists) and Quick Word Test*

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denominational Classification</td>
<td>17.53</td>
<td>2</td>
<td>8.76</td>
<td>0.08</td>
</tr>
<tr>
<td>Error</td>
<td>17153.98</td>
<td>150</td>
<td>114.36</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>17171.51</td>
<td>152</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 14  
Analysis of Variance Summary Table for Denominational Classification (Including Agnostics and Atheists as Liberals) and Quick Word Test

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denominational Classification</td>
<td>175.66</td>
<td>2</td>
<td>87.38</td>
<td>0.72</td>
</tr>
<tr>
<td>Error</td>
<td>23059.05</td>
<td>190</td>
<td>121.36</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23234.71</td>
<td>192</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Hypothesis 2 was partially supported for differences in socioeconomic status among denominational classifications. There were significant differences in mean Duncan Socioeconomic Index scores across denominational classifications, both when the Liberals group excluded Agnostics/Atheists, Table 15, $F(2, 133)=5.98$, $p=.003$, and when Agnostics/Atheists were added to the Liberals group, Table 16, $F(2, 167)=6.05$, $p=.003$. In both cases, the Fundamentalists had significantly lower Duncan Socioeconomic Index scores than the other two groups while the other two groups did not differ significantly from each other, see Tables 17 and 18.
Table 15
Analysis of Variance Summary Table for Denominational Classification (Excluding Agnostics and Atheists) and Socioeconomic Index

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denominational Classification</td>
<td>4723.61</td>
<td>2</td>
<td>2361.30</td>
<td>5.98**</td>
</tr>
<tr>
<td>Error</td>
<td>52514.77</td>
<td>133</td>
<td>394.85</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>57238.38</td>
<td>135</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**p < .01

Table 16
Analysis of Variance Summary Table for Denominational Classification (Including Agnostics and Atheists as Liberals) and Socioeconomic Index

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denominational Classification</td>
<td>4773.39</td>
<td>2</td>
<td>2386.69</td>
<td>6.05**</td>
</tr>
<tr>
<td>Error</td>
<td>65844.89</td>
<td>167</td>
<td>394.28</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>70618.28</td>
<td>169</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**p < .01
Table 17

*Least Squares Means of Socioeconomic Status for Fundamentalists, Moderates, and Liberals (Excluding Agnostics/Atheists)*

| Denominational Classification | SES LSMEAN | Std Err LSMEAN<sup>a</sup> | Pr > |T| H0: LSMEAN(i) - LSMEAN(j) |
|------------------------------|------------|----------------------------|------|-----------------------------|
| i/j                          |            |                            |      |                             |
| (1) Fundamentalists         | 37.60      | 3.97                       |      |                             |
| (2) Moderates               | 52.45      | 2.24                       | .0014|                             |
| (3) Liberals                | 53.55      | 3.51                       | .0031| .7915                        |

SES = Socioeconomic Status;
<sup>a</sup>Standard Errors are estimated based on a pooled error term and the sample size of the group contributing to each mean. See Table 4 for standard deviations.

Table 18

*Least Squares Means of Socioeconomic Status for Fundamentalists, Moderates, and Liberals Plus Agnostics/Atheists*

| Denominational Classification | SES LSMEAN | Std Err LSMEAN<sup>a</sup> | Pr > |T| H0: LSMEAN(i) - LSMEAN(j) |
|------------------------------|------------|----------------------------|------|-----------------------------|
| i/j                          |            |                            |      |                             |
| (1) Fundamentalists         | 37.60      | 3.97                       |      |                             |
| (2) Moderates               | 52.45      | 2.23                       | .0014|                             |
| (3) Liberals Plus           | 52.68      | 2.44                       | .0015| .9438                        |

SES = Socioeconomic Status;
<sup>a</sup>Standard Errors are estimated based on a pooled error term and the sample size of the group contributing to each mean. See Table 4 for standard deviations.
This finding is consistent with past reports of significant differences between Liberals and Fundamentalists; however, Moderates did not follow the pattern I expected. Catholics made up 72% of the Moderates (see Table 5). In past research Catholics were reported to have lower socioeconomic status than more liberal denominations (Argyle and Beit-Hallahmi, 1975; Spilka, Hood, and Gorsuch, 1985). Perhaps Moderates have actually improved their socioeconomic status over the last decade relative to Fundamentalists and Liberals. Or, as before with intelligence, one might speculate that part of the reason why Moderates were not significantly lower in socioeconomic status than Liberals is that this sample was restricted to college students. The families of college Moderates may have a higher socioeconomic status than Moderates in general, many of whom perhaps could not afford to send their children to college.

Of course one might ask why this argument regarding Moderates does not apply to Fundamentalists as well. How can it be logical for me to argue that Moderates in this sample come from a higher socioeconomic status than Moderates in general, but not apply this same reasoning to Fundamentalists? My response is that perhaps, consistent their significantly lower socioeconomic status, Fundamentalists in this sample depended more on grants, scholarships, and loans than did Moderates. Perhaps fundamentalist students need to be a bit more independent from their parents, in terms of financial and social support, in order to attend college. In terms of financial support, given their lower socioeconomic status
(Argyle, 1975; Chalfant, Beckley, & Palmer, 1981; Spilka, Hood, and Gorsuch, 1985), one might speculate that fundamentalist students are more likely to receive less financial help from their parents, and to be eligible for, and motivated to seek, more need-based financial aid. In terms of social support, Rhodes and Nam (1970) reported that elementary and secondary school children of more fundamentalist mothers had lower aspirations for higher education, even after IQ and socioeconomic status were statistically controlled for. Perhaps these lower student aspirations were created by lower parental expectations for their children's education. Thus, it is possible that fundamentalist parents on the whole are less supportive of higher education for their children. In summary, liberal and moderate subjects may have had more financial support from home, which is consistent with their higher socioeconomic status, plus more social support, while fundamentalist subjects may have been "on their own," which is consistent with their lower socioeconomic background. Again, this is only speculation; but it is consistent with the available data.

Another possibility is that the mean of the Moderates may have been artificially raised while the mean of the Liberals was artificially lowered through a number of job titles that seemed to best fit under the Duncan Socioeconomic Index title of "Managers and administrators, not elsewhere specified." This one category leaves some room for error and accounted for 31 (18 %) of the 176 Duncan Socioeconomic Index scores assigned in this study, a higher number than any other category. It is scored 50.89, which is quite close to the means of the
Moderates and Liberals. Twenty of the job titles coded were straightforward. Examples include: "vice president of Kroeger Manufacturing," "vice president of major international company," and "Manager of phone company." The 20 straightforward job titles seem to leave some room for error in that distinctions are not made between low, middle and upper level management. Eleven job titles were more difficult to categorize. Examples include: "sales representative (owns his own company)," "builds houses; contractor—owns business," "owns and runs a radio station," and "musician: business of music sales, self-employed."

To avoid this potential error variance in my measure of socioeconomic status, perhaps the criteria for "Managers and administrators, not elsewhere specified" could have been narrowed by coding as "other" the socioeconomic status scores for subjects whose fathers had somewhat vague job titles such as "owns his own business", "is an entrepreneur", or "is self-employed as . . . ." This would have eliminated a number of subjects from being scored on the socioeconomic status variable. The N would have been smaller, but the socioeconomic status scores purer. Unfortunately, the smaller N on the socioeconomic status variable would also decrease the N possible in the regression models that included socioeconomic status (i.e., those in hypotheses 8, 9, 10, and 11). Of course one would want to check to see if subjects who were not scored due to vague job titles were significantly different on important variables from those who were scored. If these two groups differed significantly on important
variables, it would threaten the generalizability from the sample to the population. Possibly, an alternate method of measuring socioeconomic status would have been better in this study. On the other hand, perhaps this is much ado about nothing. Other researchers in cognitive moral development have used the Duncan Socioeconomic Index as a measure of the covariate socioeconomic status without reporting difficulties in categorizing occupations or concerns about reliability (e.g., Basinger, Gibbs, & Fuller, 1992; Getz, 1985).

**Hypothesis 3**

Intelligence, as measured by the Quick Word Test, socioeconomic status, as measured by the Socioeconomic Index "MSEI2" (Men's Socioeconomic Index, version 2), and year in school will be significantly positively correlated with moral judgment as measured by the Defining Issues Test P score.

Hypothesis 3 was partially supported. As Table 19 indicates intelligence was significantly correlated ($r = .34, N = 199, p < .001$) with the Defining Issues Test P score. Socioeconomic status, however, did not reach significance ($r = .09, N = 176, p = .21$), neither did year in school ($r = .11, N = 199, p = .12$). When the same zero-order correlations were run on the sample excluding Agnostics/Atheists, the results were similar, except year in school became marginally significantly correlated with P score ($r = .15, N = 159, p = .06$), see Table 20. Intelligence remained significantly correlated with P score ($r = .37, N = 159, p < .001$), and again, socioeconomic status was not
significantly correlated with P score ($r = .14, N=142, \rho = .10$), though still in the hypothesized direction.

The time of the school year may have influenced one of the variables. Year in school might have been more highly correlated with P score if this study had been run Autumn quarter (assuming Autumn is the only quarter when the majority of the first year students are new to the college environment). Because it was run in the last 6 weeks of Spring quarter, the first year students were already more like second year students in that the majority of them could have been influenced by the college atmosphere for three quarters. Thus, running the study in Spring may have decreased the range of P scores in this sample compared to running it in Autumn.

Table 19
Zero-order Correlations between Defining Issues Test P score and IQ, Socioeconomic Status, and Year in School for Sample Including Agnostics/Atheists

<table>
<thead>
<tr>
<th></th>
<th>IQ</th>
<th>SES</th>
<th>Year in School</th>
</tr>
</thead>
<tbody>
<tr>
<td>P score</td>
<td>.34***</td>
<td>.10</td>
<td>.11</td>
</tr>
</tbody>
</table>

SES = Socioeconomic status.
N = 199 for IQ and year in school; N = 176 for SES
***$\rho < .001$
Table 20

Zero-order Correlations between Defining Issues Test P score and IQ, Socioeconomic Status, and Year in School for Sample Excluding Agnostics/Atheists

<table>
<thead>
<tr>
<th></th>
<th>IQ</th>
<th>SES</th>
<th>Year in School</th>
</tr>
</thead>
<tbody>
<tr>
<td>P score</td>
<td>.37***</td>
<td>.14</td>
<td>.15</td>
</tr>
</tbody>
</table>

SES = Socioeconomic status.
N = 159 for IQ and year in school; N = 142 for SES
***p < .001

Hypothesis 4

Higher scores (more conservative) on the Short Christian Orthodoxy Scale will correlate negatively with P score on the Defining Issues Test.

Hypothesis 5

Higher scores (more conservative) on the Scriptural Literalism Scale will correlate negatively with P score on the Defining Issues Test.

Hypotheses 4 and 5 were not supported. Based on results from zero-order Pearson Product Moment Correlations in Table 21, I was not able to reject the null hypotheses that $r = 0$ between the two measures of Christian religious beliefs and the Defining Issues P score. The correlations for the sample excluding Agnostics/Atheists were nearly zero (see Table 21). Based on my literature review, I anticipated the Scriptural Literalism Scale to correlate with P score

---

Appendix I contains regression equations for these variables. Cohen (1994) recommended reporting "raw" regression coefficients because, unlike correlation coefficients, they do not "change with the degree of variability of the variables they relate" (p. 1001).
about -.44, and the Short Christian Orthodoxy Scale to correlate with P score perhaps to as high as -.47. Before asking for Introduction to Psychology subjects I did a power analysis (Cohen, 1988). Power may be defined as the "probability of rejecting the null hypothesis" (Cohen & Cohen, 1983, p.59). With power set at the recommended level of .80, alpha set at .05 and effect size (r) estimated (conservatively) to be .20, I determined N should be 200 subjects. Based on an anticipated subject loss rate of approximately 25% I asked for, and received, 268 subjects for this study. The final sample was 199. Therefore, power was adequate for hypotheses 4 and 5, which were the foundational hypotheses of this study. However, as mentioned above, the internal reliability of the P score for this sample was only .40, considerably lower than the expected .76. This could partly explain why the correlation was lower than expected.

Table 21
Zero-Order Correlations between Defining Issues Test P score and Short Christian Orthodoxy and Scriptural Literalism Scales

<table>
<thead>
<tr>
<th></th>
<th>Short Christian Orthodoxy Scale</th>
<th>Scriptural Literalism Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>P score (entire sample N=199)</td>
<td>-.05</td>
<td>-.05</td>
</tr>
<tr>
<td>P score (without Agnostics/Atheists, N=159)</td>
<td>.02</td>
<td>.02</td>
</tr>
</tbody>
</table>
Another possible explanation for the failure of P score and the measures of Christian beliefs to exhibit the usual significant negative correlation is related to the low overall mean P score for my sample (M=30.98, N=199, SD=16.63). It was significantly below, t(198) = -9.63, p < .01, the average P scores from past research studies on college students (M=42.3, SD=13.2, N=1,734, with a range of most sample means from 37 to 46; see Rest, 1979b, p. 7.2)\(^2\). Although an average college sample P score of only 31 may be significantly lower than the average college sample, it is not unique. Rest (1979a) reported that:

The samples having the lowest P scores are from the Southern United States, areas of the country usually noted for their conservative and traditional outlook. Two college samples from Georgia and Virginia have P% averages of 24.5 and 34.0 respectively, the two lowest averages in the combined college sample, with a mean of 41.6. Possibly the low scores from the South samples reflect the effects of a conservative intellectual milieu. (p. 115)

Holley (1991) had a similar average P score for his sample (M=30.7, SD=12.51) as I did for mine, along with similar results: "none of the religious measures used in this study correlated significantly with the P-score" (p.330). And Clouse (1991) reported a mean P score of

\(^2\)Rest (1979b) gathered these data from studies sent to him by other researchers from all parts of the country. But not all studies reported a standard deviation, therefore the standard deviation was based on N=1,734. The mean of the college sample of N=1,734 for which standard deviations were reported was equal to the mean of the the large sample of N=2,674.
29.32,\textsuperscript{25} (but no standard deviation was reported) plus a statistically significant, though quite small, positive correlation ($r = .12, N = 335, p = .017$) between P score and a measure of conservative Christian belief that failed to reach significance in a stepwise regression.

This lower mean P score (and the tendency toward a concomitant lack of a significant negative zero-order correlation between P score and conservative Christian beliefs) in these studies could be occurring for a number of reasons. It could be a geographical difference similar to lower P scores in the Southern United States (Rest, 1979a). Aside from the unclear and questionable results of Brown and Annis (1978) in Mississippi, all four studies (including this study) with significantly low P scores and lack of significant negative correlations between P score and conservative Christian beliefs were gathered from state universities in Indiana and Ohio. Clouse's (1991) sample appears to have been drawn from students attending Indiana State University. Clouse (1985) reported "students in this midwestern university [Indiana State University] come from a conservative section of the United States" (p. 195). Holley's (1989, 1991) sample appears to have come from Kent State University, in Ohio. On the other hand, the five studies that reported a significant and unambiguously negative association between P score and conservative religious beliefs for post high school subjects (Cady, 1982; Ernsberger & Manaster, 1981; Getz, 1985; Lawrence, 1979; Volker, 1979) appear to

\textsuperscript{25}Clouse (1991) actually reported a P score of 8.797 (p. 346) which was a raw mean score, not a percentage mean score. I converted this raw mean score to a percentage mean score.
have been from the greater Minneapolis, Minnesota area. And four of these (all but Volker's clergy members) were college samples that appear to have been gathered at the University of Minnesota. Of Rest's (1979b) college samples on which mean college P scores were based, none were reported as located in Indiana or Ohio; however, for many samples, university affiliation was not specified. Perhaps if samples were drawn from state universities in Indiana and Ohio 10-20 years prior to this study they also would have had lower P scores. Or it could be a recent development that is unique to Indiana and Ohio. To make these determinations one would need current data from a variety of regions across the nation.

Perhaps the two studies with lower P scores (Clouse 1991; Holley, 1989, 1991) also had low internal reliability on the P score as I did with this sample. Holley (1989) suspected reliability problems:

It seems apparent from looking at the average P-scores for this sample, that they were, on the whole, unsophisticated, and perhaps, unreliable and unmotivated in their assessments (sic) of the queries (both standard and religious). This sample's most salient statistical feature is the overall low P-score (i.e., \( X = 30.7 \)). In fact, comparing the present study with past research showed that the complexity of reasoning this sample manifested was equivalent to a High School level but far below the college level (e.g., Rest, 1979a, 1979b). In fact, P-scores as high as 54 have been obtained by samples of first year college students (Rest, 1979, p. 7.11) ... This sample may not have had sufficient variability and sufficient reliability to examine the hypotheses generated in the introduction. (pp. 77-78)
Neither Clouse (1991) nor Holley (1989, 1991) reported Cronbach coefficient alphas for their samples. If Defining Issues Tests are hand scored it takes a great deal of extra effort to obtain this internal reliability measure and it seems to be rarely reported. Given the potential problems in P score reliability, however, it may be wise to obtain Cronbach’s coefficient alpha of internal reliability in future studies. While Holley did not report an internal reliability coefficient alpha, he did report for his religious queries version a relatively low test-retest reliability for the P score ($r = .57$) compared with the typical test-retest reliability for the Defining Issues Test P score of .70 to .80. Also he reported an extremely high purge rate (28%) and concluded that his sample was unsuitable for testing his hypotheses.

