Can ethics be taught? A quasi-experimental study of the impact of class size on the cognitive moral reasoning of freshmen business students

Author: Ethan Sullivan

Persistent link: http://hdl.handle.net/2345/1840

This work is posted on eScholarship@BC, Boston College University Libraries.

Boston College Electronic Thesis or Dissertation, 2011

Copyright is held by the author, with all rights reserved, unless otherwise noted.
CAN ETHICS BE TAUGHT?
A QUASI-EXPERIMENTAL STUDY OF THE IMPACT OF CLASS SIZE ON THE
COGNITIVE MORAL REASONING OF FRESHMEN BUSINESS STUDENTS

Dissertation

by

ETHAN A. SULLIVAN

submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

May 2011
DISSERTATION ABSTRACT

CAN ETHICS BE TAUGHT? A QUASI-EXPERIMENTAL STUDY OF THE IMPACT OF CLASS SIZE ON THE COGNITIVE MORAL REASONING OF FRESHMEN BUSINESS STUDENTS

Ethan A. Sullivan
Boston College Lynch School of Education
Dissertation Chair: Dr. Karen Arnold

The purpose of this study was to determine the impact of a business ethics course on the cognitive moral reasoning of freshmen business students. The sample consisted of 268 college students enrolled in a required business ethics course. The students took Rest’s Defining Issues Test – Version 2 (DIT2) as a pre-test and then post-test (upon completing the course). Descriptive analyses, t-tests, ANOVA, and multiple regression were employed to compare the pre-test and post-test scores of the students and to determine the relationship, if any, between the variables of gender, class size, instructor, class time of day, SAT scores, and students’ GPAs; and the dependent variable of moral reasoning (N2) scores.

Descriptive analyses showed that taking this kind of a course made a difference for virtually everyone. While women had higher pre-test scores, post-test scores, and overall gains in moral reasoning than men, men also had gains across the board. Gains were also found across all instructors, in both small and medium class sizes and
regardless of what time the class occurred. Further, compared to the national sample of college students, the study sample was at the level of college seniors.

Results of ANOVA testing showed that the gains in moral reasoning scores were statistically significant ones. However, the moral reasoning gains of students enrolled in the smaller sections (n = 19) were not statistically significantly different than students enrolled in medium sized (n = 27) sections of the same course. Finally, the independent variable of class time had the most statistically significant relationship with gains in moral reasoning scores.

The findings of this study suggest several practices for institutions of higher education. First, certain curricular conditions should be considered by institutions with ethics courses. The content should be explicit and pedagogical strategies should include role-taking, the discussion of moral dilemmas, reflection, active learning, and cognitive disequilibrium. Second, increasing class sizes by eight students can help to maintain moral growth while still being financially efficient. Third, these findings may inform administrators when planning class times (the earlier, the better).
ACKNOWLEDGEMENTS

First, I have been inspired and mentored by Dr. Karen Arnold these past nine years and I have been grateful for her support, guidance and work as the chair of my committee. We first connected on a shared interest in nature and hiking (maybe one day I too will hike the AT), and continued to discuss academic and personal interests throughout the dissertation work. I cannot imagine a better dissertation director, and will always be grateful for her help along the way. Dr. Jessica Greene offered her expertise in research design and analysis. I have been eternally thankful for her thoughtful reading of several drafts, guidance in the analysis process, and support in the final stages of writing. Jess, you’ve been my part-time life coach, part-time stats consultant, and full-time friend. Slushies on me this summer! Dr. Richard Spinello offered his knowledge the history and development of business ethics. Business ethics has been a relatively new field in higher education, and Rick has been able to guide me through its history and purpose. He also has been a calming presence in an otherwise chaotic time of my life!

I have also been grateful to Dean Andy Boynton and Dean Richard Keeley at Boston College for their willingness to allow me to combine professional and academic interests, and their continual support of my work. You both have created
a professional environment in which I have been able to thrive. Thank you both for your vision, leadership, and encouragement.

