

MACROMORALITY AND MORMONS: A PSYCHOMETRIC INVESTIGATION
AND QUALITATIVE EVALUATION OF THE
DEFINING ISSUES TEST-2

by

Daniel R. Winder

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Instructional Technology

Approved:

J. Nicholls Eastmond, Ph.D.
Major Professor (co-chair)

Joanne P. H. Bentley, Ph.D.
Committee Member

Matthew J. Taylor, Ph.D.
Major Professor (co-chair)

Michael K. Freeman, Ph.D.
Committee Member

Richard P. West, Ph.D.
Committee Member

Byron R. Burnham, Ed.D.
Dean of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2009

Copyright © Daniel R. Winder 2009

All Rights Reserved

ABSTRACT

Macromorality and Mormons: A Psychometric Investigation and
Qualitative Evaluation of the Defining Issues Test-2

by

Daniel R. Winder, Doctor of Philosophy

Utah State University, 2009

Major Professors: J. Nicholls Eastmond, Ph.D, Matthew J. Taylor, Ph.D.
Department: Instructional Technology

In 1988, P. Scott Richard's dissertation research at the University of Minnesota asserted that the Defining Issues Test (DIT), a widely accepted paper-and-pencil test of moral reasoning, exhibited item bias against religiously orthodox persons. Since 1988 (when Richard's data were reported), new methods of differential-item functioning (DIF) have developed, a new DIT has emerged (the DIT-2), as well as a Neo-Kohlbergian framework based upon moral schemas derived from Kohlberg's Piagetian-like six stages. With new methods, new tests, and unanswered questions, this study's results imply: (1) that DIT-2 items exhibit differential item functioning for religiously orthodox persons in statistically significant but not as practically significant ways as Richards' earlier findings, (2) that religious orthodoxy does influence macromoral reasoning as measured by the DIT-2, (3) that the maintaining norms schema is insufficient to explain the variables that contribute to why religiously orthodox persons score the way they do. This

study implies that the maintaining norms schema may be misnamed because it appears to be measuring a different construct than maintaining norms macromoral reasoning.

(197 pages)

ACKNOWLEDGMENTS

I first want to acknowledge my wife and family for their encouragement and support. Dr. J. Nicholls Eastmond has been an inspiring example of how intellectualism, humility, and excellence can be melded in academia. Dr. Matthew J. Taylor has been invaluable in helping me clearly articulate, synthesize, and edit my writing. My entire committee has helped me to learn to work with others in a spirit of fellowship. I also wish to acknowledge the many professors from Brigham Young University's Instructional Psychology and Technology Department, Religious Education Department, and Institutional Assessment Department, including Dr. Richard Sudweeks, Dr. Dennis Wright, Dr. Doug Brinley, and Dr. Eric Jensen for each contributing in significant ways to this study. In addition, the funding for this study was made possible by a grant from Brigham Young University's Religious Studies Center. Finally, I wish to acknowledge my supervisors in the Latter-day Saint Church Educational System for being so accommodating and encouraging in helping me meet my higher education goals.

Daniel R. Winder

CONTENTS

	Page
ABSTRACT.....	iii
ACKNOWLEDGMENTS	v
LIST OF TABLES	viii
LIST OF FIGURES	x
CHAPTER	
I. INTRODUCTION	1
II. REVIEW OF LITERATURE	5
The Concept of MMR in the DIT	6
The Measurement of MMR in the DIT	6
Uses of the DIT-2.....	9
Confounding Variables of the DIT MMR Scores.....	10
Religious Orthodox Variables Confound DIT Scores	12
Theories of Why Religiously Orthodox Persons Score Differently on the DIT	17
A Call for Replication Studies on a New Defining Issues Test.....	22
Conclusion and Transition to Study.....	24
III. METHODS	28
Purpose of Study and Research Questions.....	28
Independent, Dependent, and Moderating Variables.....	28
Research Design.....	30
Population and Sample	30
Instrumentation	32
Data Collection Procedures.....	35
Pilot Studies	37
Data Analysis	40
Conclusion	52
IV. RESULTS	53
Analysis #1 Descriptive Statistics Description of Sample.....	54

	Page
Analysis #2 Correlations Between RO and MMR.....	59
Analysis #3 Differential Item Functioning Person Reliability	62
Analysis #4 <i>t</i> Tests and Correlations for Other Potential Confounds with MMR.....	67
Analysis #5 Partial Correlations Between RO and MMR Accounting for Other Potential Confounds.....	75
Analysis #6 Qualitative Analysis of Open-Ended Survey Responses	76
Overall Qualitative Analysis of the Differences Between BYU Sample Participants.....	86
 V. DISCUSSION.....	 90
Purpose of Study and Research Questions.....	90
Summary of Major Findings.....	91
Discussion Points	95
Delimitations.....	117
Suggestions for Further Research	119
Conclusion	123
 REFERENCES	 127
 APPENDICES	 140
Appendix A: History of the DIT.....	141
Appendix B: Three Pilot Studies and Qualitative Interview Notes.....	146
Appendix C: Demographic and Survey Questions.....	157
Appendix D: Informed Consent Document.....	159
Appendix E: Clarifying the DIT-1 Instructions	162
Appendix F: DIT-2 Story Moral Dilemmas.....	165
Appendix G: Qualitative Data Analysis and Tables.....	168
 VITA.....	 183

LIST OF TABLES

Table	Page
1. Findings and Methodological Characteristics of Studies That Examined the Relation Between Principled Moral Reasoning (P-score) and Religiously Orthodox Ideology (RO).....	15
2. Rationale for the Choice of Variables for Testing	25
3. A Comparison of a Measure of Religious Orthodoxy by School (Brown and Low Inventory of Religious Belief).....	39
4. Linacre (Winsteps) and ETS Guidelines for Interpreting DIF Contrast	48
5. Descriptive Variable Frequencies	54
6. DIT-2 Score Descriptions Used in Chapter IV	56
7. Descriptive Statistics Comparing a BYU Average Scores to National Norms	57
8. Descriptive Statistics by Educational Level	58
9. Pearson Correlation—Religious Orthodoxy and DIT-2 Scores	60
10. DIF Contrast Charts with Statistical and Practical Significance.....	64
11. DIT-2 Scores Relationships with Demographic Variables	68
12. BYU Sample DIT-2 Story Choice Correlations.....	72
13. Statistically Significant DIT-2 Mean Differences of Those Who Endorsed a Doctor-Assisted Overdose Compared to Those Who Did Not.....	74
B-1. A Comparison of DIT Internal Measures of Religious Orthodoxy by School.....	149
B-2. A Comparison of a Measure of Religious Conservatism by School (Brown and Low Inventory).....	150
G-1. Cancer Story Codes for Qualitative Data	170

Table	Page
G-2. Relative Percent Counts of Coded Comments by DIT-2 Cancer Story Decision	172
G-3. Coded Comments and Category Examples/Descriptions.....	173
G-4. Description of Coded Comments in DIT-2 Famine Story Qualitative Comments.....	175
G-5. Relative Percent Counts of Coded Comments by DIT-2 Famine Story Decision.....	176
G-6. Reporter Story Decision Coded Comments and Category Examples/ Descriptions.....	177
G-7. Relative Percent Counts of Coded Comments by DIT-2 Reporter Story Decision and Category Examples/Descriptions	178
G-8. Coded Comments from Story #3 with Category Examples/Descriptions.....	179
G-9. Relative Percent of Coded Comments by Story #3 Decision and Category Examples/Descriptions	180
G-10. Illustrative Examples of Themes in the Cancer Story Comments.....	181
G-11. Illustrative Examples from the Reporter Story Comments	182

LIST OF FIGURES

Figure	Page
1. Turiel suggests that religion and moral judgment domains will overlap while Rest claims that they are separate domains on the DIT tests	21
2. Differential item functioning graph.....	47
3. Person DIF plot of total BYU (B) and UA (U) samples	63
4. Relative percent of coded qualitative comments by DIT-2 cancer story choice.....	79
5. Relative percent counts of coded comments by DIT-2 famine story decision.....	81
6. Relative percent of coded comments by DIT-2 reporter story decision.....	83
7. Relative percent of coded comments by DIT-2 school board story decision..	84
8. This study supported Turiel and Walker’s claims of overlapping domains within the DIT-2 stories and items	110
9. Religion and macromoral judgment domains may overlap in complex ways on the DIT-2 items	117
10. Hierarchal order of further research	119

CHAPTER I

INTRODUCTION

The measurement of macromoral reasoning¹ has had an extensive history with copious amounts of research supporting various viewpoints (Bebeau & Thoma, 2003). The late Lawrence Kohlberg, a Harvard University professor, in an iterative process of research, became the father of the modern measurement of macromoral reasoning (Haggbloom et al., 2002). Kohlberg's measurement of macromoral reasoning ability (called macromorality) has roots in Jean Piaget's stage theory of cognitive development (Kohlberg, 1958). Kohlberg took Piaget's theories and concepts of a hierarchal stage development and sought to apply them to the measurement of individual levels of macromorality. After extensive interview research, Kohlberg built a framework for measuring macromorality based upon six stages of moral development. He further developed a set of vignettes involving macromoral reasoning, interview questions, and extensive scoring guides to operationally discriminate between each stage of macromorality (see Appendix A for a description of each stage and the rationale in classifying each stage).

Adherents to Kohlberg's theory sought to simplify his moral judgment interview and scoring process by developing an additional pioneering instrument in measuring macromoral reasoning, the Defining Issues Test (the first edition of this test is commonly

¹ To define macromoral reasoning (sometimes called macromorality or moral judgment), moral philosophers borrow similar terminology from the field of economics. That is, that the field of moral judgment discriminates between a larger broader level of morality (macromorality) and a smaller, more personal level of morality (micromorality).

referred to as the DIT-1). Convergent validity studies performed with Kohlberg's interview and the DIT-1 found correlations between the two test's scores to be in the low .80s (Rest, Narvaez, Bebeau, & Thoma, 1999b; see Appendix A for further discussion of validation of both of the DIT tests). Since Kohlberg's interview had to be administered orally and individually, it was very time consuming for both the interviewer and the interviewee. In addition, scoring the responses was also very time consuming. Thus, having a paper-and-pencil instrument (the DIT-1) with a high correlation with Kohlberg's interview was considered a breakthrough by many moral philosophers.

The DIT-1 quickly established itself as the most respected assessment of macromoral reasoning. By 1999, over 450 studies had been conducted on the DIT-1 since its implementation (Rest, 1999). These studies have helped to refine and validate the scoring of the DIT-1 as well as to help provide the framework for the construct of macromoral reasoning.

Several convergent validity and reliability studies have also helped to establish the DIT-1 as a respected instrument for measuring macromorality (see Appendix A). The DIT-1 also claimed to be universally applicable and, like Kohlberg's interview, culturally invariant (Kohlberg, 1967, 1971, 1976, 1981; Rest et al., 1999b).

Although there were advantages to using the DIT-1, research on the validity of the DIT-1 indicated that religious orthodoxy² within a Mormon sample confounded the assessment of macromoral reasoning (Blackner, 1975; Richards, 1988). Further, several

² Religious Orthodoxy is defined as one who has "a firm, literal belief in scriptural teachings" and "a belief that God, and God's word as revealed in the scriptures, are legitimate sources of moral authority (Richards, 1988, p. 18). In addition, the terminology used to describe religious orthodox persons includes terms such as: religious conservative, fundamentalism, traditional religious values, and the religious right.

researchers reported that religiously orthodox patrons from various faiths consistently scored lower than nonreligiously orthodox patrons when using this instrument (Brown & Annis, 1978; Clouse, 1979; Ernsberger & Manaster, 1981; Getz, 1985; Harris, 1981; Kay, 1998; Lawrence, 1979; McGeorge, 1976; Radich, 1982; Sanderson, 1974; Volker, 1979; Wahrman, 1980). Thus, the culturally invariant claim was called into question.

As research evidence showed that there were negative correlations between scores on the DIT-1 and levels of religious orthodoxy, moral philosophers worked with measurement experts to perform differential item functioning (DIF) analyses on the DIT-1 items. These DIF studies were conducted by comparing national norms of item-response patterns with a religiously orthodox group item-response patterns *after* carefully matching both groups on their overall ability to macromorally reason. The logic followed that if certain items were performing differently for religiously orthodox groups, then the overall claims of the DIT-1 as a measure of macromoral reasoning ability would be suspect due to a confounding variable.

Reacting to these and other research claims, the makers of the DIT-1 revised their theoretical framework and instrument to formulate a less culturally sensitive instrument, the Defining Issues Test-2 (DIT-2; Rest et al., 1999b). The Center for the Study of Ethical Development (2007), has called for replication studies, but there has been no differential item functioning analysis examining the use of the DIT-2 on a sample of religiously orthodox persons. Consequently, there are no data currently available to determine if the confound is still affecting DIT-2 scores.

This study proposes to collect data from a new Mormon sample to systematically

replicate the earlier research (Richards, 1988) on the new version of the DIT. The implications of this study are that if the Mormon sample's item-response patterns do not vary from a national item-response pattern on the DIT-2, then the DIT-2 items may be a more valid measure of macromoral reasoning for religiously orthodox groups in general than the DIT-1. More specifically, the goal of this study was to examine whether the items on the DIT-2 performed similarly to the DIT-1 items for a religiously orthodox group when compared to a group that more closely adhered to U.S. national norms for DIT-2 scores and attempted to answer the following research questions.

1. Does religious orthodoxy confound the measurement of macromoral reasoning as measured by the DIT-2?
2. What other variables predict both macromoral reasoning and religious orthodoxy and thus could be potential confounds?

As a final point of introduction, this study is important to the field of Instructional Technology and Learning Sciences (ITLS), because many instructional models for character education and moral education use the DIT-2 to measure the differences between treatment and control groups based upon an instructional treatment or use it in a gain study fashion (Rest et al., 1999b). Thus, ensuring that the DIT-2 instrument is correctly measuring the learning outcomes and gains in ethical reasoning that it purports to measure, relates to the field of ITLS. Finally, studying the construct validity of the DIT-2 test is important to the ITLS field, because of the way the DIT-2 and its derivative tests have been and are being proposed to be used in fields of education, business, government, and other practices (see Chapter II, Uses of the DIT-2).

CHAPTER II

REVIEW OF LITERATURE

This review of literature will point out the DIT concept and measurement of macromoral reasoning (MMR) as well as how DIT-2 scores are used. I will then point out variables that are thought to confound the measurement of MMR on the DIT, specifically pointing out pertinent studies relating to the measurement of MMR for religiously orthodox groups. Various explanations of religiously orthodox person's MMR scores will be reviewed followed by new theories and rationales for the variables used in this study. The justification for inclusion of findings in this review of literature was based upon three criteria: (a) whether the findings were from refereed journals or referenced by the Center For The Study of Ethical Development (CSED), (b) whether the writing pertained to the measurement methodologies for identifying item bias or a potential confounding variable, and (c) whether the literature led up to the current research in the fields of MMR or current religiously orthodox views on MMR.

Much of the literature in this review was found from the CSED's DIT-2 scoring guide (Bebeau & Thoma, 2003), which gives an extensive bibliography of research involving both versions of the DIT tests. In addition, many references to religiously orthodox studies in the book *Postconventional Moral Thinking: A Neo-Kohlbergian Approach* (Rest et al., 1999b) were used as starting points for the literature review. To obtain some of the most pertinent literature on the DIT-2 tests, the Brigham Young University library services were used such as: Journal Finder, EBSCO, JSTOR, and the Electronic Thesis and Dissertation Database. Finally, several face to face meetings as

well as phone and e-mail conversations were used to acquire articles and additional information from DIT researchers and experts in the field of macromoral reasoning.

The Concept of MMR in the DIT

To define macromorality, moral philosophers borrow similar terminology from the field of economics. That is, that the field of moral judgment discriminates between a larger broader level of morality (macromorality) and a smaller, more personal level of morality (micromorality). Rest and colleagues (1999b) explained these concepts.

It is useful to see Kohlberg's theory as primarily addressed to the formal structures of society (laws, roles, institutions, general practices) instead of the personal, face-to-face relationships in particular, everyday dealings with people...that are involved in making cooperation possible at a society level (in which not just kin, friends, and long-known acquaintances are interrelated, but strangers, competitors, and diverse clans, ethnic groups, and religions are as well). Examples of the special concerns of macromorality include the rights and responsibilities of free speech, due-process rights of the accused, nondiscriminatory work practices, freedom of religion, and equity in economic and educational opportunity. (p. 2)

On the other hand, micromorality concerns the "particular face-to-face relations that people have in everyday life" such as courtesy, caring, punctuality, and empathy (Rest et al., 1999b, p. 291).

The Measurement of MMR in the DIT

The DIT framework has been used as a measure of MMR for nearly 30 years. It was designed based upon Lawrence Kohlberg's *Piagetian-like* stages of moral development. Briefly, Kohlberg's theory of moral development included three levels and six hierarchal stages of moral reasoning. His levels and stages are described as:

Level One: Preconventional

1. *The punishment and obedience orientation.* The physical consequences of action determine its goodness or badness.
2. *The instrumental-relativist orientation.* Right action consists of that which instrumentally satisfies one's own needs and occasionally the needs of others.

Level Two: Conventional

3. *The interpersonal concordance or "good boy-nice girl" orientation.* Good behavior is that which pleases or helps others and is approved by them.
4. *The "law and order" orientation.* Right behavior consists of doing one's duty, showing respect for authority, and maintaining the given social order for its own sake.

Level Three: Postconventional

5. *The social-contract legalistic orientation.* Right actions tend to be defined in terms of individual rights and standards which have been critically examined and agreed upon by society.
6. *The universal-ethical principle orientation.* Right is defined by the decision of conscience in accord with self-chosen ethical principles appealing to logical comprehensiveness, universality, and consistency. (Kohlberg, 1973, pp. 631-632)

Kohlberg eventually came to view his stages of moral reasoning analogously to cognitive development in that he proposed that a person of lower moral development would not have the necessary skills to function at a higher stage of moral reasoning (Kohlberg, 1984).

Kohlberg's followers developed a Neo-Kohlbergian framework based upon moral schemas that are tapped into by presenting hypothetical ethical scenarios (Rest, Narvaez, Bebeau, & Thoma, 1999a). However, the new framework still uses the stage scores to determine the schemas and claim that as higher moral schemas gain use, the lower ones diminish in use (for a more detailed description of MMR as measured by the DIT see Appendix A). It was as if these new schema scores were merely a changing of names and

collapsing of stages into Kohlberg's former three levels. For example, stages 2-3 are the Personal Interest Schema (PI-stages 2-3), stage 4 is the Maintaining Norms Schema (MN-stage 4), and stages 5-6 are the Postconventional Schema (PC-stages 5-6). These three levels have almost identical descriptions as Kohlberg's original three levels.

The overall concept of measurement of MMR can be explained via the framework for the DIT-2 items. These items propose to be measuring how people interrelate to each other through laws, rules, roles, and institutions to form the system of society (Rest et al., 1999b). If a person's moral schema is based upon abstract, impartial, non-partisan principles of justice to establish principles of society-wide cooperation, then the person highly endorses items that tap into the highest of the DIT-2 schemas. It is presumed that the stories and items "activate moral schemas to the extent that person has developed them" (Rest, 1999, p. 6). As the person reads an item that makes sense to their moral schema, they rate that item as being more important to their decisions.

In addition, the "just community" is a common concept in the DIT-2 framework. The just community is based upon the concept of "right rather than the concept of good" (Rest, 1999, p. 14). In the DIT-2 framework, the maintaining norms schema (MN-stage 4) justifies a decision based upon rules, formal institutions, customs, or religious codes based upon established ways of knowing. These are seen as duty-based decisions that contain these specific elements. A person who is primed to be thinking in a MN-stage 4 schema appeals to "what is" as "what ought to be." Contrast that with a post-conventional schema (PC-stages 5-6) which rather than appealing to laws, customs, and so forth, they argue that what is right is right for the sake of the entire society. To further

clarify, while the MN-stage 4 schema appeals to the current social order, the PC-stages 5-6 schema appeals to sharable impartial ideals that are argued on the basis that the act would “respect other people,” serve the common shared values or goals, and “optimize the welfare of all participants,” while still being open to scrutiny and debate (Rest, p. 54). This higher schema realizes that laws are debatable social contracts and can be changed to suit an ideally organized society’s needs. In a DIT-2 just society, all insights are appreciated but scrutinized by the participants affected. Because of a belief in full reciprocity, it is believed that the laws and rules will not favor one member at the expense of another. Therefore, if a person feels the laws are not favoring their rights due to some sort of injustice, they can and should demand their rights. As emphasized, Kohlberg’s justice-based MMR framework somewhat lends itself to a social rights-based philosophy.

Uses of the DIT-2

Some DIT-2 proponents strongly advocate its use for discriminatory decisions. Thus, the practical significance of this study is validated by the proposed use of the DIT-2. Because religious orthodoxy transcends many of the boundaries in academia, professional life, and because MMR is valued in our society, this study has significance to many fields other than MMR and religion. For example, proponents suggest that the DIT-2 test scores can be used to predict success in the following areas: military and governmental leadership (Atwater, 1998; Olsen, 2006; Stephenson & Staal, 2007; Wang et al., 2008), organizational leadership (Sims & Keon, 1997), businesses (O’Fallon & Butterfield, 2005), sales jobs (Verbeke, 1996), law firms (Bebeau, 2002), accounting

(Abdolmohammadi, Read, & Scarbrough, 2003), general ethical decision training (Bebeau, Rodriguez, & Maeda, 2002; Rest, 1999; Yeap, 1999), and medical/dental school selection as well as practicing medical clinicians (Bebeau; Benor, Notzer, Sheehan, & Norman, 1982; Latif & Dunn, 2000; Self, 2000). Most of the proposals to use the DIT-2 outcomes for these selection and evaluative purposes rely on research of predictor variables of successful persons in the various fields or training. The assumptions rest on a notion that since the successful experts in a particular field show patterns in DIT-2 scores, that individuals who also show similar patterns prior to some sort of selection process, will also have higher success rates in that field. In addition, several researchers have used the DIT-2 in measures of convergent validity for later tests of moral reasoning (Rest, Narvaez, Bebeau, & Thomas, 1999c; Stewart, 2001; White, 1997).

However, if the DIT-2 items are somehow tainted by a systematic confounding variable, then one of the most fundamental assumptions for studying individual differences and instructional outcomes is not being satisfied (Niell, 2007; Osterlind, 2007; Reigeluth, 1999). That fundamental assumption is that the construct of interest (in this study MMR), *can and is* being appropriately measured for different individuals. Consequently, if the DIT-2 scores are systematically biased against religiously orthodox persons due to confounding variables other than MMR, then there are legal ramifications if it is used as selection or evaluative criteria.

Confounding Variables of the DIT MMR Scores

Some researchers argue there is much more to MMR than Kohlberg's cognitive

stages. In fact, the result of several research studies is that MMR, as measured by the DIT, *is* confounded by other variables. For example, De Casterle (1998) suggested that the Kohlberg's justice-based principles are lacking an element of caring responsibility for those one is responsible for. Other studies have found that the framework's focus on cold cognition and rational thought does not consider the impact on the environment (Nokes, 1989). In addition, Triandis and Hui (1986) showed that cultural views of collectivism and individualism affect one's MMR. Shweder, Mahapatra, and Miller's (1987) research showed that one's view of the nature of moral realities (ontology) lead to one's moral epistemologies (the nature of knowledge)—a similar point that several prominent educational researchers reiterate within educational research domains (Mertens, 2005; Schwandt, 2000). Several other research studies agree that the Kohlberg and his follower's framework (the Neo-Kohlbergian framework) favors those who focus on societal individual rights over societal collective responsibilities (De Casterle; Nokes; Shweder, 1991; Shweder et al.; Shweder, Much, Mahapatra, & Park, 1997; Triandis & Hui). Finally, religious orthodoxy has also been purported to be a confounding variable to MMR on the DIT (Richards, 1988).

Unyielding to these research studies findings, the Center for the Study of Ethical Development (CSED-the makers of the DIT) still claimed that the DIT MMR scores are validly measuring a trait that is culturally invariant. They maintained that “no variable accounts for the trends in the validity criteria better than” the DIT outcome scores themselves (Rest, 1999, p. 108).

To explain these proposed confounding variables, the makers of the DIT tests

claimed that individual MMR development and a discussion of an orthodox versus progressive view of societal roles, are being reflected on the DIT scores (Rest et al., 1999b). Further, their explanation stopped just short of referring to those with an orthodox or traditional view of society as macromorally handicapped (Rest, 1999; Richards, 1988, Shweder et al., 1987). In the DIT framework, orthodox views of society, whether they come from religion or elsewhere, are seen as inhibitors to MMR development (Rest).

This response of orthodox views of society inhibiting MMR development was too simplistic and somewhat offensive for many who valued MMR, especially religiously orthodox persons. Most of this abrasion occurred because religiously orthodox persons felt that their religious values were valid and moral tools that could be implored to judge the “rightness” or “wrongness” of a given macromoral situation (Lloyd, 2008; Nelson, 2004; Neuhaus, 1992; Oaks, 1992, 2001; Scott, 2007a, 2007b). In stark contrast to Rest (1999), these religious values are seen as catalysts, rather than inhibitors, to high moral judgment.

Religious Orthodox Variables Confound DIT Scores

The DIT-1 framework, items, and resulting scores have had a long tradition of being suspect to a systematic favoring of nonreligious participants when compared to religious participants (Blackner, 1975; Brown & Annis, 1978; Clouse, 1979; Ernsberger & Manaster, 1981; Getz, 1985; Harris, 1981; Kay, 1998; McGeorge, 1976; Radich, 1982; Richards, 1988; Richards & Davison, 1992; Sanderson, 1974; Volker, 1979; Wahrman,

1980).

Most of this suspicion of favoring non-religious persons is due to the fact that religiously orthodox persons consistently score lower on PC-stages 5-6 levels of MMR and higher on MN-stage 4 levels of MMR. In the DIT framework, this would indicate that religiously orthodox persons defer to societal norms, rules, and laws of society over carefully scrutinizing what is best for society as a whole. Further, it also indicates that they are less developed in their MMR abilities.

These provocative findings led to further research to discover why religiously orthodox persons scored lower on MMR outcomes. Early qualitative studies called Kohlberg's cultural invariance claim into question. Blackner was the first researcher to use Mormon or Latter-day Saint (LDS) subjects to suggest that the DIT items and scores may perform differently for religiously orthodox subjects: "The DIT may be designed in such a manner that those involved in religious education may respond to certain items as *most important* without first considering other alternatives. In other words, a person may be able to think in higher terms but have a cultural set in responding to certain stimuli" (Blackner, 1975, p. 64). In 1979, Lawrence found that Protestant seminarians strongly endorsed stage 4 items and showed a lack of endorsement of post-conventional items. From her follow-up interviews, she concluded that the endorsement of stage 4 items was not due to a deficit in moral reasoning *abilities*, but rather a loyalty to their religious beliefs over their own moral reasoning. These two early studies were especially controversial because Kohlberg claimed that religious orientation was independent of moral development (1967, 1971, 1976, 1981).

Follow-up studies in the 1980s focused on empirically proving whether a religious and cultural bias existed in the DIT-1 scores and why. For example, Getz (1984) and Richards (1988) both claimed that religiously orthodox people tended to score lower than average on the principled moral reasoning score (*P score, stages 5-6 proportion of items endorsed*). Most all of these studies empirically showed a systematic bias against religiously orthodox groups when compared to nonreligious orthodox groups.

However, researcher's reviews of these studies disagreed in their separate explanations of *why* these differences existed. Getz proposed that the negative correlation between religious orthodoxy and *P scores* was due to stifled development of moral judgment (1984) while Richards proposed the correlation existed due to test bias (1988). This disagreement evolved into research on the DIT scores to see if these lower scores for religiously orthodox persons was a true effect or if it was due to poor reliability and validity of the DIT scores. Richards reviewed Getz' studies and proposed that poor research methodology of the studies in her review led Getz to faulty conclusions. Richards cited: (a) sampling problems (e.g., only 2 of the 10 were randomly sampled), (b) failure to control for extraneous variables (e.g., not matching subjects education level—the highest predictor of DIT scores), (c) unreliable and invalid measures (many of the studies did not report reliability nor validity measures), and (d) lack of clear definitions (failure to define “conservative religiosity” or “religious orthodoxy”). Richards' conclusion was that due to “the methodological shortcomings of the studies reviewed..., Getz's conclusion should be viewed tentatively” (Richards, p. 21; also see Table 1).

Table 1

Findings and Methodological Characteristics of Studies That Examined the Relation Between Principled Moral Reasoning (P-score) and Religiously Orthodox Ideology (RO)

Study	Methodology							RO & P-score correlations		
	1	2	3	4	5	6	7	Negative	No relation	Positive
Brown & Annis (1978)	N	?	N	Y	Y	Y	Y	1	0	0
Clouse (1979)	N	Y	N	Y	N	Y	N	1	0	0
Ernsberger & Manaster (1981)	Y	N	N	NA	NA	Y	Y	1	0	0
Getz (1985)	Y	N	Y	Y	Y	Y	Y	1	0	0
Harris (1981)	N	N	Y	Y	N	Y	Y	0	2	0
McGeorge (1976)	N	?	N	N	N	Y	Y	0	1	0
Radich (1982)	N	?	Y	N	N	Y	Y	0	1	0
Sanderson (1974)	N	Y	Y	Y	N	Y	NA	3	0	0
Volker (1979)	N	N	N	N	N	Y	?	1	1	0
Wahrman (1981)	N	?	N	N	N	Y	?	0	2	0
Totals								8	7	0

Note. 1 = randomly selected from population, 2 = response rate adequate (> 80%), 3 = extraneous variables controlled, 4 = reliability reported for religious measure, 5 = evidence of validity reported for religious measure, 6 = adequate sample size for methodology used, 7 = long form DIT used, N = No, Y = Yes, ? = Unknown/not reported, NA = Not applicable. (Richards, 1988, p. 13).

On the other hand, Richards' research conducted a carefully matched differential item functioning (DIF) analysis and sought to control for several extraneous variables (such as college education). In addition, Richards clearly defined the term *religious conservatives*³ by reviewing several religious research reports and concluding that these reports had two common themes for their definitions of a religious conservative: (a) "a firm, literal belief in Biblical/scriptural teachings" and (b) "a belief that God, and God's

³ Richards and many others use the terms "religious conservative" and "religious orthodox" interchangeably. For this study, I have chosen to use the term "religiously orthodox" to describe the same type of trait.

word as revealed in the scriptures are legitimate sources of moral authority” (Richards, 1988, p. 18). Several researchers have included similar definitions to define a religious conservative (Ammerman, 1982; Beatty & Walter, 1988; Dobson, 1987; Hunter, 1981; Kellstedt, 1989; Knuckey, 1999; Leege, 1989; Moore, 1995; Smidt, 1989; Smidt & Kellstedt, 1987; Wilcox, 1986).

Richards’ sample consisted of 677 undergraduates at BYU who were members of the LDS church and were taking psychology and counseling courses. His findings showed that with the various methods he used, 16 to 25 items systematically performed differently for LDS college students when using various methods of DIF. These systematic differences were all in a negative direction (i.e., all these items contributed to scoring LDS college subjects with a lower moral reasoning score).

Richards and others findings spurred some religious orthodox researchers to label the Kohlberg framework as “secularism that rejects divine authority” (Rest et al., 1999b, p. 173). Most of this criticism arose because the Kohlberg model of macromorality places conformity to transcendent religious laws into the fourth stage of morality or mid-level. This stage is referred to as the “maintaining norms stage” (MN-stage 4) of moral reasoning and is akin to one who obeys the laws because they are “the rules,” not because they are moral. Shweder went as far to accuse Kohlberg of having an antireligious agenda as he assigned religious expressions to MN-stage 4 rather than higher stages because “Kohlberg does not believe in superior beings who have privileged access to truths about natural laws” (as cited in Rest, 1999, p. 173).

Rest’s (1999) reply to this criticism was to point out that religious thinking can be

scored at every stage in Kohlberg's framework.

At Stage 1, God is awesomely powerful as creator and miracle worker and commands obedience. At Stage 2, you offer sacrifices to God and abide by His commandments so that God will be good to you. At Stage 3, God is a friend and benefactor who is interested in you and knows your every thought and deed; therefore, you want to be your best because you don't want to disappoint God. At stage 4, religious law supersedes civil law and is also the law of nature. At stage 5, God is seen as the "energizer" of a just society and a force for autonomous personhood. After Stage 5, religious faith becomes Kohlberg's "stage 7" (skipping stage 6), answering the question "Why be moral?" At Stage 7, the person is affirmed in leading the moral life, and religious faith confirms moral thinking. (Rest, 1999, p. 173)

In theory, Kohlberg's reply (which Rest also advocates) is soothing to the religiously orthodox ideology. In practice today, the Stage 7 concept is lacking because neither of the DIT tests refer to, score Stage 7 ideologies, nor ask questions regarding how one's understanding of God and one's intricate relationship with God affects one's decisions (Walker & Frimer, in press).

Since it is impossible to score a Stage 7 response on either of the DIT tests, Kohlberg and Rest's response is considered a moot point when it comes to DIT scores for religiously orthodox persons. It is as though they had designed a special solution to an underlying problem but then prohibited access to that solution. Thus, the bottom line remains—religiously orthodox persons score systematically lower on MMR as measured by the first Defining Issues Test.

Theories of Why Religiously Orthodox Persons Score

Differently on the DIT

Richards theorized three reasons why religiously orthodox people appear morally

less adequate from DIT scores—the framework, the items, and doctrinal beliefs. First, the Kohlberg MMR framework was seen as insufficient because it failed to differentiate between man’s laws/authority and God’s laws/authority. This lack of differentiation between God and man’s laws was seen as a serious drawback, which confounded a person’s principal moral judgment score by measuring a theological construct rather than a social justice construct (macromorality). Second, the items were assumed to be devoid of ideological or theological content and therefore assumed to uniformly measure the same construct across all cultures. Richards and Davison (1992) as well as Lawrence (1979), cited two items that were clearly *not* devoid of theological connotations.

- a. “Whether the Christian commandment to love your fellow man applies in this case” is coded as a stage 4 item.
- b. “Isn’t God the ultimate source for who should live and who should die?” is also coded as a stage 4 item.

The third reason why the DIT was viewed as systematically biased against an LDS sample was because Mormons deferred to law based on one of their 13 basic tenets of their faith—“We believe in honoring, obeying, and sustaining the law” (Richards, 1988).

Richards’ research pioneered a differential item function method of analysis in MMR. He concluded that the MMR scores for the BYU group were systematically in the MN-stage 4 (middle level) of Kohlberg’s MMR framework due to the 16 to 25 supposedly biased items (statistical bias was dependent upon the method of analysis used). He further concluded that the BYU sample participants could reason at a theoretical higher level but chose to respond based upon their religious understandings and principles and therefore achieved lower moral reasoning scores. Therefore, the overall DIT test scores, which had then been used in moral education to measure *ability*

to reason morally, were concluded to be systematically biased against BYU students.

In 1998, Kay further generalized Richards' biased item claim in a study involving a Protestant sample. Kay also pointed out three main theories for these systematic biases against religious conservatives for MMR scores on the DIT-1. First, *the developmental theory* suggests that conservative religiosity and ideology blocks moral development so that conservatively religious persons are developmentally less able to morally reason at post conventional levels. Second, *the bias theory* states that theological biases exist in both Kohlberg's theory and in the instruments used by Neo-Kohlbergians. Third, *the preemption theory* suggests that conservatively religious individuals purposely obstruct processing needed for principled moral reasoning because it is seen as less morally adequate than conventional items on moral reasoning measures. Most of this blocking is due to a belief that transcendent beings who have access to privileged truths about morality are the ultimate source of what *is* moral, and religiously orthodox persons seek to remain consistent in belief and actions with this religious ideology (Kay).

Kay found evidence that the preemptive theory and bias theory were supported. However, in Kay's study, the Protestant sample did not show as much item bias as reported by Richards and Davison's religiously orthodox group (i.e., LDS).

Rest's reply to these studies can be seen as a matter of crossing domains—using one domain to answer the questions of another. To summarize his argument, he claimed that simply because someone draws upon a rationale from a particular domain to answer a question from an entirely separate domain, does not mean that the domains overlap. For example, if a multiple choice science question is asked to someone who does not have the

enough knowledge/ability to correctly answer the question, they may draw upon another domain to answer it (such as counting the number of A, B, C, or D options and then guessing). Rest's argument continued by citing the danger in logic of crossing domains in science to answer questions in religion. For example, using a theory from the field of Geology to determine the spiritual laws of mankind would be fallacious logic—geological theories are best left to determine the history of the geological formations. Similarly, Rest proclaimed, “When religion defines how we in this world are to relate to each other, then religion serves to define morality. In other words, the questions of one domain (e.g., how can people organize cooperation?) are answered by another domain (be faithful to the transcendent being's will!)” (Rest et al., 1999b, p. 163). Further, persons who are “primed to be thinking in terms of transcendent matters may answer the moral questions in terms of thinking from the religious domain” (Rest et al, p. 163-164). Thus, persons who use religion to answer social justice questions on the DIT-2 are crossing domains: “morality deals with this world; religion deals with the transcendent... a person who is primed to be thinking in terms of transcendent matters (the religious domain is activated) may answer the moral questions in terms of thinking from the religious domain... [however] the domain of the DIT is social justice even if some people respond as if a religious question had been asked” (Rest et al., p. 163-164).