The purge rate for the Defining Issues Test based on the two reliability measures was also somewhat high in my sample (19%) compared to the expected rate of 5-15%. Furthermore, it seems that students finished their packets too quickly. I had expected them to take between 45 and 60 minutes, based on past research reports. But most were finished and out the door in 40-45 minutes. It is possible that the expected correlations failed to arise due to the potential lack of carefulness exhibited by the subjects. This interpretation is consistent with the low reliability of the P score in this sample (Cronbach’s coefficient alpha = .40).

Another possible factor contributing to the low correlations between the P score and the measures of religious belief are the religious measures. The Short Christian Orthodoxy Scale, which had
the best reliability and validity for a short scale, had never been used before to correlate with P score. So perhaps its particular set of items does not carry the same degree of relationship with the P score as other measures of Christian orthodoxy. The Scriptural Literalism Scale was only used once in past research, and it is still unclear whether its correlation of .44 was positive or negative.

What can be concluded from these results? It is difficult to pinpoint what factors are responsible for the failure of recent studies among secular university students to exhibit statistically significant negative correlations between conservative Christian beliefs and the P score. Have college students changed over time, or is it merely a geographical issue, or both, or some other moderating factor such as differences in intelligence and socioeconomic status, or merely unmotivated and unreliable samples, etc.? It could be that Ohio and Indiana samples may parallel the samples of the southern United States of America in their conservative milieu, thus pulling down the scores of the entire sample. Among such secular university populations where most students have conventional levels of moral development, it is possible that the correlations between P score and religious beliefs are genuinely so small as to be trivial.

Because the recent studies of secular universities with below average P scores were all from Ohio and Indiana and there are recent studies from Christian liberal arts colleges with P scores in the average range, it seems inappropriate to suggest that P scores are dropping across colleges throughout the country or that conservative Christians
are incapable of scoring in the average range on the P score. However, it is important to remember a distinction between fundamentalist conservatives and some evangelical conservatives—a literal interpretation of Scripture. Perhaps the extreme literalism of the fundamentalists does, in fact, prevent them from scoring in the average range for college students on the P score, while other conservative evangelicals, such as many of those attending the Christian College Consortium, are able to score in the average range of the P score.

Also, it seems that P scores of at least the rejectors of conservative Christian beliefs in recent studies at Indiana and Ohio state universities are lower than those relatively older studies at the University of Minnesota, Minneapolis. While I have no good evidence to conclude that conservative Christian beliefs no longer moderate P score on the Defining Issues Test, it must remain a possibility. Three recent studies of the topic (Clouse, 1991; Holley, 1989, 1991; and this study) have reported a failure to find significant correlations between the P score and various measures of Christian beliefs. However, this failure to reject the null hypothesis in no way allows us to accept it as true.
Hypothesis 6

There will be a significant interaction between Intrinsic Scale scores and the Short Christian Orthodoxy Scale scores in predicting P score. As intrinsic religiousness increases, those who score higher on the Short Christian Orthodoxy Scale will score lower on P score, while those who score lower on the Short Christian Orthodoxy Scale will score higher on the P score. However, this will not apply to those who do not consider themselves to be of the Christian faith (i.e. agnostics, atheists, Hindus, Muslims, Jews, etc.) As intrinsic religiousness decreases, the differences noted on the Short Christian Orthodoxy Scale as a predictor of P score will decline.

Hypothesis 7

There will be a significant interaction between Intrinsic Scale scores and the Scriptural Literalism Scale scores in predicting P score. As intrinsic religiousness increases, those who score higher on the Scriptural Literalism Scale will score lower on P score, while those who score lower on the Scriptural Literalism Scale will score higher on the P score. However, this will not apply to those who do not consider themselves to be of the Christian faith (i.e. agnostics, atheists, Hindus, Muslims, Jews, etc.) As intrinsic religiousness decreases, the differences noted on the Scriptural Literalism Scale as a predictor of P score will decline.

I was also not able to reject the null hypotheses that the interactions of my two measures of religious belief and the Intrinsic Scale added a zero amount of variance beyond that already accounted for by the main effects in predicting P score for the sample excluding Agnostics/Atheists. Therefore, hypotheses 6 and 7 were also not supported. However, a number of unexpected, interesting, and
significant relationships were noted related to these hypotheses due to the suppression effects of the Intrinsic scale on the measures of Christian beliefs. Because none of the studies reviewed above included the Intrinsic scale in a regression model with a measure of Christian beliefs (thus preventing the discovery of suppression, even if it existed), and because suppression is somewhat rare, I had not predicted it, nor was I looking for it. This suppression was discovered in the first phase of the planned analyses, where the main effects were analyzed prior to analyzing the interaction effect. Employing the protected $t$ test procedure, I proceeded to analyze the effects of individual variables in a set only if the set's effect was statistically significant. Once the suppression effect of the Intrinsic Scale was incorporated into the model, the measures of Christian beliefs became significant and followed the pattern predicted in Hypotheses 4 and 5. Therefore, because these findings are theoretically meaningful and statistically significant I report them below.

Zero-order correlations among the variables in hypothesis 6 appear in Table 22. Interestingly, the Intrinsic Scale accounted for a significant amount of variance in P score ($r = .19, N=142, p < .05$) for the sample excluding Agnostics/Atheists. What could explain this positive correlation? One tenable theory alluded to above is that "intrinsics try harder." Recall that Kahoe (1974) reported no significant correlation between Intrinsic religious orientation and ACT scores, but Intrinsic religious orientation was positively correlated with first year college GPA (.25) with or without statistical controls for
college aptitude. Thus it seems that intrinsics try harder than non-intrinsics. This same finding was replicated in my sample when IQ was substituted for ACT score. Intrinsic Scale correlated positively with GPA both in a zero-order correlation \( r = .21, N = 156, p < .009 \), and when IQ was statistically controlled for \( sr = .17, N = 155, p < .03 \). Cohen (1994, p. 1001) recommended reporting "raw" regression coefficients because, unlike correlation coefficients, they are not susceptible to variance changes across populations represented in different studies. For the independent variable of GPA in this study, the regression equation in raw units is reported in Appendix I.

How is the concept of intrinsics trying harder related to higher P scores? According to Rest (1988), there is evidence that students' academic orientation correlated significantly and positively with P score, as did the amount of "intellectual stimulation" in the environment, even after initial differences in high school P score were controlled (p. 191). Those who did not try as hard in college tended to have lower P scores. If such a tendency for lower than expected P scores among upper level students who do not "try hard" and prefer taking entry level classes existed in this sample, it may also help explain the lack of a significant correlation between P score and year in school.

Another interesting result in Table 22 is the somewhat unusual pattern of correlations. According to Cohen and Cohen (1983) because \( r_{y1} (-.013) \) is less than the product of \( r_{y2} (.191) \) and \( r_{12} (.133) \), i.e., less than .025, there is evidence of suppression among these variables.
(P score, Intrinsic Scale, and Short Christian Orthodoxy Scale), albeit a small amount.

Table 22
Zero-Order Correlations among Defining Issues Test P score, Short Christian Orthodoxy Scale, Intrinsic Scale, and the Product of Short Christian Orthodoxy Scale, and Intrinsic Scale for Sample Excluding Agnostics/Atheists

<table>
<thead>
<tr>
<th></th>
<th>P score</th>
<th>SCO</th>
<th>I</th>
<th>SCO x I</th>
</tr>
</thead>
<tbody>
<tr>
<td>P score</td>
<td>----</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCO</td>
<td>-.013</td>
<td>----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>.191*</td>
<td>.455***</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td>SCO x I</td>
<td>.133</td>
<td>.780***</td>
<td>.838***</td>
<td>----</td>
</tr>
</tbody>
</table>

Mean 31.34 6.12 3.22 20.05
St. Deviation 17.48 1.05 0.69 6.38

P score = the Defining Issues Test P score; SCO = Short Christian Orthodoxy Scale; I = Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; SCO x I = product of Short Christian Orthodoxy Scale and Intrinsic Scale which carries information for the interaction of these two variables. N = 142.

* p < .05  *** p < .001

In Table 23, when the Short Christian Orthodoxy Scale and Intrinsic Scale were entered together before their interaction term they accounted for a statistically significant 4.9% of variance in P score, $R^2 = .049$, $F(2, 139) = 3.58$, $p = .03$. Their interaction was not significant, increment in $R^2 = .005$, $F(1, 138) = 0.76$, $p = .38$. In Table 24 we see that the Intrinsic Scale accounted for a significant amount
of this variance in P score, $\hat{r}^2 = .049, t(139) = 2.67, p = .009$, while the Short Christian Orthodoxy Scale was not significant, $\hat{r}^2 = .013, t(139) = 1.35, p = .178$.

Table 23

*Intermediate Regression Analysis Regressing Defining Issues Test P score on Short Christian Orthodoxy Scale, Intrinsic Scale, and the Interaction of Short Christian Orthodoxy Scale, and Intrinsic Scale for Sample Excluding Agnostics/Atheists*

<table>
<thead>
<tr>
<th>IVs in model</th>
<th>Cum. $R^2$</th>
<th>F</th>
<th>df</th>
<th>IVs added in $R^2$</th>
<th>t</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCO, I</td>
<td>.049*</td>
<td>3.58</td>
<td>2, 139</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>SCO x I</td>
<td>.054*</td>
<td>2.64</td>
<td>3, 138</td>
<td>SCO x I</td>
<td>.005</td>
<td>0.88</td>
</tr>
</tbody>
</table>

SCO= Short Christian Orthodoxy Scale; I=Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; SCO x I=product of Short Christian Orthodoxy Scale and Intrinsic Scale which carries information for the interaction of these two variables. Subjects who did not complete all measures used in these and subsequent regression analyses were dropped. The resulting N was 142. Dash indicates not applicable. * $p < .05$
Table 24

Simultaneous Regression Analysis Regressing Defining Issues Test P score on Short Christian Orthodoxy Scale and Intrinsic Scale for Sample Excluding Agnostics/Atheists

<table>
<thead>
<tr>
<th>Predictor</th>
<th>sr²</th>
<th>pr²</th>
<th>Beta</th>
<th>B⁰</th>
<th>t</th>
<th>Prob. &gt; t</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCO</td>
<td>0.013</td>
<td>0.013</td>
<td>-0.126</td>
<td>-2.09</td>
<td>-1.35</td>
<td>0.178</td>
</tr>
<tr>
<td>I</td>
<td>0.049</td>
<td>0.049</td>
<td>0.248</td>
<td>6.29</td>
<td>2.67**</td>
<td>0.009</td>
</tr>
</tbody>
</table>

A Intercept = 23.82; SCO = Short Christian Orthodoxy Scale; I = Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; Subjects who did not complete all measures used in these and subsequent regression analyses were dropped. The resulting N was 142. Degrees of freedom = 139 for all t tests. ** p ≤ 0.01

I did not expect Agnostics/Atheists to respond to the Age Universal I/E-Revised Scales in a reliable or valid manner, based on Ernsberger and Manaster's (1981) report that Unitarian Universalists had difficulty with the theistic beliefs assumed of subjects in the Allport-Ross Religious Orientation Scale, and Kirkpatrick's (1989) admonition to remove nonreligious subjects from secular university samples. However, there was support for the reliability of Agnostics/Atheists' responses to the Age Universal I/E-Revised Scales. Cronbach's coefficient alpha of internal reliability (see Table 4) was actually slightly higher in the sample with Agnostic/Atheists (.81) than it was without them (.76). Perhaps this was due to the greater range of scores (inclusion of lower scores) for the sample including Agnostics/Atheists.
Given their adequate reliability, I decided to look at the results both including and excluding Agnostics/Atheists for all my remaining hypotheses to see if including them in the sample increased the relationships with important variables. Some correlations were increased, others were decreased, and some (for the sample including Agnostics/Atheists) were opposite what I had predicted for the sample excluding Agnostic/Atheists. L. Fabrigar (personal communication, January 17, 1995), suggested that to determine if the Intrinsic-Extrinsic Scales were equally valid for Agnostics/Atheists one could do a factor analysis of the scales comparing the self-reported Christians with the self-reported Agnostics/Atheists.

It turns out that several factor analyses of various Intrinsic-Extrinsic Scales have included Agnostics/Atheists. Kirkpatrick (1989) included only subjects who indicated "that they were Christian and at least moderately religious" (p. 6). However:

It is also worth noting, at this point, that for many of the analyses of student samples reported below, parallel analyses were conducted with the nonreligious respondents included. In every case, the results were nearly identical to those reported here. To the extent that results did differ, they tended to be stronger than those in the restricted samples, that is, factors were "cleaner" and more distinct, and beta weights were larger and more likely to be statistically significant. (p. 7)

This may help explain Kirkpatrick's (1989, p.7) seeming ambivalence when he recommended removing nonreligious subjects from
heterogeneous secular university samples, except when doing so would drastically restrict the range of the Intrinsic Scale.

While the original Age Universal I-E Scale was developed on all-Christian samples (Gorsuch & Venable, 1983, pp. 182-183), the Age Universal I/E-Revised Scales, which were used in this study, were developed at both secular and Christian colleges in Southern California with no mention of removing nonreligious subjects (Gorsuch & McPherson, 1989, p. 349). Gorsuch confirmed to me the presence of nonreligious subjects in the sample used to develop the Age Universal I/E-Revised Scales and stated he felt they were able to appropriately answer the items (personal communication, January 18, 1995). He believed when nonreligious subjects encounter questions that inappropriately presume their religiosity they merely assume the test writer didn't do a good job in writing the items and answer as best they can, which is to say, low on both the Intrinsic and Extrinsic scales. From their sample including nonreligious subjects, Gorsuch and McPherson (1989) found the same three factors (I, Es, and Ep) as Kirkpatrick (1989). Thus, it seemed appropriate to analyze the data with Agnostics/Atheists in the sample.

In Table 25 we see that neither of the zero-order correlations of Short Christian Orthodoxy Scale or Intrinsic Scale with P score were significant for the sample including Agnostics/Atheists. Yet when included together in Table 26 they accounted for a statistically significant 4.3% of the variance in P score, \( R^2 = .043, \ F(2, 173) = 3.93, p < .05 \). In Table 27 we see that their squared semi-partial correlation
coefficients ($\beta^2$) are both statistically significant, both larger than their zero-order correlations with P score, and when added together ($11 \cdot 0.032 + 0.036 - 0.068$) their sum is greater than $R^2 = 0.043$. This situation is another example of suppression. The relationship between the Short Christian Orthodoxy Scale and the Intrinsic Scale is hiding or suppressing their real relationship with P score which would be larger were these two independent variables not correlated. It is desirable to include the Intrinsic Scale (i.e. the suppressor, $X_2$) in the regression equation because it removes (suppresses) the unwanted variance in the Short Christian Orthodoxy Scale ($X_1$) and thereby enhances the relationship between the Short Christian Orthodoxy Scale and the P score by means of $B_{Y1.2}$. Alternatively, it is desirable to include the Short Christian Orthodoxy Scale in the regression with the Intrinsic Scale because it removes (suppresses) the unwanted variance in the Intrinsic Scale and thereby enhances the relationship between the Intrinsic Scale and the P score by means of $B_{Y2.1}$. That is, while factors measured by the Intrinsic Scale are working to increase the P score, factors measured by the Short Christian Orthodoxy Scale are simultaneously working to decrease the P score.
Table 25
Zero-Order Correlations among Defining Issues Test P score, Short Christian Orthodoxy Scale, Intrinsic Scale, and the Product of Short Christian Orthodoxy Scale and Intrinsic Scale for Sample Including Agnostics

<table>
<thead>
<tr>
<th></th>
<th>P score</th>
<th>SCO</th>
<th>I</th>
<th>SCO x I</th>
</tr>
</thead>
<tbody>
<tr>
<td>P score</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>SCO</td>
<td>-.086</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>I</td>
<td>.108</td>
<td>.564***</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>SCO x I</td>
<td>.041</td>
<td>.851***</td>
<td>.905***</td>
<td>----</td>
</tr>
</tbody>
</table>

Mean       | 31.67   | 5.69 | 3.06 | 17.99
St. Deviation | 16.85   | 1.43 | 0.73 | 7.34

P score = the Defining Issues Test P score; SCO = Short Christian Orthodoxy Scale; I = Intrinsic subscale (reverse scored) from the Age Universal I/E-Revised Scales; SCO x I = product of Short Christian Orthodoxy Scale and Intrinsic Scale which carries information for the interaction of these two variables. N=176.