Sarah Cabral, Joe Cioni, and Amy LaCombe have been my friends and colleagues throughout this final push. Your willingness to let me into your classrooms, and your help in collecting data were instrumental in this study. Your exemplary work is what inspired this topic, and I am not surprised at the great difference you make in the lives of your students. More importantly, your kindness, support, and camaraderie have kept me balanced. Joe, your turn now!

My friends, family, and colleagues – too many to mention by name, but too important to not mention at all – I regret that I haven’t had more time for you these past few years. I really was working on the dissertation all of those times you thought I was blowing you off. May I now have more time to re-connect with what is truly one of life’s most important and precious gifts – friendship.

Finally, my immediate family has been most instrumental in this achievement. It does take a village to finish a dissertation. My parents, Tom and Elaine Sullivan have made so many sacrifices for all of their children. You have always supported me and instilled in me the value of education and being a well rounded successful person. My siblings, Brett Sullivan and Claire Zukauskas have provided unconditional love to me and my own family. It is comforting to know
you are always there for me. My grandparents – Frank and Jane Sullivan and
Gerald and Elizabeth Kunz (who continues to amaze me every day!) – your moral
leadership has served as an inspiration to our entire clan. My in-laws – Maryann
and Gene Kiely – your prayers worked! Thanks for your ever-present care and
know it wasn’t easy that dad came home late so many nights during the past few
years. Your patience, exuberance, and love always kept me focused and full of joy,
even during my own times of self-doubt. And, most of all, to Jen: I’ll take the P.,
you take the H. and we’ll split the D. You deserve this as much as I do. You have
sacrificed so much to allow me the time and space to finish this task. You
motivated me when times were tough. You have inspired me with your love and
ability to be such a wonderful wife and friend to me. You are my best friend and
my true love. We did it!
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Acknowledgements</th>
<th>iv</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xiii</td>
</tr>
</tbody>
</table>

**Chapter 1. Introduction**  
Focus of the Study  
Research Questions  
Theoretical Framework  
Significance of the Study  
Design of the Study  
Limitation of the Study  
Definition of Terms  
Overview of the Study  

<table>
<thead>
<tr>
<th>Chapter 2. Review of the Literature</th>
<th>22</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moral Education in American Higher Education</td>
<td>23</td>
</tr>
<tr>
<td>Theories of Moral Development</td>
<td>26</td>
</tr>
<tr>
<td>Kohlberg</td>
<td>26</td>
</tr>
<tr>
<td>a. Critiques of Kohlberg</td>
<td>31</td>
</tr>
<tr>
<td>b. Neo Kohlbergian Theories</td>
<td>33</td>
</tr>
<tr>
<td>c. James Rest and the Defining Issues Test</td>
<td>35</td>
</tr>
<tr>
<td>d. The Four Phases of the DIT</td>
<td>38</td>
</tr>
<tr>
<td>e. Empirical Findings of the DIT</td>
<td>43</td>
</tr>
<tr>
<td>f. Student Characteristics</td>
<td>46</td>
</tr>
<tr>
<td>g. Curricular Interventions</td>
<td>50</td>
</tr>
<tr>
<td>Business Ethics</td>
<td>50</td>
</tr>
<tr>
<td>a. Impediments to the Study of Business Ethics</td>
<td>56</td>
</tr>
<tr>
<td>b. Business Ethics Pedagogy – Best Practices</td>
<td>62</td>
</tr>
<tr>
<td>Conditions for Moral Development (Class Size)</td>
<td>64</td>
</tr>
<tr>
<td>Conclusion</td>
<td>67</td>
</tr>
</tbody>
</table>
Chapter 3. Research Methodology
   Research Questions 69
   Research Design 70
   Treatment 72
   Sample 74
   Instrument 77
   Scoring the Instrument: The P Score 81
   Scoring the Instrument: The N2 Score 83
   Validity 85
   Variables 87
   Data Analysis Procedures 90
   Limitations 91