Rest and colleagues do acknowledge that many religious teachings do prescribe moral relationships and seem to cut across both domains and also claim that the DIT tests are not anti-religious. Further, they claim that all religious thinking is not automatically scored as maintaining norms MMR. However, they assert that the DIT test's items are

designed to stay in the domain of social justice, and thus religious thinking on the DIT-2 items is a matter of crossing domains.

The implication of Rest's argument is that religion and social justice should not overlap within the DIT-2 test. Domains that are mutually exclusive in MMR research are referred to as hard domains. Conversely, those that do overlap are referred to as soft domains.

Some research suggests that Rest's reply of separate domains on the DIT test's items is not seen in everyday morality. Turiel suggested that religion and social justice are not hard domains but rather soft domains that overlap (Turiel, 1978; see Figure 1). Walker researched this claim and found that when interviewing 80 persons about morality and real-life moral problems that a significant number of people referred to God or religion when making moral decisions (Walker, Pitts, Hennig, & Matsuba, 1995). Spurred by his findings, Walker studied the intertwinement of religion, spirituality, and morality (Walker, 2002). His findings suggest that many attributes of moral, religious, and spiritual domains cross all three domains while some are unique.

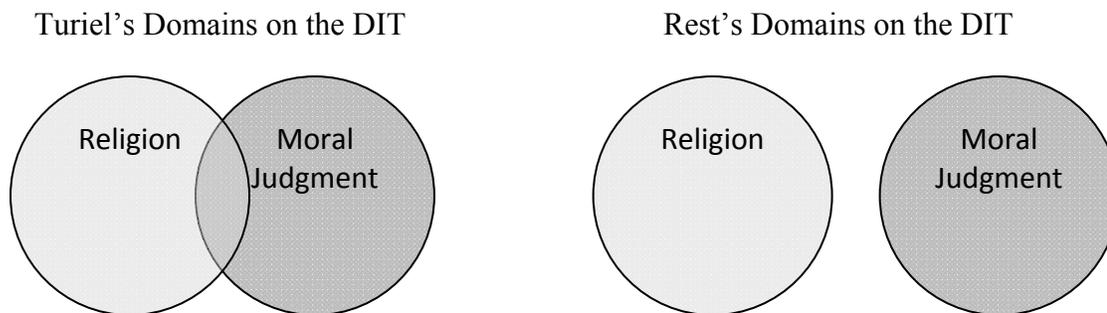


Figure 1. Turiel suggests that religion and moral judgment domains will overlap while Rest claims that they are separate domains on the DIT tests. Walker's findings (Walker et al., 1995; Walker & Pitts, 1998) support Turiel's claims.

For example, “devout” is shared between the religious and the spiritual domains while the concept of “just” is unique to the moral domain and “traditional” is unique to the religious domain (Walker & Pitts, 1998). Walker concluded that Kohlberg and his follower’s model are insufficient for handling notions of religion, spirituality, and faith when dealing with real-life moral problems (Walker & Pitts; Walker et al., 1995).

A Call for Replication Studies on a New Defining Issues Test

Reacting to several research findings, Rest and colleagues at the Center for the Study of Ethical Development (CSED) at the University of Minnesota,⁴ revised the DIT-1 and its underlying framework to introduce the DIT-2 and Neo-Kolbergian framework (see Appendix A). The CSED is the developers of both DIT tests and has invited new research using the revised and less culturally sensitive DIT-2 test (Rest et al., 1999b).

No research has replicated the DIF study that Richards performed in 1988 with the DIT-2 to examine if the previous bias still exists against religiously orthodox persons. The finding that there is a .83 correlation between the DIT-1 and DIT-2 scores suggests that 90% of the variance on one of the tests can account for the other.

Thus if the old test contained a cultural bias, the new, highly correlating DIT-2 test may contain the same bias. In addition, a new generation of BYU students, 20 years removed from Richard’s study, also invites this replication study.

A further need for this study relates to advancements in the methodology of conducting a DIF study. While Richards and Kay cited several measurement researchers

⁴ The CSED moved in 2008 to the University of Alabama. At the time of this writing it was still located at the University of Minnesota.

who give credence to the alternative chi-square methods used in his study, which were based upon group means and were acceptable in the 1980s (Murray & Mishra, 1983), there are more powerful methods of performing a DIF study using more appropriate polytomous IRT methods and new computer programs for more powerful item analysis today (Embretson & Reise, 2000).

Further, Richard's post-doctoral research confirmed that the DIT-1 scores only underestimated moral reasoning for "some but not all conservative religious individuals" (Richards, 1991, p. 364). His demonstration consisted of removing the proposed biased items from the DIT-1 scores in his BYU sample and then showing that only about half of the BYU subjects' moral reasoning scores changed. Therefore, if the majority of the participants saw religious content in the DIT-1 proposed biased items, but only half of his sample scores were affected by the removal of the proposed biased items, something other than religious content of the item could have been affecting the responses.

Put another way, the proposed theological confounding variable that was producing the item bias may be something other than group ideology or doctrinal beliefs. If there are confounding variables with certain types of BYU students (and hence certain types of religiously orthodox persons) this study may yield important information about which confounding variables produce an item bias within a religiously orthodox sample.

The confounding variables of interest in this study include: whether a person had served an LDS mission, whether they have children, marital status, association with criminals, views on the purpose of suffering, mother's education level, life experiences, doctrinal influences, personal philosophies, and other relevant issues not mentioned in the

DIT-2 items (see Table 2 for a rationale for including each of these proposed confounding variables). These demographics of interest and survey questions that will augment this administration of the DIT-2 and have been derived from interviews with BYU students, think-aloud protocols, and previous research (see Appendix B for the interview notes, Appendix C for the derived interview questions, and Appendix E for the usability instructions).

Conclusion and Transition to Study

In conclusion, many persons, corporations, agencies, educational institutions, government organizations, and religions consider MMR as a desirable trait (Abdolmohammadi et al., 2003; Bebeau, 2002; Benor et al., 1982; Latif & Dunn, 2000; O'Fallon & Butterfield, 2005; Self, 2000; Sims & Keon, 1997; Verbeke, 1996). Because of its desirability in society, many organizations provide instruction and application of ideologies to MMR under the broader umbrella of character education, ethics training, and teachings of morality (Bebeau).

The makers of the DIT-2 claim that the domain of moral judgment for macromoral issues is similar enough across different individuals and cultures that the DIT-2 can effectively measure this domain. However, there is ample evidence that the DIT-1 items were not performing equally for religious conservative groups when compared to national norms.

Since 1986 (when Richards' data were collected), new methods and computer software for DIF have developed for polytomous items (such as Likert scales), which are

Table 2

Rationale for the Choice of Variables for Testing

Independent variable	Rationale	Dependent variable
Missionary service	Deemer reported a .68 correlation between rich social life experiences and DIT-2 scores (Deemer, 1986, 1989). Nelson (2004) reported that scores on the Standardized Bible Content Test had a strong positive correlation with the stage 6 moral reasoning scores on the DIT-1 (0.57, $p < .001$) as well as a positive correlation with Principled Moral Reasoning scores (0.38, $p < .05$). In a small pilot study by Winder (see Appendix B), returned missionaries scored higher on post-conventional moral judgment scores (the highest stage of moral judgment) than non-returned missionaries.	DIT-2 scores and item-response patterns.
Marriage status	Deemer's study about the richness of social experiences (1989). Story #1 on the DIT-2 asks if a man should steal food for his family, almost like the dilemma faced by Jean Valjean in Victor Hugo's <i>Les Miserables</i> . Qualitative interviews lead me to believe that student's decisions on the DIT-2 change for this scenario dependent upon whether they feel a need to provide for a spouse or children (Appendix B). BYU is unique among US institutions of higher education in the percentage of married undergraduate student with 22% of them being married (Hall, 2005).	DIT-2 scores & item-response patterns.
Association w/ criminals	In story #2, a person's view about a person who was caught shoplifting but has since changed may be influenced by the person's having had a life experience that correlates with response patterns on the items associated with story #2.	DIT-2 scores & item-response patterns.
Purpose of suffering	Story #4 and qualitative interviews (Appendix B) warrant a question about an examinee's view of whether or not there are positive benefits from suffering. In a pilot study, a .60 correlation was reported on the euthanasia decision of story #4 with the person's response to a survey question about whether or not a patient can have a doctor "pull the plug" for a suffering patient. More BYU students advocated a doctor "pulling the plug" over not "pulling the plug" when death was inevitable. In qualitative interviews the factor in this decision seems to be whether or not a person views suffering as having a purpose. Therefore, an open-ended question about whether there is purpose in suffering was designed and inserted on the survey.	DIT-2 scores & item-response patterns.
Mother's education level	Mother's education level has been shown to predict educational performance for the mother's posterity (Eagle, 1989). Since studies show that education is a large contributor to the DIT scores, explaining some 30 to 50% of the variance (Bebeau & Thoma, 2003, p. 30), and mother's education level is a high predictor of a child's education level, perhaps this demographic (and the previously mentioned ones) will reveal important trends for future research in the BYU sample.	DIT-2 scores & item-response patterns.
Life experiences influence	In qualitative think-aloud protocols, BYU students would frequently refer to a specific life-experience when making a decision for what should be done on the DIT-2 (see Appendix B). Life experiences included academic training, work-related experience, social experiences, and family experiences (see Appendix G).	DIT-2 scores & item-response patterns.

(Table continues)

Independent variable	Rationale	Dependent variable
Doctrinal influence	In qualitative think-aloud protocols, BYU students would frequently refer to a scriptural teaching when making a decision for what should be done on the DIT-2 (see Appendix B). Doctrinal influence including referencing teachings that were clearly scriptural in nature (see Appendix G).	DIT-2 scores & item-response patterns.
Personal philosophy	In qualitative think-aloud protocols, BYU students would frequently refer to a personal philosophy when making a decision for what should be done on the DIT-2 (see Appendix B). These included personal values that define what is moral for an individual (see Appendix G).	DIT-2 scores & item-response patterns.
Identifying more relevant issues than what the DIT-2 scenario allows	In qualitative think-aloud protocols, BYU students would frequently refer to what they thought was a more relevant issue when making a decision for what should be done on the DIT-2. For example, in story #5, college students take over the administration building in a protest and halt university operations. Most students mentioned that demonstrating the way these students were was “a waste of time.” So time-management was more of an issue to them than to protest in this manner (see Appendix B and Appendix G).	DIT-2 scores & item-response patterns.

more powerful for performing DIF than the original methods used by Richards. In addition, this study will seek to account for variables that covary with MMR and religious orthodoxy by using partial correlations—an overlooked method of analysis in the previous studies. A new version of the DIT has emerged (the DIT-2—see Appendix A) as well as a Neo-Kohlbergian framework based upon moral schemas rather than Kohlberg’s Piagetian-like six stages (Rest et al., 1999a). New methods and computerized tools for analyzing differential item functioning have also been introduced which are more reliable than Richard’s methods used in the 1980s. Thus the scholarly significance of this study is justified by the new definitions of the DIT-2 framework, the new findings and tools to incorporate more advanced methodology, and also suggested potential confounds that have surfaced in the literature since 1988 when the initial replicable research was conducted. In addition, the CSED, the developers of both DIT tests, has also called for replication studies as well as invited new research using the revised and

less culturally sensitive DIT-2 test and analysis methods (Rest et al., 1999b). With the call for replication studies as well as new tests, new scoring methods, new frameworks, new DIF methods for polytomous items, new research pointing out other potential confounds to MMR, and a new generation of BYU students, a new DIF study is needed to examine if the previously identified bias still exists.

CHAPTER III

METHODS

Purpose of Study and Research Questions

The purpose of this study was to explore whether the DIT-2 solved a problem of item bias against religiously orthodox persons and to explore if other variables affect the measured relationships between religious orthodoxy and macromoral reasoning (MMR). Specifically, the research questions were as follows.

1. Does religious orthodoxy confound the measurement of macromoral reasoning as measured by the DIT-2?
2. What other variables predict both macromoral reasoning and religious orthodoxy and thus could be potential confounds?

Independent, Dependent, and Moderating Variables

The independent variable was religious orthodoxy. As emphasized in Chapter II, the general consensus among DIT and social researchers is that religious orthodoxy is a measure of how literally one adheres to scriptural teachings and believes they are legitimate sources for moral reasoning (Richards, 1988). Several research reports use the term “religious orthodoxy” interchangeably with the term “religious conservative” (Ammerman, 1982; Beatty & Walter, 1988; Knuckey, 1999; Moore, 1995; Wilcox, 1986). The DIT-2 has an internal measure of religious orthodoxy that will be described later in this chapter (see instrumentation section).

The dependent variable was macromoral reasoning (MMR) as measured by the DIT-2 items. To briefly describe macromorality, it is how people interrelate to each other through laws, rules, roles, and institutions to form the system of a just society (Rest et al., 1999b). This is different from micromorality, which consists of the correctness of actions at a personal level such as being loyal, caring, and considerate. Contrastingly, the overall concept of measurement of MMR pertains to the “formal structures of society (laws, roles, institutions, general practices)...that are involved in making cooperation possible at a society level” (Rest et al., p. 2). In the DIT-2, as the person reads an item that makes sense to their MMR schema, they rate this item as being more important to their overall decisions they made involving a MMR case study (e.g., should Mustaq steal during a time of famine?). Thus, the DIT-2 presumes that the stories and items “activate moral schemas to the extent that person has developed them” (Rest, 1999, p. 6).

There were several moderating variables that this study explored. Essentially these variables were chosen based on previous theories and research, pilot study interviews, and a review of literature. The moderating variables of interest included: returned missionary status, marital status, association with criminals, views on the purpose of suffering, mother’s education level, life experiences, doctrinal influences, personal philosophies, and other relevant issues not mentioned in the DIT-2 items (see Chapter II, Table 2 for a rationale for each of these proposed moderating variables).

Previous research has pointed out that the following variables were not confounding DIT-2 scores: gender (Rest et al., 1999b; Thoma, 1986), GPA (Rest, 1999), or any other cognitive ability (Rest, 1999; Rest et al., 1999b). In addition, from a pilot

study I ruled out “several religiosity-based problem-solving styles or orientations” based on Pargament’s Religious Problem Solving Styles instrument (Pargament et al., 1988, 1999, p. 347).

Research Design

This study was a replication of Richard’s (1988) study of religious orthodoxy and DIT-1 scores. However, this study sought to validate the second edition of the DIT, the DIT-2, for religiously orthodox persons. The research design was a cross-sectional correlational study used in conjunction with a reference group for comparison purposes. The reference group was employed to check for DIF between a group known to be religiously orthodox (BYU students) and a reference group that adhered more to national norms of religious orthodoxy. This DIF comparison with a reference group showed whether the DIT-2 items were measuring a unidimensional trait across cultures and gave evidence as to whether the DIT-2 claim of cultural invariance was valid for a religiously orthodox group by exposing any confounding items to overall DIT-2 MMR scores. In addition, I analyzed correlations of DIT-2 scores with additional moderating variables that Richards did not account for.

Population and Sample

The target population to which the study sought to generalize was U.S. religiously orthodox persons. BYU students were an appropriate sample to represent religiously orthodox persons because in past studies they consistently scored very high on measures

of religious orthodoxy (see Bergin, Masters, & Richards, 1987; Bergin, Stinchfield, Gaskin, Masters, & Sullivan, 1988; Cornwall, Albrecht, Cunningham, & Pitcher, 1986; Shepherd & Shepherd, 1984; Shupe & Heinerman, 1985). This higher than average measure of religious orthodoxy was verified again in this study as BYU DIT-2 mean religious orthodoxy score was 1.26 points higher than a national average (BYU sample mean was 6.23, 4.97 is the national average—scores range from 1-9). In the DIT-2 framework, this higher than average score means that the typical BYU student is more likely to defer their MMR based on their religious ideology.

Another reason why BYU students were an appropriate sample was because Richards' original claim of DIT item-bias against religiously orthodox persons consisted entirely of a BYU student sample (Richards, 1988). Since this was a replication study I sought to use a similar sample. Doubtless, the BYU Mormon culture has somewhat changed in the last 20 years. However, pilot study measures of religious orthodoxy show they are still a very religiously orthodox sample (see Appendix B).

To briefly describe this sample of BYU students (i.e., the focal group), they had much variation in educational level (38 freshmen, 112 sophmores, 108 juniors, 71 seniors, 2 MS, 6 missing). Because BYU requires religion credit, there was also a variety of academic majors enrolled in these courses.

An analysis of variance found that DIT-2 overall scores did not statistically vary between educational level groups. Levene's test of equality of error variance indicated that DIT-2 score's variance were equal across freshmen-senior educational levels. Thus, it was deemed appropriate not to parse out educational level further than undergraduate

college students for many of the analyses'. Therefore both reference and focal groups consisted of entirely undergraduate students in the DIF study.

The reference group came from the research of Dr. Steve Thoma, CSED director of research and the University of Alabama (UA) professor of educational psychology. He offered 888 individual undergraduate student DIT-2 response files for a reference group. No additional demographics of interest were included in this sample as it was solely used a reference group for the differential item functioning analysis.

Instrumentation

The instruments used in this study were two questionnaires. They consisted of the DIT-2 and an additional demographic questionnaire.

The DIT-2 is a 65-item form where students make decisions based on five macromoral dilemma case studies and then rate and rank, which items are of most importance to their decision (5-point Likert-like scale ranging from *great importance* to *no importance*). For example, the cancer story asks a person to decide if a doctor should administer an overdose of painkiller to an insisting cancer patient who wants to “end her suffering even if it means ending her life. Should the doctor give her an increased dosage?” (see Appendix F for a complete description of all five moral dilemmas). Test participants then rate how important 12 items are to their overall decision (e.g., Item 10- “Shouldn’t only God decide when a person’s life should end?”).⁵

Cronbach’s reliability coefficients for DIT-2 research range from the upper .70s to

⁵ The cancer story and item 10 are used as examples because these items are also used in determining one’s DIT-2 religious orthodoxy scores. Very few of the items are so overt in regards to religiosity issues.

the lower .80s. Test-retest reliability is about the same. The correlation of DIT-1 scores with DIT-2 scores is .79, nearly the same as the test-retest reliability of the DIT-1 scores with itself (Bebeau & Thoma, 2003; Rest et al., 1999c). The DIT tests have undergone 8 validation criteria ranging from correlational studies with moral comprehension scales to 10-year longitudinal studies (see Appendix A for a more detailed description of DIT validity).

Scores on the DIT-2 are similar to scores on the DIT-1. The major difference is that the stage scores have been grouped and changed to schema scores. Notice, however, that in the schema descriptions the stage scores are still used to describe the schemas:

Personal Interest Schema Score (PI-stages 2-3) represents the proportion of items selected that appeal to Stage 2 and Stage 3 considerations. Stage 2 considerations focus on the direct advantages to the actor and on the fairness of simple exchanges of favor for favor. Stage 3 considerations focus on the good or evil intentions of the parties, on the party's concern for maintaining friendships and good relationships, and maintaining approval.

Maintaining Norms Schema Score (MN-stage 4) represents the proportion of items selected that appeal to Stage 4 considerations. Stage 4 considerations focus on maintaining the existing legal system, maintaining existing roles and formal organization's structure.

Postconventional Schema Score (PC-stages 5-6) represents the proportion of items selected that appeal to Stage 5 and 6 considerations...[These stages] focus on organizing a society by appealing to consensus-producing procedures (such as abiding by majority vote), insisting on due process (giving everyone his day in court), and safeguarding minimal basic rights, ...organizing social arrangements and relationships in terms of intuitively appealing ideals. (Bebeau & Thoma, 2003, pp. 18-19; see Chapter II measurement of MMR section for a more detailed description of what each schema score means)

Each of these schema scores is the proportional rating of how each type of item was endorsed compared to the other items. As a proportion, these scores range from 0 to 1. However, they are converted to a whole number by multiplying by 100. Consequently,

each score ranges from 0-99 points. In addition, an overall MMR score that considers all responses on the DIT-2 ranges from 0-95 and is named the N2 score (Bebeau & Thoma, 2003).

The DIT-2 also collects information regarding how religiously orthodox a person is. The religious orthodoxy score ranges from 1-9 and has an estimated national average of 4.97 (Bailey, Phillips, & Scofield, 2005, C. Bailey personal communication, March 2008). This score is collected by assessing how much locus of control God has given to man in regards to life and death matters and by asking participants to rank how important God's locus of control of life and death is to their overall decision about euthanizing a suffering cancer patient. The DIT-2 scoring guide reports that this score correlates highly with Brown and Lowe's Religious Inventory of Belief Scale (Bebeau & Thoma, 2003; Brown & Lowe, 1951). The CSED makers claim this score also is an indicator of how religiously orthodox a person is.

In addition, other scores of interest on the DIT-2 are the NUMCD, HumLib, and ConLib scores. The NUMCD score indicates the degree to which one cannot decide what to do on the DIT-2 moral dilemmas and ranges from 1-5. Higher scores indicate that participants are more decisive on the DIT-2 story choices and thus are more consolidated in their macromoral reasoning (i.e., more "set" in their macromoral reasoning). The HumLib score is a reflection of the number of times a choice on the DIT-2 matches academic moral philosopher's choices from the fields of political science and philosophy. Scores range from 1-5 and high scores indicate agreement with moral philosopher's choices on the DIT-2. The HumLib score is inversely related to the RO index (i.e., RO

persons generally do not agree with the moral philosophers in some macromoral issues). The ConLib score indicates how politically conservative or liberal a person rates themselves on a scale of 1-5 (1 = very liberal, 5 = very conservative).

The second instrument used in this study was a demographic questionnaire that augmented the DIT-2. These questions were derived from interviews with BYU students, think-aloud protocols, and previous research (see Appendix B for the interview notes, Appendix C for the demographic questionnaire, and Appendix E for the usability instructions). To briefly describe the creation of the demographic questionnaire, in a pilot study, think-aloud protocol responses were coded into four general categories that influenced BYU student's decisions on the DIT-2: (a) life experiences, (b) doctrinal influences, (c) personal philosophies, and (d) other relevant issues perceived (see Appendices B and C). In the final data collection phase for this study, a fictional vignette with several examples of the four categories was provided (see Appendix E). In addition, comparing LDS returned missionary status, marital status, as well as the demographic of mother's education level revealed some interesting findings due to the nature of the macromoral reasoning vignettes (see Chapter II, Table 2 for a rationale for including these demographics). The CSED scoring and interpretation guide mentions that additional instructions, items, and instruments do not affect the reliability of DIT-2 scores (Bebeau & Thoma, 2003).

Data Collection Procedures

BYU students were given the DIT-2 test and were asked to respond to the test and

questionnaire as individuals outside of class during the second to fourth weeks of the Winter 2008 semester (see Appendix C). After having the purposes of the study explained, students could choose to participate for a small amount of extra credit or not. Those students who chose not to participate were given the option to complete a different extra credit assignment as outlined in my IRB proposal (see Appendix D). DIT-2 tests were ordered from the CSED, administered as a paper and pencil scantron, collected by myself, and sent to the CSED for scoring. The CSED returned all tests and an SPSS spreadsheet of the data.

The demographic questionnaire was administered online during the same weeks as the DIT-2 test. Students were e-mailed and given a printed copy of the website link to take the online questionnaire. Three reminder e-mails were sent out during this time. All these data collection procedures were in-line with the DIT-2 scoring guide guidelines (Bebeau & Thoma, 2003). The online questionnaire data were exported into Microsoft excel. The data was then coded into categories using Microsoft excel. SPSS and Winsteps were used for statistical analysis.

The reference group had already been collected by Dr. Steve Thoma, CSED director at the University of Alabama. He also followed the DIT-2 scoring guidelines to ensure reliability and validity. To ensure that the reference group was similar to the national average of RO, persons with scores between 4-6 on the DIT-2 religious orthodoxy variable were selected. The group mean for this group was 4.96 ($n = 272$), a very comparable religious orthodoxy mean to Bailey's national estimate of 4.97 (Bailey, 2005; C. Bailey, personal communication, March 2008).

Pilot Studies

I completed three unpublished pilot studies in 2007 with a 30-person BYU sample, a 0-person BYU sample, and a 135-person BYU sample. The first study looked at the correlations between styles of religious problem solving and DIT-2 scores among BYU students. The second study was a think-aloud interview as students took the DIT-2. The third study looked at religious orthodoxy measures of current BYU students. These studies are described in greater detail in Appendix B. However, only the findings that pertain to my research questions, methods, and analysis are reported here.

Pilot Study 1—Many Religiously Orthodox Persons Self-Direct Rather than Defer to Religion

The first pilot study looked at a claim by the Center for the Study of Ethical Development (CSED). Specifically, the CSED had claimed that religiously orthodox persons *defer* their MMR based upon a maintaining norms schema (MN-stage 4) that is related to religious laws and customs (Rest et al., 1999b). Therefore, I administered the DIT-2 along with a test that measures “several religiosity-based problem-solving styles or orientations” (Pargament et al., 1988, 1999, p. 347) to discover if religious orthodoxy scores correlated with Pargament’s scores. Specifically, Pargament’s test measures whether someone defers their decision making to their God or their religion, or is a self-director.

This pilot study showed that the 15 LDS returned missionaries (those with higher than average Stage 4 scores and religious orthodoxy scores) were considerably *more* self-

directing in their views towards solving life's problems than 15 non-returned missionaries (5 points difference on a 15-75 point scale, $p < .05$). Pargament suggested that strong self-directing scores are typical of persons that "emphasize the freedom God gives people to direct their own lives" (Pargament et al., 1988, p. 91).

Since a subgroup that was scoring higher on Stage 4 scores and religious orthodoxy scores was more self-directing as well as lower in deferring scores involving their religious problem-solving styles, it seemed that something other than *deferring* to God and religion seemed to be accounting for LDS returned missionaries' higher MN-Stage 4 scores. Rest (1999) and Richards (1991) proposed that this other confound was a propensity to defer their MMR to the laws of the land, which led to pilot study #2.

Pilot Study 2—Religiously Orthodox Person's Justifications for DIT-2 Responses

The purpose of this study was to discover what types internal thinking and external experiences religiously orthodox persons referred to as they reasoned through the DIT-2 moral dilemmas. Pilot study #2 involved interviews, discussions, and think-aloud protocols with nine BYU students. Four of these were conducted individually while two of them were conducted in focus groups of two and three. The students very infrequently mentioned the law or a societal norm as a reason to justify their decisions, a contrary finding from Richards' (1988) research and Rest and colleagues claims (1999c). However, this study found that students referred to life experiences, personal interpretations of doctrinal teachings, and philosophies of life as they made their DIT-2 decisions. In addition, some felt that the test limited their responses as they thought of

other relevant issues. This study was the basis for the demographic questionnaire. The goal of this questionnaire was to seek to find other variables that account for MMR or religious orthodoxy.

An additional finding from the first and second pilot studies also led to additional data collection and analysis. The first pilot study showed that the BYU mean DIT-2 religious orthodoxy scores were 6.23. The UA mean religious orthodox scores were 6.64 (these had already been collected by Dr. Thoma). Both group scores were normally distributed. Independent samples *t* tests led me to conclude that these differences were not statistically significant (see Table 3). Therefore, the UA and BYU samples did not differ in aggregate measures of religious orthodoxy on the DIT-2 RO scale and possibly in other aggregate matters of religious conservatism. This lack of contrast led to data collection to find an estimated national norm for the DIT-2 religious orthodoxy variable from previous study's data (Bailey et al., 2005).

Pilot Study 3—BYU and UA Measures of Religious Orthodoxy

The third pilot study further validated the assumption that BYU and UA students did not differ much in religious orthodoxy. As two measures of religious orthodoxy

Table 3

A Comparison of a Measure of Religious Orthodoxy by School (Brown and Low Inventory of Religious Belief)

School	<i>N</i>	Mean	<i>SD</i>	Std. error mean	Sig. (2-tailed)
UA	423	58.8180	10.20605	.49624	0.379
BYU	135	59.3481	3.96688	.34141	

(Brown and Lowe, and the internal DIT measure) were not statistically significant between UA and BYU samples, it was therefore determined that as a group, BYU and UA students did not significantly differ in measures of religious orthodoxy.

This finding posed a problem in my original data collection plan because I was seeking to compare a group that was very religious orthodox to a reference group that was closer to a national average of religious orthodoxy. A bright spot was that the variance of the reference group was much greater than the BYU sample. The increased variance of the reference group led to the filtered sampling approach from the UA sample based upon national averages of religious orthodoxy described in Chapter IV of my study. It also was the basis of using partial correlations as a method of data analysis.

Data Analysis

This study consisted of six analyses. These six analyses sought to answer the research questions: (1) Does religious orthodoxy confound the measurement of macromoral reasoning as measured by the DIT-2? (2) What other variables predict both macromoral reasoning and religious orthodoxy and thus could be potential confounds?

Analysis #1 Descriptive Statistics

These statistics were provided to summarize the characteristics of BYU sample participants and to provide evidence that this group represented the religiously orthodox, BYU students, LDS members, and provided a reasonable sample for replication.

Specific descriptive analysis included information on demographics of the BYU sample in the terms of frequency counts and percentages, means, and standard deviations

for the demographics such as: age, gender, U.S. citizenship, primary language, political leanings, whether a person had served an LDS mission, mother's education level, marital status, whether they had children, association with persons who shoplifted or were convicted of other crimes, and opinions of whether there is a purpose in suffering. These descriptive statistics helped to determine the appropriate methods of analysis based on the amount, distribution, and type of data collected.

Descriptive statistics also pointed out some differences between a religiously orthodox group's DIT-2 scores (BYU) and a reference group's DIT-2 scores (National Norms). In addition, DIT-2 scores were separated by educational level and compared.

Analysis #2 Correlations between Religious Orthodoxy and Macromoral Reasoning

These statistics provided initial evidence concerning the degree to which religious orthodoxy (RO) was associated with the measurement of macromoral reasoning (MMR). These associations provided evidence to answer research question 1: *Does religious orthodoxy confound the measurement of macromoral reasoning as measured by the DIT-2?* Statistically significant correlations here provided evidence that RO and DIT-2 MMR scores are related to each other and thus, RO is a *potential* confound to the measurement of MMR for an LDS population.

Because many DIT-2 scores were normally distributed, Pearson's r could be employed. These associations informed me on how strong religious orthodoxy is related to MMR and whether these variables had positive or negative associations. In addition, these correlations were squared to find the coefficient of determination (Kachigan, 1986)

on variables with high correlations. This coefficient explains the proportion of the variance in one variable that can be accounted for in the other (Kachigan). One caution to be noted is that when interpreting correlation coefficients, Cohen (1988) pointed out that it is important to remember that the practical significance of the coefficient is somewhat relative to our own values. In addition, there were no specific DIT-2 guidelines given to determine large or small DIT-2 correlations. In fact, some DIT studies even report correlation coefficients as small as $r = -.16$ (Olsen, 2006). Therefore, much of the interpretation of practical significance will be based on judgment from the observer and by using multiple approaches (t tests, qualitative data, literature reviews) to determine practical significance. If there was not a statistically significant positive or negative correlation between RO and DIT-2 scores (independent and dependent variables), then I concluded RO was not a likely confound.

Analysis #3 Differential Item Functioning

These statistics provided additional information about how MMR measurement might be effected by RO for both the BYU sample, and a representative subpopulation that score similarly on DIT-2 to the BYU sample but are not as high scoring on the DIT-2 RO score. The results of these analyses were also used to address research question one.

A DIF analysis' falls under the item-response theory (IRT) methods of analysis. To more clearly explain IRT, I borrow a high-jumping analogy from Bond and Fox (2007). In high jumping, there are two *main* factors that determine the probability of success (jumping over the bar)—a person's ability (how high they can jump), and the difficulty of the task (how high the bar is raised). For example, a person with an ability to

jump 6 feet has a high probability of clearing a bar that is only at 5 feet. The probability of success is intuitively calculated by subtracting their ability from the difficulty (how high they can jump—how high the bar is = probability of success). However, this same person with a 6-foot leaping ability would have a lower probability of clearing the bar as it nears and even surpasses the 6 foot height.

Similarly, IRT extends the concept of success, or endorsement of an item, based upon person ability *and* item difficulty. IRT places both of these aspects (person ability and item difficulty) on the same equal interval scale via a log odds transformation of the probability of success on an item. Therefore, by subtracting the item difficulty from the person ability one can obtain the probability of success (or endorsement) of an item (or category) according to an item response model. How accurate the model is (called the likelihood response model) also factors into the predictability of the item endorsement equation.

One of the benefits of IRT over classical test theory analysis is that the concept of reliability is based upon the person ability and the accuracy of the likelihood response model for each person. Thus, IRT “extends the concept of reliability by including the person ability factor. In IRT, reliability refers to the degree of precision at different values of person ability” rather than assuming that reliability (as based on the standard error of measurement in classical test theory) is a constant across all persons and all abilities (Xu, Iran-Nejad, & Thoma, 2008, p. 16).

In IRT, each item also has an item-fit statistic. The log odds transformation allows IRT to give the item location on an equal interval scale of item difficulty. Thus, an item

with a difficulty of 4 is twice as difficult to endorse as an item with a difficulty of 2. Item difficulties are in terms of logits that usually range from -6 to +6 with a mean of 0. A negative logit indicates that the item was easy to endorse (or easier) while a positive logit indicates that the item was harder to endorse (or more difficult).

In addition, in IRT each item also has a residual score which is the difference between the observed response and the model's expected response. Items with high residuals indicate a misfit. A misfitting item on a unidimensional latent trait scale indicates that the item does not fit within the assumptions of the IRT model's measurement for a unidimensional trait (i.e., for this study, something other than the latent trait of MMR is confounding the item). This is an important aspect for a DIF study.

Essentially, a DIF study involves "deciding whether the items on a particular instrument provide invariant measurement across these two groups" (Embretson & Reise, 2000, p. 252). After the two groups have been carefully matched on overall ability for the construct that the items measure, item endorsement patterns are analyzed to detect if an item performs differently for a focal and a reference group. To get an accurate IRT model, Bond and Fox (2007) recommended using large sample sizes (above 200) to have high internal consistency to make the inferences.

Another aspect of the concept of a DIF study is the application of one of measurement theorists fundamental claims—the claim that the measurement of a latent trait or construct (in this case, moral judgment) is the *true value of the person trait* + *error* (Osterlind, 2006). Error may be random or systematic. Item bias exists when a test exhibits systematic error, not random error.

To more concretely define item bias in a DIF study, an item bias exists when an item on a test “unfairly favors one group over another” (Clauser & Mazor, 1998, p. 31). For example, if a math reasoning question on the ACT is easier for males than for females who have been equally matched in their overall math reasoning ability, then the item functions differentially. If the test is unidimensional (i.e., truly measuring one trait or ability), then the items should perform relatively the same for the two groups when matched on overall ability. If they perform differently on certain items after being carefully matched on ability, then there is evidence that there is another ability that may be producing the difference in responses. “The question then becomes whether that second ability is relevant to the purpose of the testing” (i.e., is this ability part of the construct we are measuring?) (Clauser and Mazor, p. 31).

Person reliability. DIF studies yield a statistic similar to Cronbach’s coefficient alpha as it ranges from 0-1 and higher scores indicate greater reliability. This is the person reliability statistic. The person reliability statistic also relates the reproducibility of person ordering in a DIF study (i.e., whether those who score highest on the logit scale would remain there and those that score lowest would remain the lowest). This statistic also gives an estimate of how well a test separates persons into levels of ability (called person separation index). Higher reliabilities and separation indexes indicate that the test can discriminate between more levels of the trait.

Graphed IRT item DIF contrasts. The most precise way to analyze whether an item exhibits DIF is to view the projected item response models for each item and person and compare the amount of difference between the observed examinee trait level and

expected value response rates based upon the item response maximum likelihood model (Kim, Cohen, Alagoz, & Kim, 2007; Potenza & Dorans, 1995). Presumably, if the latent trait that the test proposes to measure is reliable by each item, these statistical model measures should perform similarly for persons who have been matched on overall ability. When a significant number of individuals within a particular group have similar unlikely response patterns for an item, this finding reflects that the item performs differentially for that group (Embretson & Reise, 2000). These differences are shown quite well in a DIF contrast plot (see Figure 2). In this graph, one line represents the focal group while the other represents the reference group. Because IRT uses a log-linear transformation of item endorsement probabilities, the item difficulty levels for each group are on equal intervals. This means that a 2.0 item difficulty is twice as difficult as a 1.0 item difficulty. The general concept of IRT rests upon the assumption that a test score represents a score of unidimensional trait. Therefore, if persons are matched on the overall test score, and the test indeed measures a unidimensional trait, then persons of the same ability (as measured by the overall test score) will respond in probabilistic manners to the same items. If there is a significant contrast in group's response patterns as shown by the contrast on the graph, then there may be an ancillary confound in that item—in other words, the item is exhibiting differential item functioning or DIF.