***p < .001
Table 26

Intermediate Regression Analysis Regressing Defining Issues Test P score on Short Christian Orthodoxy Scale, Intrinsic Scale, and the Interaction of Short Christian Orthodoxy Scale and Intrinsic Scale for Sample Including Agnostics/Atheists

<table>
<thead>
<tr>
<th>IVs in model</th>
<th>Cum. $R^2$</th>
<th>$F$</th>
<th>df</th>
<th>IVs added</th>
<th>incr. in $R^2$</th>
<th>t</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCO, I</td>
<td>.043*</td>
<td>3.93</td>
<td>2, 173</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>SCO x I</td>
<td>.070*</td>
<td>4.08</td>
<td>3, 172</td>
<td>SCO x I</td>
<td>.027</td>
<td>2.21*</td>
<td>172</td>
</tr>
</tbody>
</table>

SCO= Short Christian Orthodoxy Scale; I=Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; SCO x I=product of Short Christian Orthodoxy Scale and Intrinsic Scale which carries information for the interaction of these two variables. Subjects who did not complete all measures used in these and subsequent regression analyses were dropped. The resulting N was 176. Dash indicates not applicable. * $p<.05$
Table 27

*Simultaneous Regression Analysis Regressing Defining Issues Test P score on Short Christian Orthodoxy Scale and Intrinsic Scale for Sample Including Agnostics/Atheists*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$sr^2$</th>
<th>$pr^2$</th>
<th>Beta</th>
<th>$B^a$</th>
<th>$t$</th>
<th>Prob. $&gt; t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCO</td>
<td>.032</td>
<td>.032</td>
<td>-0.126</td>
<td>-2.54</td>
<td>-2.49*</td>
<td>.018</td>
</tr>
<tr>
<td>I</td>
<td>.036</td>
<td>.036</td>
<td>0.230</td>
<td>5.33</td>
<td>2.55*</td>
<td>.012</td>
</tr>
</tbody>
</table>

*a Intercept = 29.85; SCO = Short Christian Orthodoxy Scale; I = Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; Subjects who did not complete all measures used in these and subsequent regression analyses were dropped. The resulting N was 176. Degrees of freedom=173 for all $t$ tests.*

Table 28

*Simultaneous Regression Analysis Regressing Defining Issues Test P score on Short Christian Orthodoxy Scale, Intrinsic Scale, and their Interaction for Sample Including Agnostics/Atheists*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$sr^2$</th>
<th>$pr^2$</th>
<th>Beta</th>
<th>$B^a$</th>
<th>$t$</th>
<th>Prob. $&gt; t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCO</td>
<td>.041</td>
<td>.042</td>
<td>-0.994</td>
<td>-11.69</td>
<td>-2.74**</td>
<td>.007</td>
</tr>
<tr>
<td>I</td>
<td>.015</td>
<td>.016</td>
<td>-0.740</td>
<td>-17.17</td>
<td>-1.66</td>
<td>.100</td>
</tr>
<tr>
<td>SCO X I</td>
<td>.026</td>
<td>.028</td>
<td>1.557</td>
<td>3.57</td>
<td>2.21*</td>
<td>.028</td>
</tr>
</tbody>
</table>

*a Intercept = 86.44; SCO = Short Christian Orthodoxy Scale; I = Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; Subjects who did not complete all measures used in these and subsequent regression analyses were dropped. The resulting N was 176. Degrees of freedom=173 for all $t$ tests.*

* $p < .05$ ** $p < .01$
In addition to suppression, there was also a significant interaction between the Intrinsic Scale and Short Christian Orthodoxy Scale for the sample including Agnostic/Atheists which increased the variance accounted for in P score from 4.3% to 7%, Table 26, increment in $R^2 = .027$, $F(1,172) = 4.90$, $p < .05$. Thus, an important conditional relationship exists for the sample including Agnostics/Atheists: the regression of P score on conservative Christian beliefs as measured by the Short Christian Orthodoxy Scale varies with the level of intrinsic religious motivation as measured by the Intrinsic Scale, or alternatively, the regression of P score on intrinsic religious motivation as measured by the Intrinsic Scale varies with the level of conservative Christian beliefs as measured by the Short Christian Orthodoxy Scale. To make this conditional relationship clear requires some simple algebraic manipulation of the regression equation

(2) \[ Y = B_1 X_1 + B_2 X_2 + B_3 X_3 + A. \]

Because the interaction variable ($X_3$) is equal to the product of $X_1 X_2$, we can substitute as follows:

(3) \[ Y = B_1 X_1 + B_2 X_2 + B_3 X_1 X_2 + A. \]

Equation 3 can be expressed in either of the following ways:

(4) \[ Y = (B_1 + B_3 X_2) X_1 + (B_2 X_2 + A). \]

(5) \[ Y = (B_2 + B_3 X_1) X_2 + (B_1 X_1 + A). \]

According to Cohen and Cohen (1983):

taking [Equation 5] for concreteness, it can be viewed as a family of equations of regression lines whose slope is $B_2 + B_3 X_1$ and whose Y intercept is $B_1 X_1 + A$. Each different value we assign to $X_1$ defines a different line, thus justifying the
conditional formulation: the regression of Y [P score] on X
[Intrinsic scale] depends on the value of X_1 [Short Christian Orthodoxy Scale]. (pp. 322-323)

Substituting in the raw regression coefficients and intercept for Equation 3 we have:

\[ Y = (-11.69) X_1 + (-17.17) X_2 + (3.57) X_1 X_2 + 86.44 \]

This can be expressed as in equation (4):

\[ Y = (-17.17 + 3.57 X_1) X_2 + (-11.69 X_1 + 86.44) \]

Finally, three representative members of this family of lines can be determined by substituting in three different values for X_1 (Short Christian Orthodoxy Scale). One for a "low" value of X_1 (1 standard deviation below the mean), one for an "average" value of X_1 (at the mean), and one for a "high" value of X_1 (1 standard deviation above the mean), see Table 25 for means and standard deviations. However, because the distribution of the Short Christian Orthodoxy Scale was skewed left, the value for one standard deviation above the mean (\[ 5.69 + 1.43 = 7.12 \]) was greater than the highest possible score of 7.00. Therefore 7.00 was chosen for the high value, instead of 7.12.

Solving Equation 7 for these three representative lines:

\[ Y_L = -1.97 X_2 + 36.64 \]
\[ Y_A = 3.14 X_2 + 19.92 \]
\[ Y_H = 7.82 X_2 + 4.61 \]

Note that these regression lines differ both in slope and intercept. For this sample including Agnostics/Atheists, when Short Christian Orthodoxy Scale score is low there is a slight negative linear
regression of $P$ score on Intrinsic Scale. When Short Christian Orthodoxy Scale score is average there is a slight positive linear regression of $P$ score on Intrinsic Scale. But when Short Christian Orthodoxy Scale score is high there is a relatively larger positive linear regression of $P$ score on Intrinsic Scale. By solving each equation for two different values of $X_2$ listed in Table 29, we can obtain the coordinates needed to graph the three representative lines, thus visually demonstrating the interaction in Figure 2. Figure 3 provides a comparison demonstrating how misleading the regression of $P$ score on Intrinsic scale is without the interaction term, especially when Short Christian Orthodoxy Scale is low. Equations and coordinates used to graph Figure 3 may be found in Appendix J.

Table 29

<table>
<thead>
<tr>
<th>Low ($SCO = 4.261$)</th>
<th>Average ($SCO = 5.69$)</th>
<th>High ($SCO = 7.00$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_2$ $P$ score</td>
<td>$X_2$ $P$ score</td>
<td>$X_2$ $P$ score</td>
</tr>
<tr>
<td>0</td>
<td>36.64</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>26.79</td>
<td>5</td>
</tr>
</tbody>
</table>
With Interaction Term

High SCO (7.00): $Y = -7.82 (X) + 4.61$
Average SCO (5.69): $Y = -3.14 (X) + 19.92$
Low SCO (4.26): $Y = -1.97 (X) + 36.64$

Figure 2
$P$ Score Regression on Intrinsic Scale Moderated by Short Christian Orthodoxy Scale, with Interaction Term for Sample Including Agnostics/Atheists
Without Interaction Term

**High SCO (7.00):** $Y = 5.33 (X) + 17.71$

**Average SCO (5.69):** $Y = 5.33 (X) + 13.63$

**Low SCO (4.26):** $Y = 5.33 (X) + 9.90$

---

Figure 3

*P Score Regression on Intrinsic Scale Moderated by Short Christian Orthodoxy Scale, without Interaction Term, for Sample Including Agnostics/Atheists*
A view of this interaction from the other perspective, Intrinsic Scale moderating P score regression on Short Christian Orthodoxy Scale ($X_1$), is enlightening on hypothesis 6. We can use Equation 4 and solve for three representative lines with low, average and high values on the Intrinsic Scale.

(12) "Low" (1-2.33): $Y_L = -3.37 X_1 + 46.43$

(13) "Average" (1-3.06): $Y_A = -0.77 X_1 + 33.90$

(14) "High" (1-3.79): $Y_H = 1.84 X_1 + 21.37$

Again, note that these regression lines differ in both slope and intercept. For this sample, which included Agnostics/Atheists, the largest negative slope of the P score regression on Short Christian Orthodoxy Scale occurs when the Intrinsic Scale is low. This is the opposite of the interaction I had predicted for the sample excluding Agnostics/Atheists in hypothesis 6 (which was not supported as noted above).

As before, by solving each equation for two different values of $X_1$, we can obtain the coordinates needed (see Table 30) to graph the three representative lines. Compare Figure 4 below with Figure 1 in chapter 3 for a visual representation of these interactions. Thus, it seems to be subjects who are scoring average or low on the Intrinsic Scale, yet still claim a high level of orthodox Christian beliefs who are accounting for the lower P scores\(^\text{26}\). It is hard to imagine who this

\(^{26}\)Interestingly in the sample including Agnostics/Atheists, this interaction was similar and even more pronounced when the combined Intrinsic and Extrinsic Scales were substituted for the Intrinsic Scale alone. Also interesting, the Intrinsic and Extrinsic Scales without any other variables in the model
might describe, except those claim to have orthodox Christian beliefs but are not extremely motivated by their faith. In summary, the subjects who are most likely to have higher P scores are those who both deny orthodox Christian beliefs and deny any commitment to Christianity (perhaps the Atheists, but not necessarily the Agnostics), as well as those who assert a high level of orthodox Christian beliefs combined with an intense commitment to that faith. If the goal is predicting midwestern secular university students' principled level moral judgement as measured by the Defining Issues Test P score, one might apply the criterion Jesus used: it's better to be hot or cold, but not lukewarm (Revelation 3: 15-16).

This interpretation, however, must be kept in perspective. Only 7% of the variance in P score is explained by the main and interaction effects of orthodox Christian beliefs and commitment to those beliefs. Yet, had the reliability of the P score been better for this sample, it may also have increased the correlations reported above. Also, these results may only apply to samples with similar mean P scores. Figure 5 provides a comparison demonstrating how misleading the regression of P score on Short Christian Orthodoxy Scale is without the interaction term, especially when Intrinsic scale is high. Derivation of equations and coordinates used to graph Figure 5 may be found in Appendix J.

 exhibited suppression with each other in predicting P score. See Appendix I for the regression equations.
Table 30
Coordinates for Graphing Three Lines demonstrating P Score Regression on Short Christian Orthodoxy Scale Moderated by Intrinsic Scale, with Interaction Term, for Sample Including Agnostics/Atheists

<table>
<thead>
<tr>
<th>Low (1=2.33)</th>
<th>Average (1=3.06)</th>
<th>High (1=3.79)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X₁ P score</td>
<td>X₁ P score</td>
<td>X₁ P score</td>
</tr>
<tr>
<td>0 46.43</td>
<td>0 33.90</td>
<td>0 21.37</td>
</tr>
<tr>
<td>7 22.84</td>
<td>7 28.51</td>
<td>7 34.25</td>
</tr>
</tbody>
</table>
With Interaction Term

High Intrinsic (3.79): $Y = 1.84 (X) + 21.37$
Average Intrinsic (3.06): $Y = -0.77 (X) + 33.90$
Low Intrinsic (2.33): $Y = -3.37 (X) + 46.43$

Figure 4
$P$ Score Regression on Short Christian Orthodoxy Scale Moderated by Intrinsic Scale, with Interaction Term, for Sample Including Agnostics/Atheists
Figure 5

$P$ Score Regression on Short Christian Orthodoxy Scale Moderated by Intrinsic Scale, without Interaction Term, for Sample Including Agnostics/Atheists
As mentioned above, hypothesis 7 also was not supported. I was not able to reject the null hypothesis that the interaction between the Scriptural Literalism Scale and the Intrinsic Scale added a significant amount of variance beyond that already accounted for by their main effects in predicting P score for the sample excluding Agnostics/Atheists. However, there were some interesting and significant relationships among the variables related to hypothesis 7. Zero-order correlations among the variables in hypothesis 7 are recorded in Table 31. To repeat, the Intrinsic Scale accounted for a significant amount of variance in P score ($r = .19$, N=142, $p < .05$).

Table 31
Zero-Order Correlations among Defining Issues Test P score, Scriptural Literalism Scale, Intrinsic Scale, and the Product of Scriptural Literalism Scale and Intrinsic Scale for Sample Excluding Agnostics/Atheists

<table>
<thead>
<tr>
<th></th>
<th>P score</th>
<th>SLS</th>
<th>I</th>
<th>SLS x I</th>
</tr>
</thead>
<tbody>
<tr>
<td>P score</td>
<td>----</td>
<td>----------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>SLS</td>
<td>.004</td>
<td>----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>.191*</td>
<td>.648***</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td>SLS x I</td>
<td>.120</td>
<td>.884***</td>
<td>.918***</td>
<td>----</td>
</tr>
</tbody>
</table>

Mean: 31.34, 3.51, 3.22, 11.63

St. Deviation: 17.48, 0.68, 0.69, 4.33

P score = the Defining Issues Test P score; SLS = Scriptural Literalism Scale; I = Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; SLS x I = product of Scriptural Literalism Scale and Intrinsic Scale which carries information for the interaction of these two variables. N=142.

* $p < .05$  
** $p < .01$  
*** $p < .001$
In Table 32, when the Scriptural Literalism Scale and Intrinsic Scale were entered together before their interaction term they accounted for a statistically significant 6.1% of variance in P score, $R^2 = 0.061$, $F(2, 139) = 4.52$, $p = .01$, and evidenced suppression. Their interaction was not significant, increment in $R^2 = 0.001$, $t(138) = 0.47$, $p = .64$. In Table 33 we see that the Intrinsic Scale accounted for a significant amount of this variance in P score, $sr^2 = 0.061$, $t(139) = 3.01$, $p = .003$, while the Scriptural Literalism Scale was marginally significant, $sr^2 = 0.025$, $t(139) = 1.91$, $p = .058$. This relationship of suppression indicates that while factors measured by the Intrinsic Scale are working to increase the P score, factors measured by the Scriptural Literalism Scale are simultaneously working to decrease the P score. Therefore, to the extent that the Intrinsic Scale and the Scriptural Literalism Scale are positively correlated, they cancel each other out. Thus, it is important to include the Intrinsic Scale in the regression model with the Scriptural Literalism Scale (and vice versa) so that their real relationships with P score will not be hidden.
Table 32
*Intermediate Regression Analysis Regressing Defining Issues Test P score on Scriptural Literalism Scale, Intrinsic Scale, and the Interaction of Scriptural Literalism Scale and Intrinsic Scale for Sample Excluding Agnostics*

<table>
<thead>
<tr>
<th>Cum. IVs in model</th>
<th>$R^2$</th>
<th>$F$</th>
<th>df</th>
<th>IVs added in $R^2$</th>
<th>$t$</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLS, I</td>
<td>.061</td>
<td>4.52*</td>
<td>2, 139</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>SLS, I, SLS x I</td>
<td>.062</td>
<td>3.07*</td>
<td>3, 138</td>
<td>SLS x I</td>
<td>.001</td>
<td>0.47</td>
</tr>
</tbody>
</table>

SLS = Scriptural Literalism Scale; I = Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; SLS x I = product of Scriptural Literalism Scale and Intrinsic Scale which carries information for the interaction of these two variables. Subjects who did not complete all measures used in the regression analyses were dropped. The resulting $N$ was 142. Dashes indicate not applicable. * $p < .05$, ** $p < .01$.