Chapter 4. Presentations of the Findings 95
   Section One: Population, Sample, and Variables 95
   Descriptive Analysis of the Participants 100
   N2 Scores 101
      a. Norms for DIT N2 Scores 102
      b. N2 Scores by Gender 104
      c. N2 Scores by Instructor 105
      d. N2 Scores by Class Size 107
      e. N2 Scores by Class Time 109
      f. N2 Scores by SAT and GPA 110
   Section Three: An Analysis of the Research Questions 113
   Findings for Research Question One 114
   Findings for Research Question Two 117
   Findings for Research Question Three 119
   Summary of Findings 127

Chapter 5. Discussion of the Findings 129
   Summary of the Findings 129
   Interpretation of the Findings 132
   Comparison between Portico Students and the National Sample 133
   Class Size 136
   Class Time 138
   Implications for Policy and Practice 141
List of Tables

Table 1.1 Kohlberg’s Six Stages of Moral Reasoning ............................................. 10
Table 2.1 Kohlberg’s Six Stages of Moral Reasoning ............................................. 29
Table 2.2 DIT P Score Means and Standard Deviations ........................................ 44
Table 2.3 DIT P Scores for Low Education and High Education Groups .......... 45
Table 2.4 Number of Business Programs in the U.S. ............................................. 53
Table 2.5 Number of Bachelor Degrees Conferred in Business from 1921 to 2007 .......................................................... 54
Table 2.6 A Comparison of Estimated Costs Between Small and Medium Sized Sections of a Course ........................................ 68
Table 3.1 Student Gender, Academic Aptitude and First Year Grades .............. 76
Table 3.2 Number of Students by Instructor and Class Type .............................. 77
Table 3.3 Weights for Calculating the P Score of the DIT2 ................................. 82
Table 3.4 A List of Variables by Name and Type ................................................... 88
Table 4.1 Student gender, academic aptitude, and grades, by group .......... 98
Table 4.2 A list of variables by name and type, with a brief description .......... 99
Table 4.3 Gender, SAT scores, and Freshmen GPA of Portico students, by class size ................................................................. 100
Table 4.4 Gender, SAT scores, and Freshmen GPA of Portico students, by instructor ................................................................. 101
Table 4.4a  DIT2 Means and Standard Deviations for N2 Scores by Educational Level .............................................................................. 103

Table 4.5  N2 scores for Portico students, compared to the national sample .. 103

Table 4.6  N2 Score Means and Standard Deviations for Portico students, by gender ................................................................................. 105

Table 4.7  N2 Score Means and Standard Deviations for Portico students, by instructor ............................................................................ 107

Table 4.8  N2 Score Means and Standard Deviations for Portico students, by class size ............................................................................. 108

Table 4.9  N2 Score Means and Standard Deviations for Portico Students, by class time ............................................................................ 110

Table 4.10  N2 Score Means and Standard Deviations for Portico students, by SAT terciles ........................................................................ 111

Table 4.11  N2 Score Means and Standard Deviations for Portico Students, by GPA terciles ................................................................. 111

Table 4.12  Paired Sample T Test of the Difference in N2 scores from Pre-test to Post-test .............................................................................. 117

Table 4.13  Levene's Test of Equality of Error Variances .............................. 118

Table 4.14  Two-way ANOVA testing of N2 difference scores by instructor, class size, and their interaction effect ................................................ 119

Table 4.15  Rank Biserial Correlation Coefficients of Difference in N2 Score compared to SATs, GPAs, Class Size, and Class Time ..... 120
Table 4.16  Test of Homogeneity of Variances for ANOVA of N2 Difference score by Class Time ............................................................. 121

Table 4.17  One-way ANOVA of N2 Difference score by Class Time ............... 122

Table 4.18  Post-hoc Scheffe Testing of the Multiple Comparisons of N2 Difference Scores, by Class Times ............................................. 123