DIF statistical significance. In addition to graphing the contrast of the reference and focal group, the contrast of performance between groups is computed by the computer analysis program used. The statistical significance of the DIF for each item is computed to give a DIF contrast statistic.

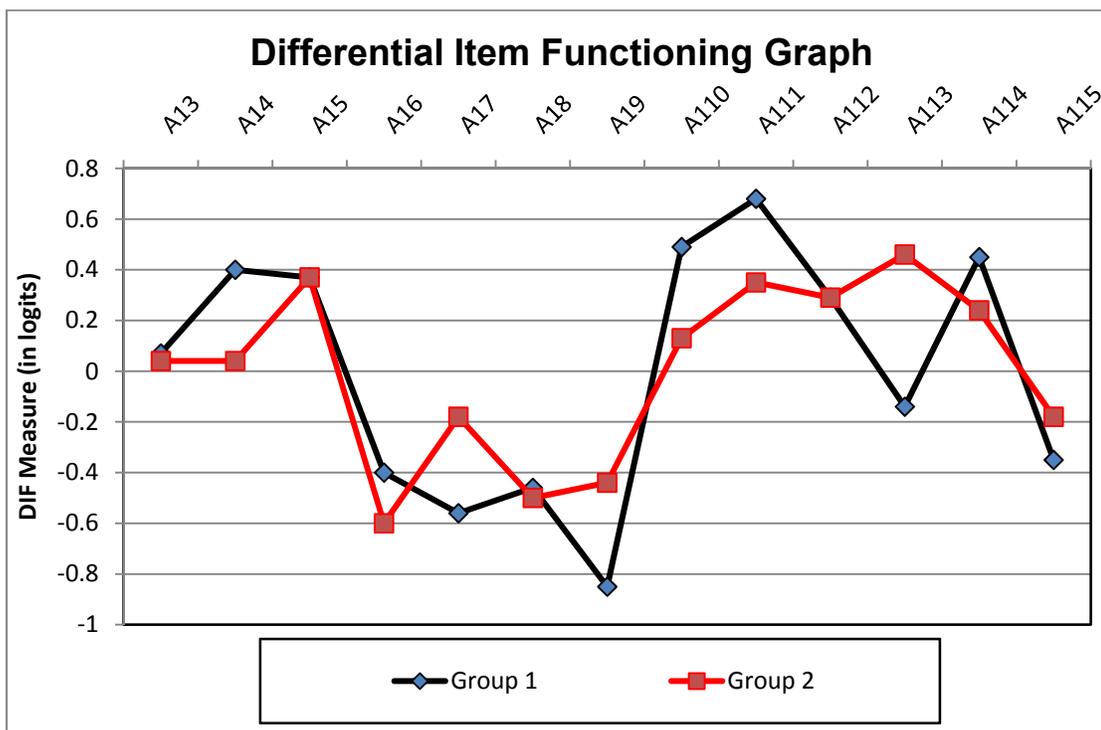


Figure 2. Differential item functioning graph. The y-axis is the overall endorsability of the item. Items higher up on the y axis are harder to endorse (agree to) while items lower on the y axis are easier to endorse. The two group's item endorsability patterns are graphed in an overlay showing differences in their individual likelihood response patterns when matched on overall ability. If the test is truly measuring a unidimensional latent trait, the DIF contrast should be relatively small between groups.

Once the DIF analysis was conducted in this study, the types of items that were exhibiting DIF were viewed for any noticeable patterns. For example, if there was more of a particular level of DIT-2 item (e.g., maintaining norms-stage 4) that consistently exhibited DIF this would indicate that these items or the definition of the trait was creating an item bias. The DIF contrast statistic is also used to interpret the statistical as well as practical significance of the DIF item.

DIF practical significance. One caution in interpreting DIF is that because statistical power to detect DIF is dependent upon sample size, large samples will

sometimes yield statistically significant but not practically significant DIF (Kim et al., 2007; Zwick, Thayer, & Mazzeo, 1997). Educational Testing Services (ETS) cautions that in DIF studies one ought to view DIF analysis in view of the degree of DIF according to three categories—negligible, moderate, and large. These categories are based on the absolute value of the DIF contrast between a reference group and a focal group. For example, suppose that on item 1 of the DIT-2 the focal group (BYU students) has a difficulty level of 1.0 and the reference group has a difficulty of 2.0. The DIF contrast absolute value would be 1.0 and this would be classified as an item that exhibits large DIF. Some authors view these DIF contrasts analogously to effect size statistics (Young & Sudweeks, 2005; see also Sudweeks & Tolman, 1993). Large DIF items are the only items that are classified as functioning differentially and are recommended for removal or revision (Sudweeks & Tolman). Although there are differing classifications of magnitude of DIF contrasts depending upon the methods used (Wright & Douglas, 1975, 1976), ETS and Linacre’s Rasch guidelines (which are the same) are commonly used in literature (Zwick, Thayer, & Lewis, 1999; see Table 4).

Table 4

Linacre (Winsteps) and ETS Guidelines for Interpreting DIF Contrast

ETS DIF Category	DIF Effect Size (Logits)
C = moderate to large	$ DIF \geq 0.64$
B = slight to moderate	$ DIF \geq 0.43 \leq 0.64$
A = negligible	$ DIF \leq 0.43$

Analysis #4 t Tests and Correlations for Other Potential Confounds with MMR

These statistics provided evidence concerning the degree to which other variables might affect the measured relationships between MMR and RO, specifically to answer research question 2—*what other variables predict both macromoral reasoning and religious orthodoxy and thus could be potential confounds?* Statistically significant results here provided evidence that another variable is a potential confound to the DIT-2 MMR scores.

This analysis consisted of two main sections. The first consists of correlations and *t* tests of the other demographic variables, specifically: returned missionary status, marital status, parent status (have children), association with criminals, mother's education level. This pointed out which of these variables should be studied more in depth (such as using partial correlation analysis) to determine if the variable is confounding MMR measurement.

The second analysis consisted of correlations and *t* tests based on DIT-2 story choices. For example, *t* tests of DIT-2 scores were conducted based on whether persons endorsed a doctor assisted overdose or not. These scores were used to show evidence of a pattern of rating the stories a certain way could predict DIT-2 scores. Thus, this analysis showed whether the content of the stories could be a potential confound to MMR. This also pointed out which of the story choices should be studied more in depth (such as partial correlation analysis) to determine if the variable is confounding MMR measurement.

Analysis #5 Partial Correlations Between RO and MMR Accounting for Other Potential Confounds

These statistics provided evidence concerning the relationship between RO and MMR after the variance of other confounds were removed. Statistically significant results may indicate that RO was still a confound to DIT-2 scores, but that the true effect was masked by other moderating variables that correlated with MMR and RO. Results here provided evidence to answer both research questions in concert.

A methodology that takes into account the relationships among the concomitant variables [variables that effect both dependent and independent variables] is partial correlations (Kachigan, 1986). Essentially, partial correlation analyses' were used to factor out existing variables that effect the relationship between RO and MMR so as to measure RO's sole effect on MMR. Therefore, I sought to account for these potential confounds and then measure the relationship between RO and MMR after the effects of these moderating variables were removed.

Analysis #6 Qualitative Analysis of Open-Ended Survey Responses

The qualitative data provided a more detailed examination of how RO and MMR were perceived by the BYU sample. Exploring these data provided more evidence to answer research question 1 by showing patterns of how experiences, philosophies, and religious teachings influence DIT-2 responses. In addition, the open-ended items collected data which provided evidence for other variables that might have confounded the measurement of MMR, yet were not specifically measured by the DIT-2 or demographic questionnaire.

The qualitative comments were collected based on the four open-ended questions (i.e., what life experiences, doctrinal teachings, philosophies of life, or other relevant issues influenced your DIT-2 decision—see Appendix B and G for a detailed description of each of these four categories). These comments were read several times. Categories of types of responses were then created. The comments were then coded by these categories (see Analysis #3 and Appendix G). This coding was done independent of any decision made and independent of any of their other comments made by the participant. These coded comments were then placed in relative frequency tables based on the four categories that elicited these responses to determine if life experiences, doctrinal teachings, philosophies of life, or other relevant issues influenced their DIT-2 story decision.

After the comments were coded, they were organized by frequency of the type of response (i.e., life experiences, doctrinal teachings, philosophies of life, other relevant experiences). They then were sorted based on DIT-2 story decisions. For example, if a participant chose the “should give overdose” option on the cancer story (see Appendix E), their qualitative comments were coded and contrasted with those who chose the “should not give overdose” option to explore patterns to see what variables, if any, were influencing the decision they made for the story. These were compared based on relative frequency of the type of coded comments. The reason that DIT-2 story decisions were used in this reporting is because research has shown that these are reliable estimates of overall DIT-2 scores and they had relatively high correlations with DIT-2 overall scores in this study and others (Xu et al., 2007).

Differences were reported as frequency counts if chi-square significance tests showed significant differences between the relative percentages of types of comments made when viewed by DIT-2 decision made. Since the chi-square statistic does not point out magnitude of a difference but only significance, relative percentages of types of responses that were graphed next to each other for visual contrasts. Descriptions were briefly described with a few illustrative examples to demonstrate the types of comments and how they were coded (see Appendix B and Appendix G for more detailed descriptions). By looking at patterns in all of the DIT-2 stories I was able to see general trends. I used the qualitative responses, the general trends noticed, and reviewed additional literature to determine why any contrasting relationships existed and discussed these in greater detail in Chapter V.

Conclusion

In summary, these methods of analysis shed light on whether religious orthodoxy had an effect on MMR as measured by the DIT-2 and if other variables also had an effect on either of these variables. Additionally, the implications of the DIF portion of this study are that if BYU student's individual item-response patterns do not vary from the UA filtered national item-response pattern on the DIT-2, then the DIT-2 items may be a valid measure of macromoral judgment for religiously orthodox persons in general.

CHAPTER IV

RESULTS

Research on the validity of the Defining Issues Test-1 (DIT-1) indicated that religious orthodoxy confounded the DIT-1 measurement of macromoral reasoning (MMR). The DIT-1 has since undergone a revision and now uses the less culturally sensitive version, the DIT-2. Since there are no data currently available to determine if the religious orthodoxy confounds are still in play, the aim of this chapter is to show an analysis of data collected from a religiously orthodox sample to answer the following two research questions: (1) *does religious orthodoxy confound the measurement of macromoral reasoning as measured by the DIT-2?* and (2) *What other variables predict both macromoral reasoning and religious orthodoxy and thus could be potential confounds?*

The analyses' in this chapter showed that, similar to Richards' study, BYU students are more religiously orthodox and politically conservative when compared to national scores. The analyses' mirrored what Richards' (1988) study found on lower and middle levels of MMR but contrasted what he found on upper levels of MMR. New methods of DIF analysis were employed as well as recognized standards for interpreting DIF contrast statistics and practical significance. Correlations were studied in more depth than previous studies by examining relationships between religious orthodoxy and DIT-2 scores when the effects of potential confounds were removed from this relationship. In addition, qualitative analyses' illuminated understanding of many of the patterns of religiously orthodox student responses and decisions made on the DIT-2.

Analysis #1 Descriptive Statistics Description of Sample

The descriptive statistics showed that the BYU sample was a good representation of U.S. college students who were religiously orthodox (RO) and that the BYU sample had a similar distribution of age, gender, and political viewpoints as Richards' 1988 sample. For example, Richards' average student age was 21.8 years while the average age of this sample was 21.4 years (99% were between the ages of 17-29). 138 of the 336 participants were male. Three hundred twenty-one of participants were U.S. citizens with over 96% of them reporting English as their primary language. As a group the sample was politically conservative as manifest by an average score of 4 (standard deviation, .71) on a scale of 1-5 (1 = very liberal political viewpoints, 5 = very conservative political viewpoints, see Table 5).

Table 5

Descriptive Variable Frequencies

Descriptive variable	Frequency ($n = 336$)
Age	99% between 17-29 years old (21.4 mean)
Gender	41% male, 59% female
U.S. citizen	96% U.S. citizen, 4% other
Primary language	96% English primary language, 4% other
Political leanings	83% conservative, 12% liberal, 5% neutral
Returned missionary	37% returned missionary, 63% not returned missionary
Mother's education level	92% had some college, 65% had bachelor's degree or more
Marital status	17% married, 1% separated, 82% single
Have children	Less than 2% have children
Shoplifting	7% close to someone arrested for shoplifting
Convicted of a crime	18% close to someone convicted of a crime
Purpose in suffering	90% yes—there is a purpose in suffering, 10% no purpose

Additional demographics showed that this sample consisted of 125 persons who served one and a half to 2-year missions for the LDS church (90% of these 125 persons were male). In addition, participant's mothers were highly educated (92% had some college, 65% had a 4-year degree or more), most participants were single (only 56 were married, four divorced/separated, while the rest were single or did not respond), and very few had children (only six reported that they had children). Twenty-four of the participants reported that they were close to someone who had been convicted of shoplifting while 61 reported that they were close to someone who had been convicted of a crime. Ninety percent of respondents said they felt there is a purpose in human suffering (see Table 5). These additional demographics were used to analyze subpopulations of this religiously orthodox sample.

To aid in understanding the interpretation of DIT-2 scores in this chapter, I have included a table here that describes the score, the acronym, range, calculation, and an interpretation of a several DIT-2 scores used in this chapter (see Table 6).

In this analysis I will show that my sample mirrored what Richards' (1988) study found on lower and middle levels of MMR but contrasted what he found on upper levels of MMR. In addition, similar to Richards' study, BYU students continue to be more religiously orthodox and politically conservative when compared to the national score.

These statistics were similar to Richards' (1988) comparisons with a national reference group for personal interest scores (PI-stages 2-3). For example, BYU students scored lower on PI-stages 2-3 levels of macromoral reasoning (MMR) as shown in Table 7, row 2 and Table 8, column 2. This finding indicates that BYU students are less likely

Table 6

DIT-2 Score Descriptions Used in Chapter IV

Full name	Acronym	Range and calculation of scores	Interpretation
Personal Interest Schema Score A.k.a Stage 2-3 score. The lowest of the DIT-2 levels of macromoral reasoning.	PI-stages 2-3	0-99. This score is a proportion of how much these items are endorsed over the other types of items. The score is multiplied by 100 to get a whole number.	High scores indicate that a person is likely to: <ol style="list-style-type: none"> 1. Allow direct advantages to themselves influence their macromoral reasoning. 2. View justice in regards of the fairness of exchanging favors for favors. 3. Allow the good or evil intentions of the others influence their macromoral reasoning. 4. Allow a desire to maintain approval or good relationships affect their macromoral reasoning.
Maintaining Norms Schema Score. A.k.a. Stage 4 score or conventional. The middle of the DIT-2 levels of macromoral reasoning.	MN- stage 4	0-99. This score is a proportion of how much these items are endorsed over the other types of items. The score is multiplied by 100 to get a whole number.	High scores indicate that a person is likely to focus on maintaining existing laws, rules, and formal structures of society to define their macromoral reasoning. These persons think that <i>what is</i> , is what <i>should be</i> . High scores also indicate that a person defers their macromoral reasoning to formal rules (secular or spiritual or otherwise).
Post-conventional Schema Score. A.k.a. Stage 5-6 schema scores. The highest level of macromoral reasoning.	PC-stages 5-6	0-99. This score is a proportion of how much these items are endorsed over the other types of items. The score is multiplied by 100 to get a whole number.	High scores indicate that a person is likely to macromorally reason with sharable, impartial, reciprocal, and debatable principles of justice. Persons macromorally reasoning focuses on organizing a society based on: consensus-producing (majority vote), due process of law (court), safeguarding rights, & socially reciprocal ideals. This score is a proportion of how much these items are endorsed over the other types of items.
N2 score. A.k.a DIT-2 overall macromoral score.	N2	0-95. The degree to which stage 5-6 items are ranked higher relative to stage 2-3 items.	Higher scores indicate that a person has a higher ability to macromorally reason and uses higher stages of macromoral reasoning.

Table 7

Descriptive Statistics Comparing a BYU Average Scores to National Norms

DIT score	Sample	Mean	SD
1. Overall DIT-2 Score (N2)	BYU	34.63	12.85
	Norm	33.43	15.23
2. Personal interest (Stages 2-3)	BYU	24.69	12.16
	Norm	27.09	12.49
3. Maintaining norms (Stage 4)	BYU	37.42	12.48
	Norm	32.73	14.00
4. Principled moral reasoning (Stages 5-6)	BYU	33.92	12.51
	Norm	34.72	14.67
5. DIT-2 Religious orthodoxy	BYU	6.19	2.67
	Norm	4.97	2.93
6. Inventory of religious belief	BYU	64.63	3.96
	Norm	47.16	12.33

BYU = Brigham Young University Undergraduates.

Norm = Average for English-speaking college students (Bebeau & Thoma, 2003).

to allow direct advantages to themselves influence their MMR when compared to the average college student (see Table 6 for more interpretation of PI-stages 2-3 scores).

However, these PI-stages 2-3 score differences were not as stark when comparing upper classmen. For example, BYU Freshmen scores were 4 points lower than the national PI-stages 2-3 average while BYU Senior's scores were nearly the same.

Another similar finding to Richards' (1988) study was that the BYU student's maintaining norms (MN-stage 4) scores were much higher for BYU students when compared to the current national averages (see Table 7, row 3 and Table 8, column 3). The DIT-2 framework indicates that persons with higher MN-stage 4 scores defer to laws, customs, rules, and formal structures of society to define their MMR.

Table 8

Descriptive Statistics by Educational Level

Educational level	Measure 1 N2 score		Measure 2 PI stages2-3		Measure 3 MN stage 4		Measure 4 PC stages 5-6		Measure 5 sample size	
	BYU	Norm	BYU	Norm	BYU	Norm	BYU	Norm	BYU	Norm
Senior	34.27	36.85	24.07	24.8	38.23	32.4	33.43	37.84	71	572
Junior	36.29	32.65	23.83	27.36	37.58	32.93	34.7	34.45	108	362
Sophomore	33.75	31.24	26.34	29.27	36.44	32.36	33.44	32.62	112	249
Freshmen	34.51	31.05	24.45	28.53	37.27	33.57	34.27	32.32	39	645

Measure 1 = Overall DIT-2 Score (N2)—degree to which stage 5-6 items are ranked higher relative to stages 2-3 items.

Measure 2 = Personal Interest (Stages 2-3)—proportion of items selected that appeal to direct advantages to the participant, focuses on good or evil intentions of the parties, or focuses on desires to maintain friendships or approval.

Measure 3 = Maintaining Norms (Stage 4)—proportion of items selected that focus on maintaining the existing system, existing roles, and/or formal organization structure.

Measure 4 = Post-Conventional (PC) or Principled Moral Reasoning (Stages 5-6)—proportion of items selected that focus on organizing a society based on: consensus-producing (majority vote), due process of law (court), safeguarding rights, & socially reciprocal ideals.

Measure 5 = Sample size (n).

BYU = Brigham Young University Undergraduates.

Norm = Average Norms for English speaking college students based on DIT-2 Scoring Guide (Bebeau & Thoma, 2003).

The descriptive statistics diverged from Richards' findings when comparing upper levels of moral reasoning. In Richards' studies, BYU students were 8 points lower on post-conventional (PC-stages 5-6) scores. This score indicates that one's MMR is built upon sharable, impartial, reciprocal, and debatable principles of justice (see Table 6 for more detailed interpretation of these scores). Although BYU students PC-stages 5-6 scores were lower, they were less than a one-point difference from a national norm

average (see Table 7, row 4 and Table 8, column 4).

As a final point of the descriptive statistics analysis, the DIT-2 has an internal measure of religious orthodoxy (RO), a measure that claims to correlate highly with RO measures (Navarez, Getz, Thoma, & Rest, 1999). This measure is referred to as the RO index and ranges from a score of 1 to 9, with 1 being low and 9 high. The descriptive statistics also showed that the BYU sample was more religiously orthodox than national averages (see Table 7 rows 5 and 6). This higher RO score indicates that BYU students were more likely to refer to scriptural teachings as a legitimate source for their MMR.

In summary, this analysis showed that BYU students were similar to Richards' 1988 sample in regards to demographics and lower to middle levels of MMR. It also showed evidence that BYU students are more religiously orthodox when compared to a national average.

Analysis #2 Correlations Between RO and MMR

This analysis showed evidence that RO scores correlated with some DIT-2 scores (see Table 9). The first column depicts a negative relationship with the humanitarian liberalism (HumLib) measure. The HumLib measure indicates the amount of agreement one's answers have with moral philosopher's and political scientists, which BYU students slightly disagree with as a whole—emphasized by the negative correlation. However, this finding is not new as the DIT-2 RO and HUMLIB scores have been known to yield strong inverse relationships (Bebeau & Thoma, 2003).

Table 9

Pearson Correlation—Religious Orthodoxy and DIT-2 Scores

Measure	1-HumLib	2-PI stages 2-3	3-MN stage 4	4- ConLib	5-Cancer story	6-Famine story	7-PC stages 5- 6	8-N2 overall score
Religious Orthodoxy	-0.18	-0.19	0.26	0.11	0.26	0.15	-0.08	0.01
Significance (two-tailed)	$p < 0.05$	$p < 0.05$	$p < 0.05$	$p < 0.05$	$p < 0.05$	$p < 0.05$	0.158	0.918
Coefficient of determination	.03	.04	.07	.01	.07	.02		

Note. The coefficient of determination of all of these scores is relatively low and therefore the proportion of variance in these variables doesn't account for too much of the proportion of variance in RO.

- 1 = Humanitarian Liberalism (humlib)
- 2 = Personal Interest (stages 2-3)
- 3 = Maintaining Norms (Stage 4)
- 4 = ConLib Scale
- 5 = Cancer story choice
- 6 = Famine story choice
- 7 = Post Conventional (stages 5-6 or p-score)
- 8 = Overall DIT-2 score (N2 score).

The second column in Table 9 shows that, overall, BYU student scores were negatively correlated with PI-stages 2-3 scores. This overall negative correlation coupled with the high RO scores of BYU students gives evidence that RO persons, in general, are less likely to focus on their own self-interests and how a macromoral decision affects only them. For example, in the famine story a person must decide if a rich man “deserves to be robbed” because he is hoarding food to make a profit during a time of severe famine. A person who has a high PI-stages 2-3 score has a higher probability of endorsing this item because their MMR framework justifies this line of thought on the basis that *if someone is unfair to me, I can treat them unfairly*. The negative correlation gives evidence that this schema of MMR does not appeal to RO persons.

The third column in Table 9 was the strongest correlation and reflected the correlation between RO and MN-stage 4 scores. The positive correlation indicates that the more RO a person is, the more likely they are to endorse stage 4 items as having great importance to the decision they made on the DIT-2. For example, in the cancer story, the question is raised as to whether the doctor is obligated by the same laws as anyone else even if he knows that giving an overdose would kill her? A person who has a high MN-stage 4 score is more likely to endorse this item as being important to their overall decision.

The fourth column displays a positive correlation with RO and the ConLib score. The ConLib score indicates how politically conservative or liberal a person rates themselves on a scale of 1-5 (1 = very liberal, 5= very conservative). The positive correlation shows that the higher RO score, the more likely a person is to rate themselves as politically conservative.

Columns 5 and 6 indicate that the more RO a person is, the more likely they are to endorse not stealing food in a famine and not euthanizing a suffering cancer patient. Thus far, all of these findings are similar to previous studies involving the DIT-2 and RO.

However, this study diverges on the final two correlations in Table 9. In previous DIT-1 studies, PC-stages 5-6 levels of MMR and N2 (overall) scores were significantly correlated for RO persons. Contrastingly, this study found there was not a statistically significant correlation or mean difference between these scores and RO scores. The lack of any correlation with RO scores and upper levels and overall DIT-2 scores suggests that the makers of the DIT-2 may have corrected this issue on the DIT-2. However, when the

BYU and UA total samples were combined ($n = 1,224$), Pearson correlations showed that RO scores had statistically significant negative correlations with PC-stages 5-6 scores and N2 scores. However, when viewed in practical terms, these larger sample correlations were still very small (all less than .09) and have very small coefficients of determination.

Analysis #3 Differential Item Functioning Person Reliability

Item-response theory (IRT) yields a statistic similar to Cronbach's coefficient alpha as it ranges from 0-1 and higher scores indicate greater reliability. When the focal and reference groups were combined (BYU & UA) the person reliability was .84, a score which indicates a high level of consistency in person responses and that the test can separate into 2-3 levels of proficiency.

Graphed IRT Item DIF Contrasts

Several iterative analyses' were examined within Winsteps. The first of these analyses' included a filtered UA sample matched with a BYU sample on total DIT-2 scores (N2). For example a 40 N2 score on the BYU sample was matched with a 40 N2 sample in the UA RO filtered sample. Item response patterns were compared on a logit scale and placed on an equal interval scale for comparison. Other than items 17-18 (from the reporter story), item response patterns for the BYU sample did not drastically differ from the UA sample (see Figure 3).

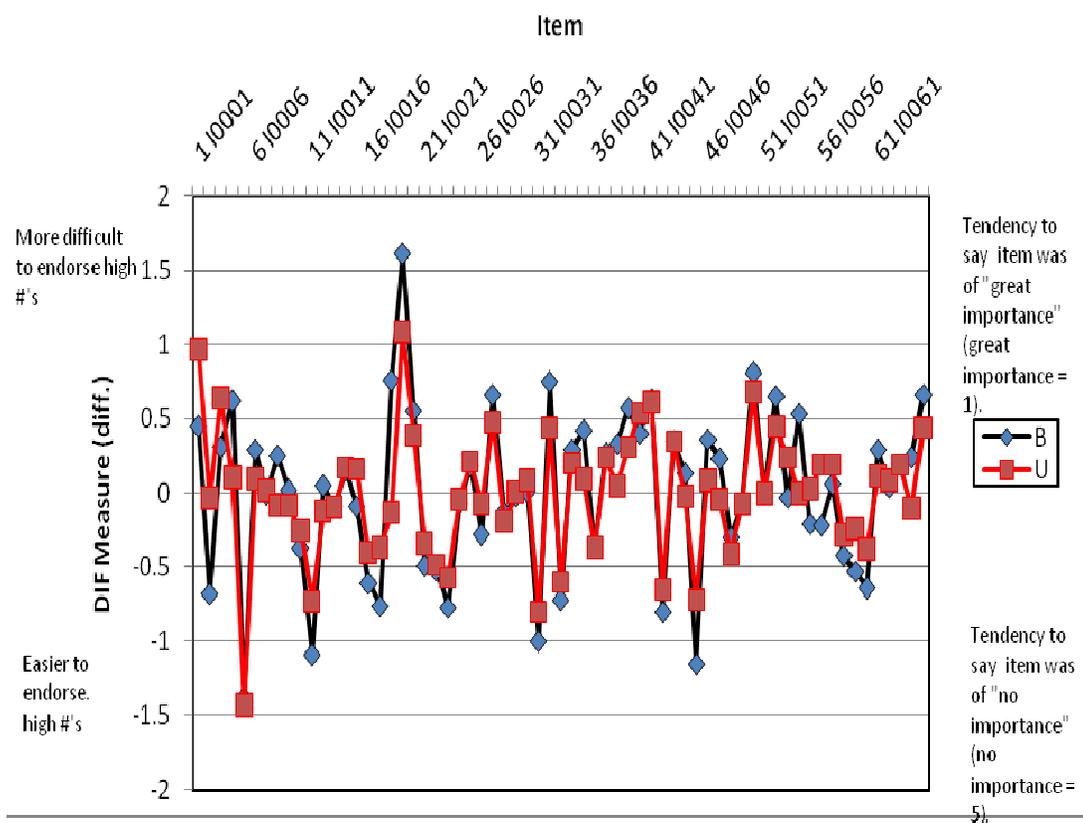


Figure 3. Person DIF plot of total BYU (B) and UA (U) samples. Notice that the relative sizes of the differences are not very large. In a DIF study, when items are easier to endorse for one group we ask if this discrepancy is due to a greater ability of that group or a systematic bias. This figure does not show any large difference patterns between the groups.

DIF Statistical Significance

In addition to viewing the item response patterns graphically, the following table summarizes the DIF statistics of this study (see Table 10). Notice that 30 of the 65 items exhibited statistically significant DIF.

Some patterns are worth noticing from the statistically significant DIF items. For example, seven of the 13 items which are easier for BYU students to endorse over UA

Table 10

DIF Contrast Charts with Statistical and Practical Significance

Item #	DIF contrast in terms of BYU	Joint standard error	Statistical probability of true difference	Practical significance	Item easier for __ to endorse (i.e., rated greater MMR importance for __)	Probable MMR stage of item
FC	-.49	.11	<.05	Slight/moderate	UA	DIT-2 choice
F1	-.51	.08	<.05	Slight/moderate	UA	PI (2-3)
F3	.48	.09	<.05	Slight/moderate	BYU	MN (4)
R4	.87	.09	<.05	Moderate/large	BYU	Anti
R5	.75	.14	<.05	Moderate/large	BYU	P (5-6)
R12	-.34	.08	<.05	Small/negligible	UA	MN (4)
S5	.42	.1	<.05	Small/negligible	BYU	MN (4)
S8	.42	.09	<.05	Small/negligible	BYU	MN (4)
S11	.38	.09	<.05	Small/negligible	BYU	P (5-6)
C5	-.42	.1	<.05	Small/negligible	UA	Meaningless
D1	.37	.09	<.05	Small/negligible	BYU	MN (4)
R1	-.32	.08	<.05	Small/negligible	UA	P (5-6)
D12	.35	.09	<.05	Small/negligible	BYU	MN (4)
R3	-.28	.09	<.05	Small/negligible	UA	PI (2-3)
F5	.27	.08	<.05	Small/negligible	BYU	MN (4)
F7	.27	.08	<.05	Small/negligible	BYU	P (5-6)
D3	-.23	.08	<.05	Small/negligible	UA	P (5-6)
R10	-.22	.08	<.05	Small/negligible	UA	MN (4)
C1	.24	.09	<.05	Small/negligible	BYU	MN (4)
CC	-.25	.1	<.05	Small/negligible	UA	DIT-2 choice
DC	-.21	.09	<.05	Small/negligible	UA	DIT-2 choice
F10	-.22	.1	<.05	Small/negligible	UA	Anti
D2	-.19	.08	<.05	Small/negligible	UA	PI (2-3)
C2	-.21	.09	<.05	Small/negligible	UA	P (5-6)
F2	-.21	.09	<.05	Small/negligible	UA	PI (2-3)
D11	.18	.08	<.05	Small/negligible	BYU	P (5-6)
S12	.19	.09	<.05	Small/negligible	BYU	P (5-6)
D5	-.18	.08	<.05	Small/negligible	UA	PI (2-3)
R7	-.17	.08	<.05	Small/negligible	UA	MN (4)
R9	-.17	.09	<.05	Small/negligible	UA	Meaningless

PI (2-3) = Personal Interest Schema which is the lowest stages of MMR.

MN (4) = Maintaining Norms Schema which is mid-level stages of MMR.

P (5-6) = Post-conventional Schema which is the highest stages of MMR.

Anti = Anti-establishment item which measures a DIT-2 index of whether someone has anti-establishment attitudes towards society's systems of justice.

Meaningless = measures whether someone is just filling in the bubbles—does not pertain to MMR.

DIT-2 Choice = What choice did a person make on the DIT-2 moral dilemma (1-3 scale).

students appear⁶ to be MN-stage 4 items. What the IRT model tells us is that these items are easier for BYU students to mark as important to their overall MMR when compared with UA students with the same overall MMR score. Thus, if the overall DIT-2 score reflects a unidimensional latent trait (MMR), and BYU and UA students are correctly matched on an overall ability of that trait, then something other than that trait appears to be confounding response patterns. The presence of DIF would suggest that there is something misaligned with the items or the definition of the trait.

A second pattern worth noticing is that five of the items that are easier for a UA student to endorse appear to be PI-stages 2-3 items. These items are statistically easier for UA students to mark as important when matched with BYU students who have the same overall MMR score.

These two patterns are worth noticing because PI-stages 2-3 and MN-stage 4 items appeared to be performing differently via descriptive and correlational analyses'. The DIF analysis may shed light on why that is the case (i.e., these items were systematically easier for one group to endorse).

Other noticeable findings showed that two of the five meaningless items were easier for the UA student to endorse. Rating these items as important to one's MMR indicates that students do not understand or are not thoughtful about the moral dilemma. This trend is important because DIT-2 designers have asserted that RO persons are not as capable of reasoning through the complex MMR dilemmas. The fact that the meaningless items are statistically easier to endorse for the reference group (UA students), may be

⁶ The CSED does not release their item coding nor their scoring methods.

evidence that this sample of RO persons are *more* cognitively capable of reasoning through the DIT-2 items.

A final pattern worth noticing is that 3 of the 5 story decisions performed differently for the UA students when compared to a BYU matched sample. Specifically, this shows that it was statistically easier for UA students to advocate stealing during a famine, euthanizing a suffering cancer patient, and protesting to a war by taking over a college administration building.

Two additional Winsteps analyses' were run using a random sample from both groups (BYU and UA) as well the full 1,224 sample of both groups and found the same results. The reason for these similar results is due to item-response theory's robust assumption of yielding sample-independent measures of item difficulty.

DIF Practical Significance

While the statistical significance of a DIF analysis can show important response patterns, DIF researchers caution to not place an overemphasis on statistical significance. This caution is because DIF studies require large samples and statistical significance is a function of sample size (Gall, 2001). From a practical standpoint, very few of these DIF items have a large magnitude of DIF contrast (analogous to effect size). In other words, 30 of the items did exhibit statistically significant DIF contrasts in some interesting patterns. However, there was not an overly excessive amount of it due to the large sample sizes.

In this study, the current ETS guidelines showed only two items were worth noting as large to heavy in practical significance (DIF Rasch analysis contrast $> .64$). In

addition, the direction of the DIF (+ or -) is not uniform across items (i.e., 13 were easier to endorse for BYU students compared to 17 for UA students). Wright and Douglas (1976) claim that when DIF is not consistently in one direction, when most DIF item contrasts are below ± 0.50 , and when the test has more than 20 items, the impact of DIF on person measurement is generally small. However, moderate and small DIF contrasts on the remaining 28 items could yield a cumulative effect on test scores. This cumulative effect may be why BYU PI-stages 2-3 and MN-stage 4 scores differ—because something other than the trait of MMR is confounding response patterns, even when comparing equally matched groups.

In applying Mantel-Hanzen ETS DIF contrasts to previous research data (Richards, 1988), only one of the 25 reported biased items appeared to be practically significant in how it performed. However, the items in Richards DIT-1 studies were all uniform in the direction of DIF (more difficult for BYU students to endorse).

Analysis #4 *t* Tests and Correlations for Other Potential

Confounds with MMR

This analysis focused on finding potential confounds to MMR DIT-2 scores. Distributions of DIT-2 N2 scores were normally distributed. This distribution allowed many parametric statistical analysis methods to be employed. However, other DIT-2 scores were skewed in their distributions, so nonparametric tests, specifically the Mann-Whitney test and chi-square methods, were also used secondarily to verify that the violation of true random sampling assumptions of *t* tests were not producing any of the

effects (Boneau, 1960; Gall, 1996).

The reporting of this analysis consists of two main sections. The first consists of correlations and *t* tests of the other demographic variables such as returned missionary and marital status (see Table 11) while the second section consists of correlations and *t* tests based on DIT-2 story choices.

Returned missionaries. Dummy variables were used for this analysis (returned missionary = 1, others = 0). There was a slight tendency for returned missionaries to endorse MN-stage 4 over stage 5 items on all the stories. The combined differences produced a significant difference on the total MN-stage 4 and 5 scores. Specifically, a -0.13 correlation ($p < .05$) was observed between returned missionaries and stage 5 moral reasoning scores. The *t* tests also showed that returned missionaries stage 5 means scores were nearly 2 points lower when compared with non-returned missionaries ($p < .05$). In

Table 11

DIT-2 Scores Relationships with Demographic Variables

Variable	<i>t</i> test statistical significance on DIT-2 scores	Significant <i>r</i> with DIT-2 scores	Description
Returned missionary	Yes	Yes	1.9 points higher on stage 4 scores ($p < .05$), 1.5 points lower on stage 5 scores ($p < .05$), more frequently reported a purpose in suffering (.015 Pearson chi Square). Positive <i>r</i> w/stage 4 scores (.12, $p < .05$) and negative correlation w/stage 5 scores (-.13, $p < .05$). However, there was no significant <i>r</i> when stages 5-6 were combined.
Marital status	No	No	
Have children	No	No	
Association w/criminals	No	No	
Mother's education	No	No	

other words, returned missionaries, as a group, were scoring lower on upper levels of MMR than nonreturned missionaries. However, when stages PC-stages 5-6 were combined, there was no statistically significant difference.