Table 33
*Simultaneous Regression Analysis Regressing Defining Issues Test P score on Scriptural Literalism Scale and Intrinsic Scale for Sample Excluding Agnostics/Atheists*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$sr^2$</th>
<th>$pr^2$</th>
<th>Beta</th>
<th>$B^a$</th>
<th>$t$</th>
<th>Prob. &gt; $t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLS</td>
<td>.025</td>
<td>.026</td>
<td>-.206</td>
<td>-5.30</td>
<td>-1.91</td>
<td>.058</td>
</tr>
<tr>
<td>I</td>
<td>.061</td>
<td>.061</td>
<td>.324</td>
<td>8.23</td>
<td>3.01*</td>
<td>.003</td>
</tr>
</tbody>
</table>

$a$ Intercept = 23.43; SLS = Scriptural Literalism Scale; I = Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; Subjects who did not complete all measures used in these and subsequent regression analyses were dropped. The resulting $N$ was 142. Degrees of freedom = 139 for all $t$ tests. ** $p < .01$.
As noted above, there is support for the reliability and validity of the Age Universal I/E-Revised Scales for nonreligious subjects and hence no reason to refrain from reporting and interpreting results for the sample that included Agnostics/Atheists. Zero-order correlations for the model in hypothesis 7 are reported in Table 34. For this sample the correlation between Intrinsic Scale and P score was lower than for the sample excluding Agnostics/Atheists and not statistically significant ($r = .11, N=176, p = .15$).

Table 34
Zero-Order Correlations among Defining Issues Test P score, Scriptural Literalism Scale, Intrinsic Scale, and the Product of Scriptural Literalism Scale and Intrinsic Scale for Sample Including Agnostics/Atheists

<table>
<thead>
<tr>
<th></th>
<th>P score</th>
<th>SLS</th>
<th>I</th>
<th>SLS x I</th>
</tr>
</thead>
<tbody>
<tr>
<td>P score</td>
<td>______</td>
<td></td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>SLS</td>
<td>-.067</td>
<td>----</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>I</td>
<td>.108</td>
<td>.711***</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>SLS x I</td>
<td>.046</td>
<td>.903***</td>
<td>.929***</td>
<td>______</td>
</tr>
</tbody>
</table>

Mean | 31.67 | 3.31 | 3.06 | 10.52 |
St. Deviation | 16.85 | 0.79 | 0.73 | 4.59 |

P score - the Defining Issues Test P score; SLS - Scriptural Literalism Scale; I - Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; SLS x I - product of Scriptural Literalism Scale and Intrinsic Scale which carries information for the interaction of these two variables. N=176.

*** $p \leq .001$
In Table 35, when the Scriptural Literalism Scale and Intrinsic Scale were entered together before their interaction term they accounted for a statistically significant 5.3% of variance in P score, $R^2 = .053$, $F(2, 173) = 4.87, p = .009$, and evidenced suppression. Their interaction was not significant, increment in $R^2 = .014$, $t(172) = 1.62, p = .108$. In Table 36 we see that the Scriptural Literalism Scale accounted for a significant amount of this variance in P score, $sr^2 = .042, t(173) = -2.75, p = .007$, as did the Intrinsic Scale, $sr^2 = .049, t(173) = 3.00, p = .003$. This relationship of suppression indicates that factors measured by the Scriptural Literalism Scale are working to decrease the P score, while factors measured by the Intrinsic Scale are simultaneously working to increase the P score. Therefore, to the extent that the the Scriptural Literalism Scale and the Intrinsic Scale are positively correlated, they cancel each other out. Thus, it is important to include the Intrinsic Scale in the regression model with the Scriptural Literalism Scale (and vice versa) so that their real relationships with P score will not be hidden.
Table 35

Intermediate Regression Analysis Regressing Defining Issues Test P score on Scriptural Literalism Scale, Intrinsic Scale, and the Interaction of Scriptural Literalism Scale and Intrinsic Scale for Sample Including Agnostics

<table>
<thead>
<tr>
<th>IVs in model</th>
<th>Cum. R²</th>
<th>F</th>
<th>df</th>
<th>IVs added</th>
<th>Incr. in R²</th>
<th>t</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLS, I</td>
<td>.053</td>
<td>4.87**</td>
<td>2, 173</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>SLS, I, SLS x I</td>
<td>.067</td>
<td>4.15**</td>
<td>3, 172</td>
<td>SLS x I</td>
<td>.014</td>
<td>1.62</td>
<td>172</td>
</tr>
</tbody>
</table>

SLS=Scriptural Literalism Scale; I=Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; SLS x I=product of Scriptural Literalism Scale and Intrinsic Scale which carries information for the interaction of these two variables. Subjects who did not complete all measures used in the regression analyses were dropped. The resulting N was 176. Dashes indicate not applicable.

** p < .01

Table 36

Simultaneous Regression Analysis Regressing Defining Issues Test P score on Scriptural Literalism Scale and Intrinsic Scale for Sample Including Agnostics/Atheists

<table>
<thead>
<tr>
<th>Predictor</th>
<th>sr²</th>
<th>pr²</th>
<th>Beta</th>
<th>Bᵃ</th>
<th>t</th>
<th>Prob. &gt; t</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLS</td>
<td>.042</td>
<td>.042</td>
<td>-.290</td>
<td>-6.20</td>
<td>-2.75**</td>
<td>.007</td>
</tr>
<tr>
<td>I</td>
<td>.049</td>
<td>.049</td>
<td>.314</td>
<td>7.29</td>
<td>3.00**</td>
<td>.003</td>
</tr>
</tbody>
</table>

⁽ᵃ⁾ Intercept = 29.90; SLS=Scriptural Literalism Scale; I=Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; Subjects who did not complete all measures used in these and subsequent regression analyses were dropped. The resulting N was 176. Degrees of freedom=173 for all t tests.

** p < .01
**Hypothesis 8**

Once socioeconomic status, year in school and intelligence are partialled out, higher scores (more conservative) on the Short Christian Orthodoxy Scale will still correlate negatively with P score on the Defining Issues Test, although to a lesser degree.

**Hypothesis 9**

Once intelligence, year in school and socioeconomic status are partialled out, higher scores (more conservative) on the Scriptural Literalism Scale will still correlate negatively with P score on the Defining Issues Test, although to a lesser degree.

**Hypothesis 10**

There will be a significant interaction between Intrinsic Scale scores and the Short Christian Orthodoxy Scale scores in predicting P score. As intrinsic religiousness increases, those who score higher on the the Short Christian Orthodoxy Scale will score lower on P score, while those who score lower on the the Short Christian Orthodoxy Scale will score higher on the P score. However this will not apply to those who do not consider themselves to be of the Christian faith (i.e. agnostics, atheists, Hindus, Muslims, Jews, etc.) As intrinsic religiousness decreases, the differences noted on Short Christian Orthodoxy Scale as a predictor of P score will decline. Once socioeconomic status, year in school and intelligence are partialled out, the above relationships will still hold, although to a lesser degree.

**Hypothesis 11**

There will be a significant interaction between Intrinsic Scale scores and the Scriptural Literalism Scale scores in predicting P score. As intrinsic religiousness increases, those who score higher on the the Scriptural Literalism Scale will score lower on P score, while those who score lower on the the Scriptural Literalism Scale will score higher on the P score. However this will not apply to those who do not consider themselves to be of the Christian faith (i.e. agnostics, atheists, Hindus, Muslims,
Jews, etc.) As intrinsic religiousness decreases, the differences noted on the Scriptural Literalism Scale as a predictor of P score will decline. Once socioeconomic status, year in school and intelligence are partialled out, the above relationships will still hold, although to a lesser degree.

Hypotheses 8, 9, 10 and 11 were based on hypotheses 4, 5, 6 and 7 being supported. However, hypotheses 4, 5, 6 and 7 were not supported. One might at first assume it illogical and fruitless to partial out any variables from a correlation that did not reach significance. Usually this assumption would be correct, however as we have observed above, in the case of suppression it is logical and may even be fruitful to partial out variables from a correlation that was too small to reach significance. Under conditions of suppression with three or more independent variables the sum of the squared semi-partial correlations may be "less than, greater than, or equal to \( R^2 \), depending on the size of the contribution of the variables involved in the suppression and the amount of redundancy among these and other variables" (Cohen and Cohen, 1983, p. 103). In addition, according to Cohen and Cohen, if one were to perform stepwise regression "when there is suppression between two variables neither may reach the criteria for entrance to the equation, although if both were entered they would make a useful contribution to \( R^2 \)" (p. 125). Thus, if a situation of suppression existed between the covariate set (IQ, socioeconomic status, and year in school) and the set of independent variables in this study, it would be logical to analyze the data further to see if the insignificant zero-order correlations between the
dependent and independent variables were misrepresenting a relationship that was perhaps significantly larger when more than one covariate/independent variable was considered at a time.

In order to check if suppression was occurring in this data, one could follow Cohen and Cohen's (1983) example: "When any one of three correlations, \( r_{y1} \), \( r_{y2} \), or \( r_{12} \) is less than the product of the other two, the relationship is what is commonly referred to as suppression" (p. 94). Results from calculating these relationships, when \( Y \)- Defining Issues Test P score, \( X_1 \)- the Short Christian Orthodoxy Scale or the Scriptural Literalism Scale indicated that suppression was occurring when \( X_2 \)- Quick Word Test, year in school, or socioeconomic status. Suppression occurred with all combinations of pairs of \( X_1 \) and \( X_2 \).

Unfortunately, while Cohen and Cohen (1983) mention statistically significant suppression, they do not elaborate how to test for this, nor do they define a minimum effect size of suppression below which further exploration would be fruitless. Therefore, in this vacuum of guidelines, if any degree of suppression occurred, it would seem to be the conscientious researcher's duty to investigate further. Given the existence of potentially relevant suppression among the variables in hypotheses 8, 9, 10 and 11, it seemed incumbent upon me to run the regression models of hypotheses 8, 9, 10 and 11.

While this exercise was logical, no significant improvements were made: neither the Short Christian Orthodoxy Scale nor the Scriptural Literalism Scale, by themselves, added any significant increase to the variance explained in the Defining Issues Test P score.
according to the models in hypotheses 8 and 9. Similarly, the interaction between the Intrinsic Scale and the Short Christian Orthodoxy Scale and the interaction between the Intrinsic Scale and the Scriptural Literalism Scale both failed to add any significant increase to the variance explained in the Defining Issues Test P score according to the models in hypotheses 10 and 11. However, as before, the Intrinsic Scale continued to be a suppressor variable, thus increasing the variance explained by the Short Christian Orthodoxy Scale and the Scriptural Literalism Scale after the covariate set of socioeconomic status, year in school, and IQ had been partialled. Therefore I will present these models below, as before, with the sample excluding and the sample including Agnostics/Atheists.

Zero-order correlations for the model including the Short Christian Orthodoxy Scale and Intrinsic Scale are presented in Table 37. The model is then analyzed somewhat hierarchically by sets, or what Cohen and Cohen (1983, p. 148) refer to as intermediate analysis, in Table 38. Before proceeding to interpret the results we must check the last set to see if there is any evidence to support a significant interaction between the covariates and the other independent variables. If the variance accounted for by this set is statistically significant then there is evidence against the assumption of homogeneity of regression of Y on the covariates across the other independent variables. If this were the case it would be improper to
continue with an analysis of partial variance; instead one should
analyze the interactions.

Fortunately, the interaction set was not significant, increment in
$R^2 = .031$, $F(6, 130) = 0.96$, $p > .05$. For the first set, the covariates
accounted for 19.7% of the variance in P score, $R^2 = .197$, $F(3, 138) =
11.30$, $p < .0001$. The second set, Short Christian Orthodoxy Scale and
Intrinsic Scale, accounted for a significant 5.5% of variance in P, and
6.9% of the variance in what remained in P score once the portion of
variance in P score associated with the covariates was removed (i.e.,
squared multiple partial correlation was .069). In general, I had
expected a significant amount of redundancy between the variance in
P score accounted for by the covariate set and the variance in P score
accounted for by the religious variables. I expected once the
covariates were partialled out that the variance in P score accounted
for by the set of religious variables would be cut in half. But it wasn't.
In fact, it increased from 4.9% (see Table 23) to 5.5% (see Table 38).
Therefore, it seems — at least for this sample, and perhaps for other
conventional college samples with similarly low P scores— that the
religious variable set of Short Christian Orthodoxy Scale and Intrinsic
Scale explained a significant amount of variance even after the
covariates of socioeconomic status, year in school and IQ were
partialled out. Because of the significant overall $F$ tests for these sets
I went on to analyze the variables within the sets.
Table 37
Zero-order Correlations among Defining Issues Test P Score, Socioeconomic Status, Year in School, IQ, Short Christian Orthodoxy Scale, and Intrinsic Scale for Sample Excluding Agnostics/Atheists

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>SES</th>
<th>Year</th>
<th>IQ</th>
<th>SCO</th>
<th>Intrinsic</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES</td>
<td>.139</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>.148</td>
<td>-.009</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQ</td>
<td>.419***</td>
<td>.002</td>
<td>.232**</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCO</td>
<td>-.013</td>
<td>.081</td>
<td>.003</td>
<td>.145</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>Intrinsic</td>
<td>.191</td>
<td>-.124</td>
<td>.102</td>
<td>.127</td>
<td>.455***</td>
<td>-----</td>
</tr>
</tbody>
</table>

Mean 31.34 49.73 1.72 52.41 6.12 3.22
SD 17.48 20.43 0.93 10.78 1.05 0.69

P=P score; SES=Socioeconomic status; Year=Year in school; SCO=Short Christian Orthodoxy Scale; Intrinsic=Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; Subjects who did not complete all measures used in subsequent regression analyses were dropped. The resulting N was 142.

** p<.01  *** p<.001
Table 38
Intermediate Regression Analysis Regressing Defining Issues Test P score on Three Sets each Adding the Following Variables: 1) Socioeconomic Status, Year, IQ, 2) Short Christian Orthodoxy Scale, Intrinsic Scale, and 3) All Interactions between variables in Set 1 and Set 2 for Sample Excluding Agnostics/Atheists

<table>
<thead>
<tr>
<th>IVs in model</th>
<th>Cum R²</th>
<th>F</th>
<th>df</th>
<th>IVs added</th>
<th>Incr. in R²</th>
<th>F</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES, Year, IQ</td>
<td>.197</td>
<td>11.30***</td>
<td>3, 138</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>SES, Year, IQ, SCO, I</td>
<td>.252</td>
<td>9.15***</td>
<td>5, 136</td>
<td>SCO, Iα</td>
<td>.055</td>
<td>4.95**</td>
<td>2, 136</td>
</tr>
<tr>
<td>SES, Year, IQ, SCO, I, (6 IVs: SCO x SES to I x IQ)</td>
<td>.283</td>
<td>4.65***</td>
<td>130</td>
<td>6 IVs</td>
<td>.032</td>
<td>0.96</td>
<td>6, 130</td>
</tr>
</tbody>
</table>

α For the set (SCO, I) the squared multiple partial correlation = .069. SCO=Short Christian Orthodoxy Scale; I=Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; SES=Socioeconomic Status; Subjects who did not complete all measures used in these regression analyses were dropped. The resulting N was 142. Dash indicates not applicable. **p < .01 ***p < .001

In Table 39 we see that, of the three covariates in the model, only IQ accounted for a statistically significant amount of variance (15.6%) in the P score, $\sigma_r^2 = .156$, $t(138) = 5.17$, $p < .0001$. In Table 40 when Short Christian Orthodoxy Scale and Intrinsic Scale were added into the model, all the variables added a statistically significant amount of variance, except year in school.
Table 39
Simultaneous Regression Analysis Regressing Defining Issues Test P score on Socioeconomic Status, Year, and IQ for Sample Excluding Agnostics/Atheists

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$sr^2$</th>
<th>$pr^2$</th>
<th>Beta</th>
<th>$B^a$</th>
<th>$t$</th>
<th>Prob. $&gt; t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES</td>
<td>.019</td>
<td>.023</td>
<td>.141</td>
<td>0.11</td>
<td>1.82</td>
<td>.071</td>
</tr>
<tr>
<td>Year</td>
<td>.003</td>
<td>.004</td>
<td>.056</td>
<td>1.03</td>
<td>0.70</td>
<td>.486</td>
</tr>
<tr>
<td>IQ</td>
<td>.156</td>
<td>.162</td>
<td>.411</td>
<td>0.66</td>
<td>5.17***</td>
<td>.001</td>
</tr>
</tbody>
</table>

$a$ Intercept $= -10.81$; SES=Socioeconomic Status; Subjects who did not complete all measures used in these regression analyses were dropped. The resulting $N$ was 142. Degrees of freedom $=138$ for all $t$ tests. $*** p<.001$

Table 40
Simultaneous Regression Analysis Regressing Defining Issues Test P score on Socioeconomic Status, Year, IQ, Short Christian Orthodoxy Scale, and Intrinsic Scale for Sample Excluding Agnostics/Atheists