Returned missionaries also exhibited a statistically significant positive correlation with MN-stage 4 responses ($r = 0.11, p < .05$) and a 1.5 point greater mean score on stage 4 scores ($p < .05$). Returned missionaries also tended to respond that there is a purpose in suffering more than the rest of the BYU sample (mean difference .134, $p < .05$). The single returned missionary response patterns did not statistically differ from the married returned missionary's response patterns.

Marital Status and Children

The previous findings of slightly higher MN-stage 4 and lower stage 5 scores were also manifest when viewed by marital status. However, none of these were statistically significant. The higher descriptive MN-stage 4 and lower stage 5 scores may be due to a secondary effect of the returned missionary variables because it came as no surprise that returned missionaries significantly differed from non-returned missionaries in marital status at BYU. In the LDS faith, young men are taught that their lives will be blessed by serving a mission *prior* to marriage (Hinckley, 1998; Kimball, 1979).

Analyses were also conducted to study whether having children would impact DIT-2 scores. However, only six of the participants reported having children in their home. Chi-square categorical tests showed that participants who had children said that it would be best to call off the riotous and violence-threatened school board meeting and advocated not to steal food during a famine. These findings are suspect because of the

scarcity of data or variance ($n = 6$) to run further tests with the parenting variable.

Association with Criminals

The next survey response that was analyzed showed that those BYU students who knew someone who had been caught shoplifting seemed to be more sympathetic towards the person with a history of shoplifting in the reporter story #2. In this moral dilemma, a reporter needs to decide whether to report a story about a man running for political office who had shoplifted 20 years ago but had since changed his life. Those who reported that someone close to them had shoplifted tended to endorse *not* to report the story more than those who endorsed to report the story because the DIF analysis showed statistically significant likelihood response patterns between those who were close to someone who had shoplifted and those who were not. The following comment from the qualitative data illustrates the reason for the different response patterns.

My brother shop lifted for a long time when he was a child, minor things like candy bars and lighters from drug stores, but has since grown up since then. He is one of the most incredible people, extremely honest and good. His actions as a young boy, if anything, have kept him from making poor decisions now because of the guilt and pain he felt for stealing all those years ago.

In so many words, this DIT-2 decision performs in a systematically unexpected manner for persons who are close to someone who has been convicted of shoplifting (i.e., it was easier for them to endorse not to report the story). However, an association with criminals had no other effect on overall DIT-2 stage or RO scores.

Mother's Education Level

To determine the combined predictive strength of the demographics, nominal

logistic regression analyses were used with Minitab software. Although no strong least squares regression formulas were discovered, there was an association with mother's education level that surfaced as the only statistically significant variable of the regression model. However, this variable was weak in practical significance (less than 1% of variance explained). In addition, Chi-square and correlational analyses did not find any statistical significance that mother's education level had on DIT-2 scores or RO scores.

DIT-2 Story Choices

This section consists of an analysis of DIT-2 story choice correlations followed up by a *t* test based on the cancer story decision (groups who did or did not endorse doctor-assisted overdose). The cancer story decision shows the most contrast between groups and is appropriate to use as the items pertaining to this story are used to compute the religious orthodoxy variable.

To begin this analysis, all DIT-2 story choices were dummy coded. (Do you favor the action of taking the food? Do you favor the action of reporting the story? Do you favor calling off the next open meeting? Do you favor the action of giving more medicine? Do you favor the action of demonstrating this way?) (see Appendix F for a full description of DIT-2 stories). The number 1 was used for those who advocated a DIT-2 decision such as: should steal in a famine, report a story, keep having the unruly open meetings, euthanize a patient, and demonstrate by taking over the administration building. The number 2 was used for "can't decide." The number 3 was used for those who did *not* advocate: stealing in a famine, reporting the story, holding the unruly open meetings, and so forth.

When missing responses were deleted, correlation and chi-square statistics for category responses showed that the cancer story and famine story choices were positively correlated (see Table 12). This correlation means that a person's decision on whether or not to euthanize a suffering cancer patient could be used to predict their response to whether a person should steal during a famine. In addition, group statistics indicated that persons who advocated stealing food in a time of famine and euthanizing a suffering patient were likely to have higher PI-stages 2-3 scores, lower MN-stage 4 scores, and lower RO scores. Many of the other story choices correlated with these scores also. In addition, only the Cancer story choice correlated with PC-stages 5-6 scores ($r = -0.17$, $p < .01$). This means that persons who advocated euthanizing the patient had higher PC-stages 5-6 scores. None of the other story choices correlated with PC-stages 5-6 scores nor N2 scores.

These correlations with the RO, PI-stages 2-3, and MN-stage 4 scores show

Table 12

BYU Sample DIT-2 Story Choice Correlations (n = 291)

DIT-2 story choice	1.	2.	3.	4.	5.	6.	7.
1. Famine							
2. Reporter	0.11						
3. School board	0.06	0.09					
4. Cancer	0.22*	-0.03	0.05				
5. Demonstration	-0.07	0.07	-0.06	0.12*			
6. Personal interest (Stage 2/3)	-0.32*	0.17*	0.02	-0.22*	-0.06		
7. Maintain norms (Stage 4)	0.32*	-0.23*	0.03	0.35*	0.05	-0.50*	
8. RO	0.15*	0.19	-0.13	0.26*	0.10	-0.19*	0.26*

* $p < .01$

evidence of a consistent pattern of the way people rate the stories. These consistent correlation patterns are evidence some of the DIT-2 story decision choices are predictors of whether someone's MMR focuses on a PI-stage 2-3 or a MN-stage 4 framework for their MMR as well as whether a person is more or less RO. These correlations add to the evidence that there is an underlying construct being measured by the DIT-2 stories that may be confounded by RO or another confounding variable.

DIT-2 Story Decisions Correlations and t Tests

Since the DIT-2 story decisions were correlated with PI-stages 2-3, MN-stage 4, and RO scores, further analysis was performed based upon the DIT-2 story decisions. The cancer story gives the strongest illustration of a general pattern discovered (see Table 13). In addition, the cancer story ranking items are used to help determine the RO score so it was appropriate to separate persons based on this story decision for analysis in this study.

Key statistically significant findings of this analysis are summarized in Table 13. For example, those within the BYU who advocated not giving an overdose (non-OD) to a suffering cancer patient (non-OD group) scored 5 points higher on stage four reasoning when compared to those who advocated giving an overdose (OD group). The non-OD group was also 2 points higher on the average religious orthodoxy score (1-9 scale). The non-OD group's scores were also lower on the NUMCD variable, a variable that indicates the degree to which one cannot decide what to do on the DIT-2 moral dilemmas (1-5 scale). This score indicates that non-OD BYU students were more decisive on the DIT-2 stories and is related to the amount of consolidation in one's macromoral

Table 13

Statistically Significant DIT-2 Mean Differences of Those Who Endorsed a Doctor-Assisted Overdose Compared to Those Who Did Not ($p < .01$)

Score	Should give OD	Score description
RO	-2.11	RO is a measure of religious orthodoxy and ranges from 1-9.
HumLib	1.11	Humlib is a reflection of the number of times a choice on the DIT-2 matches academic moral philosopher's choices in political science and philosophy. Scores range from 1-5.
PI schema (stages 2-3)	2.63	PI Schema represents the proportion of items that represent lower levels of moral reasoning and focus on the direct advantages to the actor and the fairness of simple exchanges of favor for favor. This is converted to a whole number and ranges from 1-99.
MN schema (stage 4)	-5.15	MN Schema represents the proportion of items selected that appeal to a stage four, maintaining norms consideration. These focus on maintaining the existing legal system, maintaining existing roles and formal organizational structures. This is converted to a whole numbers and ranges from 1-99.
Post-conventional schema (stages 5-6)	2.34	Post-conventional Schema score represents the postconventional schema score. This score reflects the proportion of items selected that appeal to the two highest stages of moral reasoning, stages 5-6, which were selected over the two lowest stages (2-3). This is converted to whole numbers and ranges from 1-99.

reasoning. In other words, the non-OD BYU students are a little more “set” in their macromoral reasoning. The HumLib variable is the degree to which one’s DIT-2 scores conform with content experts (moral philosophers) in the field of moral reasoning. This score is usually inversely related to the RO index so it is no surprise that the non-OD BYU students do not score as high on this measure (i.e., RO persons tend to not agree with the moral philosophers in some macromoral issues).

Overall, these scores reflect that as a group, those who do not advocate euthanizing a suffering cancer patient tend to give high endorsements to MN-stage 4

items. Conversely, those who advocate euthanizing a suffering cancer patient also tend to give low endorsements to MN-stage 4 items. The DIT-2 story advocacy and item endorsement patterns were consistent across all stories (e.g., those who advocated stealing in a famine, reporting a story, etc., had lower MN-stage 4 item scores). These patterns suggests that the actual content of the story and initial story decision, not the DIT-2 items (which the test purports to derive its scores from), may also be producing a confound to MMR scores.

Analysis #5 Partial Correlations Between RO and MMR Accounting for Other Potential Confounds

The goal of this study was to analyze the effects that RO has on DIT-2 scores. The purpose of this analysis was to verify that extraneous variables were not confounding MMR scores. Recall that RO scores were correlating with PI-stages 2-3 and MN-stage 4 scores. In addition, this study has shown that whether or not someone has served an LDS mission, the cancer and famine story choice, political adherence (ConLib), marriage status, how well a person agrees with moral philosopher choices (HumLib), whether they feel there is a purpose in human suffering, and how consolidated they are in the DIT-2 story decisions (NUMCD), all affect RO, PI stage 2-3, and MN stage 4 DIT-2 scores. Therefore, each of these variables was viewed in partial correlation analyses' to view the effects of RO on DIT-2 stages 2-4 scores.

When missionary status, marriage status, purpose in human suffering, famine choice, NUMCD, ConLib, and HumLib scores were individually partialled out of the

correlation between RO and PI-stages 2-3 and MN stage-4 scores, none of the adjusted correlations between RO and stages 2-4 changed more than three hundredths. The consistency in RO and DIT-2 score relationships is evidence that these variables are not significantly confounding variables in the RO and stages 2-4 relationship. However, the cancer story decision slightly affected the RO and stages 2-4 relationship more than the rest of these extraneous variables. Specifically, the RO—MN-stage 4 correlations decreased from 0.26 to 0.21. In addition, the RO—PI-stages 2-3 correlations decreased from -0.19 to -0.15. This decrease is to be expected because the cancer story decision is actually used to calculate the RO score. In other words, DIT-2 researchers assume this relationship already exists. In essence, the partial correlations showed that religious orthodoxy can predict PI stages 2-3 and MN stage-4 scores regardless of the extraneous variables that this study sought to control for. Thus we can rule out the effects of these other measured variables upon the DIT-2 MMR and RO relationship.

Analysis #6 Qualitative Analysis of Open-Ended Survey Responses

The final analysis involved viewing open-ended survey responses. The purpose of this analysis was to discover any uncontrolled variables as well as to verify the DIT-2 framework. For example, to verify the DIT-2 framework, persons with high PI-stages 2-3 scores should have qualitative comments that focus on direct advantages to themselves and conversely, persons with low scores should not focus on direct advantages to justify their MMR decisions. The findings reflected that BYU students had very few PI-stages 2-3 justifications. In addition, persons with high MN-stage 4 scores should have

qualitative comments reflecting that they are unreflectively deferring to societal or religious norms. However, the MN-stage 4 framework was not verified in qualitative data. Contrastingly, student comments were very reflective of personal values which happen to be religious.

To briefly describe how qualitative comments were exported, coded, and analyzed I describe two phases: an interpretation of qualitative responses and a frequency description of total relative responses (see Appendix G for a complete description of the interpretation of qualitative responses and frequencies of responses). The data illuminated understanding of many of the patterns of BYU student responses and decisions made on the DIT-2.

Relative Percent of Response Rates

As noted earlier, the decisions made for the DIT-2 stories had significant correlations with DIT-2 scores. Therefore, it was anticipated that examining the qualitative data by decisions made on the DIT-2 story decisions would illuminate significant patterns that would illustrate how BYU students differed in their macromoral judgment as well as discover any additional potential confounds to MMR.

After student responses were interpreted and coded, the results were reported as frequency counts and descriptive narratives to look for differences between the relative amounts of different types of responses on the macromoral dilemmas. For example, on the cancer story, persons who did and did not endorse euthanizing the suffering patient both wrote that mercy was a relevant issue in this story. However, patterns of statistical significance emerged as relative frequencies of expected and observed responses by DIT-

2 decision made were compared. For example, by looking at the relative frequency amount of responses involving mercy as an issue that influenced a person's cancer story decision, those who endorsed a doctor-assisted overdose mentioned this factor of mercy much more frequently than those who did not endorse a doctor-assisted overdose. It was interesting to graph the differences in the relative percent of coded qualitative data by the decision made for the DIT-2 stories (see Figures 4-7 shown and described separately below). These showed some remarkably insightful contrasts for what these individuals reported was influencing their macromoral decisions.

Cancer Story

The cancer story was of most interest to this study because the internal RO measure is derived from an item pertaining to this moral dilemma. As comments were coded and graphed, it was clear that as one type of comment increased for those who endorsed a story decision, they decreased for a group who did not endorse a story decision. For example, it was quite clear that those who endorsed the doctor-assisted overdose mentioned the following influences on their decision more often than those who did not endorse this action: the right to the avoidance of suffering, that individual rights are greater than family rights, and the God-given right to chose or agency.⁷ Contrastingly, those who did not endorse the doctor assisted overdose mentioned the following influences on their decisions more often than those who did endorse this action: God's dominion over life (God has not given the locus of control in life/death matters to man),

⁷ Agency is an LDS doctrine that involves the right to individual choice. Agency was coded as a locus of control issue in this story as participants indicated that God gives man control to decide for himself.

life is a sacred gift—therefore, one has a responsibility to fight for it, the doctor’s actions are an irresponsible act of killing, and that there is purpose/ benefits in suffering so one has a responsibility to endure it (Figure 4).

In analyzing these contrasting findings in the relative percent of qualitative responses by DIT-2 story choice, there appeared to be two dominating factors influencing the cancer decision: a person’s view of macromoral rights versus macromoral responsibilities, and God versus Man’s locus of control in matters of life and death. However, after an additional reflection of the qualitative comments, the God versus Man’s locus of control may also be related to a person’s view of macromoral rights and macromoral responsibilities (i.e., does God give man a *right* to chose or is there a *responsibility* to Him in matters of life and death?). A few illustrative examples of

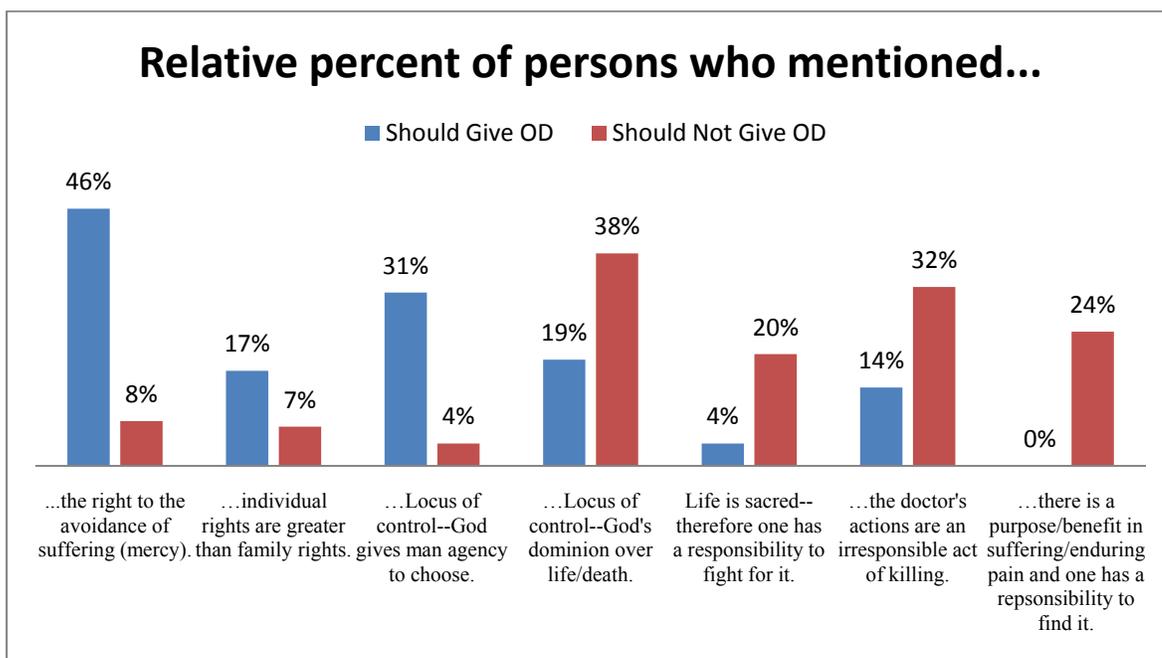


Figure 4. Relative percent of coded qualitative comments by DIT-2 cancer story choice.

macromoral rights-based comments included the right of a patient and the right to the avoidance of suffering. For example:

“The patient has every right to declare the comfort of her death before it actually happens in practical means.”

“No one should have to suffer more than they can handle.”

“If we can help others avoid suffering we should do that.”

Many of the above macromoral rights-based rationales seemed to exist within a vacuum that focused on individual macromoral rights with very little expressions of reciprocity of the macromoral code.

Contrastingly, the individuals who did not endorse the doctor assisted overdose frequently focused on an individual, family, and societal macromoral responsibilities and reciprocity of their macromoral code. A few illustrative examples of the responsibilities-based comments included a focus on the macromoral responsibility to loved ones, to life itself, and to find meaning through experiences. For example:

“My friend’s father had a brain tumor that caused him pain and suffering, however, he suffered with it until the end so he could spend more time with her.”

“The situation above is tough one, however, my thoughts on it is that we shouldn’t help others destroy their life.”

“I believe Victor Frankl was right when he said that if there is any meaning in life there must be meaning in suffering.”

Famine Story

The relative percentage of qualitative comments in the Famine Story⁸ followed

⁸ It should be pointed out that this story may have been confounded by LDS teachings regarding self-reliance. LDS leaders teach that one way to be more self-reliant is to have at least a 3 month supply of food storage in case of emergencies, job loss, or other extenuating circumstances. Thus, a rich man hoarding food may not be seen as inherently immoral if it was viewed as an act of self-reliance. The findings of this study did not confirm nor rule out the impact of the LDS teachings on self-reliance and food storage.

similar patterns as the cancer story—that is, as one type of comment increased for those who endorsed a “should steal” decision, they decreased for a “should not steal” decision (see Figure 5).

Again, the most common theme for the “should not steal” patrons was a theme of macromoral responsibility/accountability for one’s actions while the “should steal” patrons focused on their comments on the macromoral right to life. There was also a differential prescribed locus of control between God and Man in the “should not steal” patron’s comments as evidenced by the relative percent of “choose the right” comments and that some of the laws we are accountable to are the 10 commandments, specifically “Thou shalt not steal.” Illustrative examples focused on a macromoral weighting of the right to life versus the responsibility to respect others rights to own property. For example, the focus on macromoral responsibilities and reciprocity of the moral code to

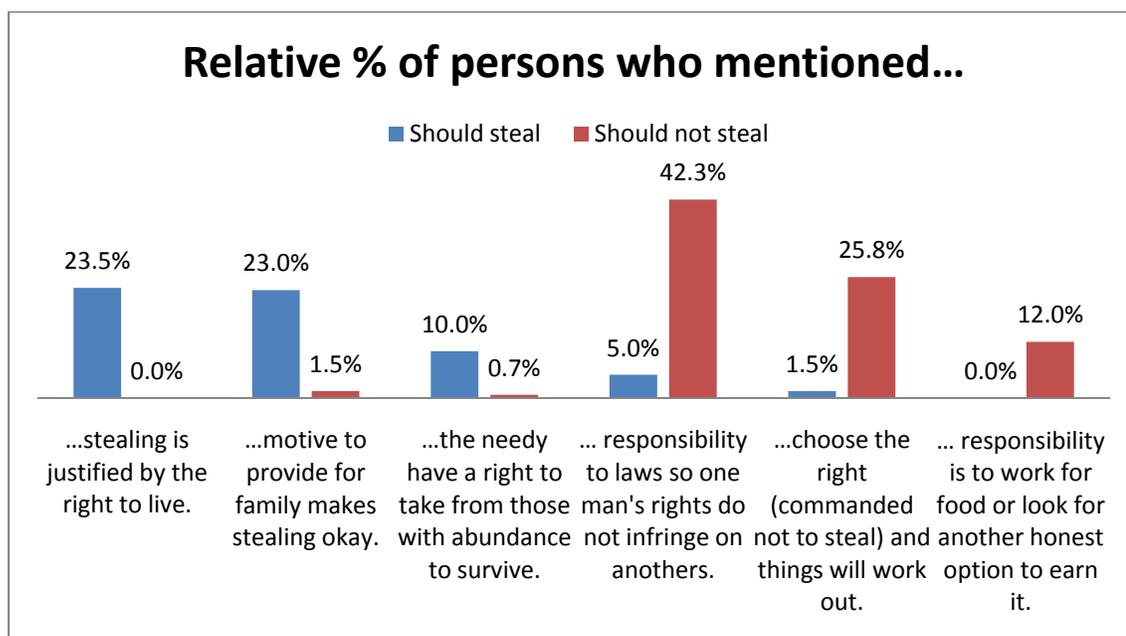


Figure 5. Relative percent counts of coded comments by DIT-2 famine story decision.

protect both party's rights to property are reflected in the following illustrative comments.

"The person who is rich still has a right to his property regardless of what his intentions are."

"If all of Mustaq Singh's neighbors had the same idea he did and thought that the food they took wouldn't be missed, then a significant amount of the rich man's food would be taken... I think it is human nature to think that we are the exception and that we are justified in what we do."

"Although it is unfair the rich man is hoarding food and being greedy does not mean Mustaq should also sink to his level to steal."

Contrastingly, the focus on macromoral rights comments, were similar to the following:

"I place a higher priority on human life than money. Sometimes laws are not moral."

"The right to life [influenced my decision]."

"It is far more important to keep someone alive than to worry about stealing."

"This was confusing because it's an obvious commandment that we shouldn't steal, but then again, I look at it as being the letter of the law vs. the spirit of the law. If you are doing it to save your family and you aren't doing anyone much harm I feel that I would still be able to justify it, even though it may not be right."

Reporter Story

The third qualitative data analyzed came from the reporter story. In this story, a young reporter (Molly) has to decide if she will report a story about a candidate whom she supports that shoplifted when he was young but has since changed his life. The data from this story also showed similar inverse relationships among comments as the first two (i.e., those who endorsed reporting the story and those who did not tended to have opposite frequencies of coded qualitative data). For example, 85% of the relative amount

of comments of those who endorsed not reporting the story mentioned a macromoral responsibility towards the subject while only 31% of those who said to report the story mentioned this similar theme (see Figure 6). The following comments illustrate the crux of the macromoral rights of the public to get the full story contrasted with the macromoral responsibility to the subject:

“Her job is to report the news to the public.... She gives the community the information/story and they decide for themselves.”

“I learned one time of President Truman falling down one time and a young reporter tried taking a picture of him and another reporter hit his camera out of his hand and broke it because of the ideal that reporters then had of respecting their subjects.”

A second major factor of a differential prescribed locus of control between God and man was manifest in the disproportionate number of references to the assumed

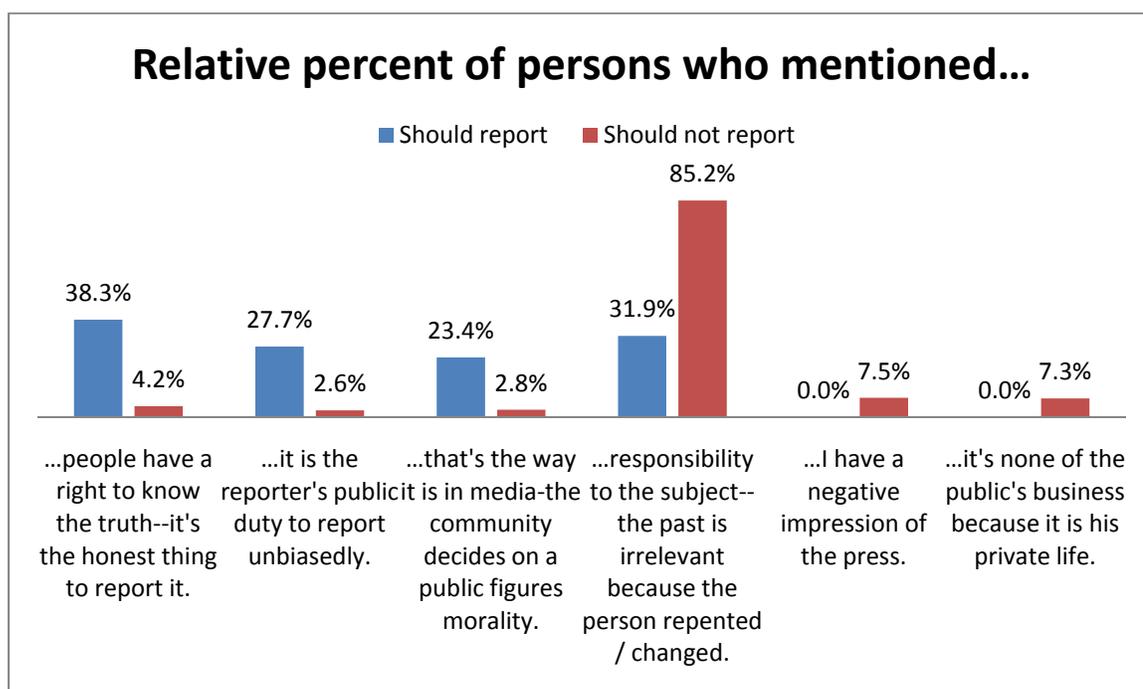


Figure 6. Relative percent of coded comments by DIT-2 reporter story decision.

“repentance” of the politician mentioned in “should not report” comments (see Figure 6).

School Board Story

The fourth story analyzed was the school board story. The macromoral issue was whether or not to continue to hold a public school board meeting, despite the unruly attendees and threats made to the board members. BYU students responded in a more diverse way to this story than they did with the others as shown in the balanced amount of respondents to each extreme (102 comments “should call off the next open meeting” and 114 comments “should not call of the next open meeting”). Perhaps this more diverse response pattern is due to the lack of religious content in the story (i.e., scriptures do not speak much on the morality of holding unruly public school board meetings).

However, similar inverse relationships were also found in this story in the coded comments (see Figure 7). For example, comments from those who advocated holding the

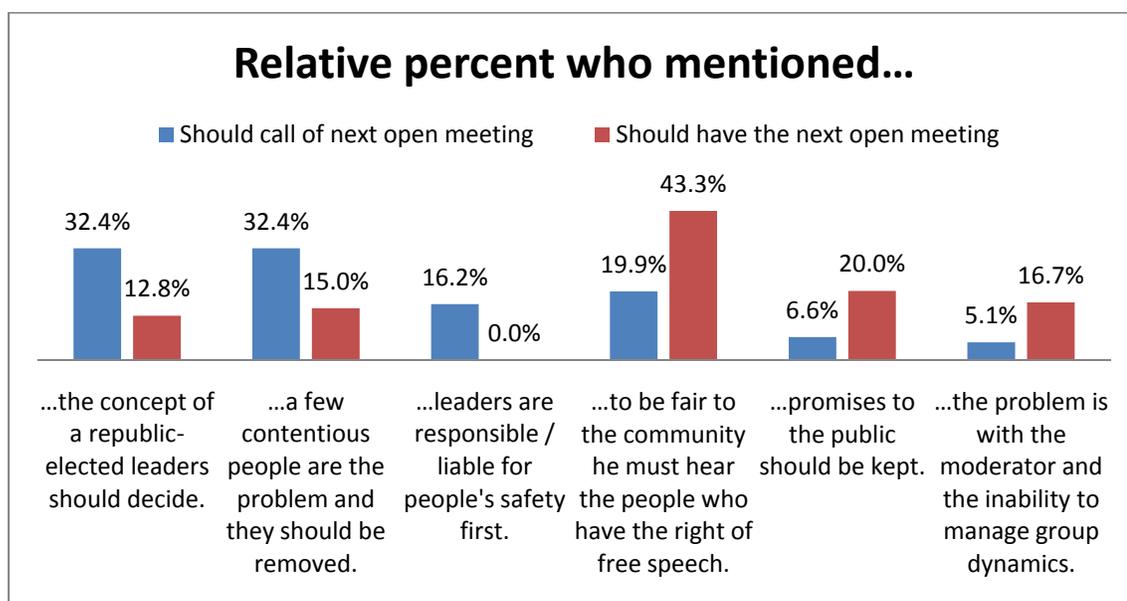


Figure 7. Relative percent of coded comments by DIT-2 school board story decision.

next meeting focused more on the macromoral right to be heard (free speech), that the problem was with the leadership more than the unruly people, and the right to expect a leader to keep their word. For example, “We all have the freedom of speech” and “I think people opinions are important in a democracy. If possible they should be able to express their opinions.”

Those who advocated calling off the next meeting focused on the macromoral responsibility of the leaders to make decisions (the notion of a republic), the responsibility of leaders to keep the board members and others safe, and the irresponsibility of the unruly people as a reason for forfeiting their rights to be heard (thus hinting that the macromoral right of free speech is attached to a macromoral responsibility). One student reported, “If the people cannot conduct themselves properly, they should not be allowed in the meetings.” Another reported, “If the community can’t discuss the school closure with open-minds and respect to all, then they should not have the opportunity to.”

Demonstration Story

The demonstration story raises the issue of whether it is right for college students to take over their administration building in protest of a war. This DIT-2 story was seen as irrelevant and unappealing to the vast majority of BYU students. Qualitative comments related a sentiment that most of them thought it was not worth their time and energy to cognitively reason much about it. Comments also reflected a cost/benefit analysis in terms of student’s time management, student’s responsibility to other’s rights in society, and the ineffectiveness of results of protesting in this way—all reflective of a

pragmatic philosophical approach to life. For example:

“We cannot transcend the agency of others, like the School.”

“Responsibilities over rights.”

“There are usually more progressive resolutions to problems but many are just too lazy to look or wait for those solutions.”

There was also little variance in the comments and decisions on this story. Only seven persons said the students should keep demonstrating in these ways, 17 were undecided, 281 were against it and the rest of the sample did not respond to this item. The lack of interest and lack of a perceived macromoral issue made this story irrelevant in analysis.

Overall Qualitative Analysis of the Differences Between BYU Sample Participants

Recall that the purpose of this analysis was to identify if other variables could be potential confounds that were influencing MMR and RO on the DIT-2 stories and to verify if the DIT-2 framework was appropriately describing MMR for a RO person. This analysis consisted of viewing what life experiences, doctrinal teachings, personal philosophies, and other relevant issues were influencing respondent’s decisions on the DIT-2 stories (see Appendix G for a more detailed description of these four components).

When groups were compared based on the DIT-2 story decisions made (e.g., should or should not give an overdose), the groups of students did not differ in the amount or depth of life experiences that related to any of the stories—relatively the same percentage of comments were made about family experiences, work-related experience,

educational related experiences, and general life experiences for both groups. For example, on the Cancer story, 33% of the comments of those who endorsed a doctor assisted overdose focused on family or close friends who had cancer. Similarly, 31% of the comments of those who did not endorse a doctor assisted overdose focused on family or close friends who had cancer. The “other relevant issues” responses were either common to both parties or unique to individual comments. Thus, the influences of external experiences (in the form of life experiences) appeared to be quite balanced for the two groups. However, the extent to which doctrinal teaching or which philosophy of life a person felt applied to the macromoral issue were very imbalanced when viewed by which story decision a person endorsed (see Figures 4-7). Both the doctrinal teachings and philosophy of life comments appeared to be related more to internal reasoning rather than external events in the person’s life. These findings led the researcher to conjecture that the internal interpretation of external life experiences may be what influences DIT-2 responses rather than an external experience alone. However, it should be pointed out that the mere balanced amount of external life experiences in the qualitative data is not sufficient grounds for assuming that *all* external life experiences do not influence one’s macromoral schema. Nor is the balanced amount of external life experiences sufficient grounds for assuming that all internal views of macromorality are the sole basis for the DIT-2 scoring. In addition, this study did not fully verify that doctrinal teachings and philosophy of life comments were based solely on internal reasoning. Nonetheless, the balanced quantity of references to life experiences is a relevant finding in seeking to discover patterns of what other variables influence DIT-2 responses.

Another relevant finding that the qualitative data showed was that two variables that appeared to be influencing the RO sample's MMR responses and that a third variable was not. The first variable that was influencing the RO sample's MMR responses was actually a verification of the DIT-2 schema scores. That is, the vast majority of the BYU students did not appeal to personal interests (PI-stages 2-3) to justify their MMR story decisions. In other words, there was a lack of personal interest comments justifying why RO persons chose the way they did on the DIT-2 stories. This scarcity of personal interest comments may be why BYU student mean PI-stages 2-3 scores were significantly lower.

The second variable influencing decisions made on the stories was dependent upon differing doctrinal interpretations of God versus man's locus of control and a general weighting of whether one focused on macromoral rights versus macromoral responsibilities. In one way, God versus man's locus of control could be couched into a view of macromoral rights and macromoral responsibilities. For example, a man who believes that he has a God-given right to dictate his own macromorality, may respond in this way in order to emphasize his preference for an individual right over a responsibility to a Transcendent Truth or Being. This unexpected variable of weighting macromoral rights and macromoral responsibilities appeared to be stronger than any other quantified demographic variable of this study and should be considered in future studies.

The third variable that did not seem to be influencing the RO sample's MMR responses was that they were not overly deferring to laws, customs, cultural or religious norms, or societal norms for the reasons behind their MMR. This lack of deferring to law was also shown in focus group interviews as students would say things such as,

“Euthenasia is against the law but that’s not what makes it wrong.” These types of comments contrasted with what the DIT-2 framework suggests is the reasons that RO persons have higher MN-stage 4 scores.

In essence, the qualitative analysis showed that within a BYU sample there are differing philosophies that influence a student’s views on Kohlberg’s justice based concepts that form the basis of his view of macromorality (what is right for society). Specifically, a minority and majority group emerged. These groups of students had differing philosophies on macromoral rights and macromoral responsibilities when it comes to macromoral choices, differ in opinions on whether a macromoral right and a macromoral responsibility exist separately or are conjoined, and appeal to differing doctrines as influencing their macromoral choices (such as agency and mercy). In addition, the framework for rating PI-stages 2-3 items appears to be valid for RO persons while the MN-stage 4 framework appears to be misaligned when the qualitative data are viewed. These findings will be elaborated upon in Chapter V.

CHAPTER V

DISCUSSION

Purpose of Study and Research Questions

The purpose of this study was to explore whether the DIT-2 solved a problem of item bias against religiously orthodox persons and to explore if other variables were related to the measurement of religious orthodoxy and macromoral reasoning (MMR). Specifically, the research questions were as follows.

1. Does religious orthodoxy confound the measurement of macromoral reasoning as measured by the DIT-2?
2. What other variables predict both macromoral reasoning and religious orthodoxy and thus could be potential confounds?

The short answer to these questions is that this study presented evidence that RO does confound MMR as measured by the DIT-2. However, the type of DIT-2 scores affected by RO and the reasons for why RO confounds DIT-2 scores differed from previous DIT-1 research. The following is a summary of major findings. These are followed by three major discussion points that will explain: first, which DIT-2 scores were confounded by RO; second, what previous theories for why RO persons score the way they do were supported or not supported; and third, propose that the DIT-2 maintaining norms (MN-stage 4) schema is insufficient to explain the variables that contribute to why RO persons score the way they do. These discussion points are followed by delimitations in this study and suggestions for further study.

Summary of Major Findings

1. The BYU Sample Was an Appropriate Sample to Generalize to RO Persons

There were four reasons that led me to conclude that a BYU student sample was an appropriate sample to represent RO college students in the United States (US). First, the BYU sample had a DIT-2 RO mean score that was 1.22 points higher than the national averages (1-9 scale). Previous research has pointed out that the DIT-2 RO score strongly correlates with other measures of RO (Rest, 1999). Second, on Brown and Lowe's Inventory of religious belief (1-75 scale), the BYU sample mean scores were almost 15 points higher than the national averages. In addition, the BYU sample standard deviations were significantly smaller on both RO scales, indicating that the BYU sample was a more homogeneous group in their RO ideology. Third, qualitative data showed that BYU students frequently used religious vernacular as they explained their MMR on a questionnaire (e.g., the doctrine of agency, mercy, compassion, forgiveness, repentance, etc—see also Appendix G). Fourth, recall that RO is defined by how one uses scriptural teachings. As ancillary evidence that Mormons are more religiously orthodox, recent research has pointed out that Mormons study what they consider the word of God far more often than most U.S. religions (Pew Forum on Religion & Public Life, 2008). Thus, a BYU student sample is an effective one for generalizing to RO college students in the US.