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$sr^2$</th>
<th>$pr^2$</th>
<th>Beta</th>
<th>$B^a$</th>
<th>$t$</th>
<th>Prob. $&gt; t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES</td>
<td>.033</td>
<td>.042</td>
<td>.188</td>
<td>0.13</td>
<td>2.45*</td>
<td>.016</td>
</tr>
<tr>
<td>Year</td>
<td>.001</td>
<td>.001</td>
<td>.030</td>
<td>0.33</td>
<td>0.39</td>
<td>.700</td>
</tr>
<tr>
<td>IQ</td>
<td>.154</td>
<td>.171</td>
<td>.415</td>
<td>0.63</td>
<td>5.30***</td>
<td>.001</td>
</tr>
<tr>
<td>SCO</td>
<td>.031</td>
<td>.040</td>
<td>-.205</td>
<td>-4.70</td>
<td>-2.37*</td>
<td>.019</td>
</tr>
<tr>
<td>I</td>
<td>.047</td>
<td>.059</td>
<td>.245</td>
<td>7.03</td>
<td>2.93**</td>
<td>.004</td>
</tr>
</tbody>
</table>

$a$ Intercept $= -15.16$; SES=Socioeconomic Status; SCO=Short Christian Orthodoxy Scale; I=Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; Subjects who did not complete all measures used in these regression analyses were dropped. The resulting $N$ was 142. Degrees of freedom $=138$ for all $t$ tests. $* p<.05$ $** p<.01$ $*** p<.001$
Proceeding to the sample including Agnostics/Atheists, in Table 41 are zero-order correlations among variables in the model that included Short Christian Orthodoxy Scale and Intrinsic Scale. The model is then analyzed by sets in an intermediate analysis in Table 42. Because the last set is not statistically significant, increment in $R^2 = .045$, $F(6, 164) = 1.59$, $p > .05$, there is no evidence against the assumption of homogeneity of regression of $Y$ on the covariates across the other independent variables and we may proceed with the analysis of partial variance. For the first set, the covariates accounted for 14.3% of the variance in $P$ score, $R^2 = .143$, $F(3, 172) = 9.57$, $p < .0001$. This was 5.4% less variance in $P$ score than accounted for by the same covariates in the sample excluding Agnostics/Atheists. The second set, Short Christian Orthodoxy Scale and Intrinsic Scale, accounted for a significant 3.3% of variance in $P$ score, $R^2 = .033$, $F(2, 170) = 3.42$, $p < .05$, and 3.9% of the variance in $P$ score once the portion of variance in $P$ score associated with the covariates was removed (i.e., squared multiple partial correlation was .039). Because of the significant overall $F$ tests for these sets I went on to analyze the variables within the sets in Tables 42 and 43.
Table 41
*Zero-order Correlations among Defining Issues Test P Score, Socioeconomic Status, Year in School, IQ, Short Christian Orthodoxy Scale, and Intrinsic Scale for Sample Including Agnostics/Atheists*

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>SES</th>
<th>Year</th>
<th>IQ</th>
<th>SCO</th>
<th>Intrinsic</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----------</td>
</tr>
<tr>
<td>SES</td>
<td>.095</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----------</td>
</tr>
<tr>
<td>Year</td>
<td>.122</td>
<td>.003</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----------</td>
</tr>
<tr>
<td>IQ</td>
<td>.371***</td>
<td>.065</td>
<td>.261**</td>
<td>-----</td>
<td>-----</td>
<td>-----------</td>
</tr>
<tr>
<td>SCO</td>
<td>-.087</td>
<td>-.022</td>
<td>-.149*</td>
<td>-.091</td>
<td>-----</td>
<td>-----------</td>
</tr>
<tr>
<td>Intrinsic</td>
<td>.108</td>
<td>-.121</td>
<td>-.007</td>
<td>.004</td>
<td>.564***</td>
<td>-----</td>
</tr>
</tbody>
</table>

Mean | 31.67 | 50.14 | 1.81 | 53.57 | 5.69 | 3.06 |
SD   | 16.85 | 20.32 | 0.99 | 11.01 | 1.43 | 0.73 |

P-P score; SES-Socioeconomic status; Year-Year in school; SCO-Short Christian Orthodoxy Scale; Intrinsic-Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales. Subjects who did not complete all measures used in subsequent regression analyses were dropped. The resulting N was 176.

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

* * * p < .001
Table 42

Intermediate Regression Analysis Regressing Defining Issues Test P score on Three Sets each Adding the Following Variables: 1) Socioeconomic Status, Year, IQ, 2) Short Christian Orthodoxy Scale, Intrinsic Scale, and 3) All Interactions between variables in Set 1 and Set 2 for Sample Including Agnostics/Atheists

<table>
<thead>
<tr>
<th>IVs in model</th>
<th>Cum IVs</th>
<th>IVs</th>
<th>Increment. (in $R^2$)</th>
<th>Cum F</th>
<th>F</th>
<th>df</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES, Year, IQ</td>
<td>.143</td>
<td>9.57***</td>
<td>3, 172</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>SES, Year, IQ, SCO, I</td>
<td>.176</td>
<td>7.27***</td>
<td>5, 170</td>
<td>SCO, Ia</td>
<td>.033</td>
<td>3.42*</td>
<td>2, 170</td>
</tr>
<tr>
<td>SES, Year, IQ, SCO, I, (6 IVs: SCO x SES to I x IQ)</td>
<td>.222</td>
<td>4.24***</td>
<td>164</td>
<td>6 IVs</td>
<td>.045</td>
<td>1.59</td>
<td>6, 164</td>
</tr>
</tbody>
</table>

For the set (SCO, I) the squared multiple partial correlation $r^2 = .039$. SCO-Short Christian Orthodoxy Scale; I-Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; SES-Socioeconomic Status; Subjects who did not complete all measures used in these regression analyses were dropped. The resulting N was 176. Dash indicates not applicable. ** $p < .01$ *** $p < .001$

In Table 43 we see that among the covariate set, only IQ accounted for a statistically significant amount of variance (11.9%) in P score, $sr^2 = .119$, $t(172) = 4.89$, $p < .0001$, which was nearly 4% less than had been accounted for in the sample excluding Agnostics/Atheists. The other two covariates, socioeconomic status and year in school, did not even approach significance. In Table 44 when Short Christian Orthodoxy Scale and Intrinsic Scale were added into the model, they, along with IQ, added a significant amount to the variance explained in P score. The model in this sample, unlike the
sample excluding Agnostics/Atheists, did not show a significant
amount of variance explained in P score by socioeconomic status.

Table 43
*Simultaneous Regression Analysis Regressing Defining Issues Test P score on Socioeconomic Status, Year, and IQ for Sample Including Agnostics/Atheists*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$sr^2$</th>
<th>$pr^2$</th>
<th>Beta</th>
<th>$B^a$</th>
<th>t</th>
<th>Prob. &gt; t</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES</td>
<td>.005</td>
<td>.006</td>
<td>.072</td>
<td>0.06</td>
<td>1.01</td>
<td>.315</td>
</tr>
<tr>
<td>Year</td>
<td>.001</td>
<td>.001</td>
<td>.029</td>
<td>0.48</td>
<td>0.39</td>
<td>.698</td>
</tr>
<tr>
<td>IQ</td>
<td>.119</td>
<td>.122</td>
<td>.367</td>
<td>0.55</td>
<td>4.89***</td>
<td>.001</td>
</tr>
</tbody>
</table>

---

$a$ Intercept = -1.57; SES = Socioeconomic Status; Subjects who did not complete all measures used in these regression analyses were dropped. The resulting N was 176. Degrees of freedom = 172 for all t tests. 

*** $p<.001$
Table 44
Simultaneous Regression Analysis Regressing Defining Issues Test P score on Socioeconomic Status, Year, IQ, Short Christian Orthodoxy Scale and Intrinsic Scale for Sample Including Agnostics/Atheists

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$sr^2$</th>
<th>$pr^2$</th>
<th>Beta</th>
<th>$B^a$</th>
<th>$t$</th>
<th>Prob. &gt; $t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES</td>
<td>.009</td>
<td>.010</td>
<td>.096</td>
<td>0.08</td>
<td>1.34</td>
<td>.181</td>
</tr>
<tr>
<td>Year</td>
<td>.000</td>
<td>.000</td>
<td>.008</td>
<td>0.13</td>
<td>0.10</td>
<td>.918</td>
</tr>
<tr>
<td>IQ</td>
<td>.110</td>
<td>.118</td>
<td>.354</td>
<td>0.53</td>
<td>4.77***</td>
<td>.001</td>
</tr>
<tr>
<td>SCO</td>
<td>.020</td>
<td>.023</td>
<td>-.175</td>
<td>-2.03</td>
<td>-2.01*</td>
<td>.046</td>
</tr>
<tr>
<td>I</td>
<td>.031</td>
<td>.036</td>
<td>.215</td>
<td>5.00</td>
<td>2.52*</td>
<td>.013</td>
</tr>
</tbody>
</table>

$^a$ Intercept = -4.56; SES-Socioeconomic Status; SCO-Short Christian Orthodoxy Scale; I-Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; Subjects who did not complete all measures used in these regression analyses were dropped. The resulting N was 176. Degrees of freedom=170 for all $t$ tests. * $p < .05$ *** $p < .001$

Zero-order correlations for the model including the Scriptural Literalism Scale and Intrinsic Scale are presented in Table 45. The model is then analyzed by intermediate regression analysis, in Table 46. Because the last set is not statistically significant, increment in $R^2$ - .028, $F(6, 130) = 0.82$, $p > .05$, there is no evidence against the assumption of homogeneity of regression of $Y$ on the covariates across the other independent variables and we may proceed with the analysis of partial variance.

For the first set, the covariates again accounted for 19.7% of the variance in P score, $R^2 = .197$, $F(3, 138) = 11.30$, $p < .0001$. The second set, Scriptural Literalism Scale and Intrinsic Scale, accounted for a significant 4.2% of variance in P, and 5.3% of the variance in what
remained in P score once the portion of variance in P score associated with the covariates was removed (i.e., squared multiple partial correlation was .053). This was consistent, in general, with my expectation of a significant amount of redundancy between the variance in P score accounted for by the covariate set and the variance in P score accounted for by the religious variables. I expected once the covariates were partialled out that the variance in P score accounted for by the Scriptural Literalism Scale would decrease and perhaps be cut in half. It did decrease for the set including both Scriptural Literalism Scale and Intrinsic Scale from 6.1% (see Table 32) to 4.2% (see Table 46) so there was in fact some redundancy among the covariates and the these two religious variables. However, according to Lee Fabrigar (personal communication, February 23, 1995) no formal tests exist to determine if these differences are statistically significant. Materially, this difference seems rather small, only a drop of 1.9%, however, this is nearly a decrease of a third of the variance explained in P score (6.1%) by the set of Scriptural Literalism Scale and Intrinsic Scale without partialling the covariate set. Because of the significant overall $F$ tests for these sets I went on to analyze the variables within the sets.
Table 45

Zero-order Correlations among the Defining Issues Test P Score, Socioeconomic Status, Year in School, IQ, Scriptural Literalism Scale, and Intrinsic Scale for Sample Excluding Agnostics/Atheists

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>SES</th>
<th>Year</th>
<th>IQ</th>
<th>SLS</th>
<th>Intrinsic</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----------</td>
</tr>
<tr>
<td>SES</td>
<td>.139</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----------</td>
</tr>
<tr>
<td>Year</td>
<td>.148</td>
<td>-.009</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----------</td>
</tr>
<tr>
<td>IQ</td>
<td>.419***</td>
<td>.002</td>
<td>.232**</td>
<td>-----</td>
<td>-----</td>
<td>-----------</td>
</tr>
<tr>
<td>SLS</td>
<td>.004</td>
<td>-.091</td>
<td>-.072</td>
<td>.058</td>
<td>-----</td>
<td>-----------</td>
</tr>
<tr>
<td>Intrinsic</td>
<td>.191</td>
<td>-.124</td>
<td>.102</td>
<td>.127</td>
<td>.648***</td>
<td>-----</td>
</tr>
</tbody>
</table>

Mean 31.34 49.73 1.72 52.41 3.51 3.22
SD 17.48 20.43 0.93 10.78 0.68 0.69

P=P score; SES=Socioeconomic status; Year=Year in school; SLS=Scriptural Literalism Scale; Intrinsic=Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; Subjects who did not complete all measures used in subsequent regression analyses were dropped. The resulting N was 142.

**p < .01  ***p < .001
Table 46
Intermediate Regression Analysis Regressing Defining Issues Test P score on Three Sets each Adding the Following Variables: 1) Socioeconomic Status, Year, IQ, 2) Scriptural Literalism Scale, Intrinsic Scale, and 3) All Interactions between variables in Set 1 and Set 2 for Sample Excluding Agnostics/Atheists

<table>
<thead>
<tr>
<th>IVs in model</th>
<th>Cum R²</th>
<th>F</th>
<th>df</th>
<th>IVs added</th>
<th>Incr. R²</th>
<th>F</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES, Year, IQ</td>
<td>.197</td>
<td>11.30***</td>
<td>3, 138</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>SES, Year, IQ, SLS, I</td>
<td>.240</td>
<td>8.57***</td>
<td>5, 136</td>
<td>SLS, Ia</td>
<td>.042</td>
<td>3.79*</td>
<td>2, 136</td>
</tr>
<tr>
<td>SES, Year, IQ, SLS, I (6 IVs): SLS x SES to I x IQ)</td>
<td>.267</td>
<td>4.31***</td>
<td>130</td>
<td>6 IVs</td>
<td>.028</td>
<td>0.82</td>
<td>6, 130</td>
</tr>
</tbody>
</table>

a For the set (SLS, I) the squared multiple partial correlation = .053.
SLS = Scriptural Literalism Scale; I = Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; SES = Socioeconomic Status; Subjects who did not complete all measures used in these regression analyses were dropped. The resulting N was 142. Dash indicates not applicable. ** p < .01 *** p < .001

In table 38 we saw that, of the three covariates in the model, only IQ accounted for a statistically significant amount of variance (15.6%) in the P score, $r^2 = .156$, $t(138) = 5.17$, $p < .0001$. In Table 47 when Scriptural Literalism Scale and Intrinsic Scale were added into the model, socioeconomic status, IQ and Intrinsic Scale each added a statistically significant amount of variance while Scriptural Literalism Scale dropped from a statistically significant 2.5% of variance in P score explained without covariates and with Intrinsic Scale in the
model (Table 33) to add a statistically insignificant 1.9% ($p = .069$) with covariates and Intrinsic Scale in the model.

Thus, while the Short Christian Orthodoxy Scale increased from 1.3% (Table 24) to 3.1% (Table 40) of variance explained in P score when the covariates were added, the Scriptural Literalism Scale decreased from 2.5% to 1.9% of variance explained in P score when the covariates were added (with all models including the Intrinsic Scale, a suppressor variable to both measures of Christian beliefs). Again, these are rather small changes and there is no formal test to determine if they are statistically significant (Lee Fabrigar, personal communication, February 23, 1995). They would need to be replicated for generalizability to other conventional college samples with relatively low P scores, and it would also be interesting to see what pattern of results emerged from a less conventional college sample exhibiting P scores in the average range for college samples. Given these cautions, please note that these findings are consistent with some of the expectations on which this study was based: namely, that socioeconomic status would be positively associated with P score and negatively associated with fundamentalism (see Tables 17 and 18), whose chief marker is a literal interpretation of Scripture. Thus, in predicting P score, the Scriptural Literalism Scale would be likely to exhibit more partial redundancy with the covariate set of year in school, socioeconomic status, and IQ than would the Short Christian Orthodoxy Scale, because the Scriptural Literalism Scale is more
sensitive to this marker of fundamentalism—literal interpretation of Scripture.

Table 47
Simultaneous Regression Analysis Regressing the Defining Issues Test P score on Socioeconomic Status, Year, IQ, Scriptural Literalism Scale, and Intrinsic Scale for Sample Excluding Agnostics/Atheists

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$sr^2$</th>
<th>$pr^2$</th>
<th>Beta</th>
<th>$B^a$</th>
<th>$t$</th>
<th>Prob. &gt; $t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES</td>
<td>.024</td>
<td>.031</td>
<td>.158</td>
<td>0.13</td>
<td>2.07*</td>
<td>.040</td>
</tr>
<tr>
<td>Year</td>
<td>.000</td>
<td>.000</td>
<td>.018</td>
<td>0.33</td>
<td>0.22</td>
<td>.825</td>
</tr>
<tr>
<td>IQ</td>
<td>.142</td>
<td>.157</td>
<td>.395</td>
<td>0.63</td>
<td>5.04***</td>
<td>.001</td>
</tr>
<tr>
<td>SLS</td>
<td>.019</td>
<td>.024</td>
<td>-.182</td>
<td>-4.70</td>
<td>-1.83</td>
<td>.069</td>
</tr>
<tr>
<td>I</td>
<td>.042</td>
<td>.053</td>
<td>.272</td>
<td>7.03</td>
<td>2.75**</td>
<td>.007</td>
</tr>
</tbody>
</table>

*a Intercept = -15.16; SES=Socioeconomic Status; SLS = Scriptural Literalism Scale; I=Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; Subjects who did not complete all measures used in these regression analyses were dropped. The resulting N was 142. Degrees of freedom=138 for all t tests.