2. RO Persons' Scores Differ When Compared to a National Reference Group

The first evidence that RO person's scores differ when compared to a national reference group was found in the descriptive data. Specifically, RO persons mean scores were lower on PI-stages 2-3 scores, higher on MN-stage 4 scores, and lower on PC-stages 5-6 scores. Second, *t* tests verified that PI-stages 2-3 scores and MN-stage 4 scores were significantly different ($p < .05$), while PC-stages 5-6 were not. Third, RO scores for the BYU sample correlated with PI-stages 2-3 and MN-stage 4 scores but not with PC-stages 5-6. Fourth, when the effects of other potential confounds were removed via partial correlation analyses, the RO-DIT-2 correlations still remained intact. In essence, the partial correlations showed that religious orthodoxy can predict PI stages 2-3 and MN stage-4 scores regardless of the extraneous variables that this study sought to control for. Thus, we can rule out the effects of these other measured variables upon the DIT-2 MMR scores (e.g., mother's education level, association with criminals, etc.). However, as we view these correlations, it is important to point out that correlation is not causation. Thus, I concluded that either RO does predict DIT-2 scores *or* it is being masked by other variables not quantifiably measured in this study. Qualitative data showed evidence of the latter and will be elaborated on in the discussion section.

3. DIT-2 Items Appear to Measure Internal Schemas of Macromoral Reasoning

The DIT-2 claims it is measuring internal schemas of MMR (Rest et al., 1999a). Findings from this study lend support that the DIT-2 is measuring internal schemas of MMR. Specifically, the high person reliability and item reliability scores of the RO group

shows that RO persons responded quite consistently to the items. Thus, I conclude there was not a reliability measurement issue confounding the relationships between RO persons and DIT-2 scores.

A second point of support for the claim that the DIT-2 measures internal schemas of MMR, is that BYU students IRT response patterns were not showing a consistently strong pattern of item bias. Although there were 30 of the 65 items that showed statistically significant DIF, the DIF contrast statistics only showed that there were two items that exhibited a practically significant amount of DIF contrast levels against the RO group. The remaining 28 items exhibited only moderate or small amounts of DIF contrast. Further, the moderate and small amounts of DIF contrasts were not consistently in one direction (thus, there was no favoring or unfavoring either the RO or reference group). Thus, item bias theory does not appear to be having a strong influence on DIT-2 scores. Since item bias is the only existing theory that focuses on external measurement issues, and since this study did not strongly support an item-bias theory, the reasons that RO persons have differing DIT-2 scores may be more due to internal reasoning, as the developers of the DIT-2 suggest.

A third point of support that the DIT-2 is measuring internal schemas of MMR, comes from the amount of qualitative comments regarding external events (life experiences) that were influencing person's DIT-2 choices. Specifically, the relative amount of qualitative comments about life experiences, that influenced persons DIT-2 decisions were mostly equal when contrasted by DIT-2 story choice. For example, in the cancer story, 33% of the relative comments from persons who endorsed a doctor-assisted

overdose mentioned that someone close to them had been diagnosed with cancer. Thirty-one percent of the relative comments from persons who did not endorse a doctor-assisted overdose mentioned that someone close to them had been diagnosed with cancer. As anecdotal evidence of the balanced amount of life experience comments, an interesting experience happened in pilot study focus group interview. Two persons, who had coincidentally both had cancer treatment, used their experiences to justify or denounce euthanasia and respectfully disagreed with each other's conclusions drawn from their life experience. This similar pattern of life experiences being balanced when viewed by DIT-2 choices was in all of the DIT-2 stories (see Appendix G). The balancing of external experiences lends support that there are internal schemas for how people macromorally reason and it may be the internal interpretation of life experiences that determine how life experiences are used to support or oppose a macromoral choice. However, the mere balancing of life experiences that influenced DIT-2 story decisions is not sufficient evidence to conclude that the DIT-2 stories are completely independent from life experiences and *only* measure internal schemas. Literature on moral development also shows that there are internal rationales in moral development (Emsberger & Manaster, 1981).

A fourth point of support that the DIT-2 is measuring internal schemas of MMR also comes from the qualitative data. In contrast to the finding of equal amounts of external life-experiences rationales for their decisions, persons who chose differently on the DIT-2 stories cited differing internal rationales to justify their decisions (see Chapter IV, Figures 4-7). The patterns of relatively equal amounts of external rationales (life

experiences) to justify DIT-2 story decisions, while noticing differing amounts of internal rationales (such as doctrinal teachings and philosophies of life) that were used to justify DIT-2 story decisions, led me to conjecture that the DIT-2 stories may tap into internal schemas of MMR. However, the MN-stage 4 schema does not appear to be sufficient to explain RO persons' internal schema of MMR, which seems to be masked by other confounding variables. This possible masking will be my third discussion point.

To summarize, the three major findings of this study I found that first, a BYU sample was an appropriate representation of RO persons, second, that RO persons' scores differed in predictable ways, and third, that there is support that the DIT-2 is consistently measuring an internal schema. However, this study also found evidence that even though the DIT-2 may be consistently measuring MMR in predictable ways for RO persons, the validity of the MN-stage 4 framework is brought into question. These findings will be elaborated on in the following discussion points.

Discussion Points

Discussion Point #1-RO and DIT-2 Scores RO and PI-stages 2-3 Negative Correlations

The analyses of Chapter IV showed that the BYU sample was similar to Richards' 1988 sample in regards to lower levels of personal interest (PI-stages 2-3) scores when compared to a national norm-referenced group (BYU score = 24.7, UA score = 27.0). In addition, RO scores were negatively correlated with PI-stages 2-3 scores.

In the DIT-2 framework the lower PI-stages 2-3 scores indicate that BYU students are less likely to allow direct advantages and their own interests to influence their MMR

when compared to the average college student. Qualitative comments also verified that most BYU students were not focusing on their own interests in their MMR. Thus, I conclude, as earlier researchers conclude, that RO persons in general are less likely to focus on their own self-interests and how a macromoral decision affects only them. The tendency to not focus on their own self-interests appears to be a valid variable, which explains why RO persons score lower on the PI-stages 2-3 score. Thus, I also conclude that the PI-stages 2-3 schema is a valid framework for explaining RO person's DIT-2 scores.

The finding that RO persons are less likely to focus on direct advantages to themselves may also be why the DIF analysis of the DIT-2 story decisions on famine and cancer stories were moderately easier for the UA students to endorse—the very nature of these stories is self-serving. For example, in the famine story a person must decide if it is appropriate to steal from a rich man who is hoarding food to make a profit during a time of severe famine. In fact, one of the PI-stages 2-3 items asks if this rich man “deserves to be robbed” because he is doing hoarding food.⁹ Someone with a high PI-stages 2-3 score has a higher probability of endorsing this item as important to their decision because their MMR schema justifies this line of reasoning (i.e., if someone is unfair to me, I can treat them unfairly). Not surprisingly, the DIF study showed that this item was moderately easier for the reference group to endorse over the BYU sample. This is just one example

⁹ Again, it should be pointed out that this story may have been confounded by LDS teachings regarding self-reliance. LDS leaders teach that one way to be more self-reliant is to have at least a 3-month supply of food storage in case of natural disaster emergencies, job loss, or other extenuating circumstances (which could include a famine). Thus, a rich man hoarding food may not be seen as inherently immoral if it was viewed as an act of self-reliance. The findings of this study did not confirm nor rule out the impact of the LDS teachings on self-reliance and food storage.

that RO persons are less likely to allow their personal interests to interfere with their MMR. This expected finding mirrors what earlier research studies have found (Richards, 1988).

RO and MN-stage 4 scores positive correlations. The analyses' also showed that the BYU sample was similar to Richards' 1988 sample in regards to higher levels of maintaining norms (MN-stage 4) scores when compared to a national norm-referenced group (BYU score = 37.4, UA score = 32.7). In addition, RO scores were positively correlated with MN-stage 4 scores indicating that the more RO a person is, the more likely they are to endorse stage 4 items as having great importance to the decision they made on the DIT-2.

RO persons have a more complex MMR schema than rest suggests. The DIT-2 framework suggests that persons with higher MN-stage 4 scores defer to laws, customs, rules, and formal structures of society to define their MMR. For example, in the cancer story, the question is raised as to whether the doctor is obligated by the same laws as anyone else even if he knows that giving an overdose would kill his patient. A person who has a high MN-stage 4 score is more likely to endorse this item as being important to their overall decision.

No quantified variable in this study accounted for higher MN-stage 4 scores in the BYU sample so I turned to the qualitative data to verify that the DIT-2 framework was correctly identifying the reasons for why a RO person would have higher MN-stage 4 scores (i.e., were they indeed deferring their MMR to laws, rules, customs, etc.). However, qualitative data did not show that RO persons were excessively deferring to

societal roles, customs, formal structures, nor laws in their MMR. In fact, the law was only cited as a reason for their MMR decision in 8% of the relative comments on the cancer story and was practically nonexistent in the other stories (less than 2%).

In addition, I should point out that the mere correlation of personal values to U.S. legislature values (laws) or religious laws does not mean one is deferring to law. For some RO persons, in an Aristotle-like manner, they view U.S. laws and the end of politics as inherently moral (Barnes, 1984; Clayton, 2006) but only when the laws align with their own values. This Aristotle-like view of laws was reflected in interviews as students would report, “Well, the law says... but that’s not what makes it right.” By Kohlberg’s own reasoning, rights derived from laws alone are insufficient (Rest, 1999) because laws can be manipulated to support any end (Denvir, 1999).

Therefore, RO persons agree with Kohlberg on this point—rights and what is just is not derived from U.S., state, or city legislature laws alone. An item, such as “should the community’s laws be upheld,” is not discriminating enough to determine whether someone is deferring their MMR to the law, or if the law just happens to align with their own MMR schema or values. Thus, a positive correlation on the MN-stage 4 score with RO, may not truly be measuring what it claims to be measuring.

Similarly, rights derived from an isolated doctrinal teaching alone can also be used to support any end. In addition, religious laws were mentioned more frequently in the qualitative data because I asked if religious teachings applied. However, students who scored high and low on MN-stage 4 scores mentioned religious teachings that applied to the DIT-2 story. Very few seemed to be deferring to these teachings or laws

alone for their decision. Rather than deferring to the religious teaching alone, they showed reflective thoughts regarding how to apply their chosen teaching that was influencing their DIT-2 decision. The application of the differing doctrines appeared to be based on personal values, some of which happen to be religious, to the given macromoral story.

In addition, DIT-2 decisions appeared to inform a theological weighting of differing LDS doctrines to justify the DIT-2 decision. This weighting of differing LDS doctrines was unexpected because these students all came from the same theological ideologies yet they focused on differing LDS doctrines to justify their decisions—a contrast from the previous research which suggested that the same LDS doctrines were confounding student responses (Richards, 1988). This weighting of differing LDS doctrines was illustrated as the qualitative data showed that the most RO and least RO in the BYU sample appeared to be weighing differing doctrinal teachings and philosophies of life as justifications for their MMR decision. As a more concrete example, recall that those who endorsed the doctor-assisted overdose mentioned the God-given right to choose or agency¹⁰ and the doctrine of mercy considerably more than those who did not endorse a doctor-assisted overdose. Contrastingly, those who did not endorse the doctor-assisted overdose mentioned God's dominion over life (i.e., God has not given the locus of control in life/death matters to man) and that there is a divine purpose/benefit in suffering (see Chapter IV, Figures 4-7). Thus, the DIT-2 story decision appeared to be informing a theological weighting of differing LDS doctrinal teachings as justifications to

¹⁰ Agency is an LDS doctrine that involves the right to individual choice. Agency was coded as a locus of control issue in this story as participants indicated that God gives man control to decide for himself.

support their macromorality.

Interestingly, the most RO and least RO within the BYU sample showed complex and logical MMR rationales stemming from their chosen premises' of which doctrine or philosophy they weighted as important. For example, if providing for one's family was macromorally valued more than respecting other's property then a person logically chose for Mustaq to steal during the famine. Thus, illustrative statements of "allowing one's family to die is far worse than stealing" contrasted with "I believe in respecting other's property" and were seen throughout the famine story qualitative data.

Thus, I conclude that RO persons have much more complex and logical MMR schema rather than a simple deferring to societal norms. In addition, these complex patterns of weighting differing philosophies and theological doctrines appear to be a viable reason why RO persons generally have higher MN-stage 4 scores, that is they subscribe to macromoral responsibilities over macromoral rights. This unexpected finding and mislabeling of MN-stage 4 will be elaborated later on as discussion point #3.

As final applicable evidence that RO persons have a more complex MMR schema than the DIT-2 suggests, I cite the finding from the DIF analysis that two of the five meaningless items were moderately easier for the reference group to endorse rather than for the RO group. Endorsing these meaningless items shows whether a person understands and is reasoning through each item or if they are merely marking an item as important because it sounds complex. I raise this finding from the DIF study as ancillary evidence that the RO group was *more* carefully reasoning through each item than the reference group. However, this finding may have been due to cognitive abilities and

verbal abilities rather than MMR abilities. Nonetheless, it is related as it shows that RO students were less likely to endorse an irrelevant-to-MMR-item as important to their overall MMR. The lack of endorsement of meaningless items adds to my conclusion that RO persons are not simple deferrers to external laws, rules, and societal roles in their MMR schema but come to their MMR conclusions from their own logical and complex MMR schemas—schemas which the DIT-2 framework is deficient in measuring.

RO and PC-stages 5-6 and N2 scores neutral correlation. The post conventional (PC-stages 5-6) and overall DIT-2 MMR score (N2) clearly diverged from Richards' findings. The lack of any practically significant correlation with RO scores and PC-stage 5-6 and N2 scores suggests that the makers of the DIT-2 may have corrected a scoring issue that existed on the DIT-1 in their new version, the DIT-2. This lack of correlation of RO and PC-stages 5-6 scores was unexpected because in past studies, PC-stages 5-6 scores were significantly lower for RO persons. These scores indicate that one's MMR is built upon sharable, impartial, reciprocal, and debatable principles of justice. The only finding that came close to verifying these earlier findings was that LDS returned missionaries, a more RO subgroup, had a small negative correlation with stage 5 scores. However, when stage 5-6 scores were combined into one score (as the DIT-2 current framework for scoring does) or when correlations were analyzed by RO alone, there were no significant differences or correlations.

In summary of discussion point #1, RO did correlate strongly with PI-stages 2-3 scores. The reasons why these scores were lesser for the RO group can be explained very well by the DIT-2 framework—RO persons do not focus on direct advantages to

themselves in the MMR. This finding was verified in the qualitative responses showing very little PI-stages 2-3 reasoning among RO persons.

RO was confounding MN-stage 4 scores. The reasons why these scores were greater for the RO group cannot be explained empirically by the quantified variables that this study measured (e.g., mother's education level, returned missionary, political orientations, etc). In addition, the qualitative responses do not support that DIT-2 framework's answer for why a person would score higher on MN-stage 4 MMR. That is, RO person's qualitative responses do not show a simple deferring to laws, customs, nor societal norms as reasons for their MMR but rather a weighting of macromoral responsibilities over macromoral rights and differing LDS doctrines as reasons for their MMR.

There were no practically significant differences in PC-stages 5-6 scores or N2 scores for the RO group when compared with the reference group. This lack of difference between the RO and reference group on PC-stages 5-6 scores adds evidence that the DIT-2 is consistently measuring this schema of MMR for RO persons and others.

*Discussion Point #2-How Previous Theories
of RO and MMR Correlations Were
Supported or not Supported*

Historically, RO groups have scored lower on PC-stages 5-6 levels of macromoral reasoning measured by DIT-1 outcomes. As mentioned, this study found that the DIT-2 measures this level of MMR consistently. In addition, historically, RO groups score higher on maintaining norms scores (MN-stage 4) of moral reasoning. This second pattern was consistently strong in this study. In past research studies, three main theories

have been proposed for why stages 4, 5, and 6 levels of DIT-2 items perform differently for RO persons: the developmental theory, the preemptive theory, and the bias theory.

Developmental theory. This study's findings are consistent with others in rejecting the developmental theory for why RO persons do not score well on the higher stages of moral reasoning (Kay 1998; Richards, 1988; Richards & Davis, 1992). Recall that the developmental theory suggests that RO ideology blocks moral development so that RO persons are developmentally less able to morally reason at higher levels. Rest claimed that some religious ideologies block developmental reasoning by casting a sinful eye towards the questioner and those who wish to scrutinize a code of morality (Rest et al., 1999b). In addition, Rest claimed that if an adolescent was caught up into a religious ideology with a strong MN-stage 4 development phase of life, that this ideology could block moral development in later life (Rest et al.).

However, several studies found evidence against this theory (see Chapter II). For example, Kay (1998) found evidence against this theory as well as Lawrence (1979) by showing that RO persons *could* reason at higher levels of moral reasoning but *chose not to* based upon religious reasoning. Emler also simulated a study in which he told conservative persons to answer like a radical libertarian and found that persons could morally reason in different schemas, but chose to adhere to a particular one (Emler, Palmer-Cantion, & St. James, 1995; see also Emler, Resnick, & Malone, 1983). Quantitative evidence in this study showed that RO persons did understand the complex moral issues and, therefore, were not developmentally disabled in their MMR.

Rest further claimed that RO persons had high MN-stage 4 scores because they

simply deferred to external laws to solve complex moral problems (Rest et al., 1999a). However, pilot study data using Pargament's Religious Problem Solving Scale (Pargament et al., 1988) gave evidence that a simple deferring is not the case (see Chapter III). In fact, returned missionaries had very high self-directing scores on this scale. Pargament suggests that strong self-directing scores are typical of persons that "emphasize the freedom God gives people to direct their own lives. This approach appears to be an active coping orientation which stresses personal agency" (Pargament et al., p. 91). This style is not anti-religious but compatible with a self-directing approach to life espoused by RO persons. This finding correlates with prominent LDS writings which assert that the existence of God and absolute truth does not lead one to *defer* their moral reasoning, but rather leads to a belief in freedom with accountability. For example:

It was Dostoyevsky's character Ivan Karamazov who believed that if God is dead, then everything is allowed. Well, both the premise and the conclusion are misleading. Neither God, nor law, *tell* you what you must do. That is a fiction. They tell you what the inevitable consequences will be of what you do do. (Madsen, 2000)

From the pilot study finding and the comments made in qualitative responses, RO persons are very much into a self-directing style of problem solving rather than a deferring one. This finding contrasts with what Rest and colleagues have claimed by suggesting that RO persons *defer* their moral judgment based upon a maintaining norms schema (Rest et al., 1999b) and is also evidence against the developmental "moral blocking" theory. In essence, Pargament's RPSS outcome scores, the lack of contrast in PC-stages 5-6 scores, and qualitative comments reflected that RO persons were not post-conventionally deficient or superior in this schema of MMR skills—some other factor

was accounting for differences in DIT-2 scores. In addition, qualitative data showed that MN-stage 4 scores were not high due to a RO person's deferring their MMR.

Preemptive theory. The second theory to explain the DIT-2 scoring patterns of RO persons is the Preemptive Theory. This theory suggests that RO persons purposely impede processing needed for stage 5-6 levels of principled moral reasoning because it is seen as less morally adequate than stage 4 conventional moral reasoning. The theory proposes that most of these preemptive strategies are employed due to a belief that transcendent beings who have access to privileged truths about morality are the ultimate source of what *is moral* and RO persons seek to endorse this ideology (Kay, 1998). The way this theory currently exists was not supported in this study.

Previous support for the Preemptive Theory has been that religious persons refer to their religious ideology to provide at least some of the reasons for why they chose what they did on the DIT-2 story (Kay, 1998; Richards, 1988). However, the mere existence of religious content in a story does not logically support that this is the *reason* for preemptive thinking. In fact, a larger finding of this study is that persons from the same religious organization applied religious doctrines differently or weighted religious doctrines differently to justify their DIT-2 endorsement patterns. For example, 46% of doctrinal comments of persons who endorsed a doctor-assisted overdose focused on the doctrine of agency (i.e., the ability of the person to make personal life and death choices). Those who did not endorse a doctor-assisted overdose mentioned the doctrine of agency in only 8% of their comments. Similar inverse patterns were seen by the relative percentage of persons who focused on God's dominion over life. Both those that did and

did not support a doctor-assisted overdose showed familiarity with these two doctrines (agency and God's dominion), but in their replies it was the *application* of the doctrine that appeared to be contributing to their different responses. In so many words, different religious teachings were emphasized according to the DIT-2 decisions made.

These findings led me to conjecture that if one is going to accept that pre-emptive thinking stifles stage 5-6 moral reasoning then they must also accept that pre-emptive thinking will stifle stage 4 moral reasoning. The theory has never been explained in this multidimensional fashion. For example, RO persons who chose to focus on differing doctrines may also pre-empt their thinking based on the application of religious doctrines that stifle various levels of MMR schemas. In the past, it has only been proposed that the preemptive theory *stifles* higher moral reasoning (stages 5-6). Perhaps it is even the case that non-RO persons preempt stage 4 thinking based on their philosophies of life. In essence, the mere existence of religious content in an item is not evidence to support the existing Preemptive Theory and thus the way this current theory exists, was rejected.

Bias theory. The final major theory for why RO persons score lower on the DIT-2 test is the bias theory. In the previous two theories, the explanation of the scoring trends for RO persons is explained as part of the individual's internal personalogical variables. Contrastingly, the bias theory focuses on the content of the item. The overall crux of Bias Theory is that the DIT-2 scores are actually being confounded by other constructs rather than macromoral judgment that stem from the item content (Kay 1998; Richards, 1988). This theory was weakly supported in this study, but not in the traditional ways that previous studies have supported it.

Traditionally, the confounding construct has been traced to the existence of theological content¹¹ (Richards, 1988). However, this study concludes that the mere existence of religious content in the items is not sufficient evidence to support a theory as strong as Bias Theory. After all, non-RO persons may also claim there is religious content in the stories. Also, RO persons who scored high and low on DIT-2 upper levels of moral reasoning both mentioned that religious doctrines influenced their decisions but not in the ways that Richards proposed. For example, Richards proposed that his RO persons sample scored higher on the maintaining norms stages due to the LDS doctrinal ideology of “honoring, obeying, and sustaining the law”—LDS Article of Faith #12 (Richards, 1988). While this tendency was sometimes mentioned as a theological reason, it was extremely infrequent. In fact, for participants who made a decision on DIT-2 stories, none of them mentioned the law as influencing their decision in more than 8% of the relative comments. For most of the stories, obeying, honoring, and sustaining the law was mentioned in less than 2% of the comments.

Bias Theory has also been supported in the past by showing statistically significant amounts of DIF levels. However, past studies have ignored practical significance as evidenced by the DIF contrast statistic. This study found *statistically significant* DIF for 30 of the 65 items. However, in this study, the DIF contrast guidelines showed only two items were worth noting as large to heavy in practical significance. In addition, the direction of the DIF (+ or -) was not uniform across items

¹¹ Recall from Appendix G that theological content for MMR include doctrinal statements that are clearly scriptural in nature (e.g., Thou shalt not steal, Agency is a Divine gift, Repentance) while philosophical content for MMR includes all other statements that refer to a person’s moral code (e.g., two wrongs don’t make a right, the ends don’t justify the means, anarchy is not an acceptable way to protest).

and thus was not favoring or disfavoring any particular group (i.e., 13 were easier to endorse for BYU students compared to 17 for UA students). Wright and Douglas (1976) claim that when DIF is not consistently in one direction, when most DIF items are below ± 0.50 , and when the test has more than 20 items, the impact of DIF on person measurement is generally small. The lack of practical significance, lack of favoring a particular group, and the fact that the DIT-2 has a large number of items on this test, does not strongly support DIF item bias.

The final way in which bias theory was not supported comes from the .84 person reliability statistic (similar to Cronbach's alpha). In essence, if the persons are responding to similar types of items in a consistent manner, then there is some evidence that the items are not misperforming in a systematic manner. However, reliability is different from validity, so a reliable score does not mean that the trait being consistently measured is indeed macromoral reasoning.

*Discussion Point #3—The Insufficiency of the DIT-2
MN-Stage 4 Schema for RO Persons: A Possible
Masking of the RO and MN-Stage 4 Relationship*

So if the reason that RO persons score systematically higher on MN-stage 4 scores is not due to developmental deficiencies, preemptive thinking, nor item-bias, then the question remains—is the DIT-2 MN-stage 4 framework's explanation valid? Recall that the DIT-2 framework indicated that persons with higher MN-stage 4 scores defer to laws, customs, rules, and formal structures of society to define their MMR. In reviewing the findings of this study, as well as additional literature, I have found three aspects of the relationship between MMR and RO that do not support the DIT-2 MN-stage 4 schema's

framework and may be masking the relationship between RO and DIT-2 MN-stage 4 scores.

The three aspects I mention are, first, this study supports Turiel and Walker's findings that religion and macromorality are overlapping domains, even within the DIT-2 items. Thus, applicable religious thought on some of the DIT-2 stories is appropriate and macromorality is masked by religious orthodoxy because the domains overlap.

Second, RO persons have an internal schema that involves ownership of thought and values, some of which may align with existing laws and social structures. However, the mere alignment of values with existing structures does not necessarily mean that RO persons are deferring to external religious rules to justify MMR. Thus, the RO and DIT-2 score relationship may be masked due to personal reflective values that happen to align with existing social structures.

Third, there is a much more complex overlapping of religion and macromoral judgment on the DIT-2 items than has been suggested. The overlapping appears to be influenced by a preference of macromoral shared responsibilities over macromoral individual rights and a preference of religious doctrines used to justify a person macromoral decisions. Overall these three points do not support the DIT-2 MN-stage 4 framework but show a possible masking of the relationship between RO and MN-stage 4 scores.

Religion and macromoral reasoning are overlapping domains. As discussed in chapter 2, Turiel (1978) suggested that religion and macromoral judgment domains overlap while Kohlberg and Rest claim they are separate domains on the DIT-2 (see

Figure 8). Walker's findings (Walker & Pitts, 1998; Walker et al., 1995) supported Turiel's claims. This study supports Turiel and Walker's claims for DIT-2 MN stage-4 levels of MMR and thus does not support the DIT-2 assumption that religion and MMR are separate domains. Thus, if the domains overlap on DIT-2 items, then RO persons may not be deferring to a religious domain for their MMR reason.

The most overt evidence of overlapping domains in this study was the existence of correlations between MN-stage 4 scores and RO scores. These correlations remained intact when partial correlation analyses' removed potential quantified confounds. Thus, there is an association between RO and MMR scores. However, qualitative data showed that the relationship between RO and MMR scores may be masked by other unquantified variables discovered in this study.

Second, the support of overlapping domains pertains to the application of the LDS doctrine of agency on the DIT-2 stories, a prominent concept mentioned in the RO student's qualitative responses. In the LDS literature, the power to choose is one of four concepts to explain the doctrine of agency and choice is seen as method to act morally (Church of Jesus Christ of Latter-day Saints, 2001, p. 4; Oaks, 1992, 2001; Packer, 1995,

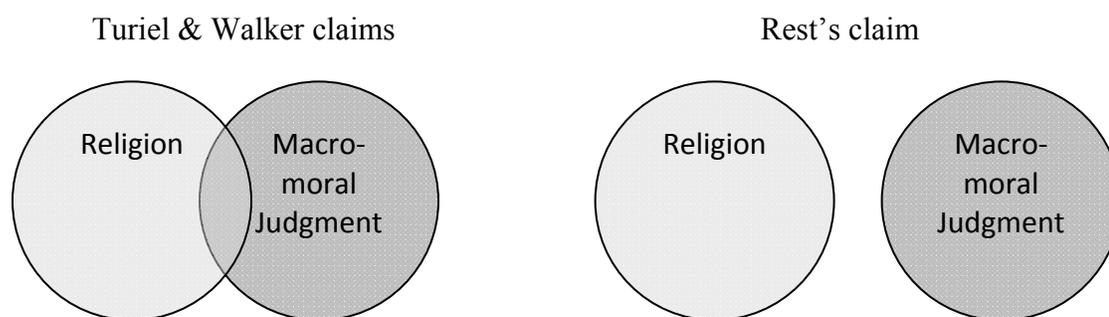


Figure 8. This study supported Turiel and Walker's claims of overlapping domains within the DIT-2 stories and items.

2000). Agency pertains to the content of all the DIT-2 stories because they all ask one to *choose* the most moral course of action. Because some macromoral choices support or hinder one's spiritual aspirations, the alignment or misalignment to a religiously orthodox person's own values leads to choices being viewed as moral or immoral. However, these aspirations come from ownership and reflective values, not deferrals to religious ideology or laws (shown by the lack of comments that simply defer to law as reasons for DIT-2 decisions). Moral choices are usually based upon what a person values (religious or nonreligious). Alignment with existing social structures or even with religious laws is not sufficient evidence to conclude a deferring to existing social structures.

Further, interviews and qualitative data reflected the concept that choice was used as a method that supported one's personal macromoral values. Consequently, the very nature of the DIT-2 (i.e., make a moral choice or, in RO terms, use one's agency) overlaps with the domain of religion because choice is a method one uses to be moral *and* a method one uses to be religious (Oaks, 1992, 2001).

A third point that supports overlapping domains pertains to the DIT-2 items themselves. Recall that Rest (1999) claimed that the items measure MMR, not transcendent truths from divine being. However, the content of some of the items violate his claim. As a specific example of a DIT-2 item that has religious content in it, one of the items asks if only God should decide if a person should live or die? This item clearly is a crossing of domains. In other words, if Rest claims that religion and macromorality are separate domains, then he ought not to cross those domains in his items by attempting to predict religious orthodoxy by mentioning a Transcendent Being (i.e., God).

A fourth point stems from the finding of general patterns—those who advocated euthanizing a suffering cancer patient also tended to give low endorsements to MN-stage 4 items. Conversely, those who did not advocate euthanizing a suffering cancer patient tended to give high endorsements to MN-stage 4 items. The DIT-2 story advocacy and item endorsement patterns were consistent across all stories (e.g., those who advocated stealing in a famine, reporting a story, etc., had lower stage 4 item scores than those who did not advocate these actions). These patterns suggest that the actual content of the story and initial story decision, not the DIT-2 items (which the test purports to derive its scores from), may overlap with RO person’s MMR schema.

Ownership of thought versus deferral of thought. Much of Kohlberg and Rest’s argument that religion and macromorality are separate domains relies on the assumption that religious persons defer their MMR to the domain of religion rather than use their own reflective thought. However, the existence of religious values that are used to make a moral choice is not necessarily deferring to an external custom, but rather is most likely an internal alignment with religious goals and values that one espouses. John Richard Neuhaus, a prominent Catholic priest and opinion leader for many RO persons, articulated this concept quite well:

In a democracy that is free and robust, an opinion is no more disqualified for being “religious” than for being atheistic, or psychoanalytic, or Marxist, or just plain dumb. There is no legal or constitutional question about the admission of religion to the public square; there is only a question about the free and equal participation of citizens in our public business. Religion is not a reified “thing” that threatens to intrude upon our common life. Religion in public is but the public opinion of those citizens who are religious. (Neuhaus, 1992, p. 13; see also Neuhaus, 1997)

The decisions made on the DIT-2 by the majority of BYU students suggested

similar thinking patterns (i.e., I can apply my personal religious views as I decide why a macromoral decision is right/wrong and that they are just as valid as any other reason because they are my own reflective views). The qualitative data as well as informal discussions with the RO sample (after data was collected), both showed that respondents clearly take ownership for their macromoral views and even find it offensive when persons claim they are deferring, stifling, preempting, or nonreflective in their thinking. They do not point to a steeple or the sky for their macromoral choices but would point inward for their macromoral reasoning reflecting ownership of thought. Furthermore, many RO persons consider it not only their right, but also their responsibility as citizens, to use their values to influence macromoral political issues (Nelson, 2007; Public Issues, 2008).

Perhaps the reason that many in this sample take such deep ownership in their macromoral thinking because they have been encouraged by their church leaders to *not* follow them blindly (Lee, 1964, as cited in Lee, 1974), but to undergo serious reflective inquiry about any teaching that claims to have come from God. Most of this argument (or reasoning) can be summarized by referring to the Church's *General Handbook of*

Instructions, a guide for church leaders:

An act is moral only if it expresses the character and disposition of the person, that is, if it arises out of knowledge, faith, love, or religious intent. Fear and force have no place in the kingdom because they do not produce moral actions, and are contrary to God's gift of free agency. (The Church of Jesus Christ of Latter-day Saints, 1963, p. ii)

This "figure-it-out-for-yourself" and "use-your-agency" philosophy leads to ownership of thought on macromoral issues for RO persons (and I would add secularly

orthodox persons too). This ownership of thought concept could explain why a subgroup of extremely RO persons (LDS returned missionaries) score high on self-directing scores in Pargament's RPSS scales, yet score high on RO and MN-stage 4 DIT-2 scores.

In essence, a religiously orthodox person *reasons with* their personal values—some of which happen to be religious—rather than *defers to* their religious values on macromoral decisions. Similarly, many secular persons *reason with* their personal values on a macromoral decision. Thus, if personal values about societal cooperation are a part of a person's religion, a pretty safe assumption, then religion may be seen as overlapping with macromoral choices—it may be merely the existence of a personal value or preference (which happens to be religious) that is used to facilitate a macromoral choice on the DIT-2. This discussion of values leads to my third and final point that supports overlapping domains of macromorality and religion and masking of the relationship between RO and MN-stage 4 scores.

Associations of macromoral individual rights, macromoral shared responsibilities, and doctrinal justifications mask the DIT-2 MN-stage 4 scores. A third way that does not support the MN-stage 4 framework, and also add evidence of masking variables, comes from an overlapping of several variables noticed by patterns in the qualitative data. Specifically, an additional argument of overlapping domains was supported was by a finding that the DIT-2 story decisions were related to differing applications of LDS doctrines used to justify why one's decision was the most moral choice (i.e., if you decided it was moral to euthanize a suffering patient, you were more likely to mention agency rather than God's dominion over life). Thus, it appeared that

students who endorsed a doctor-assisted overdose valued the doctrine of agency, in this situation, more than the doctrine of God’s dominion over life. However, this assumption of placing differing values on doctrines was not fully verified in this study—that is, this study only asked which doctrine influenced their decision, not which doctrine they valued more than another. However, as was shown in Figure 4, 46% of those who endorsed a doctor assisted overdose focused on the right to avoid suffering via the doctrine of mercy. Contrastingly, only 8% of those who did not endorse a doctor-assisted overdose mentioned this concept as influencing their macromoral decision. The differing application of LDS doctrines used to support DIT-2 story decisions was evidenced in Chapter IV, Figures 4-7, by the relative percentage of type of comments based on the DIT-2 story decision.

Further, when all the qualitative comments were sorted by DIT-2 choice, a larger theme of macromoral individual rights versus macromoral shared responsibilities were associated with which doctrinal teachings BYU students used to justify their DIT-2 decisions. For example, in the famine story, stealing from a rich man was justified by an individual’s right to avoid starving. In contrast, those who said Mustaq should not steal from the rich man mentioned a reciprocal principle of owning and allowing others to own property much more than those who chose to steal the food (e.g., “The person who is rich still has a right to his property regardless of what his intentions are”). Another concrete example exemplifies that many students expressed that the suffering cancer patient’s individual rights existed separate from other stakeholders affected by her decision: “She should have the right to decide how she wants to live her life—she has free agency and

should be able to exercise that right according to the dictates of her own conscious.”

Notice the focus on the individual rights, independence, and individualism rather than on her collectivistic relationships with loved ones—relationships which could make her macromorally responsible to others.

These juxtaposing moral issues of macromoral individual rights versus macromoral shared responsibilities pervaded all the DIT-2 qualitative data. These juxtaposing moral issues were also associated with which LDS doctrine was used to justify the macromorality of the choice. However, the qualitative data were inconclusive on whether one preference caused the other (i.e., whether the preference of a particular LDS doctrine *caused* the preference of individual rights over shared responsibilities or vice versa). In addition, the data was inconclusive on whether the actual doctrines were weighted as being more valued than another is, whether they were viewed in isolation, or whether they were even fully understood. Still, the collective data were conclusive that there was an association between LDS doctrines used to justify the DIT-2 choice and a preference of individual rights versus shared responsibilities.