* $p < .05$  ** $p < .01$  *** $p < .001$

Proceeding, finally, to the sample including Agnostics/Atheists, in Table 48 are zero-order correlations among variables in the model that included Scriptural Literalism Scale and Intrinsic Scale. The model is then analyzed by sets, as before, through intermediate analysis in Table 48. Because the last set is not statistically significant, increment in $R^2 = .051$, $F(6, 165) = 1.79, p > .05$, there is no evidence against the assumption of homogeneity of regression of $Y$ on the covariates across the other independent variables and we may
proceed with the analysis of partial variance. For the first set, the
covariates accounted for 14.3% of the variance in P score, $R^2 = .143$,
$\mathcal{F}(3, 172) = 9.57, p < .0001$. As noted above, this was 5.4% less
variance in P score than accounted for by the same covariates in the
sample excluding Agnostics/Atheists. The second set, Scriptural
Literalism Scale and Intrinsic Scale, accounted for a significant 3.3% of
variance in P score, $R^2 = .033$, $\mathcal{F}(2, 170) = 3.42, p < .05$, and 3.9% of the
variance in what remained in P score once the portion of variance in P
score associated with the covariates was removed (i.e., squared
multiple partial correlation was .039). Because of the significant
overall F tests for these sets I went on to analyze the variables within
the sets. The first set, the covariates, is located in Table 43 with
descriptive text preceding it (which I will not repeat here). The
second set is in Table 50.
Table 48
Zero-order Correlations among Defining Issues Test P Score, Socioeconomic Status, Year in School, IQ, Scriptural Literalism Scale, and Intrinsic Scale for Sample Including Agnostics/Atheists

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>SES</th>
<th>Year</th>
<th>IQ</th>
<th>SLS</th>
<th>Intrinsic</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>----</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES</td>
<td>.095</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>.122</td>
<td>.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQ</td>
<td>.371**</td>
<td>.065</td>
<td>.261**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLS</td>
<td>-.067</td>
<td>-.109</td>
<td>-.137</td>
<td>-.114</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intrinsic</td>
<td>.108</td>
<td>-.121</td>
<td>-.007</td>
<td>.004</td>
<td>.711***</td>
<td></td>
</tr>
</tbody>
</table>

Mean: 31.67  50.14  1.81  53.57  3.31  3.06
SD:  16.85  20.32  0.99  11.01  0.79  0.73

P-P score; SES-Socioeconomic status; Year-Year in school; SLS-Scriptural Literalism Scale; Intrinsic-Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; Subjects who did not complete all measures used in subsequent regression analyses were dropped. The resulting N was 176.

** p < .01
*** p < .001
Table 49

Intermediate Regression Analysis Regressing Defining Issues Test P score on Three Sets each Adding the Following Variables: 1) Socioeconomic Status, Year, IQ, 2) Scriptural Literalism Scale, Intrinsic Scale, and 3) All Interactions between variables in Set 1 and Set 2 for Sample Including Agnostics/Atheists

<table>
<thead>
<tr>
<th>IVs in model</th>
<th>Cum IVs</th>
<th>IVs added</th>
<th>Incr.</th>
<th>Cum IVs</th>
<th>IVs added</th>
<th>Incr.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>$F$</td>
<td>df</td>
<td>$R^2$</td>
<td>$F$</td>
<td>df</td>
</tr>
<tr>
<td>SES, Year, IQ</td>
<td>.143</td>
<td>9.57***</td>
<td>3, 172</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>SES, Year, IQ, SLS, I</td>
<td>.176</td>
<td>7.28**</td>
<td>5, 170</td>
<td>SLS, I</td>
<td>.033</td>
<td>3.43*</td>
</tr>
<tr>
<td>SES, Year, IQ, SLS, I</td>
<td>.227</td>
<td>4.38***</td>
<td>11</td>
<td>SLS x SES to 1 x IQ</td>
<td>.051</td>
<td>1.79</td>
</tr>
</tbody>
</table>

a For the set (SLS, I) the squared multiple partial correlation = .039.

SLS=Scriptural Literalism Scale; I=Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; SES=Socioeconomic Status; Subjects who did not complete all measures used in these regression analyses were dropped. The resulting N was 176. Dash indicates not applicable. ** $p < .01$ *** $p < .001$

In Table 50 when Scriptural Literalism Scale and Intrinsic Scale were added into the model, they along with IQ added a significant amount to the variance explained in P score. The model in this sample that included Agnostics/Atheists, unlike the sample excluding Agnostics/Atheists, did not show a significant amount of variance explained in P score by socioeconomic status.
Table 50

Simultaneous Regression Analysis Regressing Defining Issues Test \( P \) score on Socioeconomic Status, Year, IQ, Scriptural Literalism Scale, and Intrinsic Scale for Sample Including Agnostics/Atheists

<table>
<thead>
<tr>
<th>Predictor</th>
<th>( sr^2 )</th>
<th>( pr^2 )</th>
<th>Beta</th>
<th>( B^a )</th>
<th>( t )</th>
<th>Prob. &gt; ( t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES</td>
<td>.007</td>
<td>.008</td>
<td>.083</td>
<td>0.07</td>
<td>1.17</td>
<td>.244</td>
</tr>
<tr>
<td>Year</td>
<td>.000</td>
<td>.000</td>
<td>.007</td>
<td>0.12</td>
<td>0.10</td>
<td>.921</td>
</tr>
<tr>
<td>IQ</td>
<td>.105</td>
<td>.113</td>
<td>.348</td>
<td>0.52</td>
<td>4.66***</td>
<td>.001</td>
</tr>
<tr>
<td>SLS</td>
<td>.020</td>
<td>.023</td>
<td>-.203</td>
<td>-4.36</td>
<td>-2.01*</td>
<td>.046</td>
</tr>
<tr>
<td>I</td>
<td>.033</td>
<td>.038</td>
<td>.261</td>
<td>6.08</td>
<td>2.61**</td>
<td>.010</td>
</tr>
</tbody>
</table>

\( ^a \) Intercept = -4.56; SES = Socioeconomic Status; SLS = Scriptural Literalism Scale; I = Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales; Subjects who did not complete all measures used in these regression analyses were dropped. The resulting \( N \) was 176. Degrees of freedom = 170 for all \( t \) tests.

* \( p < .05 \)  ** \( p < .01 \)  *** \( p < .001 \)

Hypothesis 12

A meaningful Christian personal religious experience — as measured by Clouse's (1991) technique — will be significantly negatively associated with Stage 3 and positively associated with Stage 4 on the Defining Issues Test.

As table 50 indicates, no significant differences were found on Stage 3 or Stage 4 mean percentage scores between those who claim to have had a Christian personal religious experience that is currently meaningful to them and those who have not had such an experience.
Table 51
*T-Test Procedure of Defining Issues Test Stage 3 and Stage 4 Means for those with and those without a Currently Meaningful Christian Personal Religious Experience*

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>MEAN</th>
<th>SD</th>
<th>df</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stage 3 Mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Christian Personal Experience</td>
<td>130</td>
<td>20.46</td>
<td>12.46</td>
<td>195</td>
<td>.39</td>
</tr>
<tr>
<td>Yes Christian Personal Experience</td>
<td>67</td>
<td>18.91</td>
<td>10.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stage 4 Mean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Christian Personal Experience</td>
<td>130</td>
<td>34.51</td>
<td>14.37</td>
<td>195</td>
<td>.88</td>
</tr>
<tr>
<td>Yes Christian Personal Experience</td>
<td>67</td>
<td>34.18</td>
<td>15.45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Table 52 we see that neither Stage 3 nor Stage 4 correlated significantly with a currently meaningful Christian personal religious experience. However, these zero-order correlations are provided to give a full picture of the relationships among the variables and are not intended as a replication of Clouse (1991). In Tables 52 and 53 we see that Clouse's (1991) finding was not replicated. Neither Stage 3 nor Stage 4 were significantly correlated to a currently meaningful Christian personal religious experience after GPA was partialled.
Table 52
Zero-order Correlations among Stage 3, Stage 4, GPA, and Currently Meaningful Christian Personal Religious Experience

<table>
<thead>
<tr>
<th></th>
<th>Stage 3</th>
<th>Stage 4</th>
<th>GPA</th>
<th>Chr. Exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 3</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage 4</td>
<td>-.41***</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>-.17*</td>
<td>-.16*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chr Exp.</td>
<td>-.06</td>
<td>-.03</td>
<td>.12</td>
<td></td>
</tr>
</tbody>
</table>

N=199 for all variables except GPA. N=196 for GPA. * p<.05 *** p<.001

Table 53
Hierarchical Regression Analysis Regressing Defining Issues Test Stage 3 on GPA and Currently Meaningful Christian Personal Religious Experience

<table>
<thead>
<tr>
<th>IVs in model</th>
<th>Cum R²</th>
<th>F</th>
<th>df</th>
<th>IVs Incremental in R²</th>
<th>F</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA</td>
<td>.029</td>
<td>5.45*</td>
<td>1, 194</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>GPA, Christian Experience</td>
<td>.030</td>
<td>3.02*</td>
<td>2, 193</td>
<td>.001</td>
<td>0.24</td>
<td>1, 193</td>
</tr>
</tbody>
</table>

Subjects who did not complete all measures used in these regression analyses were dropped. The resulting N was 196. Dash indicates not applicable. * p<.05
Table 54
Hierarchical Regression Analysis Regressing Defining Issues Test Stage 4 on GPA and Currently Meaningful Christian Personal Religious Experience

<table>
<thead>
<tr>
<th>IVs in model</th>
<th>Cum $R^2$</th>
<th>$F$</th>
<th>df</th>
<th>IVs added</th>
<th>Incr. $R^2$</th>
<th>$F$</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA</td>
<td>.025</td>
<td>4.70*</td>
<td>1, 194</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>GPA, Christian Experience</td>
<td>.026</td>
<td>2.52</td>
<td>2, 193</td>
<td>Exp. .001</td>
<td>0.10</td>
<td>1.193</td>
<td></td>
</tr>
</tbody>
</table>

Subjects who did not complete all measures used in these regression analyses were dropped. The resulting N was 196. Dash indicates not applicable.

* $p < .05$
Chapter V
Summary and Recommendations

The purpose of this study was to determine to what extent IQ and socioeconomic status may be mediating the negative association, which has been reported in most post high school samples, between conservative Christian beliefs and principled moral judgement as measured by the Defining Issues Test P score.

When denominational affiliation was categorized as Fundamentalist, Moderate, and Liberal, mean score on the Short Christian Orthodoxy Scale did not differ significantly among these groups, except when Agnostics/Atheists were added to the Liberals' group. Only Fundamentalists were significantly different (higher) on the Scriptural Literalism Scale when Agnostics/Atheists were excluded from the Liberals' group. When the Agnostics/Atheists were added to the Liberals group, then all three groups were significantly different from each other, with the Liberals plus Agnostics/Atheists' group scoring the lowest. Given that a literal interpretation of Scripture is a key marker differentiating Fundamentalists from other conservative Christians (e.g., some Evangelicals), this finding makes sense.

There were no significant differences among Fundamentalists, Moderates, and Liberals in IQ, thus failing to replicate earlier studies.
of elementary and secondary students. This difference may have been due to the fact that college students are likely to represent those with the highest IQs, thus restricting the range measured in this sample. Fundamentalists had significantly lower socioeconomic status than both Moderates and Liberals who did not differ significantly from each other. The above, regarding IQ and socioeconomic status, was true both with Liberals alone and when the Agnostics/Atheists were added to the Liberals' group.

A significantly low mean P score for this sample led to an intense scrutiny of the literature that exposed a pattern among the studies I was able to find in which reliable and usable data were reported on the association between conservative Christian beliefs and P score: with one exception, all the studies in which a significant negative correlation was reported between conservative Christian beliefs and P score also had P scores in the "average" range for college samples and appear to have been conducted in Minneapolis, Minnesota, a fairly cosmopolitan city where the Defining Issues Test research center is located. On the other hand, with the same one exception noted above, all the studies in which no significant negative correlation was reported between conservative Christian beliefs and P score exhibited P scores that were significantly below the average college P score (as reported by Rest, 1979b), and were conducted more recently in Indiana and Ohio, which are probably more conservative and conventional than Minneapolis, Minnesota. The one exception, which I refer to above, was a study in Indiana in which the mean P
score was below the average college sample, but also slightly above (Clouse, 1985; mean P score=33.33) the mean P score of studies where no significant negative correlation was reported between conservative Christian beliefs and P score. This mean P score was therefore in the middle, and may have represented the cutoff score at which the mean P score is high enough for there to be a significant negative correlation with conservative Christian beliefs, given similar sample sizes.

Thus, there may be a continuum of sample mean P scores along which conservative Christian beliefs correlate with P score in different but predictable ways. At the lowest end (e.g., Clouse 1991, mean P score=29.329, no SD reported, N=335) conservative Christian beliefs are significantly positively correlated \((r=.12)\) with P score. In the middle range, there are no significant zero-order correlations between conservative Christian beliefs and P score (e.g., Holley, 1991 mean P score=30.7, SD=12.51, zero-order correlation not reported, but labeled insignificant for N=119; and this study, mean P score=30.98, SD=16.63, \(r=-.05\), N=199). At the higher range, perhaps mean P score=33 and above (or higher if sample sizes are smaller), there is the potential for significant negative correlations between conservative Christian beliefs and P score. Of course, this could all be coincidence, random fluctuations of Meehl's "crude factor" (Cohen, 1994, p. 1000) that by chance formed into the above pattern. Or the pattern could be due to poor reliability among the particular samples that lacked statistically significant correlations, as noted above. Therefore, more research across secular campuses is needed to challenge my observation that
mean sample P score dictates the direction and amount of correlation between P score and measures of conservative Christian beliefs.

An alternate method for future research, recommended by Cohen (1994), would be to include regression coefficients rather than only report correlation coefficients to measure the relationship between P score and religious beliefs:

Unlike regression coefficients, correlations are subject to vary with selection as researchers change populations. . . . The major problem with correlations applied to research data is that they can not provide useful information on causal strength because they change with the degree of variability of the variables they relate. Causality operates on single instances, not on populations whose members vary. The effect of A on B for me can hardly depend on whether I'm in a group that varies greatly in A or another that does not vary at all. It is not an accident that causal modeling proceeds with regression and not correlation coefficients. (p. 1001)

None of the above studies reported regression coefficients regressing P score on measures of conservative Christian beliefs. But I suspect if they had, the same pattern would emerge: significantly positive regression coefficients at the lowest range of mean P score, insignificant trivial regression coefficients in the middle range, and significantly negative regression coefficients in the upper range. I say this because differences in standard deviations among samples are probably not creating the pattern of differences noted in correlations above. If the differences in correlations noted above were due solely to differences in standard deviations, then those with smaller
standard deviations (i.e., restricted range) would have smaller correlations and those with larger standard deviations would have larger correlations. But this is probably not the case in the data. My sample, for example, had a standard deviation of 16.63, which is larger than the average college samples' standard deviation of 13.2 based on N=1,734 (Rest, 1979a) so there was plenty of variability, yet a small correlation between measures of conservative Christian beliefs and P score.

Given my suspicion of the changing relationship between P score and conservative Christian beliefs at different ranges of mean sample P score, I would not generalize my findings beyond secular universities with mean sample P scores from 30 to 33, quite a small range. Given my low P score reliability, any correlations with independent variables may actually have been higher had the reliability of P score been better. There are a number of interesting results that need to be replicated, and may in fact generalize to other populations.

The Intrinsic Scale was a suppressor variable with the measures of conservative Christian beliefs. While Intrinsic Scale was positively correlated with P score, the measures of conservative Christian beliefs were negatively correlated with P score, and Intrinsic Scale was positively correlated with both measures of Christian beliefs. In the sample that included Agnostics/Atheists there was also an interaction between Intrinsic Scale and Short Christian Orthodoxy Scale in predicting P score, such that higher P scores were associated with
subjects who strongly rejected orthodox Christian beliefs and any commitment to Christianity, but also with subjects who were highly committed to Christianity as an end in itself and strongly endorsed orthodox Christian beliefs. Therefore, future studies would do well to include the Intrinsic Scale along with measures of conservative Christian beliefs, and to check for suppression and interactions. If the focus were on Agnostics/Atheists one would want to design a method to separate them, because Agnostics may account for some of the lower P scores, while Atheists account for some of the higher P scores.

The covariate set (socioeconomic status, year in school, and IQ) accounted for a significant amount of variance in P score. There was more evidence of redundancy for the covariate set and Scriptural Literalism Scale than for the Short Christian Orthodoxy Scale (again with the suppressor Intrinsic Scale in the model). The Short Christian Orthodoxy Scale showed more evidence of suppression with the covariate set. Because a number of the measures of Christian beliefs that correlated significantly with P score in past studies contained items that could distinguish fundamentalism from other forms of conservative Christianity (as does the Scriptural Literalism Scale, with its greater potential for partial redundancy than Short Christian Orthodoxy Scale), it could be that their correlations with P score would have decreased significantly had the covariate set of socioeconomic status, year in school, and IQ been entered before them in a regression model.
Because the sample in this study did not exhibit a significant negative correlation between P score and measures of conservative Christian belief, it was not possible to test the main hypotheses of this study. Again, it would be useful to run this study again on a sample with a significant positive correlation between P score and measures of conservative Christian beliefs, which I theorize would need to be a sample with a higher mean P score than this sample. I suspect the redundancy between measures of socioeconomic status and measures of fundamentalist beliefs would be greater for secular college samples with higher mean P scores than those with lower mean P scores.

It may be those with fundamentalist beliefs alone, and not all conservative Christians, who are responsible for the negative correlation between P score and measures of conservative Christian beliefs in post high school samples. This is consistent with the pattern of average mean P scores among Christian liberal arts colleges and below average mean P scores reported at a Bible college, which I assume to be more fundamentalist than the Christian liberal arts colleges. However, again there are differences in IQ and socioeconomic status likely between the Bible college and Christian liberal arts college students. Perhaps Christian liberal arts college subjects are scoring higher on P score than Bible college subjects and students with fundamentalist beliefs at secular colleges because of their higher socioeconomic status and higher IQ. Therefore, socioeconomic status and IQ should both be measured in future studies.
Regarding measures of the constructs, their psychometric qualities, and their relationships to each other, the following may be helpful. The Quick Word Test seems to be an adequate measure of IQ. Researchers should be aware that Quick Word Test scores tend to increase with years of education. I think I would need to carefully reconsider how best to measure socioeconomic status, were I to run this study again. It might better be measured by parental education and perhaps income, directly, or using Hollingshead's index, than by the Duncan Socioeconomic Index. Tierney (1979) reported a significant positive correlation between parental education level and P score, however, parental income was not significantly correlated to P score. Of course the six item form of the Defining Issues Test is preferable over the three item version, if one has the time to administer it. I would highly recommend obtaining Cronbach's coefficient alpha on the P score. It may be quite low as it was in this sample. Furthermore, I would recommend doing all one can to encourage subjects to take their time and carefully complete the Defining Issues Test. Finally, I would repeat Cohen's (1994) repetition of others: please report regression coefficients and standard deviations for samples. It makes appropriate comparisons among samples possible.
Appendix A
Defining Issues Test and Directions
Directions for "Opinions About Social Problems"

1. We are interested in your own opinions about controversial social issues. Different people have different opinions.

2. You should have ample time to finish. Consider every item carefully. But pace yourself to finish the entire packet within the hour.

3. After reading each story, decide what the main character should do. Then read the 12 issues provided and rate their importance from "Great" importance to "No" importance. After rating the 12 issues' importance, pick the top four and rank them from "Most Important" to "Fourth Most Important." (See sample case).

4. Any meaningless, nonsense items should be rated as having "No" importance.

5. If you do not understand a word in a story, ask me. If you do not understand a word in one of the 12 issues, I can't help you. If you have no idea what it means, then rate it low.

6. Rate and rank the 12 issues based on how important they are to making this social decision. Although many issues may be important, ask yourself if the decision should rest on that issue.
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.

In this questionnaire you will be asked to give your opinions about several stories. Here is a story as an example. We note that this is not really a social problem, but it will illustrate our instructions.

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. Below 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?

Instructions for Part A: (Sample Question)

On the left hand side check one of the spaces by each statement of a consideration. (For instance, if you think that statement #1 is not important in making a decision about buying a car, check the space on the right.)

**IMPORTANCE:**

<table>
<thead>
<tr>
<th>Great</th>
<th>Much</th>
<th>Some</th>
<th>Little</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Whether the car dealer was in the same block as where Frank lives. (Note that in this sample, the person taking the questionnaire did not think this was important in making a decision.)

2. Would a used car be more economical in the long run than a new car. (Note that a check was put in the far left space to indicate the opinion that this is an important issue in making a decision about buying a car.)

3. Whether the color was green, Frank's favorite color.

4. Whether the cubic inch displacement was at least 200. (Note that if you are unsure about what "cubic inch displacement" means, then mark it "no importance.")

5. Whether the front fenders were differential. (Note that if a statement sounds like gibberish or nonsense to you, mark it "no importance.")

Instructions for Part B: (Sample Question)

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. (Note that the top choices in this case will come from the statements that were checked on the far left-hand side--statements #2 and #5 were thought to be very important. In deciding what is the most important, a person would re-read #2 and #5, and then pick one of them as the most important, then put the other one as "second most important," and so on.)

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>5</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Copyright, James Rest, 1979, All rights reserved.
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 $1,000, 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

IMPORTANCE:

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

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

Most Important
Second Most Important
Third Most Important
Fourth Most Important
A man has 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?

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

**IMPORTANCE:**

<table>
<thead>
<tr>
<th>Importance</th>
<th>Great</th>
<th>Much</th>
<th>Some</th>
<th>Little</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Hasn't Mr. Thompson been good enough for such a long time to prove he isn't a bad person?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Ever since someone escapes punishment for a crime, doesn't that just encourage more crime?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Wouldn't we be better off without prisons and the oppression of our legal systems?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Has Mr. Thompson really paid his debt to society?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Would society be failing what Mr. Thompson should fairly expect?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. What benefits would prisons be apart from society, especially for a charitable man?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. How could anyone be so cruel and heartless as to send Mr. Thompson to prison?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Would it be fair to all the prisoners who had to serve out their full sentences if Mr. Thompson was let off?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Was Mrs. Jones a good friend of Mr. Thompson?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Wouldn't it be a citizen's duty to report an escaped criminal, regardless of the circumstances?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. How would the will of the people and the public good best be served?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Would going to prison do any good for Mr. Thompson or protect anybody?</td>
<td></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
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 use of the military in international disputes 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 should 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 as 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

### IMPORTANCE:

<table>
<thead>
<tr>
<th>Great</th>
<th>Much</th>
<th>Some</th>
<th>Little</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

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

- Most Important
- Second Most Important
- Third Most Important
- Fourth Most Important
Appendix B

Examples of Principled Level Christian Teachings
An example of principled level thinking based on Christian teachings is provided by Yoder (1983), a Mennonite theologian who presents a Christian response to the dilemma commonly posed to pacifists: "What would you do if a criminal with a gun were about to kill your (wife, daughter, mother, etc.)?" Christian pacifists who say they would refuse to kill the attacker are often confronted with: "Perhaps as a Christian you do have the right to sacrifice your own welfare to be loving toward an attacker. But do you have the right to sacrifice the welfare of others for whom you are responsible?" (Yoder, 1983, p.20)

Yoder responds to this challenge as follows:

We must pierce through the screen of this apparent altruism and point out that it distorts the real nature of the argument. It is an altruistic form of egoism when I defend my wife or my child because they are precisely my own. This argument does not suggest that I would have the same responsibility to defend the wives and children of Vietnamese, for example, who are being attacked by my countrymen. It does not suggest any special concern for the wife or child of the attacker. The reason I should defend my wife and child in this argument is not that they are my neighbors, innocent threatened third parties, but because they are mine. Thus this becomes an act of selfishness; though covered over with the halo of service to others, it is still self-oriented in its structure.

Now self-centeredness is not all bad; "thou shalt not seek thy own interest" is not a generally accepted moral axiom. In fact, it can be argued that a certain amount of self-love is necessary for psychological health, and for one to take care of what has been entrusted to oneself. But to make self-centeredness or egoism the basis on which one chooses how to respond in any situation is not a Christian approach to a problem. Christianity relativizes the value of self and survival
as it affirms the dignity of the enemy and offender. True, the potential victim is my neighbor and deserving of my help. But the attacker also becomes at that moment a neighbor, and any attempt to distinguish between these two and say that the nearness of my family member as preferred neighbor takes precedence over that of my attacker is also a form of egoism. Again this cannot be sufficient basis for Christian ethical decision-making. (pp. 20-21)

Yoder's (1983) answer fulfills the criteria of Kohlberg's stages 5 and 6 (Rest, 1983). It is a "Prior-to-society perspective" because it is the "perspective of a rational individual aware of values and rights prior to social attachments and contracts" (Stage 5, p. 577). It also is the "perspective of a moral point of view from which social arrangements derive" because it may be considered the perspective of "any rational individual recognizing the nature of morality or the fact that persons are ends in themselves and must be treated as such" (Stage 6, p. 577). (While "any" rational person may be too comprehensive, it is at least the perspective of "a" rational person--Yoder.) Yoder's perspective on the teaching of Christ is also consistent with Rest's Stage 6 scheme of balancing interests:

A scheme of cooperation that negates or neutralizes all arbitrary distribution of rights and responsibilities is the most equilibrated, for such system is maximizing the simultaneous benefit to each member so that any deviation from these rules would advantage some members at the expense of others. (Rest 1979, p 23)
In addition to this example of Mennonite (a conservative Christian denomination) thinking, many examples of principled level teachings are found in the Bible. For example in John 7:24 (New International Version) Jesus confronted those of his day for making moral judgements at level 4 (law and order) and recommended principled level thinking, "Stop judging by mere appearances, and make a right judgement." He said this in response to the Jews being angry that he had healed a man on the Sabbath, which apparently some Jews considered a form of work and therefore against the law. Jesus conceptualized his Sabbath healing in terms of fulfilling the law by looking at the principle behind the law requiring no work be done on the Sabbath. It was to be a day of rest to benefit humans (similar to his teaching in Mark 2:27 [New International Version] that "the Sabbath was made for man, not man for the Sabbath"). At the same time Jesus did not deny that the Sabbath was to be a time set aside to honor God. Clouse (1985) wrote:

Interestingly, Jesus emphasized both the conventional and the postconventional and showed that the two are closely related. He explained in the Sermon the Mount that he did not come to destroy the law but to fulfill the law (Matthew 5:17). He then proceeded to add a principled statement to the mandates of the law. . . . The law regulates behavior; a principle monitors the cognitions that precede behavior. Both are important in the Christian life. (p. 196)
Appendix C
Quick Word Test
(First Five Items Only)
QUICK WORD TEST
Edgar F. Borgatta       Raymond J. Corsini

Directions: Fill in the answer space next to the word that means the same as the first word. If you do not know, GUESS. Work quickly. Answer all questions.

SAMPLE: happy dull () seem () glad () fast ()

1 chant ( ) dire ( ) bend ( ) ring ( ) sing ( )
2 talon ( ) hold ( ) coin ( ) claw ( ) peak ( )
3 salve ( ) mild ( ) ease ( ) oleo ( ) soft ( )
4 rivet ( ) bolt ( ) flow ( ) tray ( ) part ( )
5 brook ( ) rill ( ) fish ( ) tile ( ) pool ( )
Appendix D

Three Religious Scales and a Questionnaire

including:

Age Universal I/E-Revised Scales (items 1-14),
Short Christian Orthodoxy Scale (items 15-20),
Scriptural Literalism Scale (items 21-44)
and Demographics Questionnaire (items 45-62)
This survey includes a number of statements related to specific religious beliefs. You will probably find that you agree with some of the statements, and disagree with others, to varying extents. Please mark your opinion by circling the number below each statement according to the amount of your agreement or disagreement. Some scales will have more numbers from which to choose than others.

**Strongly Disagree**  **Neutral**  **Strongly Agree**

1. I enjoy reading about my religion.
   1  2  3  4  5

2. I go to church because it helps me to make friends.
   1  2  3  4  5

3. It doesn't matter what I believe so long as I am good.
   1  2  3  4  5

4. It is important to me to spend time in private thought and prayer.
   1  2  3  4  5

5. I have often had a strong sense of God's presence.
   1  2  3  4  5

6. I pray mainly to gain relief and protection.
   1  2  3  4  5

7. I try hard to live all my life according to my religious beliefs.
   1  2  3  4  5

8. What religion offers me most is comfort in times of trouble and sorrow.
   1  2  3  4  5

9. Prayer is for peace and happiness.
   1  2  3  4  5

10. Although I am religious, I don't let it affect my daily life.
    1  2  3  4  5

11. I go to church mostly to spend time with my friends.
    1  2  3  4  5

12. My whole approach to life is based on my religion.
    1  2  3  4  5

13. I go to church mainly because I enjoy seeing people I know there.
    1  2  3  4  5

14. Although I believe in my religion, many other things are more important in life.
    1  2  3  4  5

15. Jesus Christ was the divine Son of God.
    1  2  3  4  5  6  7

16. The Bible may be an important book of moral teachings, but it was no more inspired by God than were many other such books in human history.
    1  2  3  4  5  6  7

17. The concept of God is an old superstition that is no longer needed to explain things in the modern era.
    1  2  3  4  5  6  7
<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Neutral</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>18. Through the life, death, and resurrection of Jesus, God provided a way for the forgiveness of people's sins.</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>19. Despite what many people believe, there is no such thing as a God who is aware of our actions.</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>20. Jesus was crucified, died, and was buried but on the third day He arose from the dead.</td>
<td>1 2 3 4 5 6 7</td>
<td></td>
</tr>
<tr>
<td>21. The scriptures contain religious truths.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>22. The scriptural writers were divinely inspired.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>23. Life originated differently than suggested by the scriptures.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>24. The &quot;rules&quot; or &quot;laws&quot; found in the scriptures were conceived by man.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>25. Quotations appearing in the scriptures are accurate.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>26. We can put our trust in the teachings of the scriptures.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>27. The scriptures have only one interpretation.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>28. The miracles reported in the scriptures actually occurred.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>29. The scriptural account of creation is accurate.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>30. The scriptures describe real people.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>31. The scriptures are the ultimate truth.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>32. The scriptures should be regarded more as beautiful writing than as religious truths.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>33. The precise words spoken by God may be found in the scriptures.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>34. Some events mentioned in the scriptures never occurred.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>35. The scriptures should be taken as divinely-inspired writings.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
</tbody>
</table>
Strongly Disagree  Neutral  Strongly Agree
36. The scriptures contain God's rules for living.
   1  2  3  4  5
37. The passage of time is accurately presented in the scriptures.
   1  2  3  4  5
38. The scriptures are a product of man's imagination.
   1  2  3  4  5
39. There are no errors in the scriptures.
   1  2  3  4  5
40. There are more accurate accounts of history than the scriptures.
   1  2  3  4  5
41. Most of the writing in the scriptures should be taken literally.
   1  2  3  4  5
42. Many events from early history are recorded accurately in the scriptures.
   1  2  3  4  5
43. The scriptures are a collection of myths.
   1  2  3  4  5
44. The scriptures accurately predict future events.
   1  2  3  4  5

Please answer the following questions as honestly and completely as possible. All answers will be anonymous.

45. What is your age? _________
46. What is your gender? M  F (Circle one.)
47. What is your race? (Choose the most appropriate)
   American Indian
   African American
   Asian
   Hispanic
   White
   Other (specify) _________
48. Circle your year in college
   Fr.  Soph.  Jr.  Sr.  Other (Specify) __________________________
49. What is your current GPA (on a 4.00 scale)? _________
50. What level of education did your father complete? (Circle one.)
   grade school  junior high school  senior high school  college  graduate school
51. What level of education did your mother complete? (Circle One.)
   grade school  junior high school  senior high school  college  graduate school

For numbers 52 and 53, be as specific as you can and list the type of industry or business. Examples:
stock clerk or postal clerk (not just "clerk"); secretary, legal or medical, etc. (not just "secretary"); professor
of Engineering (not just "college professor"); sales representative, wholesale trade (not just "sales rep.")

52. What is your father's occupation? __________________________________________________________

53. What is your mother's occupation? _________________________________________________________

54. What is your father's religious affiliation? Please write religion and denomination (If appropriate).
   e.g. United Methodist, Episcopal, Missouri-Synod Lutheran, Assemblies of God, Orthodox Jew,
   Muslim, etc.). Write "none", "atheist", or "agnostic" if appropriate.

55. What is your mother's religious affiliation? Please write religion and denomination (If appropriate).
   e.g. United Methodist, Episcopal, Missouri-Synod Lutheran, Assemblies of God, Orthodox Jew,
   Muslim, etc.). Write "none", "atheist", or "agnostic" if appropriate.

56. What is your religious affiliation? Please write religion and denomination (If appropriate).
   e.g. United Methodist, Episcopal, Missouri-Synod Lutheran, Assemblies of God, Orthodox Jew,
   Muslim, etc.). Write "none", "atheist", or "agnostic" as appropriate.

57. How often do you attend religious activities (including Christian youth group meetings, Sunday
   morning worship services, non-Sunday morning Bible studies, church meetings, etc.)? Circle the most
   correct (only one) response.

4+ . . . . . 3 .... 2 .... 1 .... time(s) per week.
or 4 . . . . . 3 .... 2 .... 1 .... time(s) per month.
or less than once a month, but at least . . .
11 10 9 8 7 6 5 4 3 2 1 time(s) a year.
or less than once a year.