This larger factor of choosing to focus on macromoral shared responsibilities over macromoral individual rights also separated the most and the least RO in this study (i.e., the more RO, the more likely to focus on shared responsibilities). This finding, coupled with the evidence that LDS doctrines were used to justify differing macromoral decisions, supports Turiel and Walker’s claims of an overlapping domains of religion and macromorality but also adds to the complexity of how religiously orthodox persons macromorally reason (see Figure 9).

Delimitations

The first delimitation of this study may have been due to differing verbal abilities of students. Some students may not have been able to clearly articulate in writing their reasons for the choices they made on the DIT-2 stories. This lack of expressive ability of some students has been seen as one of the strengths of the DIT-2 when it is viewed as aggregate data. The DIT-2 claims it reduces the error of this articulation confound by creating a selected response items rather than relying on students differing abilities to

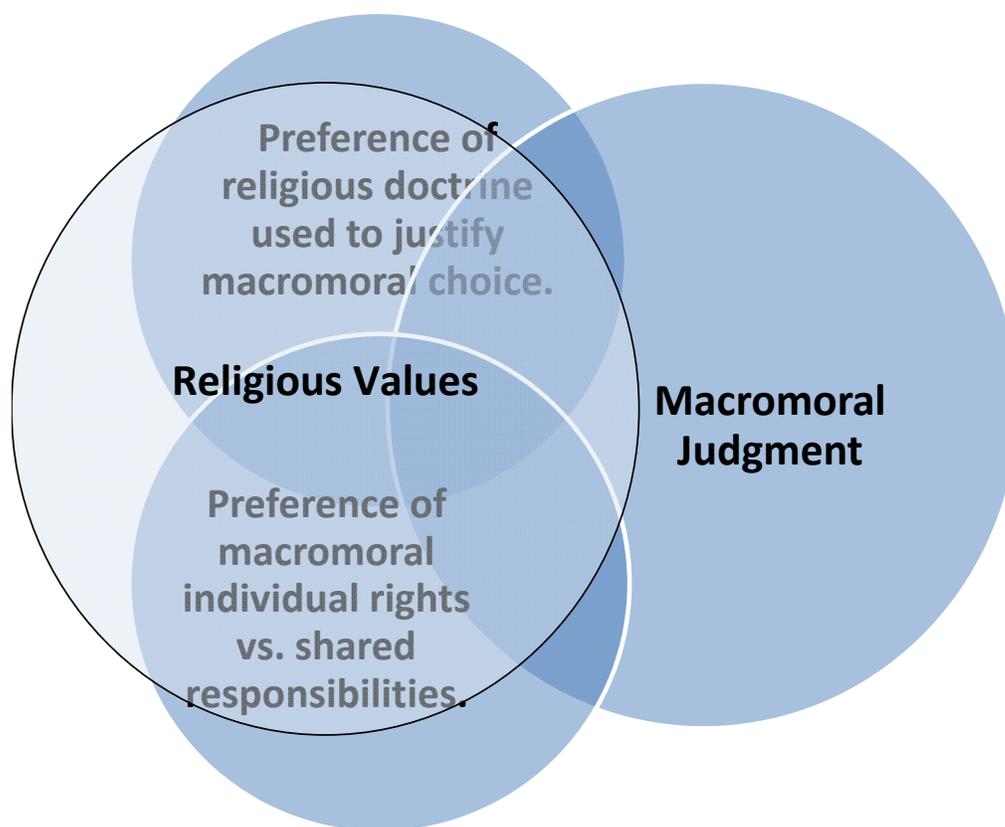


Figure 9. Religion and macromoral judgment domains may overlap in complex ways on the DIT-2 items. The qualitative data supported Turiel and Walker's findings of an overlapping of religion and macromoral judgment domains but in a more complex manner than a simple overlapping with religion.

articulate their moral schemas as in Kohlberg's MJI (Rest, 1999). While this may be true when the DIT-2 scores are viewed as aggregate data, it may not be true for individual data. However, the qualitative data was based on the assumption that BYU students were able to accurately articulate in writing the reasons for their MMR choices.

A second delimitation of this study was that the reference group was made up of persons from University of Alabama, a group whose mean scores of RO are comparable to BYU in religiosity. However, as noted earlier, the students from Alabama vary much more in measures of RO (BYU standard deviation on Brown and Lowe's Religious Inventory of Belief was 3.97, UA was 10.63). Although an estimated national norm was used to filter those who adhere to a national level of religious orthodoxy, a more religiously neutral reference group (other than the University of Alabama students) could have strengthened this study's findings.

A final delimitation pertains to the qualitative data. These data were based upon written comments, which were assumed to be honest and accurate. It was assumed that all relevant data expressed what student's perceived as relevant issues that influenced their MMR was freely available to the person's consciousness. However, many reasons for a person's morality may be in underlying sub consciousness that may have been untapped due to some psychological block that did not allow the students in this sample to explain their macromoral reasoning. Although the presence of a subconsciousness confounding variable sounds beyond the scope of measurement, we may not always tap into every readily available resource that makes up our macromorality. Nevertheless, I assumed that the sub conscious was not confounding scores or comments. In addition, because much

of the data were aggregate data, I assumed these factors varied at random and thus did not favor any particular group.

Suggestions for Further Research

There are four suggestions for further research that stemmed from the major and ancillary findings of this study. The order of these suggestions for further research is hierarchal, each one building upon each other and ruling out additional potential religious confounds to macromoral reasoning (see Figure 10).

It is anticipated that each of these studies will add insight to build a more appropriate framework for macromoral reasoning that captures the complex reasoning of religiously orthodox as well as secularly orthodox persons. In these suggestions for

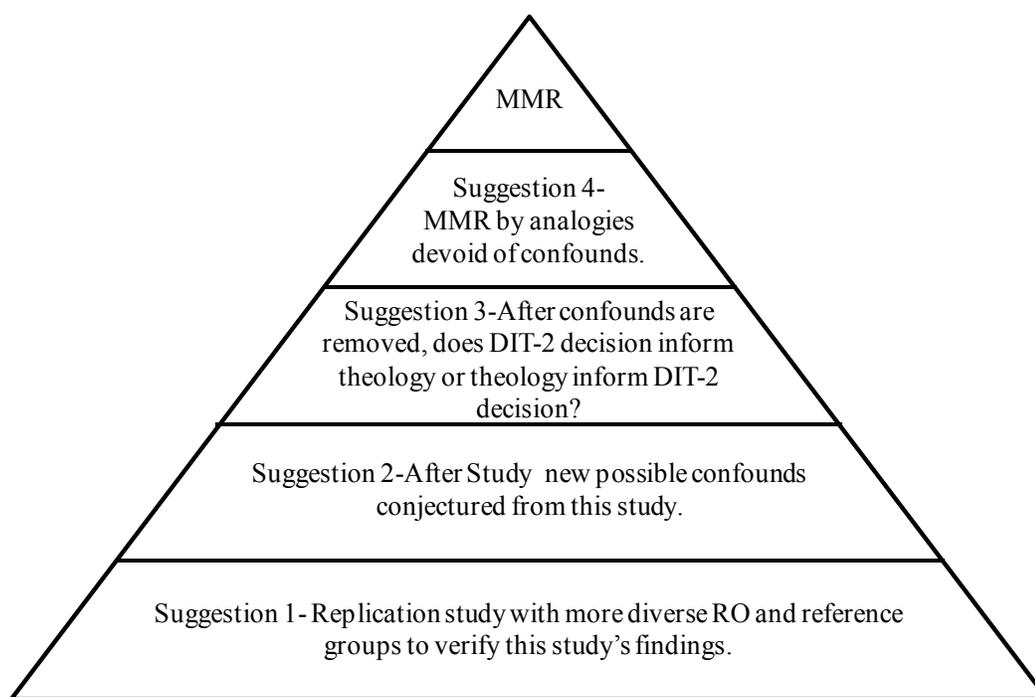


Figure 10. Hierarchal order of further research. The four suggestions for further research should build on each other's findings to build a better framework for macromoral reasoning (MMR).

further research references are provided as helpful starting points for a review of literature. However, none of the suggestions for further research have been conducted.

The first suggestion for further research involves a replication study. Recall that this study ruled out several potential confounds that could affect MMR of a religiously orthodox person. A replication study could discover how generalizable the findings of this study are with other religiously orthodox groups. In addition, this study could seek to replicate these analyses' with a more diverse reference group than those from UA as well as a different religiously orthodox group than the LDS students at BYU. Thus, the potential confounds that this study concluded were not affecting MMR could be verifiably said to not affect other religiously orthodox groups as well.

The second suggestion pertains to other potential confounds that were ancillary findings of this study. For example, in reviewing the qualitative data, several themes that did not pertain to my research questions emerged. There were findings in this study and anecdotal evidence in religiously orthodox literature that the following five factors may influence DIT-2 story choices: (a) an understanding, interpretation, and application of the doctrine of agency to macromoral issues (Lee, 1950; Oaks, 1992, 2001; Packer, 1971, 1995, 2000; Petersen 1972; Romney, 1942), (b) a schism between modernism and postmodern ontological views of truth (Greer, 2003; Hare & McLaughlin, 1998; Madsen, 2000; Mertens, 2005), (c) a preference of macromoral individual rights versus shared responsibilities (Christofferson, 2005; Hindeman, 1997; Kimball, Tanner, & Romney, 1980; Tabor & Adask, 2008), (d) RO persons concepts of a just society (Lloyd, 2008; Tabor & Adask), and (e) God versus man's locus of control on DIT-2 MMR issues

(Gabbard, Howard, & Tageson, 1986; Gorsuch, 1983; Hill & Hood, 1999; Ritzema & Young, 1983). These five aspects surfaced in reviewing responses among the most and least RO persons in this study. Confirmatory studies could seek to validate the five elements as reasons for a systematic test bias. Specifically, studies could correlate the doctrinal influence of one's understanding of agency, one's ontological views of modernism and postmodernism, one's views on macromoral rights being inherently attached to macromoral responsibilities, one's concepts of a just society, and a God versus man's locus of control, all with DIT-2 endorsement patterns. Multiple regression studies could view the interrelationship of these proposed variables to further confirm the existence of these potential confounds to MMR for a religiously orthodox person.

After these potential external confounds were verified or ruled out as confounds to religiously orthodox person's MMR, they could be controlled for so a further causal study could be performed. This second study could lead to a third suggestion for further study which also stems from the qualitative and quantitative data, which showed two distinct groups within the BYU sample, a majority and minority group. By studying and contrasting majority and minority patterns of RO person's responses on the DIT-2 along with qualitative responses that corresponded with these groups, some key findings emerged. For example, the relative amount of which doctrinal teaching was emphasized between these two groups from the same religion on the same moral dilemma was interesting (see Chapter IV, Figures 4-7). These key findings raised the question, did theology inform the DIT-2 macromoral decision or did the DIT-2 decision inform theology? Further, did a person's preference of macromoral individual rights over shared

responsibilities cause the focus on the divine gift of agency or did their interpretation of the divine gift of agency cause them to focus on a rights-based macromoral schema?

This study was not designed to fully answer these questions, although preliminary findings showed an association between these variables.

Once internal relationships between DIT-2 macromoral decisions and theology are understood, a fourth study could seek to control for all theological and religiously orthodox content from the DIT-2 items. This suggestion was conjectured due to the findings from the fourth DIT-2 story—the school board story. Perhaps the variance in these responses is due to the lack of religious content in the story (i.e., God’s word does not speak much on the morality of holding unruly public school board meetings). RO persons responded in a more diverse way with this story than they did with the others as shown in the balanced proportion of respondents to each extreme (102 comments “should call off the next open meeting” and 114 comments “should not call of the next open meeting”). If relationships between how RO and DIT-2 scores were fully understood, an analysis of covariance (ANCOVA) may be a method to control for the RO variable.

The finding of a more spread out response pattern along with the lack of religious commentary suggests that the school board story performs quite well for the most and least religiously orthodox persons in this sample. It led me to question, if the same issues in the other DIT-2 stories were presented to religiously orthodox persons in an unfamiliar context, devoid of any external and internal confounds to a religiously orthodox person’s MMR (studies 1-3), would the religiously orthodox and secularly orthodox respond in similar ways? Besides seeking to control via ANCOVA, another way the question of

whether religious content influences DIT-2 scores could be researched is to present similar moral dilemmas via analogies—“the repetition of the same fundamental pattern in two different contexts” (Emmet, 1945, p. 6). These analogies could be pilot tested with RO persons to see if they have religious content in them. This analogical method is common in moral discourse (Post & Leisey, 1995).

A study of the analogical moral reasoning of RO persons could also focus on an alterability of macromoral rules. Alterability refers to whether or not there are rules that are fixed or if those rules could or should be altered (Rest et al., 1999b). While very few RO persons will contradict the scriptures and church teachings, some may show more alterability when the stories are devoid of religious content via analogy (see Glausiusz, 2007; Hauser, 2007; Thomson, 1986 for a good starting point for such a study).

These four suggestions for further research would build upon this study’s findings. In addition, they would add to building a more robust framework for measuring religiously orthodox person’s macromoral reasoning and thus strengthen the field.

Conclusion

The purpose of this study was to determine whether there were items in the DIT-2 that perform differently for a group of religiously orthodox college students than for other American college students. Using new differential item functioning methods, correlational statistical analysis, the DIT-2 test, and examining unexplored demographic variables, this study shed light on the appropriateness of using the DIT-2 as a measure of moral judgment for religiously orthodox persons. This study is important to the field of

instructional technology because many instructional models for character education and moral education use the DIT-2 to measure the differences between treatment and control groups based upon an instructional treatment (Rest et al., 1999b). In addition, studying the construct validity of the test is important because of the way it and its derivative tests have been and are being proposed to be used (see Chapter II).

A benefit of this study is that it illuminated how a new generation of students at BYU responds to moral judgment measures (some of these students could even be the children of the previous study performed by Richards 20 years ago). This study differed from previous studies by involving a demographic and qualitative questionnaire along with the DIT-2 and quantitative methods of analysis. The rich amount of data allowed me to examine patterns of responses within the BYU sample that illustrated *how* the DIT-2 performs differently for different subgroups of students—something the previous DIT studies have not explored. In addition, the qualitative data helped to explain what life experiences, doctrinal teachings, philosophies, or other relevant issues contributed to the decisions made on the DIT-2 scenarios and items. Examining these data yielded information that will aid moral judgment researchers' understanding of elements that contribute to a religiously orthodox person's macromoral reasoning schema and are a step in understanding the macromoral reasoning of religiously orthodox persons.

Findings from this study indicate that items in this and previous studies exhibit differential item functioning in statistically significant but not practically significant ways when a religiously orthodox sample is compared with a norm-reference sample. However, quantitative data and open-ended survey responses showed general patterns

that religiously orthodox persons have a much more complex schema of macromoral reasoning than the DIT-2 schemas assert. For example, Rest and colleague's (1999b) assertion that religiously orthodox persons score higher on MN-stage 4 scores due to deferring to religious ideology does not explain why returned missionaries score very high on the self-directing items of Pargament's Religious Problem Solving Style and lower on the deferring items. In contrast, these scores suggest a belief in personal agency to direct one's life and suggest a high internal locus of control. The qualitative comments in this study also suggest that something other than deferring to religious laws appeared to be creating the MN-stage 4 differences. As explained earlier, an alternative to Rest and colleague's explanation of MN-stage 4 scores is that the MN-stage 4 items actually measure a philosophical difference based on a weighting of macromoral responsibilities and macromoral rights and RO persons gravitate towards valuing macromoral responsibilities as more important to their macromorality.

In addition, suggestions for further study invite the further validation of these findings to determine the magnitude that several facets of religiously orthodox ideology and philosophy contribute to DIT-2 scores. A further invitation that this study begs for an explanation is to study the proposed reasons of why religiously orthodox persons score higher on the DIT-2 maintaining norms (stage 4) scores. One of the challenges that the Neo-Kohlbergians must deal with are this study's finding of a far too simplistic defining of religious reasoning as deferring to societal norms. How the macromoral reasoning of a religiously orthodox person fits into their framework based on the findings of this study suggest that religiously orthodox person's MMR schema is far more complex than Rest

suggests.

An important concluding point of this study is to indicate that the findings do not debunk the DIT-2 test for religiously orthodox persons but embark on the beginning of a conversation. This conversation begins with a defining philosophical question of what the relationship is between a religious person's view of MMR and the method of how and why they use religious doctrines to justify their MMR on macromoral issues within the DIT-2 test. The further this conversation continues the more it will lend itself to the measurement of this relationship. The conversation naturally will expand beyond the circle of RO persons to study the generalizability of the measured relationship and whether it exists for all persons. The discussion should continue to the point where one can conclude whether the items of the DIT-2 tap into the measurement of this conjectured relationship between the degree of one's RO views and their MMR. Finally, the exchange of ideas in the field of macromorality that results from this dialogue may develop a stronger framework for MMR and hence a stronger test of the construct.

REFERENCES

- Abdolmohammadi, M., Read, W., & Scarbrough, D. (2003). Does selection-socialization help to explain accountants' weak ethical reasoning? *Journal of Business Ethics*, 42(1), 71.
- Ammerman, N. (1982). Operationalization of evangelicalism: An amendment. *Sociological Analysis*, 42, 363-372.
- Atwater, L. (1998). Individual attributes and leadership style: Predicting the use of punishment and its effects. *Journal of Organizational Behavior*, 19, 559-576.
- Bailey, C., Phillips, T., & Scofield, S. (2005). Does political bias in the DIT or DIT-2 threaten validity in studies of CPAs? *Behavioral Research in Accounting*, 17, 23-42.
- Barnes, J. (Ed.). (1984). *The complete works of Aristotle: The revised Oxford translation*. Princeton, NJ: Princeton University Press.
- Beatty, K., & Walter, B. (1988). Fundamentalists, evangelicals, and politics. *American Politics Quarterly*, 16, 43-59.
- Bebeau, M. (2002). The defining issues test and the four component model: Contributions to professional education. *Journal of Moral Education*, 31, 3.
- Bebeau, M., Rodriguez, M., & Maeda, Y. (2002, March). *Using DIT moral schema profiles to evaluate educational intervention effects*. Paper presented at the annual meeting of the Association for Practical and Professional Ethics, Cincinnati, OH.
- Bebeau, M., & Thoma, S. (2003). *Guide for DIT-2: A guide for using the defining issues test, version 2, and the scoring service of the Center for the Study of Ethical Development*. Minneapolis: University of Minnesota.
- Benor, D., Notzer, N., Sheehan, T., & Norman, G. (1982, March). *Moral reasoning as a criterion for admission to medical school*. Paper presented at the American Education Research Association Annual Conference, New York, NY.
- Bergin, A., Masters, K., & Richards, P. (1987). Religiousness and mental health reconsidered: A study of an intrinsically religious sample. *Journal of Counseling Psychology*, 34(2), 197-204.

- Bergin, A., Stinchfield, R., Gaskin, T., Masters, K., & Sullivan, C. (1988). Religious lifestyles and mental health: An exploratory study. *Journal of Counseling Psychology, 35*(1), 91-98.
- Blackner, G. (1975). *Moral development of young adults involved in weekday religious education and self-concept relationships*. Unpublished master's thesis, Brigham Young University, Provo, UT.
- Bond, T., & Fox, C. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Erlbaum.
- Boneau, C. (1960). The effects of violations of assumptions underlying the t-test. *Psychological Bulletin, 57*, 49-64.
- Brown, D., & Annis, L. (1978). Moral development and religious behavior. *Psychological Reports, 43*, 1230.
- Brown, D., & Lowe, W. (1951). Religious beliefs and personality characteristics of college students. *Journal of Social Psychology, 33*, 103-129.
- Center for the Study of Ethical Development. (2007). *Center for the Study of Ethical Development*. Retrieved June 21, 2007, from <http://www.centerforthestudyofethicaldevelopment.net/DIT2.htm>
- Christofferson, D. T. (2005). Allegiance to God. *Ensign, 1*, 8-13.
- Church of Jesus Christ of Latter-day Saints, The. (1963). *General handbook of instruction*. Salt Lake City, UT: Author.
- Church of Jesus Christ of Latter-day Saints, The. (2001). *Book of Mormon video guide*. (2001). Salt Lake City, UT: Author.
- Clauser, B., & Mazor, K. (1998). Using statistical procedures to identify differentially functioning test items. *Educational Measurement: Issues and Practices, 17*(1), 31-44.
- Clayton, E. (2006). *Aristotle: Politics*. Retrieved May 1, 2008, from <http://www.iep.utm.ed/a/aris-pol.htm>
- Clouse, B. (1979). *Moral judgment of teacher education students as related to sex, politics, and religion*. Unpublished manuscript, Indiana State University, Terre Huate.

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cornwall, M., Albrecht, S., Cunningham, P., & Pitcher, B. (1986). The dimensions of religiosity: A conceptual model with an empirical test. *Review of Religious Research*, 27(3), 226-244.
- De Casterle, B. (1998). An adjusted version of Kohlberg's moral theory: A discussion of its validity for research in nursing ethics. *Journal of Advanced Nursing*, 27(4), 829-835.
- Deemer, D. (1986). *Moral judgment and life experiences*. Unpublished doctoral dissertation, University of Minnesota, Duluth.
- Deemer, D. (1989). Moral judgment and life experience. *Moral Education Forum*, 14(2), 11-21.
- Denvir, J. (1999). *The law of rules*. Retrieved April 30, 2008, from <http://www.usfca.edu/pj/ruleoflaws.htm>
- Dobson, E. (1987). The Bible, politics, and democracy. In J. R. Neuhaus (Ed.), *The Bible, politics, and democracy* (pp. 1-18). Grand Rapids, MI: Eerdmans.
- Eagle, E. (1989, March). *Socioeconomic status, family structure, and parental involvement: The correlates of achievement*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Elm, D., & Weber, J. (1994). Measuring moral judgment: The moral judgment interview or the Defining Issues Test? *Journal of Business Ethics*, 13(5), 341-355.
- Embretson, S., & Reise, S. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.
- Emler, N., Palmer-Cantion, E., & St. James, A. (1995). Politics, moral reasoning and the Defining Issues Test: A reply to Barnett et al. *British Journal of Social Psychology* 37,457-476.
- Emler, N., Resnick, S., & Malone, B. (1983). The relationship between moral reasoning and political orientation. *Journal of Personality and Social Psychology*, 45, 1073-1080.
- Emmet, D. (1945). *The nature of metaphysical thinking*. New York: St. Martin's.

- Emsberger, D., & Manaster, G. (1981). Moral development, intrinsic/extrinsic religious orientation and denominational teachings. *Genetic Psychology Monographs*, 104, 23-41.
- Gabbard, C., Howard, G., & Tageson, C. (1986). Assessing locus of control with religious populations. *Journal of Research in Personality*, 20, 292-308.
- Gall, M. (1996). *Educational research: An introduction* (6th ed.). White Plains, NY: Longman.
- Gall, M. (2001, April). *Figuring out the importance of research results: statistical significance versus practical significance*. Paper presented at the American Educational Research Association, Seattle, WA.
- Getz, I. (1984). Moral judgment and religion: A review of the literature. *Counseling and Values*, 28, 94-116.
- Getz, I. (1985). *Moral reasoning, religion, and attitudes towards human rights*. Unpublished doctoral dissertation, University of Minnesota, Duluth.
- Glausiusz, J. (2007, May 10). Discover interview with Marc Hauser: Is morality innate and universal? *Discover Magazine*, 63-66.
- Gorsuch, R. (1983). Attributions of responsibility to God: An interaction of religious beliefs and outcomes. *Journal for the Scientific Study of Religion*, 22, 340-352.
- Greer, R. (2003). *Mapping postmodernism: A survey of Christian options*. Madison, WI: Intervarsity.
- Hall, S. (2005, April 29). *BYU NewsNet: Study focuses on BYU marriage perceptions*. Retrieved August 2, 2008, from <http://nn.byu.edu/story.cfm/55351>
- Haggbloom, S., Warnick, R., Warnick, J., Jones, V., Yarbrough, G., Russell, T., et al. (2002). The 100 most eminent psychologists of the 20th century. *Review of General Psychology*, 6(2), 139-145.
- Hare, W., & McLaughlin, T. (1998). Four anxieties about open-mindedness: Reassuring Peter Gardner. *Journal of Philosophy of Education*, 32 (2), 283-292.
- Harris, A. (1981). The development of moral reasoning: Some observations about Christian fundamentalism. *Journal of Psychology and Theology*, 19, 249-256.

- Hauser, M. (2007). *Discovery interview: Is morality innate and universal?* Retrieved August 25, 2009, from <http://discovermagazine.com/2007/may/the-discover-interview-marc-hauser>
- Hill, P., & Hood, R. (1999). *Measures of religiosity*. Birmingham, AL: Religious Education.
- Hinckley, G. B. (1998). Words of the living prophet. *Liahona*, 8, 16.
- Hindeman, E. (1997). *Rights vs. responsibilities: The Supreme Court and the media*. Westport, CT: Greenwood.
- Hunter, J. (1981). Operationalizing evangelicalism: A review, critique, and proposal. *Sociological Analysis*, 42(4), 363-372.
- Kachigan, S. (1986). *Statistical analysis: An interdisciplinary introduction to univariate and multivariate methods*. New York: Radius
- Kay, S. (1998). *Preemptive response style and test bias in measuring principled moral reasoning of conservatively religious groups*. Unpublished master's thesis, Biola University, Lake Mirada, CA.
- Kellstedt, L. (1989). The meaning and measurement of evangelicalism: Problems and prospects. In T. Jelen (Ed.), *Religion and American political behavior* (pp. 3-22). New York: Praeger.
- Kim, S-H., Cohen, A., Alagoz, C., & Kim, S. (2007). DIF detection and effect size measures for polytomously scored items. *Journal of Educational Measurement*, 44(2), 93-116.
- Kimball, S. (1979). Q&A: Questions and answers. *Ensign*, 5, 38-41.
- Kimball, S., Tanner, N. E., & Romney, M. G. (1980). The church and the proposed equal rights amendment: A moral issue. *Ensign*, 3, 2-3, 5-17, 19-23.
- Knuckey, J. (1999). Religious conservatives, the Republican Party and evolving party coalitions in the United States. *Party Politics*, 5(4), 485-496.
- Kohlberg, L. (1958). *The development of modes of moral thinking and choice in the years 10 to 16*. Unpublished doctoral dissertation, University of Chicago.

- Kohlberg, L. (1967). Moral and religious education and the public schools: A developmental view. In T. Sizer (Ed.), *Religion and public education* (pp. 164-183). Boston: Houghton-Mifflin.
- Kohlberg, L. (1969). Stage and sequence: The cognitive developmental approach to socialization. In D. Goslin (Ed.), *Handbook of socialization: Theory and research* (pp. 347-480). Chicago: Rand McNally.
- Kohlberg, L. (1971). From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In L. Mischel (Ed.), *Cognitive development and epistemology* (pp. 151-284). New York: Academic.
- Kohlberg, L. (1973). The claim to moral adequacy of a highest stage of moral judgment. *Journal of Philosophy*, 40, 630-646.
- Kohlberg, L. (1976). Moral stages and moralization: The cognitive-developmental approach. In T. Likona (Ed.), *Moral development and behavior: Theory, research, and social issues* (pp. 31-53). New York: Holt, Rinehart, and Winston.
- Kohlberg, L. (1981). *The philosophy of moral development*. San Francisco: Harper and Row.
- Kohlberg, L. (Ed.). (1984). *Essays on moral development: Vol. 2, the psychology of moral development*. San Francisco: Harper and Row.
- Latif, D., & Dunn, B. (2000). Moral reasoning: should it serve as a criterion for student and resident selection in pharmacy? *American Journal of Pharmaceutical Education*, 65, 119-124.
- Lawrence, J. (1979). *The component procedures of moral judgment making*. Unpublished master's thesis, University of Minnesota, Minneapolis.
- Lee, H. (1950). Conference report. In 121st semiannual conference of *The Church of Jesus Christ of Latter-day Saints*, September 29-October 1st (pp.129-132). Salt Lake City, UT: Deseret Book Company.
- Lee, H. (1974). The prophet, seer, and revelator. In H. Lee (Ed.), *Stand ye in holy places* (pp.162-163). Salt Lake City, UT: Deseret Book Company.
- Leege, D. (1989). Toward a mental measure of religiosity in research on religion and politics. In T. Jelen (Ed.), *Religion and political behavior in the United States* (pp. 45-65). New York: Praeger.

- Lloyd, D. (2008). *Social issues*. Retrieved May 15, 2008, from www.vision.org/visionmedia/printerfriendly.aspx?id=4750
- Madsen, T. (2000). The joy of the Lord is your strength. *BYU Devotional Address, 21 November 2001* (pp. 1-10). Provo, UT: Brigham Young University.
- McGeorge, C. (1976). Some correlates of principled moral thinking in young adults. *Journal of Moral Education, 5*, 265, 273.
- Mertens, D. (2005). *Research and evaluation in education and psychology* (2nd ed.). Thousand Oaks, CA: Sage.
- Moore, D. (1995). The religious right: Definition and measurement. *The Public Perspective, 6*, 11-13.
- Murray, A., & Mishra, S. (1983). Item bias in individually administered ability tests: implications for nondiscriminatory assessment. *Journal of Psychoeducational Assessment, 1*, 353-366.
- Navarez, D., Getz, I., Thoma, S., & Rest, J. (1999). Orthodoxy, progressivism and development of moral judgment. *Developmental Psychology, 35*, 478-488.
- Nelson, D. (2004). Bible knowledge and moral judgment: Knowing scripture and using ethical reasoning. *Journal of Research on Christian Education, 13*(1), 41-57.
- Nelson, R. (2007). Faith and families. *The Ensign, 3*, 36-41.
- Neuhaus, J. R. (1992). *A new order of religious freedom*. Retrieved December 2, 2008, from http://www.firstthings.com/article.php3?id_article=5238
- Neuhaus, J. R. (1997). *The naked public square* (2nd ed.). New York: Eerdmans. (Original work published 1984)
- Niell, J. (2007). *Introduction to individual differences*. Retrieved March 6, 2008, from <http://www.wilderdom.com/personality/L1-1Introduction.html>
- Nokes, K. (1989). Rethinking moral reasoning. *Journal of Nursing Scholarship, 21*(3), 172-175.
- Oaks, D. (1992, February). *Religious values and public policy*. Address to BYU Management Society presented at the BYU Management Society, Washington, DC.

- Oaks, D. (2001). The weightier matters. *The Ensign*, 1, 12-17.
- O'Fallon, M., & Butterfield, K. (2005). A review of the empirical ethical decision-making literature: 1996-2003. *Journal of Business Ethics*, 59(4), 375-413.
- Olsen, O. (2006). Moral behavior and transformational leadership in Norwegian naval cadets. *Military Psychology*, 18(suppl.), S37-S56.
- Osterlind, S. (2006). *Modern measurement: Theory, principles, and applications of mental appraisal*. Upper Saddle River, NJ: Prentice Hall.
- Packer, B. K. (1971). Obedience. *BYU speeches of the year* (pp. 1-7). Provo, UT: Brigham Young University Press.
- Packer, B. K. (1995, May). *The play and the plan*. Paper presented at the CES fireside at Kirkland Washington Stake Center, Kirkland, WA.
- Packer, B. K. (2000). Ye are the temple of God. *The Ensign*, 11, 72-74.
- Pargament, K., Kennell, J., Hathaway, W., Grevengoed, N., Newman, J., & Jones, W. (1988). Religion and the problem-solving process: Three styles of coping. *Journal for the Scientific Study of Religion*, 27(1), 90-104.
- Pargament, K., Kennell, J., Hathaway, W., Grevengoed, N., Newmand, J., & Jones, W. (1999). Religious problem-solving scale. In P. Hill & R. Hood (Eds.), *Measures of religiosity* (pp. 347-350). Birmingham, AL: Religious Education Press. (Original work published 1988)
- Peterson, M. (1972). *For righteousness sake*. Salt Lake City, UT: Bookcraft.
- Pew Forum on Religion & Public Life. (2008). *U.S. religious landscape survey—Religious affiliation: Diverse and dynamic*. Retrieved May 15, 2008, from <http://religions.pewforum.org/reports>
- Post, S., & Leisey, R. (1995). Analogy, evaluation, and moral disagreement. *Journal of Value Inquiry*, 29(1), 45-55.
- Potenza, M., & Dorans, N. (1995). DIF assessment of polytomously score items: A framework for classification and evaluation. *Applied Psychological Measurement*, 19(1), 23-27.

- Public Issues*. (2008). Retrieved May 13, 2008, from http://newsroom.lds.org/ldsnewsroom/v/index.jsp?vgnextoid=726511154963d010VgnVCM1000004e94610aRCRD&locale=0&sourceId=5203d93c8688f010VgnVCM100000176f620a____&hideNav=1
- Radich, V. (1982). *Conservatism, altruism, religious orientation, and the Defining Issues Test: With Catholic brethren and non-religious adolescents*. Unpublished manuscript, Murdoch University, Australia.
- Reigeluth, C. (1999). What is instructional-design theory and how is it changing? In C. Reigeluth (Ed.), *Instructional-design theories and models* (Vol. 2, pp. 5-30). Mahwah, NJ: Erlbaum.
- Rest, J. (1999). Designing and validating a measure of moral judgment: Stage preference and stage consistency approaches. *Journal of Educational Psychology*, 89(1), 5-28.
- Rest, J., Narvaez, D., Bebeau, M., & Thoma, S. (1999a). A Neo-Kohlbergian approach: The DIT and schema theory. *Educational Psychology Review*, 11(4), 291-324.
- Rest, J., Narvaez, D., Bebeau, M., & Thoma, S. (1999b). *Postconventional moral thinking: A neo-Kohlbergian approach*. Mahwah, NJ: Erlbaum.
- Rest, J., Narvaez, D., Bebeau, M., & Thoma, S. (1999c). DIT2: Devising and testing a revised instrument of moral judgment. *Journal of Educational Psychology*, 91(4), 644-659.
- Richards, P. (1988). *The relation between principled moral reasoning and conservative religious ideology: A critical reevaluation and investigation of test item bias in the defining issues test*. Unpublished doctoral dissertation, University of Minnesota, Duluth.
- Richards, P. (1991). The relation between conservative religious ideology and principled moral reasoning: A review. *Review of Religious Research*, 32(4), 359-368.
- Richards, P., & Davison, M. (1992). Religious bias in moral development research: a psychometric investigation. *Journal for the Scientific Study of Religion*, 31(4), 467-485.
- Ritzema, R., & Young, C. (1983). Causal schemata and the attribution of supernatural causality. *Journal of Psychology and Theology*, 11, 36-43.

- Romney, M. (1942). Conference report. In *112th annual conference of The Church of Jesus Christ of Latter-day Saints*, Saturday, April 4-6 (pp. 17-20). Salt Lake City, UT: Deseret Book Company.
- Sanderson, S. (1974). *Religion, politics, and morality: An approach to religious and political belief systems and their relation through Kohlberg's cognitive-developmental theory of moral judgment*. Unpublished doctoral dissertation, University of Nebraska at Lincoln.
- Schwandt, T. (2000). Three epistemological stances for qualitative inquiry: Interpretivism, hermeneutics, and social constructionism. In N. K. Denzin & Y.S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 189-214). Thousand Oaks, CA: Sage.
- Scott, R. G. (2007a, May). The supernal gift of prayer. *Ensign*, 5, 8-11.
- Scott, R. G. (2007b, November). Truth: The foundation of correct decisions. *Ensign*, 11, 90-92.
- Self, D. (2000). Should moral reasoning serve as a criterion for student and resident selection? *Clinical Orthopedics and Related Research*, 378, 115-123.
- Shank, M. (2005). *The impact of moral reasoning on the performance of salespeople*. Unpublished doctoral dissertation, North Carolina State University, Raleigh.
- Shepherd, G., & Shepherd G. (1984). Mormonism in secular society: Changing patterns in official ecclesiastical rhetoric. *Review of Religious Research*, 26(1), 28-42.
- Shupe, A., & Heinerman, J. (1985). Mormonism and the new Christian right: An emerging coalition? *Review of Religious Research*, 27(2), 146-158.
- Shweder, R. (1991). *Thinking through cultures*. Cambridge, MA: Harvard University Press.
- Shweder, R., Mahapatra, M., & Miller, J. (1987). The emergence of morality in young children. In L. Kagan & S. Lamb (Eds.), *Culture and moral development* (pp. 1-83). Chicago: University of Chicago Press.
- Shweder, R., Much, N., Mahapatra, M., & Park, L. (1997). The big three of morality (autonomy, community, divinity), and the big three explanations of suffering. In A. Brandt & P. Rosin (Eds.), *Morality and health* (pp. 119-172). New York: Taylor & Francis.