58. Mark which of the following statements comes closest to your religious experience:
   (a) I have never had a personal religious experience.
   (b) I have had a personal religious experience but it doesn't mean much to me now.
   (c) I have had a personal religious experience which is very important to me now.

59. If you marked (b) or (c) above, please mark (a) or (b) below:
   (a) My personal religious experience is with Christianity.
   (b) My personal religious experience is with a religion other than Christianity.

60. Is English your first, native-spoken language? Yes____ No____

61. Were you born in the United States? Yes____ No____

62. If you marked "No," where were you born? __________________________________________________
Appendix E

Socioeconomic Index Coding Guidelines
Socioeconomic Index Coding Guidelines

1a. If retired then write "retired," unless retired armed forces (see 580).
1b. If disabled then write "disabled."
1c. If student then write "student."
1d. If deceased, put "deceased."
1e. If blank, ?, unknown, or self-employed with no further information put "unknown."

2. If teacher, assume elementary school teacher unless elaborated. Choose secondary school teacher (144) if a or b is true:
   a) junior high or high school - "secondary school teacher"
   b) specific subject taught is included in the description (e.g. "English teacher") - "secondary school teacher"
   c) Substitute teacher - "Teachers, except college and universities, n. e. c." (145).
Choose College teacher if that is indicated or the word professor is used.

3. Use your discretion with phrases such as "works in/on ...". If it seems reasonable to assume clerk or laborer, then do so. For example, for "works on the farm" one could assume farm laborer (822).

4. If an occupation is clearly within a given level, but has no exact title, then use one of the "allocated" answers. For example, Professional, technical and kindred workers-allocated (196), etc. When you must choose one, use the "allocated" code instead of the "n. e. c." or "miscellaneous" codes. **The one exception** is the managerial level. In general, use "Managers and administrators, n. e. c." (245) instead of "allocated" (246). Only use "allocated" for the managerial level when the staff being managed are all below SEI-25, e.g. "Supervises truck drivers". If subject answers only "manager," assign SEI-52.90 (the average SEI of the two management "allocated" categories, 245 and 246).

5. If "executive," "owns his own business," or "self-employed" and seems to be managing the operation score as "Managers and administrators, n. e. c." (245)
6. For maintenance man use "Craftsmen and kindred workers-allocated." (586).

7. For maintenance supervisor use "Managers and superintendents, building" (216).
Appendix F

Oral Instructions to Subjects
Oral Instructions to Subjects
(five minutes)

Hello. This is experiment CK-1—Attitudes and Experiences. Is everyone here for the correct experiment? Before I begin, I would like to thank you for helping with this study. It will help me graduate. If, at any time, you become uncomfortable with the nature of the questions, and would choose not to continue, please bring your packet to me. I will sign your experiment card. You will not be penalized. Before we begin, please put your experiment cards in front of you on the desk. I will come around and sign them while you work.

This study involves answering some questionnaires. Please work as steadily as possible, but try to answer all of the questions in the packet thoughtfully and honestly. Be sure to choose one of the possible answers. Do not create your own answer.

When you have completed the packet, bring it up to me and set it in this box. Return your pencils to the plastic box. Then pick up a debriefing sheet describing the purpose of this study as you leave. It will explain why each questionnaire was in the packet and I would recommend you read it.

Please complete all portions of the packet in order, from beginning to end. Before you begin, I would like to provide more details about two sections of the questionnaire. First, please look at the Directions for "Opinions About Social Problems" and follow along.

#1. We are interested in your own opinions about controversial social issues. Different people have different opinions.

#2. You should have ample time to finish. Consider every item carefully. But pace yourself to finish the entire packet within the hour.

#3. After reading each story, decide what the main character should do. Then read the 12 issues provided and rate their importance from "Great" importance to "No" importance. After rating the 12 issues' importance, pick the top four and rank them from "Most Important" to
"Fourth Most Important. Let's look at the sample case on the next page.

There are three reasons why you would rate an item of "No" importance. First, if it seems irrelevant, such as item #1. Second, if you are unsure of the meaning of the item, such as item #4. This person didn't know what a cubic inch displacement of 200 meant so he marked it low. You should do the same if you don't understand something. And third, if an item is meaningless nonsense, such as item #6. Several items in the Opinions about Social Problems questionnaire will be meaningless nonsense, just made up words, or totally irrelevant. So if you find yourself asking what the heck this item means, be sure to mark it "No" for no importance. Turn back to the first page.

#4. Meaningless, nonsense items should be rated as having "No" importance.

#5. If you do not understand a word in a story, ask me. If you do not understand a word in one of the 12 issues, I can't help you. If you have no idea what it means, then rate it low. By the way, the actual stories have 12 issues, but the sample story only has 6.

#6. Rate and rank the 12 issues based on how important they are to making this social decision. Although many issues may be important, ask yourself if the decision should rest on that issue.

Now turn to the Quick Word Test, Form A. It's on the 6th page. If you do not know the meaning of a word, GUESS. Work quickly—ANSWER ALL THE QUESTIONS." Since these Quick Word Test sheets will be computer scored, please fill in the space completely as in the sample (in other words, don't just check it). If you wish to change an answer you must erase the unwanted mark completely. Please fill in your age, sex, and class (in other words your year in school: sophomore, junior, senior, not "Psychology"). You may turn back to the first page in your packet now.

Near the end of your packet, there will be a question about parental occupation (you don't have to look now) -- what do your parents do for a living. Please be specific. For example, if they are
self-employed, self-employed what? Do they cut hair, mow lawns, are they accountants? If they are managers, do they manage a newspaper stand or a large department store? Please be specific.

One more thing, (point to a rough sketch of DIT chart and ranking system on blackboard) on the Opinions About Social Problems Questionnaire, when ranking the four most important items, choose those from the left side before those on the right side. Ranking those on the right side above those on the left side doesn't make any sense. But I've seen it done before.

Does anyone have any questions? (wait . . . ) Does everyone have two #2 pencils and an eraser? If you need to sharpen your pencils at any time, I have a sharpener up here at the desk. You may begin now.
Appendix G
Debriefing
Debriefing

Thank you for your help. The purpose of this study is to determine the relationship between one's moral judgment and the degree and type of one's Christian beliefs and experiences. Christian beliefs or attitudes may range from very liberal to very conservative. The different religious questionnaires measured religious beliefs -- the degree of belief and type of commitment to those beliefs. The stories, ratings, and rankings measured moral judgment.

One factor that is moderately related to moral judgment, as defined in this study, is cognitive ability. Thus, you completed the "Quick Word Test". By the way, this word test was quite difficult -- first and second year college students, on average, only select one half of the correct answers. But that gives a greater degree of variability among students' scores than an easy word test. And greater variability in scores improves a measures' ability to correlate with other measures. So don't feel bad if you only got about half right. You were not supposed to do much better than that.

If you would like to know more about the nature of this study, please see me afterwards. I would be glad to discuss it with you further. If you would like a summary of the results of this study, please leave your name and campus address (or place a self-addressed, stamped envelope in my mail box in the counseling psychology student room right across the hall from Townshend 127 if you won't have a campus address this summer) so I can send you a copy when it is finished.

Please do not discuss this study with anyone as it is important that no bias occur among potential subjects.
Appendix H

Table 55

Frequency Table of Subject Denomination
Table 55
Frequency Table of Subject Denomination

<table>
<thead>
<tr>
<th>Subject's Denomination</th>
<th>Frequency</th>
<th>Percent</th>
<th>Cumulative Frequency</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>African Methodist Episcopal Church</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Apostolic Christian</td>
<td>2</td>
<td>1.0</td>
<td>3</td>
<td>1.5</td>
</tr>
<tr>
<td>Baptist, Don't know which</td>
<td>12</td>
<td>6.0</td>
<td>15</td>
<td>7.5</td>
</tr>
<tr>
<td>Bible teaching, any Christian Church</td>
<td>1</td>
<td>0.5</td>
<td>16</td>
<td>8.0</td>
</tr>
<tr>
<td>Catholic</td>
<td>65</td>
<td>32.7</td>
<td>81</td>
<td>40.7</td>
</tr>
<tr>
<td>Church of Christ</td>
<td>4</td>
<td>2.0</td>
<td>85</td>
<td>42.7</td>
</tr>
<tr>
<td>Coptic Orthodox</td>
<td>2</td>
<td>1.0</td>
<td>87</td>
<td>43.7</td>
</tr>
<tr>
<td>Episcopal</td>
<td>4</td>
<td>2.0</td>
<td>91</td>
<td>45.7</td>
</tr>
<tr>
<td>Evangelical Lutheran</td>
<td>1</td>
<td>0.5</td>
<td>92</td>
<td>46.2</td>
</tr>
<tr>
<td>Evangelical, Evangelist</td>
<td>1</td>
<td>0.5</td>
<td>93</td>
<td>46.7</td>
</tr>
<tr>
<td>Greek Orthodox</td>
<td>1</td>
<td>0.5</td>
<td>94</td>
<td>47.2</td>
</tr>
<tr>
<td>Lutheran Church, Missouri Synod</td>
<td>1</td>
<td>0.5</td>
<td>95</td>
<td>47.7</td>
</tr>
<tr>
<td>Lutheran, Don't know which</td>
<td>11</td>
<td>5.5</td>
<td>106</td>
<td>53.3</td>
</tr>
</tbody>
</table>
### Table 55 (continued)

<table>
<thead>
<tr>
<th>Subject's Denomination</th>
<th>Frequency</th>
<th>Percent</th>
<th>Cumulative Frequency</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mennonite</td>
<td>1</td>
<td>0.5</td>
<td>107</td>
<td>53.8</td>
</tr>
<tr>
<td>Methodist, Don't know which</td>
<td>2</td>
<td>1.0</td>
<td>109</td>
<td>54.8</td>
</tr>
<tr>
<td>Non-denominational</td>
<td>12</td>
<td>6.0</td>
<td>121</td>
<td>60.8</td>
</tr>
<tr>
<td>Pentecostal</td>
<td>1</td>
<td>0.5</td>
<td>122</td>
<td>61.3</td>
</tr>
<tr>
<td>Presbyterian, Don't know which</td>
<td>8</td>
<td>4.0</td>
<td>130</td>
<td>65.3</td>
</tr>
<tr>
<td>Protestant</td>
<td>3</td>
<td>1.5</td>
<td>133</td>
<td>66.8</td>
</tr>
<tr>
<td>Quaker</td>
<td>1</td>
<td>0.5</td>
<td>134</td>
<td>67.3</td>
</tr>
<tr>
<td>Southern Baptist Convention</td>
<td>1</td>
<td>0.5</td>
<td>135</td>
<td>67.8</td>
</tr>
<tr>
<td>United Church of Christ</td>
<td>6</td>
<td>3.0</td>
<td>141</td>
<td>70.9</td>
</tr>
<tr>
<td>United Methodist Church</td>
<td>16</td>
<td>8.0</td>
<td>157</td>
<td>78.9</td>
</tr>
<tr>
<td>The Way Ministry</td>
<td>1</td>
<td>0.5</td>
<td>158</td>
<td>79.4</td>
</tr>
<tr>
<td>Wesleyan</td>
<td>1</td>
<td>0.5</td>
<td>159</td>
<td>79.9</td>
</tr>
<tr>
<td>Agnostics/Atheists</td>
<td>40</td>
<td>20.1</td>
<td>199</td>
<td>100.0</td>
</tr>
</tbody>
</table>
Appendix I
Selected Raw Regression Equations and Standard Errors
not retrievable from results presented above
SCO-Short Christian Orthodoxy Scale
SLS-Scriptural Literalism Scale
I-Intrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales
E-Extrinsic Scale (reverse scored) from the Age Universal I/E-Revised Scales
$SE$-Standard Error for the variable in subscript

Equations from Sample including Agnostics/Atheists ($N=199$)

(15) $P$ score = $-0.63 (X_{SCO}) + 34.55; SE_{SCO} = 0.82$
(16) $P$ score = $-1.10 (X_{SLS}) + 34.58; SE_{SLS} = 1.53$
(17) $P$ score = $3.05 (X_I) + 21.67; SE_I = 1.64$
(18) $P$ score = $-2.97 (X_E) + 38.56; SE_E = 1.63$
(19) $P$ score = $4.82 (X_I) - 4.75 (X_E) + 28.38; SE_I = 1.74, SE_E = 1.73$
(20) $P$ score = $0.56 (X_I) - 9.99 (X_E) + 1.79 (X_{EX_I}) + 40.42$

$SE_I = 5.59, SE_E = 6.70, SE_{EX_I} = 2.22$

(21) GPA = $.017 (X_{IQ}) + .149 (X_I) + 1.38; SE_{IQ} = 0.003, SE_I = 0.053$

Equations from Sample excluding Agnostics/Atheists ($N=159$)

(22) $P$ score = $0.37 (X_{SCO}) + 28.54; SE_{SCO} = 1.30$
(23) $P$ score = $0.63 (X_{SLS}) + 28.61; SE_{SLS} = 2.05$
(24) $P$ score = $5.07 (X_I) + 14.47; SE_I = 2.03$
(25) $P$ score = $-2.35 (X_E) + 37.13; SE_E = 2.17$
(26) $P$ score = $5.57 (X_I) - 3.27 (X_E) + 21.61; SE_I = 2.04, SE_E = 2.15$
(27) $P$ score = $5.25 (X_I) - 3.68 (X_E) + 0.13 (X_{EX_I}) + 22.63$

$SE_I = 7.66, SE_E = 9.73, SE_{EX_I} = 2.96$
Appendix J

Regression Equations and Coordinates used to Graph Figures 3 and 5
Equations and Coordinates for Figure 3 (without Interaction Term)

$X_1$ is Short Christian Orthodoxy Scale (SCO)
$X_2$ is Intrinsic Scale (I)

(28) $Y = -2.85 (X_1) + 5.33 (X_2) + 29.85$

SCO: mean=5.69, SD=1.43

-1SD Low (SCO=4.26):

(29) $Y_L = -2.85 (4.26) + 5.33 (X_2) + 29.85$

(30) $Y_L = 5.33 (X_2) + 17.71$

Mean Average (SCO=5.69):

(31) $Y_A = -2.85 (5.69) + 5.33 (X_2) + 29.85$

(32) $Y_A = 5.33 (X_2) + 13.63$

Almost

+1SD High (SCO=7.00):

(33) $Y_H = -2.85 (7.00) + 5.33 (X_2) + 29.85$

(34) $Y_H = 5.33 (X_2) + 9.90$

Table 56

Coordinates for Graphing Three Lines demonstrating P.Score Regression on Intrinsic Scale Moderated by Short Christian Orthodoxy Scale, without Interaction, for Sample Including Agnostics/Atheists

<table>
<thead>
<tr>
<th></th>
<th>Low (SCO = 4.26)</th>
<th>Average (SCO = 5.69)</th>
<th>High (SCO = 7.00)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_2$</td>
<td>P score</td>
<td>$X_2$ P score</td>
<td>$X_2$ P score</td>
</tr>
<tr>
<td>0</td>
<td>17.71</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>44.36</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40.28</td>
<td>36.55</td>
</tr>
</tbody>
</table>
Equations and Coordinates for Figure 5 (without Interaction Term)

$X_1$ is Short Christian Orthodoxy Scale (SCO)
$X_2$ is Intrinsic Scale (I)

(28) $Y = -2.85 (X_1) + 5.33 (X_2) + 29.85$
I: mean=3.06, SD=0.73

-1SD Low (I=2.33):
(35) $Y_L = -2.85 (X_1) + 5.33 (2.33) + 29.85$
(36) $Y_L = -2.85 (X_1) + 42.27$

Mean Average (I=3.06):
(37) $Y_A = -2.85 (X_1) + 5.33 (3.06) + 29.85$
(38) $Y_A = -2.85 (X_1) + 46.16$

+1SD High (I=3.79):
(39) $Y_H = -2.85 (X_1) + 5.33 (3.79) + 29.85$
(40) $Y_H = -2.85 (X_1) + 50.05$

Table 57

Coordinates for Graphing Three Lines demonstrating P Score Regression on Short Christian Orthodoxy Scale Moderated by Intrinsic Scale, without Interaction Term, for Sample Including Agnostics/Atheists

<table>
<thead>
<tr>
<th></th>
<th>Low (I=2.33)</th>
<th>Average (I=3.06)</th>
<th>High (I=3.79)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td></td>
<td>$X_1$</td>
<td>$X_1$</td>
</tr>
<tr>
<td>0</td>
<td>P score 42.27</td>
<td>P score 46.16</td>
<td>P score 50.05</td>
</tr>
<tr>
<td>7</td>
<td>P score 22.32</td>
<td>P score 26.21</td>
<td>P score 30.10</td>
</tr>
</tbody>
</table>
List of References


Buier, R. M., Butman, R. E., Burwell, R., & Van Wicklin, J. (1989). The critical years: Changes in moral and ethical decision-making in


Rest, J. (1972). *Devising and validating an objective test of moral judgment*. Unpublished manuscript, University of Minnesota, Minneapolis, MN.