- Sims, R., & Keon, T. (1997). Ethical work climate as a factor in the development of person-organization fit. *Journal of Business Ethics, 16*(11), 1095-1105.
- Smidt, C. (1989). Identifying evangelical respondents: An analysis of “born again” and Bible questions used across different surveys. In T. Jelen (Ed.), *Religion and political behavior in the United States* (pp. 23-44). New York: Praeger.
- Smidt, C., & Kellstedt, L. (1987). Evangelicalism and survey research: Interpretive problems and substantive findings. In R. Neuhaus (Ed.), *The Bible, politics and democracy* (pp. 81-102, 131-167). Grand Rapids, MI: Eerdmans.
- Stephenson, J., & Staal, M. (2007). An ethical decision-making model for operational psychology. *Ethics and Behavior, 17*(1), 61-82.
- Stewart, D. (2001). Moral development in public administration. In T. Cooper (Ed.), *Handbook of administrative ethics* (2nd ed., pp. 457-480). New York: Dekker.
- Sudweeks, R., & Tolman, R. (1993). Empirical versus subjective procedures for identifying gender differences in science test items. *Journal of Research in Science Teaching, 30*(1), 3-19.
- Tabor, M., & Adask, A. (2008). *Rights vs. “rights.”* Retrieved May 15, 2008, from http://www.lawfulpath.com/ref/rights_vs_rights.html
- Thoma, S. J. (1986). Estimating gender differences in the comprehension and preference of moral issues. *Developmental Review, 6*, 165-18.
- Thomson, J. J. (1986). A defense of abortion. In W. Parent (Ed.), *Rights, restitution, and risks: Essays in moral theory* (pp. 1-19). Cambridge, MA: Harvard University Press. (Original work published in 1971)
- Triandis, H., & Hui, C. (1986). Individualism-collectivism: A study of cross cultural researchers. *Journal of Cross-Cultural Psychology, 17*(2), 225-248.
- Turiel, E. (1978). Social regulations and domains of social concepts. *New Directions for Child Development, 1*, 45-74.
- Verbeke, W. (1996). Exploring the contextual and individual factors on ethical decision making of salespeople. *Journal of Business Ethics, 15*(11), 1175-1188.
- Volker, J. (1979). *Moral reasoning and college experience*. Unpublished manuscript, University of Minnesota at Minneapolis.

- Wahrman, I. (1980). A study of the relationship between dogmatism, religious affiliation, and moral judgment development. *The Journal of Psychology, 108*, 151-154.
- Walker, L. (2002). *Nurturing morality* (T. Thorkildsen, J. Manning, & H. Walberg, Eds.). Washington, DC: CWLA Press.
- Walker, L., & Frimer, J. (in press). Being good for goodness' sake: Transcendence in the lives of moral heroes. In F. K. Oser & W. M. M. H. Veugelers (Eds.), *Getting involved: Global citizenship development and sources of moral values*. Rotterdam, The Netherlands: Sense Publishers.
- Walker, L., & Pitts, R. (1998). Naturalistic conceptions of moral maturity. *Developmental Psychology, 34*, 403-419.
- Walker, L., Pitts, R., Hennig, K., & Matsuba, M. (1995). Reasoning about morality and real-life moral problems. In M. Killen & D. Hart (Eds.), *Morality in everyday life: Developmental perspectives* (pp. 371-407). New York: Cambridge University Press.
- Wang, C., Lu, W. Chang, F., Chang, S., Hao, J., Yang, T., et al. (2008). *A preliminary survey on the cadets' moral reasoning and attitudes toward authoritarianism, collectivism, and filial piety of two military academies in Taiwan*. Unpublished manuscript, Chinese Naval Academy & Graduate Institute of Educational Psychology and Counseling, Pingtung, Taiwan.
- White, F. (1997). Measuring the content of moral judgment development: The revised moral authority scale (MAS-R). *Social Behavior and Personality: An International Journal, 25*(4), 321-334.
- Wilcox, C. (1986). Fundamentalists and politics: An analysis of the effects of differing operational definitions. *Journal of Politics, 48*(4), 1041-51.
- Wright, B. D., & Douglas, G. A. (1975). *Best test design and self-tailored testing*. Chicago: University of Chicago, Department of Education.
- Wright, B. D., & Douglas, G. A. (1976). *Rasch item analysis by hand*. Chicago: University of Chicago, Department of Education.
- Xu, Y., Iran-Nejad, A., & Thoma, S. (2007). Administering Defining Issues Test online: Do response modes matter? *Journal of Interactive Online Learning, 6*, 1-18.
- Yeap, C. (1999). *An analysis of the effects of moral education interventions on the development of moral cognition*. Unpublished doctoral dissertation, University of Minnesota at Minneapolis.

- Young, E., & Sudweeks, R. (2005). Gender differential item functioning in the multidimensional self-concept scale with a sample of early adolescent students. *Measurement and Evaluation in Counseling and Development, 38*(1), 29-44.
- Zwick, R., Thayer, D., & Mazzeo, J. (1997). Descriptive and inferential procedures for assessing differential item functioning in polytomous items. *Applied Measurement in Education, 10*(4), 321-344.
- Zwick, R., Thayer, D.T., & Lewis, C. (1999). An empirical bayes approach to Mantel-Haenszel DIF analysis. *Journal of Educational Measurement, 36*(1), 1-28.

APPENDICES

Appendix A
History of the DIT

The History of the DIT

The Defining Issues Test (DIT) has been used as a measure for moral reasoning for nearly 30 years. It was designed based upon Lawrence Kohlberg's stages of moral development framework. Briefly, Kohlberg's theory of moral development included three levels and six hierarchical stages of moral reasoning. His stages are described as:

Level One: Preconventional

7. *The punishment and obedience orientation.* The physical consequences of action determine its goodness or badness...
8. *The instrumental-relativist orientation.* Right action consists of that which instrumentally satisfies one's own needs and occasionally the needs of others...

Level Two: Conventional

9. *The interpersonal concordance or "good boy-nice girl" orientation.* Good behavior is that which pleases or helps others and is approved by them...
10. *The "law and order" orientation.* Right behavior consists of doing one's duty, showing respect for authority, and maintaining the given social order for its own sake...

Level Three: Postconventional

11. *The social-contract legalistic orientation.* Right actions tend to be defined in terms of individual rights and standards which have been critically examined and agreed upon by society...
12. *The universal-ethical principle orientation.* Right is defined by the decision of conscience in accord with self-chosen ethical principles appealing to logical comprehensiveness, universality, and consistency... (Kohlberg, 1973, p. 631-632).

For Kohlberg, the process of determining a person's moral reasoning stage was derived by lengthy interviews where the subjects discussed hypothetical vignettes in a role-playing type of interview called the Moral Justice Interview (MJI). Responses were then coded and classified using the six stages on the MJI based on Kohlberg's 1958 dissertation (Kohlberg, 1958).

In response to critics over the years, Kohlberg progressively narrowed his theory and framework for scoring from "the study of morality to the study of moral development, to restricting it to the study of moral judgment (and its correspondence with action), to restricting it to the form or cognitive-structural stage of moral judgment as embodied in judgments of justice" (Rest 1999 p. 9). Kohlberg eventually came to view the stages of moral reasoning analogously to cognitive development in that he proposed that a person of lower moral development would not have the necessary skills to function at a higher stage of moral reasoning (Kohlberg, 1984). Due to his

pioneering work on the measurement of moral reasoning and its wide use, Kohlberg has been classified in the “top 20” of the most eminent psychologists of the 20th century (Haggbloom et al, 2002).

The move towards cognitive moral reasoning has led Kohlberg’s followers to develop a Neo-Kohlbergian framework that posits that there are moral reasoning schemas (as opposed to stages of development) that are tapped into by presenting the hypothetical ethical scenarios (Rest, Narvaez, Bebeau, & Thoma, 1999). In addition, his followers focus on macromorality (societal justice concepts such as fairness for all concerned and human rights) over micromorality (the interpersonal aspects of relationships such as loyalty in relationships; Rest, 1999).

As the framework for moral reasoning shifted from developmental stages to moral schemas, the measure of moral reasoning also changed in significant ways which were motivated due to some limitations of Kohlberg’s MJI such as: (a) the 800 + page scoring guide for scoring interviewees, (b) concerns over confounding variables due to production tasks (articulating verbally one’s moral judgments), (c) variability of scorer interpretations, and (d) lack of convenience for administration.

In answer to these concerns, James Rest and colleagues developed the Defining Issues Test (DIT; Rest et al., 1999b). Similar to Kohlberg’s MJI, the DIT is a test of moral reasoning based upon hypothetical moral vignettes. However, in the DIT, each vignette is followed by twelve “item category” items rather than open-ended responses and is scored electronically, thus removing some of the previously mentioned concerns. Thus, the main difference from the MJI and the DIT was that the MJI used a production procedure (meaning that the respondent was required to produce the answers), while the DIT uses a recognition procedure (meaning that the respondent sees and chooses an appropriate measure; Elm & Weber, 1994; Shank, 2005).

During the initial 25 years of research with the first DIT (the DIT-1) Kohlberg’s theory and framework were adapted to a Neo-Kohlbergian framework. Neo-Kohlbergians attribute Kohlberg’s theories as a starting point that emphasized a person’s internal cognitive social construction of a cooperative social system based upon moral judgment (Rest et al., 1999b). Neo-Kohlbergians are more explicit in pointing out that moral judgment is only one component of moral development and that their theories refer to macromorality issues (societal justice concepts such as fairness for all concerned and human rights) rather than micromorality (the interpersonal aspects of relationships such as loyalty in relationships). In addition, they reject the notion of hard, clearly defined stages that each have specific justice operations—operationalization of specific justice operations that define specific stages (Kohlberg, 1969)—and view the stages more broadly (or loosely) than did Kohlberg—they are seen analogously like epochs in history such as the stone age, the bronze age, the industrialization age, etc. This modified position has led to an adoption of moral schemas in which as higher moral schemas gain use, the lower ones diminish in use. However, in this shift from stages to schemas, Kohlberg’s stages are not entirely abandoned. To help with the transition from stages to schematic language, Bebeau and Thoma (2003) provided the following official CSED clarification:

Personal Interest Schema Score represents the proportion of items selected that appeal to Stage 2 and Stage 3 considerations (due to the lack of stage 1 responses for persons that the DIT is recommended for, the first stages were collapsed into one category). Stage 2 considerations focus on the direct advantages to the actor and on the fairness of simple exchanges of favor for favor. Stage 3 considerations focus on the good or evil intentions of the parties, on the party's concern for maintaining friendships and good relationships, and maintaining approval.

Maintaining Norms Schema Score represents the proportion of items selected that appeal to Stage 4 considerations. Stage 4 considerations focus on maintaining the existing legal system, maintaining existing roles and formal organizations structure.

Postconventional Schema Score represents the proportion of items selected that appeal to Stage 5 and 6 considerations...[These stages] focus on organizing a society by appealing to consensus-producing procedures (such as abiding by majority vote), insisting on due process (giving everyone his day in court), and safeguarding minimal basic rights, ...organizing social arrangements and relationships in terms of intuitively appealing ideals. (Bebeau & Thoma, 2003, p. 18-19)

In this shift from stages to schemas, Kohlberg's stages are not entirely abandoned, and the schema's are still listed hierarchally as far as their adequacy of moral judgment. The final difference in the Neo-Kohlbergian and the Kohlbergian approach is methodological in that the Neo-Kohlbergians use a recognition task (the DIT dilemmas and Likert-scored questionnaire) as compared with the previous verbal production task (the MJI).

The validity of the DIT scores is based upon several validity criteria (Rest, 1997). Rest, Narvaez, Bebeau & Thoma (1999) cite over 400 published articles and include the following criteria to establish validity of the DIT scores:

1. Differentiation of various age/education groups—studies show that 30-50% of the variance of DIT scores correlate to level of education.
2. Upward movement of scores in a 10 year longitudinal study. Effect sizes of .80 were reported for freshmen to senior college students.
3. Correlations with moral comprehension—the DIT scores are significantly related to cognitive capacity measures of moral comprehension (r .60's), the ability to recall and reconstruct postconventional arguments and fill in the missing gaps of the arguments based upon one's moral schemas, the MJI scores (r .70 to .80), and other cognitive developmental measures.
4. Differentiation of experts in moral judgment (PhD's in moral philosophy and political science) from non-experts.
5. Sensitivity to moral educational interventions—effect sizes from over 50 intervention studies were: .41 (moderate gains) compared to .09 for control groups (little gain).
6. Predictability to positions on controversial public policy issues—typical correlations in this realm range from .40 to .65 (Narvaez, Getz, Rest, & Thoma, 1999).
7. The DIT is significantly linked to many pro-social behaviors and desired professional

- decision making—37 out of 47 correlations in these realms were statistically significant.
8. Predictability to various measures of “moral behavior” (Rest, 1999, pp. 46, 59-97; Bebeau & Thoma, 2003).

In addition to the validity evidence for DIT scores, Cronbach’s reliability coefficients range from the upper .70s to the lower .80s. Test-retest reliability is about the same. The DIT-2 scores show similar correlations, but with fewer items and higher levels of reliability coefficients than the DIT-1 scores. The correlation of DIT-1 scores with DIT-2 scores is .79, nearly the same as the test-retest reliability of the DIT-1 scores with itself (Bebeau & Thoma, 2003). In addition, Rest claims that the DIT-2 uses less culturally sensitive language (Rest, 1999).

Appendix B

Three Pilot Studies and Qualitative Interview Notes

Three Pilot Studies

I completed three unpublished pilot studies in 2007 with a 30-person sample, a qualitative sample of 9 BYU students, and 135 BYU student sample. The first study looked at the correlations between styles of religious problem solving and DIT-2 scores among BYU students. The second study was a think-aloud interview as students took the DIT-2. The third study looked at religious orthodoxy measures of current BYU students.

Pilot Study #1

In the first study I used Pargament's Religious Problem Solving Scale (RPSS; Pargament et al., 1988), which measures "several religiosity-based problem-solving styles or orientations" (Pargament et al., 1988, 1999, p. 347). This connection to problem solving seemed like a plausible explanation that a person's styles of using religion to problem-solve would predict DIT-2 scores. After all, the DIT-2 poses moral dilemmas and asks a person to problem-solve using their moral judgment.

The RPSS scale measures how a person uses their religion to problem-solve and cope with life by viewing two key elements underlying an individual's relationship with God: (a) the locus of responsibility for the problem-solving responsibility, and (b) the level of activity in the problem-solving process (Pargament et al., 1988). This test returns three scores: 1) Self-directing—a score that reflects a person who takes an active problem-solving stance, 2) Collaborative—a score that reflects how the person works with God to solve problems, and 3) Deferring—a score that reflects how the person passively defers life's problems to God to solve for them.

Rest claimed that higher stage 4 scoring persons (a finding of returned missionaries) means that persons are deferring their moral judgment. Therefore, I originally hypothesized that scores on the RPSS would correlate highly with DIT-2 scores. For example, if one has a high stage four score on the DIT-2 (a score that reflects a deferring to religious or societal norms) that they would have a high deferring score on the RPSS reflecting that they are deferring to God in the moral dilemmas.

The findings give evidence that those with the highest stage four scores as a group (returned LDS missionaries) were not deferring in their problem-solving. In fact, this pilot study showed that LDS returned missionaries (those with the highest stage 4 scores in this study) were considerably *more* self-directing in their views towards solving life's problems than nonreturned missionaries (5 points difference on a 15-75 point scale, $p < .05$). In addition, returned missionaries had significantly lower "deferring" scores (scores that show that one defers to God for problem-solving) than non-returned missionaries. Pargament suggests that strong self-directing scores are typical of persons that "emphasize the freedom God gives people to direct their own lives. This approach appears to be an active coping orientation which stresses personal agency" (Pargament et al., 1988, p. 91). This style is not antireligious but compatible with a self-directing approach to life espoused by these religious conservatives.

From this finding and the comments made in qualitative responses, a very religious conservative subgroup, LDS returned missionaries, are very much into a self-directing style of problem solving rather than a deferring to duty, deferring to law-abiding schemata, deferring to Deity, or deferring to a societal norms style of problem solving. This finding contrasts with what Rest and colleagues have claimed by suggesting that religious conservatives *defer* their moral judgment based upon a maintaining norms schema (Rest et al., 1999b) and is also evidence against the developmental “moral blocking” theory. However, the pilot study found no significant correlations between the three styles of religious problem-solving (self-directive, collaboration with God, deferring to God) and DIT-2 outcome scores for BYU students. However, this first pilot study also had some other surprise findings.

Richards and Davison’s studies reported that BYU undergraduate students were scoring below national averages on their overall DIT-1 scores. However, in the pilot study, the DIT-2 scores for a group of 30 BYU undergraduate students were higher than the national average (the pilot study shows a mixed sample of freshmen, sophomore, junior, and senior BYU undergraduate students average a 41.40 overall score of moral reasoning when the national average score for college seniors is 38.94).

Another interesting finding was that BYU undergraduate’s scores of religious and political conservatism were higher than the national average but their moral judgment scores were also higher than the national average. Therefore, if the current BYU undergraduate students are still scoring high on measures of religious and political conservatism (a similar finding in Richards 1988 DIT-1 study) while their moral judgment scores are above average on the DIT-2 (an opposite finding from the 1988 study), then there is evidence that the DIT-2 items may have been improved and may not be performing differently for BYU undergraduate students. However, the full research to validate that the DIT-2 items are or are not performing better for religious conservatives has not been conducted via a DIF study which requires a much larger sample than 30 persons.

One theory that attempts to explain these differences away from an instrumental improvement in measuring moral judgment is that BYU has become more selective in their admissions to their university. The inference here is that the discriminant validity of the DIT is suspect due to correlations with cognitive abilities. However, according to Thoma, Navarez, and Rest who reviewed the literature from 1977-1997, studies have not shown significant correlations between GPA and DIT scores (Rest et al., 1999b). According to their findings “no variable (verbal ability, IQ, general cognitive ability, or GPA) accounts for the trends in the validity criteria better than” the DIT outcome scores themselves (Rest, 1999, p. 108). Based upon these claims, there may be evidence against a tightening of the selection process for incoming freshman (as based upon cognitive measures such as higher GPA’s) as an explanation for why student’s DIT-2 scores increased. However, this study did not fully explore this issue nor seek to validate Thoma, Navarez, and Rest’s claim. In other words, based upon past research, I assume that something other than cognitive abilities is the reason for higher BYU DIT scores. This led to the second pilot study.

Pilot Study #2

A second pilot study involved a qualitative think-aloud protocol as BYU students took the DIT-2 test. This led to asking about life experiences, purposes in suffering, doctrinal influences, philosophies of life, and other relevant issues on the DIT-2 moral dilemmas. It also led to a clarification of DIT-2 instructions (see Appendix E). I have included these pilot study qualitative interview notes at the end of this Appendix. These first two pilot studies were the basis for my online survey (see Appendix E).

The final finding of the first pilot study found that BYU focal group-mean religious orthodox scores were 6.23. The University of Alabama reference group scores were 6.64. These scores were normally distributed. Independent samples *t* tests led me to conclude that these differences were not statistically significant (see Table B-1). Therefore, the UA and BYU samples did not differ in matters of religious orthodoxy on the DIT-2 RO scale and possibly in other matters of religious conservatism. This led to the third pilot study.

The third pilot study further validated the assumption that BYU and UA students did not differ much in religious orthodoxy. I compared scores from Brown and Lowe's Religious Inventory of Religious Belief (Brown & Lowe, 1951) and found similar trends when two items were removed from both samples. The UA Brown and Lowe Religious Inventory sample was provided by Dr. Steve Thoma. The BYU sample came from 135 BYU students enrolled in Religion C 333. In the analysis, two items were removed from Brown and Lowe's inventory. The first item that was removed was item 7: "God exists as: Father, Son, and Holy Spirit." The reason this item was removed was because while discussing this item *ex-post facto* with the BYU students they mentioned that they had interpreted this question as meaning that God was three individuals in one and consequently strongly disagreed with this statement. However, BYU students did report believing in The Father, Son, and Holy Ghost as three separate individuals. This item also did not correlate with the BYU student's item-total correlations either indicating that it did not fit with the overall construct being measured for BYU students.

In addition, for the BYU students, item 8 did not correlate with the 14 other items from BYU students, nor with their item total correlations, yet it did for UA students. Item 8, a question about the whether there are mistakes and errors in the Bible and was intended to measure a belief in the inerrancy of the *Bible*, a trait common to religious conservatives (Smidt, 1989). In

Table B-1

A Comparison of DIT Internal Measures of Religious Orthodoxy by School

School	<i>N</i>	Mean	<i>SD</i>	Std. error mean	Sig. (2-tailed)
UA	423	6.6430	2.42377	.11785	.373
BYU	30	6.2333	2.52823	.46159	

conversing with the BYU students, item 8 was interpreted by BYU students to refer to Biblical *translations* rather than the doctrinal *content* of the *Bible* that the item was intended to be measuring, so this item was removed from both samples. As noted in Table 4, there were no other significant differences on this overall measure of religious orthodoxy when these two items were removed (see Table B-2).

As both of these measures of religious orthodoxy were not statistically significant, it was therefore determined that as a group, BYU and UA students did not significantly differ in measures of religious orthodoxy. This finding posed a problem in my original data collection plan because I was seeking to compare a group that was very religious orthodox to a group that was closer to a national average of religious orthodoxy. A bright spot in this comparison sample was that the variance in the UA sample was much greater and the size was much larger than in the BYU sample. This led to the filtered sampling approach from the UA sample based upon national averages of religious orthodoxy described in chapter IV of my study. It also was the basis of using ANCOVA in the dissertation study.

Pilot Study #3 & Qualitative Interview Notes

Notes from students who took the DIT-2.

Research Goal #1: To identify any portion of the instructions, words, or phrases that are confusing to a BYU student. The DIT-2 scoring guide states that using a sheet with definitions and clarifications does not compromise the results.

Research Goal #2: To identify what kind of life experiences, doctrinal teachings, or philosophies a student uses to morally judge in the DIT-2 scenarios. This will aid in forming the qualitative aspects of my study (e.g., an essay portion, open-ended question portion).

Comments about the instructions on the DIT-2:

The “No” option. Two students pointed out that it would be better if this was explained as referring to “Not important at all.” They even suggested that it be changed on the answer sheet to “NOT” or just to remind students that the “Great, much, some, little, and no” options should

Table B-2

A Comparison of a Measure of Religious Conservatism by School (Brown and Low Inventory)

School	<i>N</i>	Mean	<i>SD</i>	Std. error mean	Sig. (2-tailed)
UA	423	58.8180	10.20605	.49624	0.379
BYU	135	59.3481	3.96688	.34141	

be followed by the word importance. Otherwise, a student may think that the item is a “yes” or “no” question and select the “No” option when they really do not mean that.

For each item it helped when I restated in my head this question “Is this important to me or not.”

Also, it was helpful to ask, “Does this pertain to why I decided what I did about this scenario?”

This [the presidential election example in the DIT-2 instructions] is a difficult example. It’s hard to understand because it’s not a story but the test is all stories.

This [which issue is the most important to you in making up your mind about...] is not stressed enough. It’s hard to realize that you are rating them [the following 12 items] as to how important they are in terms of making your decision. Perhaps underline this phrase on the instructions.

In the instructions it says to read the story and then go to the 12 issues and rate then rank them but the first step is to answer “what should he do?” not go straight to the 12 issues. That’s kind of confusing.

I would re-word the instructions with a simple example that is more similar to the actual problems. The presidential example isn’t very easy to understand.

Terms that need more clarification:

Story #1 (Famine)

Item 7: “Social cooperation”

Item 8: “epitome”, “reconcilable”, “culpability”

Item 2: I was a little unclear that he is a father. I thought he was a kid.

Story #2 (Reporter)

Item 9: “habeas corpus”

Story #4 (Cancer)

Item 5: “active heliotropic drug”

Life Experiences, Doctrinal Teachings, Philosophies:

Story #1 (Famine)

Life Experiences:

One participant served an LDS two-year mission in the Bahamas where there is a great separation of rich and poor. He was bothered by how the poor viewed the rich and felt that they could take what they wanted from them regardless of a need and mostly out of a want. He saw that as “not okay” to do ever and was bothered by it.

Another participant has lived his whole life in North America (Vancouver and the US). He

mentioned that there is always some way to provide or some other alternative to stealing.

Another participant said his feelings would probably change if he were married: “I’m not married with children but if I was this might pull a little harder on my heartstrings.”

Doctrinal Teachings:

The Ten Commandments were mentioned as a source by several participants. For example, “Thou shalt not steal” was a common phrase. Several other scriptural teachings were mentioned or paraphrased.

One participant mentioned a concept found in the Book of Mormon in regards to the Famine story: “It’s like what you see in 1 Nephi 3:7: ‘...for I know that the Lord giveth no commandments unto the children of men save he shall prepare a way for them that they may accomplish the thing which he commandeth them.’” This statement was followed up by: “When God gives a commandment, He provides a way to obey it. Therefore, there must be some sort of alternative other than stealing—there is a way.”

Another person brought up one of the ten commandments in modern language: “The commandment do not steal.”

Another concept mentioned was D&C 134:11: “‘We believe that men should appeal to the civil law for redress of all wrongs and grievances, where personal abuse is inflicted or the right of property or character infringed, where such laws exist as will protect the same; but we believe that all men are justified in defending themselves, their friends, and property, and the government, from the unlawful assaults and encroachments of all persons in times of exigency, where immediate appeal cannot be made to the laws, and relief afforded.’ [Therefore], you can fight against injustice.”

Another concept was that of an “eternal perspective”—that this life is only a small part of a greater eternal plan: “Mortality is but a small moment, therefore, death is not the worst thing that could happen to you.”

Some used scriptural stories to justify stealing the food: “David ate the temple bread [when he was desperate for food] (this is in reference to David in the Old Testament eating something that was forbidden to eat but in desperate circumstances the priests gave him this sacred bread to eat—see 1 Samuel 21:6; Matthew 12:4; Mark 2:26; Luke 6:4).

Another mentioned a story in the Book of Mormon where a prophet is commanded to slay a wicked man to save the sacred records: “The Lord told Nephi [a prophet in the Book of Mormon] ‘it is better that one man should perish than that a nation should dwindle and perish in unbelief’ (1 Ne. 4:13). The inference from these scriptural stories of David and Nephi was that sometimes it’s better to break one commandment to keep another.

Philosophies:

Several students mentioned philosophies of life that they live by or have studied. Three of the following comments were typical in several of the DIT-2 stories:

“I am a firm believer in property rights. I have studied a lot of Economics and am a proponent of ownership concepts.”

“I believe more can be done by governments.”

“Two wrongs don’t make a right.”

Story #2 (Reporter)Life Experiences:

The college students had rich life experiences that they drew on to justify their choice for the DIT-2 stories. Several of the following comments were typical:

“I have often thought that the moral character of a candidate is important to me in an election.”

“I know I have lived wrong and I don’t want the world to know about it.”

“My brother has shady past and he came to mind right away as I was reading this.”

“In High School, the media published a report about the security of government installations [places that should be secure] in the name of knowledge. I thought it was silly that they [the media] would jeopardize national security.”

Doctrinal Teachings:

These comments were not as prevalent in this story but still surfaced:

“These [doctrinal teachings] did not affect me as much on this [story].”

“I thought of the principles of repentance and forgiveness, particularly that of restitution.” “It seems that he has made restitution but it doesn’t necessarily say that.”

“I don’t think this is related to Mormon belief. It’s more of a political question about the role of media.”

Philosophies

Only one person mentioned a philosophy that influenced them in this story: “I have studied media ethics in my journalism classes and we go over things like this all the time. The answer I have decided is that it all comes down to how it is reported and the motives of the reporter.”

Story #3 (School Board)

Life Experiences:

A couple of persons mentioned that their life experiences influenced them in the following ways:

“There were two high schools in my town and we had a very controversial topic that pertained to both of them. However, we handled it very civilly in our meetings. People acted responsibly, so I know that that can be done.”

“The people are acting irresponsibly [in a citizen type of way]. My parents often took away rights when I was not acting responsibly. However, they also sometimes let me learn from my mistakes. It depended upon the consequence of letting me go my own way.”

Doctrinal Teachings:

The students responded in the following ways when they were asked, “Were there any doctrinal teachings that influenced your decision?”

“Honesty—he gave his word and he should keep his word. I’m sure there is a way he can initiate rules or regulations for the debate that will still encourage open discussion.”

“He gave his word and he tried to keep it but it didn’t work. So he tried to keep it so he’s really not accountable for going back on it—it’s for the greater good.”

“The doctrine I see relates to keeping promises.”

“Don’t fear man, just fear God.”

Philosophies

The students responded in the following ways when they were asked, “Were there any philosophies of life that influenced your decision?” “In this case the people are acting like babies so they should be treated like babies.”

“There has got to be a way of diplomacy that will work in this situation.”

“The reason we have leaders is to make decisions. If the leader can’t make decisions, why even have one?”

Story #4 (Cancer)

Life experiences:

Two students mentioned that they had gone through cancer treatment but interestingly chose to use their experience to support different DIT-2 decisions about whether to euthanize a suffering cancer patient.

“I had testicular cancer at age 19. I’ve gone through cancer treatment and been close to people with cancer and seen their suffering. Going through chemo[therapy] you are still alive but not living life. When there is no hope of improving and they [the patient] can make a conscious decision, they should be allowed to make that decision. This is a regular concept of life and living life.

“I have been through bone cancer as a child, six years old. Fortunately it was caught early. I don’t recall the radiation because I was a child and just figured that was what I was to go through. My grandma died of a brain tumor and we tried to make her comfortable but we let things take their course.”

When students were asked if this story had anything to do with euthanizing animals, students all nodded as one participant voiced the following: “I’ve had pets euthanized but having pets euthanized has nothing to do with this.”

Doctrinal Teachings:

The students responded in the following ways when they were asked, “Were there any doctrinal teachings that influenced your decision?”

“Agency—how far do you let someone go? If she wants to go through with it then she should. If not, she may find a different way to end it (suicide).”

“I know the church has probably taught something about this but I really don’t know the doctrine of the church on this issue.”

“The reason pets are not really an issue here is that God has given man dominion over animals but not [over] other men. Man does not have dominion over another man.”

“[Understanding] the purpose of life and the doctrine of life after death play into effect here.”

“In the scriptures, the Anti-Nephi-Lehis allowed their brethren to perish without exercising their abilities to preserve their lives. Was that allowing them to be killed? [this was a parenthetical question—the Anti-Nephi-Lehi’s were a group of people in The Book of Mormon who refused to go to war against an invading people because they had made sacred promises not to fight or kill their fellow men. A large number of them were killed when the invading army came in].

Philosophies:

The students responded in the following ways when they were asked, “Were there any philosophies of life that influenced your decision?”

“Assisted suicide—are you responsible?”

“It’s not my call. There are lots of things in life I would like to control but I can’t.”

“The doctor has a duty to warn and inform her of the consequences but it is her choice.”

“I think it is completely different than euthanizing an animal.”

Story 5 (Demonstration)

Life experiences:

The students responded in the following ways when they were asked, “Were there any life experiences that influenced your decision?”

“I can relate because I am in college but I think there are better ways to protest something than conflict.”

“This seems to be a lot like the Iraq conflict but the protest is the issue not the conflict.”

Doctrinal teachings:

The students responded in the following ways when they were asked, “Were there any doctrinal teachings that influenced your decision?”

Several mentioned article of faith # 12—“We believe in being subject to kings, presidents, rulers, and magistrates, in obeying, honoring, and sustaining the law.” Many of the students agreed or quoted the part about honoring and sustaining the law. Others mentioned scriptural stories that might justify such a protest: “Some people in the book of Mosiah broke laws and did ‘bad things’ to the Lamanite soldiers because they didn’t like oppressive taxes [they referred to a rebellion against a King who had high taxes to support his luxurious and impious lifestyle].”

“The [Latter-day] Saints endured persecution and went through the proper channels to effect change” [This is a reference to early Mormons who were driven from the State of Missouri, some at gunpoint. An extermination order was signed by Governor Boggs in 1838 that allowed a person to shoot a Mormon on site].

Philosophies:

The students responded in the following ways when they were asked, “Were there any philosophies of life that influenced your decision?” This kind of pragmatic argument was found in several comments.

“Whether they have a right is not as much an issue as whether it is an effective use of time.”

“Productive use of time is what I thought of. It seems like they are wasting time.”

“There’s no order here.”

“Do two wrongs make a right?”

Appendix C

Demographic and Survey Questions

Demographic & Survey Questions

- 1) Religious Affiliation _____
 - 2) Returned Missionary (Yes or No) _____
 If “Yes” when and where did you serve? _____
 - 3) Married, Single, Widowed, Divorced, Separated, Other _____
 - 4) Mother’s formal education level (Please check one—if your answer needs some explaining you may write any comments below).
 - Junior High
 - Some High School
 - High School diploma
 - Post High School Technical training (e.g., Dental Hygenist training). Please indicate
 - Some College
 - Associates Degree
 - Four Year Degree—bachelor degree
 - Some Graduate School
 - Masters Degree
 - Some Beyond Masters Degree
 - Doctoral Degree (PhD, MD, EdD, JD, etc.)
 - Post Doctoral Work
 1. Has anyone very close to you ever been convicted of shoplifting?
 2. Has anyone very close to you ever served Jail Time for a crime they committed? Yes or No.
 3. Do you feel that there is a purpose in human suffering? Yes or No. Please explain.
 4. As you made your decisions and answered the items on the DIT-2, please briefly describe if you thought of any life experiences, doctrinal teachings, philosophies of life, or other relevant issues. See the attached example for further explanation (Appendix E).
- Story #1 Famine
- a. Please describe any of your life experiences that influenced the decisions you made in regards to this story.
 - b. Please describe any doctrinal teachings that influenced the decisions you made in regards to this story.
 - c. Please describe any philosophies of life that influenced the decisions you made in regards to this story.
 - d. Please describe any other relevant issues that may not be mentioned in this story but that influenced the decisions you made in regards to this story.
- These same a-d questions were repeated for stories 2-5 also.
5. I give permission for the principal investigator, Daniel R. Winder, to contact me to ask additional questions to clarify my replies on this survey. Yes or No.
 6. If “yes” I can be reached at:

Appendix D
Informed Consent Document

Consent to be a Research Subject

Introduction

This research is being conducted by Daniel R. Winder at Brigham Young University to determine whether there are items in the Defining Issues Test-2 (DIT-2) that perform differently for BYU undergraduate students and if so, whether those items exhibit a test bias among all BYU undergraduate students or only among certain subgroups of BYU undergraduate students.

Procedures

You will be given the DIT-2 and a short answer survey form with some demographic questions about yourself. These surveys consist of moral dilemmas followed by questions about the dilemma and questions about how a person uses their religion to problem-solve. Both surveys are to be completed individually within two weeks of receiving them and returned to the religion department secretary, Cheryl Snelgrove (370 JSB) or the principal investigator, Daniel Winder (215 JSB). The total amount of time needed to take these surveys is about 60 minutes.

Risks/Discomforts

There are minimal risks for participation in this study. However, there is religious content in the demographic and qualitative survey that asks questions about one's personal beliefs.

Benefits

There are no direct benefits to the subjects. However, it is hoped that through your participation researchers will learn more about the field of moral judgment for religious persons and be able to design more effective instruments to measure moral reasoning for such persons.

Confidentiality

All information will remain confidential and will only be reported as group data with no identifying information. All information will be kept in a locked fire-safe and only be accessed by the principal investigator. Once the individual has been compensated for completing the surveys and non-identifying data has been entered into a computer program for analysis, any identifying information will be deleted.

Compensation

Participants will receive a \$5 gift certificate to the BYU bookstore or 25 points of extra credit in a course taught by the principal investigator. For those students of the principal investigator who do not wish to participate in the research, the same amount of extra credit can be earned by writing a 5 paragraph essay on a talk from a living prophet from the October 2007 General Conference.

Participation

Participation in this research study is voluntary. You have the right to withdraw at anytime or refuse to participate entirely without the jeopardy to your standing at BYU, your class status, or grade.

Questions about the Research

If you have any questions regarding this study, you may contact Daniel Winder at 422-2330, winderdr@gmail.com, or 215 JSB during office hours or the IRB Chair Renea Beckstrand at 422-3841, irb@byu.edu, A-285 ASB.

Questions about your Rights as Research Participants

If you have questions regarding your rights as a research participant, you may contact Daniel Winder at 422-2330, winderdr@gmail.com, or 215 JSB during office hours or the IRB Chair Renea Beckstrand at 422-3841, irb@byu.edu, A-285 ASB.

I have read, understood, and received a copy of the above consent and desire of my own free will to participate in this study.

Signature: _____ Date: _____

Appendix E

Clarifying the DIT-2 Instructions

Clarifying the DIT-2 Instructions

This questionnaire is concerned with how you define the issues in a social problem or dilemma. Several stories about social problems will be described. After each story, you will be asked to do the following:

1. Indicate whether you agree, disagree, or cannot decide your stance regarding the action in question.
2. There will be 12 issues or questions that an individual might ask when considering the situation. You are to rate these 12 issues on how important they are to you as you consider the situation or story.
3. Decide which four issues were the most important in your considerations.

Here is a short example of the task:

Chinese Army – (Story # 0)

Fa Zhou is an old man in central China, and because of previous war injuries, he must now walk with a cane. Because China has been attacked by the Huns, the draft requires that the family of Zhou must send a man to be a soldier. Zhou is the only male in his family and must join the army, despite his injuries and age. His daughter Mulan is a capable young woman, and is considering cutting her hair and taking her father's armor and sword to join the army in his stead, even though women are not allowed to be in the army.

Chinese Army – (Story # 0)

Do you think that Mulan should take her father's place in the army? (Mark one.)

1. Should join the army. 2. Can't decide 3. Should not join the army

Great Much Some Little No

- | | | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--|
| <input type="checkbox"/> | 1. Is it improper for Mulan to disobey her parents? |
| <input type="checkbox"/> | 2. Shouldn't the government make exceptions for the elderly? |
| <input type="checkbox"/> | 3. Is Zhou's honor more important than his life and comfort? |

Rank which issue is most important (from the issues raised above).

Most important item: 1 2 3.

Second most important: 1 2 3.

Note that some items may be irrelevant to you or not make sense—in that case, rate the item as “no.” Also, note that there will be 12 items to rate for each story, rather than 3.

When you are finished with the DIT-2 answer sheet, please fill out the “Demographic and Short Answer Questions” sheet. The demographic section is pretty self-explanatory.

The short answer portion will ask you to think back as you made your decisions for the five stories and the items following each of the five stories. It asks you to briefly describe if you thought of any life experiences, doctrinal teachings, philosophies of life, or other relevant issues as you filled out the DIT-2. For example, using the Chinese Army Story #0 previously described:

Story #0 Chinese Army

- a. Please describe any of your life experiences that influenced the

decisions you made in regards to this story.

My father is suffering from a hip replacement and I thought of him and how I would do anything to keep him out of a war in his weakened condition because I care so much for him.

OR

I served a mission in China and I understand family honor and dishonor in this culture. In this case it would be so inappropriate to risk getting caught and disgracing the family name, that her good desires in no way could overshadow the potential dishonor. For example, I had a 65 year old man refuse to be baptized because his grandmother said “no”.

- b. Please describe any doctrinal teachings that influenced the decisions you made in regards to this story.

As I was thinking about this item I thought of the commandment to honor thy father and mother. In this case it would be dishonorable to serve in a father’s place so that doctrinal teaching influenced my decisions on these items.

- c. Please describe any philosophies of life that influenced the decisions you made in regards to this story.

I have always asked myself “do two wrongs make a right?” It’s wrong for the military to ask an old man to serve and it’s wrong for a woman to serve in the military in this story.

OR

Like the patriots in the Boston tea-party, I have a philosophy that I live by: “if the rule is wrong, then don’t let it rule you.” In this case, I see the rule of not letting women serve as wrong.

- d. Please describe any other relevant issues that may not be mentioned in this story but that influenced the decisions you made in regards to this story.

As I read the story I thought that a more relevant moral issue in this case is whether or not a country can force its citizens to fight. That is what would most influence my decision in this case but it was not asked on any of the items.

The following definitions may be helpful when taking the DIT-2.

Story #1 (Famine) Item 7: “social cooperation”—refers to human society & joint organization.

Story #1 (Famine) Item 8: “epitome”—a strong example.

“reconcilable”—made compatible or consistent.

“culpability”—blameworthiness: a state of guilt.

Story #2 (Reporter) Item 9: “habeas corpus”—A writ commanding that a person be brought before a judge.

Story #4 (Cancer) Item 5: “active heliotropic drug”—a drug that actively produces effects in humans [I think that you have to include your part about plants that follow the sun]

Appendix F

DIT-2 Story Moral Dilemmas

DIT-2 Story Moral Dilemmas

Famine (Story #1)

The small village in northern India has experienced shortages of food before, but this year's famine is worse than ever. Some families are even trying to feed themselves by making soup from tree bark. Mustaq Singh's family is near starvation. He has heard that a rich man in his village has supplies of food stored away and is hoarding food while its price goes higher so that he can sell the food later at a huge profit. Mustaq is desperate and thinks about stealing some food from the rich man's warehouse. The small amount of food that he needs for his family probably wouldn't even be missed.

Reporter (Story #2)

Molly Dayton has been a news reporter for the Gazette newspaper for over a decade. Almost by accident, she learned that one of the candidates for Lieutenant Governor for her state, Grover Thompson, had been arrested for shoplifting 20 years earlier. Reporter Dayton found out that early in his life, Candidate Thompson had undergone a confused period and done things he later regretted, actions which would be very out-of-character now. His shoplifting had been a minor offense and charges had been dropped by the department store. Thompson has not only straightened himself out since then, but built a distinguished record in helping many people and in leading constructive community projects. Now, Reporter Dayton regards Thompson as the best candidate in the field and likely to go on to important leadership positions in the state. Reporter Dayton wonders whether or not she should write the story about Thompson's earlier troubles because in the upcoming close and heated election, she fears that such a news story could wreck Thompson's chance to win.

School Board (Story #3)

Mr. Grant has been elected to the School Board District 190 and was chosen to be Chairman. The district is bitterly divided over the closing of one of the high schools. One of the high schools has to be closed for financial reasons, but there is no agreement over which school to close. During his election to the school board, Mr Grant had proposed a series of "Open Meetings" in which members of the community could voice their opinions. He hoped that dialogue would make the community realize the necessity of closing one high school. Also he hoped that through open discussion, the difficulty of the decision would be appreciated, and that the community would ultimately support the school board decisions. The first Open Meeting was a disaster. Passionate speeches dominated the microphones and threatened violence. The meeting barely closed without fistfights. Later in the week, school board members received threatening phone calls. Mr. Grant wonders if he ought to call off the next Open Meeting.

Cancer (Story #4)

Mrs. Bennett is 62 years old, and in the last phases of colon cancer. She is in terrible pain and asks the doctor to give her more painkiller medicine. The doctor has given her the maximum safe dose already and is reluctant to increase the dosage because it would probably hasten her death. In a clear and rational mental state, Mrs. Bennett says that she realizes this; but she wants to end her suffering even if it means ending her life. Should the doctor give her an increased dosage?

Demonstration (Story #5)

Political and economic instability in a South American country prompted the President of the United States to send troops to "police" the area. Students at many campuses in the U.S.A. have protested that the United States is using its military might for economic advantage. There is widespread suspicion that big oil multinational companies are pressuring the President to safeguard a cheap oil supply even if it means loss of life. Students at one campus took to the streets, in demonstrations, tying up traffic and stopping regular business in town. The president of the university demanded that the students stop their illegal demonstrations. Students then took over the college's administration building, completely paralyzing the college. Are the students right to demonstrate in these ways?

Appendix G

Qualitative Data Analysis and Tables

Qualitative Data Analysis and Tables

Qualitative responses were first coded and then separated into a table of codes (see Table G-1). Data was then sorted into comments by categorical responses based upon one of the three decisions a person can make in each of the DIT-2 moral dilemmas (see Table 2). For example, notice in Table 2 that the responses are sorted by the categories “should give,” “can’t decide,” and “should not give.” This pertains to whether a doctor should give an overdose to end a suffering cancer patient’s life.

The comments were also sorted in the table by the survey question that brought forth the comment: “life experiences,” “doctrinal teachings,” “philosophies of life,” and “other relevant issues” (see Table G-2).

Life experiences included academic training, work-related experience, social experiences, and family experiences. The difference between theological and philosophical was based on whether the MMR was based on doctrinal teachings or another moral code. Much of the content for MMR include statements that are clearly scriptural in nature (i.e., Thou shalt not steal, Agency is a divine gift, Repentance) while philosophical content for MMR includes statements that refer to a person’s moral code (i.e., two wrongs don’t make a right, the ends don’t justify the means, anarchy is not an acceptable way to protest). In qualitative think-aloud protocols, BYU students would frequently refer to what they thought was a more relevant issue when making a decision for what should be done on the DIT-2. For example, in story #5, college students take over the administration building in a protest and halt university operations. Most students mentioned that demonstrating the way these students were was “a waste of time.” So time-management was more of an issue to them than to protest in this manner.

Because the qualitative questions were open-ended, some participant’s responses contained multiple thoughts. Notice on the relative totals column in Table 2 that when the percent of comments dropped below 3% that the relative total was not computed. Most of these comments did not add any information in the comparison of responses because they were specific comments pertaining to an uncommon category.

Once comments were coded and categorized, the relative percent of the comments pertaining to that category were computed. These relative percentages of coded categorical comments were compared by the decision made in the story. After these relative percentages were computed for the cancer story, a summary table was written up summarizing the relative percent of comments and the general feel of the comments from the specific category (see Table G-3). Table G-3 was a little redundant because Table G-1 and G-2 already contained this information so this table was not repeated for the other DIT-2 stories. However, the rest of the process was repeated for four of the five DIT-2 stories (see Tables G-4 to G-9). The demonstration story comments were not computed because of the lack of variance and because most students thought that this story was not a moral dilemma at all—it was actually viewed in terms of cost-benefit use of time, effort, and effectiveness or saliency of making a point. There were only a handful of comments besides the ones that discounted it a true moral dilemma.

Table G-1

Cancer Story Codes for Qualitative Data

Code	Collapsed categories	General descriptions
A	Avoid suffering	Avoid suffering for the person and their loved ones, no chance to live anyway, not same person if they are living in pain, hate to see them suffer, I wanted a (relative or close friend) to die for them, their quality of life is hindered by their suffering, why should one prolong life just to prolong it if someone is suffering--there seems to be no purpose.
AG	Anti-God	God doesn't pertain to this question.
AM	Mercy	The doctor is extending her mercy, her dying wish ought to be realized as a merciful gift to her.
B	Beneficial	It would be very beneficial to have or allow doctor assisted suicide.
C	Choose	God gave a person agency to act for themselves and take control of situations. He wants us to decide.
Capt	Capital punishment	Capital punishment has man taking man's life into their dominion/hands.
D	Doctor	The doctor has a responsibility here. There is liability, lawsuit, and legal issues. A doctors role is to sustain life--that is their duty.
DP	Defer	I have my opinions because my parents taught me this or friends taught me this.
Duty	Duty	The doctor has a duty to do it.
E	Euthanize	I have euthanized animals. My opinions of euthanasia affected me.
F	Family	A family member, relative, loved one, or self went through a similar ordeal. Life support is similar. I or a loved one have experienced similar pain and suffering.
FM	Frame of mind	Her frame of mind may prevent her from making the best decision because she is suffering so much. She is in the wrong mind frame to make this decision.
G	God	God, doctrine, religious leadership helps me understand this dilemma. A belief in God's doctrine of life after death. A doctrine that God has given man dominion over animals and plant life but not over their fellow men's lives. Faith to endure pain comes from God.
GR	Golden rule	Do unto others as you would have others do to you.
H	Helping God	This is actually helping God. He has already called her home. He is in control.
I	Individual	The individual is ready to go
K	K & OD	The doctor would be killing her and this is wrong. It is murder or suicide and one should not end life prematurely. Overdosing is wrong in any situation.
L	Life	One should fight for life. Not an option to not fight for life. Life is sacred and therefore this is serious.
LDS	LDS	Is she a member? Is she converted? Does she know?
LL	Law & Govt	The laws of the land make one accountable for this choice, the laws have defined what is legal and right in this case. The government and society rule in this matter.
M	Myself	It's her life, her body, her choice, her responsibility or my life, my choice, my body, my responsibility.
Med	Medicine	Anti-medication, had a negative experience with meds.

(Table continues)

Code	Collapsed categories	General descriptions
New	New	Make a new law.
O	Old age	The issue is easier if she is old rather than young.
P	Peaceful endurance	One should peacefully endure suffering until a natural death occurs. There may be a blessing for her or others if she endures it patiently. There is a purpose in her suffering. Endure faithfully to the end.
Pol	Political	This is a political issue and should be voted on.
R	Rights	A defining of who has the right to make this decision. Family and close friends are affected also so they should have a right.
Rel	Relative	There is no right or wrong answer in every situation like this. It is all relative.
RIP	Death	Death is painful for her loved ones.
S	Surgery	Surgery ends life prematurely at times. We take people off life support. We allow triage in certain situations.
Self	Self & suicide	The woman in this example is very selfish to ask what she is asking of the doctor. The problem is hers, not the doctors and she is deferring the responsibility for her life to him. If she wants to end her pain she should do it herself, not make someone else do it.
Time	Time	How long has she been on pain medication, meds?
Try	Try	Has she tried everything else or anything else for her pain treatment?
TV	Media	I have media exposure to similar circumstances that help me understand this dilemma.
W	Work & school	I have studied or have work experience related to this moral dilemma. (Table Continues)
Weak	Weak	Weak moral dilemma because it contradicts itself.
Z	Exception	This is an exception to God's rules for life and death.

Table G-2

Relative Percent Counts of Coded Comments by DIT-2 Cancer Story Decision

Cancer Story	Should give	Can't Decide	Should not give
# of completed cells	118No Comment, 94 Completed comment fields	187No Comment, 117 Completed comment fields	284No Comment, 260 Completed comment fields
Relative Totals of completed comments	46%A 31%C 20%F 19%G 17%R 14%K, 11% R 11% AM 7%D 6% I 6% H 5%W 4%L 3%TV	30% G 20% K 20% F 17% D 17% C 17% L 16% A 9% R 8% P 8% LL 6% FM 4% AM 4% W 4% SELF 3% TV 3% I	38%G 32%K 24%P 20%L 16%F 14%D 8%A 8% LL 7%R 5%Self 5% W 4%C 3%FM 2%DP

Table G-3

Coded Comments and Category Examples/Descriptions

Category	Description
Life experiences	Of those who wrote about life experiences, 64% of the comments were about family or a close personal friend's experience with cancer. 29% mentioned something about an experience where they wanted themselves or close associate to avoid suffering. Seventeen percent mentioned work related experiences. Fourteen percent mentioned an experience of thinking about a doctor's liability or responsibility. Fourteen percent mentioned experiences of a peaceful endurance of suffering and the benefits of it. Twelve percent of the comments referred to the sacredness of life and that it therefore should be fought for. Ten percent mentioned that a family or close friends are affected by this decision so that they should have a say in it, 6% mentioned that she was not in the right frame of mind to make such an important decision, 6% also mentioned that this would be considered killing and that killing is wrong, 6% mentioned the individual's readiness to die, and 4% mentioned that the media influenced some of their decision.
Doctrinal teachings	Of those who wrote about doctrinal teachings, 54% of the comments mentioned a doctrine of God's dominion over life and His purposes in suffering. 40% mentioned that killing or murder is wrong and that the doctor is killing in this situation. Eighteen percent mentioned that one should seek to patiently endure suffering and learn from it, 16% mentioned the sacredness of life and that life should be fought for, 13% mentioned that God gave man agency to chose for themselves, 4% mentioned the avoidance of suffering and mercy to the woman, 3.5% mentioned that the laws of the land ought to be kept, the doctors liability, that one may be assisting God by euthanasia, and that one needs to define the rights in this situation. Two percent mentioned the teaching of mercy and 2% deferred to their parents teachings.
Philosophies of life	Of those who wrote about philosophies of life, 24% of the comments mentioned that because of their belief in God they had a philosophy of life that included a purpose in suffering, 22% mentioned that man has a right and ability to chose for himself, 20% mentioned a doctor's liability here, 19% mentioned that life is sacred and should be viewed as a gift, 16% mentioned the avoidance of suffering, 14% mentioned the sacredness of life and that life should be fought for, 8% mentioned that one should follow the laws of the land,, 7% mentioned the selfishness of the woman for asking the doctor to do that, 4% mentioned that she is not in the frame of mind to make the decision.
Other relevant issues	Of those who wrote about other relevant issues, 18% of the comments mentioned the avoidance of suffering, 16% mentioned the doctor's responsibility and liability, 13% mentioned that the doctor would be killing or murdering, 13% mentioned that one ought to follow the laws of the land, 9% mentioned that the woman was selfish for asking the doctor to end her life, 8% saw flaws in the dilemma that weakened it or brought up media and a definition of rights for individual and family and friends, 6% brought up whether she has tried different things and the sacredness of life, 5% brought up surgery or other medical methods of reducing pain, 3% brought up the woman's frame of mind.

(Table continues)

Category	Description
Should give her an overdose	Of those who responded in the 1 category, 46% of the comments mentioned that the avoidance of suffering was important to their decision, 31% of the comments mentioned that God gave man agency or the ability to chose for themselves, 20% mentioned that a family member or close associate when through a similar ordeal and this influenced their response, 19% mentioned God’s doctrine on life and his teachings, 17% of the comments mentioned a defining of rights for the individual and the loved ones, 14% mentioned that the doctor would be killing, 11% mentioned that it would be a last merciful gift to the patient, 7% mentioned the doctor’s liability, 6% mentioned helping God by euthanasia and whether or not the individual was ready to die, 5% mentioned work related experiences that influenced their decisions, 4% mentioned the sacredness of life and that it should be fought for, and 3% mentioned the role of the media in influencing their decision.
Can’t decide	Of those who responded in the 2 category, 30% of the comments mentioned a belief in God’s word affected their views of life/death, 20% mentioned that the doctor would be killing, 20% mentioned that a family or close associate went through a similar experience that this influenced their decision, 17% mentioned the doctor’s liability, 17% mentioned that God gave man the ability to chose for themselves, 17% mentioned the sacredness of life and that it should be fought for, 16% mentioned the avoidance of suffering for the individual, 9% mentioned a definition of the rights of the individual and of the loved ones affected, 8% mentioned the laws of the land should be followed, 8% mentioned a peaceful endurance and the benefits of such, 6% mentioned that the woman is not in the right frame of mind to make the decision, 4% mentioned that the woman is selfish for asking the doctor to make this decision, 4% mentioned that it would be a last merciful gift to the woman, 4% mentioned work experience, and 3% mentioned the media’s influence on them and whether the individual was ready to die.
Should not give her an overdose	Of those who responded in the 3 category, 38% of the comments mentioned a belief in God’s word affected their views of life/death and His teachings, 32% mentioned that the doctor would be killing, 24% mentioned a peaceful endurance of pain and the benefits of such, 20% mentioned the sacredness of life and that it should be fought for, 16% mentioned that a family or close associate went through a similar experience that this influenced their decision, 14% mentioned the doctor’s liability, 8% mentioned the laws of the land, 8% mentioned the avoidance of suffering, 7% mentioned a definition of the rights of the individual and the loved ones affected, 5% mentioned that the woman is selfish for asking the doctor to make the decision, 5% mentioned work experience, 4% mentioned that God gave man agency or the ability to chose for themselves 3% mentioned that the woman is not in the right frame of mind to make this decision, and 2% deferred to their parent’s teachings.

Table G-4

Description of Coded Comments in DIT-2 Famine Story Qualitative Comments

Code	Collapsed Categories	General Descriptions
J	GG R TS	Justified. Life is greater than law. Greater good. Life was more precious than stealing. There was a greater good worth breaking laws to achieve. Everyone has a right to life and the right to protect that life, even if it means stealing. If stealing is done as last resource, only to survive, than it is okay.
PF	Lv	Provide for family. Comments generally agreed that if the motive was purely to provide for your family than stealing would be justified. Also, the idea that fathers have a duty to provide for their families and that commandment/duty is weightier than that of not stealing. If love is the intent, love is the motive than it is okay.
TA		It's okay to take from someone who has abundance.
L	DS	Law and obedience to laws is paramount. This most often was in reference to heavenly laws or commandments to not steal ("thou shalt not steal"), etc. Comments focused on how laws must be respected in order to maintain order in society. Laws are not to be broken--ever. Persons have a responsibility to uphold the public laws. Comments were very black and white, no one should steal (don't steal) no matter what!
CTR		Choose the right and God will bless you. Things will work out for your good--somehow God will support you and if he doesn't than you will be blessed in heaven for keeping His commandments. These comments tended to lean toward orthodox mentalities.
Wk		Work for it rather than steal. This code was also used if comments suggested that there could be another way for obtaining the food besides stealing. Look for more options.

Table G-5

Relative Percent Counts of Coded Comments by DIT-2 Famine Story Decision

Should steal <i>n</i> = 143			Can't Decide <i>n</i> = 59			Should not steal <i>n</i> = 74		
200 coded responses	296 total cells	Relative %	99 coded responses	236 total cells	Relative %	267 coded responses	573 total cells	Relative %
47	J	23.5%	29	L	29.3%	113	L	42.3%
46	PF	23.0%	21	J	21.2%	69	CT R	25.8%
20	TA	10.0%	9	WK	9.1%	32	WK	12.0%
10	L	5.0%	8	PF	8.1%	4	PF	1.5%
3	CTR	1.5%	7	TA	7.1%	2	TA	0.7%
0	WK	0	6	CTR	6.1%	0	J	0

Table 6

Reporter Story Decision Coded Comments and Category Examples/Descriptions

Code	Collapsed Categories	General Descriptions
PRI	Life (P), Doctrine (R, P, 2nd, A), Philosophies (I, R, 2nd, P, NoP), Other	Repentance, forgiveness, forsaken, length of time shows a change, changed, prophet's teachings or examples about repentance, the past is irrelevant because of the present (he has changed for the better), everyone deserves a second chance (chance to change), grace, atonement, God forgets so we should too, No mortal is perfect/all need repentance, crime is not against me so who am I not to forgive.
T	Life (T), Doctrine (T, H, D), Philosophies (T, H)	Truth exposed, truth will set you free/help you be free, it is right, honesty/dishonest, dishonest to not to report.
PD	PD, F	Public duty to report, report unbiasedly, Fairness, as long as she reports good and bad on both candidates, consistency with reporting on candidates, take a less offensive stand or be neutral in your reporting
B	B, C, Bene, Benefit, O, +	It will benefit him, chance for him to correct himself to public for a crime against public, optimism-it will work out, this is a positive thing and can be reported in a positive way, it's good that a supporter rather than an enemy of his is doing this.
JE	Life (J)	I have studied Journalism's code of ethics, I have educational experiences or have discussed this in my classes at school, I am a reporter.
TV	TV, S, O	That's the way it is with media, when you are in the spotlight this comes with it, the community does decide on a public figure's morality, opinions of public figures are made by the people after they have information on them.
N	N	Negative impression of the press
NB	NB, PR, Res, E	Not the public's business, he has a right to privacy, journalists should respect their subjects, this is an ecclesiastical issue not a public issue, this is between him and the store owner.

Table G-7

Relative Percent Counts of Coded Comments by DIT-2 Reporter Story Decision and Category Examples/Descriptions

Should Report			Can't Decide			Should Not Report		
47 comment cells	n=29 69 na cells	40.5% comment	42 comment cells	n=29 74 na cells	36.2% comment	425 comment cells	n=211 419 na cells	50.4% comment
Amount	Category	Relative % totals	Amount	Category	Relative % totals	Amount	Category	Relative % totals
18	T	38.3%	33	PRI	78.6%	362	PRI	85.2%
15	PRI	31.9%	11	T	26.2%	32	N	7.5%
13	PD	27.7%	9	PD	21.4%	31	NB	7.3%
11	TV	23.4%	4	JE	9.5%	18	T	4.2%
5	JE	10.6%	4	TV	9.5%	13	B	3.1%
4	B	8.5%	2	N	4.8%	12	TV	2.8%
0	N	0	1	B	2.4%	11	PD	2.6%
0	NB	0	1	NB	2.4%	5	JE	1.2%

Table G-8

Coded Comments from Story #3 with Category Examples/Descriptions

Code	Collapsed categories	Descriptions
L	Leadership (L) & Results (R)	Elected leaders should decide, people elect leaders for a purpose, the concept of a republic, There will be no results if everyone is fighting, this meeting does more harm than good anyway, this meeting is inefficient use of time.
C	Contention (C) & Take out (T)	Contention of people, Leader needs to deal with the few contentious, I have dealt with similar angry people, anger is wrong, contention of individuals is wrong.
F	Fair/Free	Fair to community, should give people a say if it effects them, rights, free speech, in the end this is better, if he is unfair he will lose their trust, the public has a right to know, democracy.
S	Safety	Safety should come first, the well-being of the members involved, the public's safety, leaders are liable for the safety of those they lead.
P	Promises	Promises should be kept, he should stick to his word, honesty.
M	Moderator	The moderator and the way they are dealing with this mob mentality is the problem, the group dynamics are such that there is no responsibility in large #s, the problem is that this moderator does not know how to compromise, there is no order.

Table G-9

*Relative Percent of Coded Comments by Story #3 Decision and Category**Examples/Descriptions*

Should call off next open meeting		Can't decide		Should have the next open meeting	
n=102 Category	Relative % totals	n=47 Category	Relative % totals	n=114 Category	Relative % totals
L	32.4%	L	37.8%	F	43.3%
C	32.4%	F	35.1%	P	20.0%
F	19.9%	S	27.0%	M	16.7%
S	16.2%	M	24.3%	C	15.0%
P	6.6%	P	24.3%	L	12.8%
M	5.1%	C	21.6%	S	0

Table G-10

Illustrative Examples of Themes in the Cancer Story Comments

Rights-based macromoral reasoning themes	Responsibility-based macromoral reasoning themes
<p><i>Theme 1. An individual and a society has a right or duty to avoid suffering when possible. Illustrative examples:</i></p> <p>The patient has every right to declare the comfort of her death before it actually happens in practical means.</p> <p>No one should have to suffer more than they can handle.</p> <p>If we can help others avoid suffering we should do that.</p> <p>Life should be relatively sting free.</p> <p>Life is precious but also the ability to enjoy life is so important. If you want to live then you should live. But if you are in pain and everything else has been taken from you, shouldn't you have the right to decide when and how you die?</p> <p><i>Theme 2. Individual rights trump the rights of those affected. Illustrative examples:</i></p> <p>My grandpa was sick and suffered for a couple years until finally passing away. I don't know if he ever considered pulling the plug on himself, but as a family we were strictly against it. All I can say is that it is a very difficult decision, but the individual doing the suffering, if anyone, is probably the only one who could make such a life or death call.</p> <p><i>Theme 3. The right to make their own choice. Illustrative examples:</i></p> <p>In the end, this is her life to choose what she wants to do with it.</p> <p>God has given us agency. Agency is so important to Heavenly Father that...although the doctor may not agree with the lady's wishes, he does not have authority to take away her agency.</p>	<p><i>Theme 1. Individual's responsibility to endure suffering for their own or others benefit. Illustrative examples:</i></p> <p>My friend's father had a brain tumor that caused him pain and suffering, however, he suffered with it until the end so he could spend more time with her.</p> <p>Some suffering can serve as a motivating factor for overall good.</p> <p>Both of my parents had cancer so the idea of a loved one in pain is very sensitive to me. However, the idea of rushing the death of a loved one is nearly unbearable. I could never put a doctor in those circumstances because there are more people involved than the victim.</p> <p>If you are still alive, there is a reason behind it. Discover your purpose and make the most of the remaining time on earth. Maybe you need to befriend someone, make amends, or have some other experiences before you are called to depart this life.</p> <p>I believe Victor Frankl was right when he said that if there is any meaning in life there must be meaning in suffering.</p> <p><i>Theme 2. Responsibility for individuals to fight for their life. Illustrative examples:</i></p> <p>Life is precious. Although suffering is hard and leaving this earth seems wonderful compared to the pain-our mortal life is a gift, we must always treat it like one.</p> <p>My older sister fought leukemia for 7 months and then passed away. Sometimes it was so hard for her to get through the days. I know that I would not have wanted her to give up and just settle for "comfort" instead of fighting for her life. And that's what she did.</p> <p><i>Theme 3. The responsibility that the doctors have to society—to heal life, not take it. Illustrative examples:</i></p> <p>Doctors are meant to help and sustain life, not end it faster than should naturally happen.</p> <p>I think doctors should continue to fight for their patients.</p> <p>The situation above is tough one, however, my thoughts on it is that we shouldn't help others destroy their life.</p> <p>Doctors take an oath to help and protect those they care for.</p>

(Table continues)

Rights-based macromoral reasoning themes	Responsibility-based macromoral reasoning themes
	<p><i>Theme 4. A responsibility or accountability for one's actions when they involve another person. Illustrative summaries:</i></p> <p>The doctor is accountable to: society for not considering what is best for more than the one individual, to God for taking life, to the living loved ones of the family, to the laws of the land (least frequent comment). The woman is accountable because the doctor is accountable to internal and external consequences for her wish (e.g., guilt and possible allegations/prosecutions).</p>

Table G-11

Illustrative Examples from the Reporter Story Comments

Rights-based comments	Responsibility-based comments
<p><i>Theme 1. The public has a right to know.</i></p> <p>The reason that I approved of the reporters right to go ahead and publish the story is because I have seen time and time in my life that even no matter how small the issue, the public must know the facts. If not, then we all just become subject to the one-sidedness of some man's opinion.</p> <p>It is important for people to know the truth and to be able to make informed decisions based off of the knowledge that they have. The WHOLE truth is that Thompson went through some hard times in his life, but he overcame them. If the whole truth is presented it shouldn't affect his chance of losing if he was the most likely candidate.</p> <p>In the recent presidential primaries that [are] occurring I believe that the public has the right to know everything about the future candidates, even their past lives. We should just know that they have changed.</p> <p>Her job is to report the news to the public.... She gives the community the information/story and they decide for themselves.</p>	<p><i>Theme 1. Responsibility to subject.</i></p> <p>I learned one time of President Truman falling down one time and a young reporter tried taking a picture of him and another reporter hit his camera out of his hand and broke it because of the ideal that reporters then had of respecting their subjects.</p> <p>It would not be dishonest for Molly to write the story, but it would be mean and disrespectful.</p> <p>We should treat our neighbors kindly and not expose bad things about people.</p> <p><i>Theme 2. Responsibility to forgive and move on.</i></p> <p>Once people have repented of their sins, they are forgiven and God remembers them [their sins] no more. We should also forgive.</p> <p>You can't spend your life looking in the rear view mirror. There is a reason that it is smaller than the windshield. It is more important to focus on what lies ahead.</p> <p><i>Theme 3. Ends don't justify the means.</i></p> <p>Don't hurt someone for your own gain.</p>

VITA

DANIEL R. WINDER

Address

Church Educational System
Seminaries & Institutes
Office of Research, Evaluation, and Assessment

983 COB
SLC, Utah 84150
801.240.5386
winderdr@ldschurch.org

Education

PhD Instructional Technology, Utah State University, Logan, Utah,
2002-2009

- Emphasis: Quantitative Assessment and Moral Education
- Dissertation: “Assessment of Moral Reasoning of Brigham Young University Religion Students”
- Committee: Nick Eastmond (co-chair), Matthew Taylor (co-chair), Joanne Bentley, Mike Freeman, Richard West.
- Post-graduate certifications from Texas A&M in Item-Response Theory & Factor analysis. Additionally, have taken online courses in Rasch Analysis from Mike Linacre.

MA Instructional Technology, Utah State University, Logan, Utah, 2002

- Master’s Project: An Interactive Web-based Multimedia Journey through the Plan of Salvation for Latter-day Saint Church Educational System High School Students

BA Brigham Young University, Provo, Utah, 1999

- Highlights: participated in Fall 1994 Brigham Young University Jerusalem Study Abroad program and studied Biblical Hebrew.

Research Experience

- 2008. Founder of the Church-Sponsored Educational Researcher’s Meeting (CSERP). Organized MTC, BYU Research Methodology Instructors, CES, Church Correlation Department, and BYUIAD researcher’s annual meeting. Last years speakers included: Richard Sudweeks, Cyril Figuerres, and Elder Paul Johnson.
- 2007-2008. Developed, designed, & analyzed the Seminaries and Institutes Basic Doctrines Test and Teaching Emphasis Scales of Measurement using

classical test theory and Rasch Analysis. Other projects: Assessment of special needs classrooms best practices, developing consistent portfolio rubric for church schools in the pacific, psychometric investigation and instrument development for pre-service raters.

- Currently using item-response theory graded-response model for polytomous items to measure the accuracy of the Defining Issues Test-2 as a test for the measurement of moral reasoning. Using Winsteps and SPSS software for this analysis.
- Currently studying how effective concept mapping is as an assessment of understanding doctrinal issues. Using Generalizability theory and Excel.
- Served on Brigham Young University Religion Department's Learning Outcomes Assessment Committee—New Testament Item Writing Committee. Analyzed data collected from beta and initial Religion department learning outcomes and recommended items that need improvement based upon classical test theory and item response theory. January 2007-present.
- Used Generalizability Theory statistical analysis to analyze 5 religion professor's scores on twelve student's 5-paragraph essays graded on two different occasions. Used SPSS for this analysis. Concluded that the vague grading rubric used and occasion the paper is graded on contributed to most of the variance in Brigham Young University religion professor's scores. January 2007.
- Developed Missionary Training Center Doctrinal Competency Exam—wrote over 500 items to measure LDS doctrinal understanding and consulted with psychometricians, Dr. Richard Sudweeks, and the MTC research department for how to test the items for effectiveness. June-August 2006.
- Assessment of David Wang's Interactive Learning Piano for Summerhay's Music, Orem, Utah. Beta tested and used piano pedagogy experts to make recommendations for improvement for an interactive learning system for piano instruction. January-May 2005
- Assessment of Brigham Young University Religion Student's use of The Virtual Historian Doctrine and Covenants Edition—assessed how Brigham Young University students use this instructional CD-rom in Dr. Wright's classes. August-December 2004.
- Essay grading for Dr. Richard Galbraith's Religion Classes. 2003-2006.

Teaching Experience

- Church Educational System visiting instructor in Brigham Young University's Church History and Doctrine Department. Taught: Religion C 324 Doctrine & Covenants, Sections 1-76 & Religion C 333 Teachings of the Living Prophets. June 2006-present.
- Church Educational System Instructor. Highland, American Fork, & Lehi, Utah 1998-2006
- Missionary Training Center Instructor, Provo, Utah. 1997-1998
- Full-time Missionary for The Church of Jesus Christ of Latter-day Saints. Boston, Massachusetts. 1995-1997

Professional Experience

Curriculum Development

- Developed Religion C 333 Selected Sermons of Living Apostles Packet. July 2007.
- CES New Testament Visual Curriculum Committee—submitted instructional video ideas for CES Visual Curriculum and created PDF's for online web content. June 2006.

Multimedia

- Compressed PPT and Web version of Church Educational System Doctrine & Covenants videos for Rel 324-325 as well a blackboard course that integrates these videos during out of classroom instruction time. Also developed various multimedia for Rel 333 course. June-August 2006.
- Virtual Revelation—analyzed, designed, & developed an interactive 3D Journey through John's Throne Room (Revelation 4) based upon Richard Draper's scholarly interpretations and developed an interactive 3D journey through Lehi's Dream (1 Nephi 8). Several CES instructors use these free instructional products today. August 2005-May 2006
- Developed web-sites for Utah Valley Photo, Loopy Laws, Inc., Val Vista HOA. 2003.
- Analyzed, developed, & designed an interactive journey through the plan of salvation using Macromedia Flash Software and Pinnacle Video Systems video

editing software.

- Utah Valley Photo—Professional digital photography family-run business that uses Adobe products for professional photography, photo enhancement, and photo restoration. 2002-present.

Editing

- Edited Journal of Discourses volumes for www.scriptures.byu.edu project.

Publications

“Wilford Woodruff’s Prophet from Simsbury: The Life of Robert Mason.” Forthcoming, 2009.

“Macromorality for Mormons: A Psychometric Investigation of the Defining Issues Test-2.” Forthcoming, 2009.

“The Smith Family Legacy of Prayer” In *Selections from the Religious Education Student Symposium 2007*, Religious Studies Center, Brigham Young University, Provo, Utah: 2007.

“The Smith Family Legacy of Prayer.” In *David O. McKay Essay Content 2007*, Brigham Young University, Provo, Utah: 2007.

“Incorporating Character Education into an Engineering Department.” Chapter in *The Gospel: The Foundation for a Professional Career—Moral Character in Professional Life*, Deseret Book, Salt Lake City, Utah: 2008.

“Is Educational Research Generalizable? The Need for No Child Left Behind Research Methodology Mandates.” Under Review.

Presentations

“Incorporating Character Education into an Engineering Department.” Presentation at *The Gospel: The Foundation for a Professional Career—Moral Character in Professional Life*, Brigham Young University, Provo, Utah. March 2007.

“The Smith Family Legacy of Prayer.” Paper presented at the annual *Religious Education Student Symposium*, Brigham Young University, Provo, Utah. February 2007.

Awards and Grants

Graduate Student Paper Award, Ninth Annual Brigham Young University Religious Education Student Symposium, \$300 award.

Graduate Student Paper Award, David O. McKay Essay Contest. Third Place in Graduate Division, \$700 award.

\$686 grant to study the measurement of moral reasoning among BYU religion students from the Religious Studies Center.

\$500 grant to study Robert Mason's life from the Department of Church History and Doctrine.

Professional Organizations

2009. AERA & NCME.

2008-present. AERA Ethics Committee Member.