Table 4.19  T test of N2 Difference Scores, by “Morning” and “Afternoon” Class Times .............................................................................................. 125

Table 4.20  Levene’s Test of Equality of Error Variances ............................... 125

Table 4.21  Two-way ANOVA of N2 Difference Scores by Class Time, Instructor, and Their Interaction ........................................................... 127

Table 4.22  The Null Hypotheses of This Study and Their Findings ............... 128
List of Figures

Figure 4.1  N2 Scores by Instructor ................................................................. 106
Figure 4.2  N2 Scores by Class Size ................................................................. 109
Figure 4.3  Normal Q-Q Plot of N2 Pre-test Scores ........................................ 115
Figure 4.4  Normal Q-Q Plot of N2 Post-test Scores ....................................... 116
Figure 4.5  Normal Q-Q Plot of the Difference in N2 Scores ......................... 116
Figure 4.6  Estimated Marginal Means of N2 Scores by Class Time ............... 124
CHAPTER ONE

INTRODUCTION

Ethical crises from the business arena have frequently contributed to the news of the past twenty five years. These scandals have been noteworthy due to their extensive and measurably damaging impact upon society (Davis, Ruhe, Lee, & Rajadhyaksha, 2010). For example, the savings and loan crisis cost U.S. taxpayers $125 billion (Curry & Shibut, 2000). Scandals at Enron, Arthur Anderson, WorldCom, and Tyco resulted in destructive losses of jobs and retirement funds (Wilhelm, 2005). The most recent global financial crisis is estimated to cost trillions of dollars in the United States alone (Bestani, 2009). Much within these scandals can be traced to the greed, dishonesty, and overall lack of “ethical decision making by executives involved in these debacles, many of whom were graduates of top-tier U.S. business schools” (Wilhelm, 2005, p. 202).

Business is the most popular undergraduate field of study in the United States. According to the National Center for Educational Statistics, there were 327,531 undergraduate degrees in business conferred in 2006-2007. This represents 21.5 percent of all undergraduate degrees, or more than one in five undergraduates. This number is close to double that of the social sciences and
history, the next largest field, which conferred 10.8 percent of all degrees in 2006-2007. The number of business majors rose from 225,934 degrees conferred in 1996-1997, a one decade increase of 45 percent. Two decades prior, business was the second most common degree conferred, following education. In 1976-1977, there were 115,396 business degrees conferred, which comprised 15 percent of the total degrees that year.

This steady and unslowing rise in the popularity of business throughout the last thirty years corresponds with research on student motivations and values. Using descriptive data from the Cooperative Institutional Research Program (CIRP) of the Higher Education Research Institute at UCLA, Astin (1999) found a number of conflicting trends. In 1966, being “well off financially” was “essential” or “very important” for 45 percent of surveyed undergraduate students (ranking sixth on the list). “Developing a meaningful philosophy of life” was first on the list with 80 percent of students indicating it was “essential” or “very important.” In 1996, the two traded positions. Being well off financially (74 percent) became the top value while developing a meaningful philosophy of life was sixth (42 percent). The number of students who believed that “the chief benefit of a college education is to increase one’s earning power” rose from 54 percent in 1969 to 71 percent in 1996. Likewise, one half of students in 1971 stated
that they attended college “to be able to make more money.” In 1996, almost three quarters of students indicated this to be a reason for college attendance.

Astin’s findings (1994) have been related to a rise in neoliberalism of college students (Saunders, 2007). Neoliberalism, “a socio-economic theory that...promulgates materialism, consumerism, and the commodification of many public goods” (Saunders, 2007, p. 1) is often associated with business students, and is linked to greed. And, while there is still some debate as to the cause of the most recent financial crisis, most experts cite the greed and dishonesty of Wall Street executives as the primary motive. Most of the Wall Street bankers were graduates of American business schools which encourage students to emphasize profits over ethics in order to maximize shareholders’ wealth (Davis, Ruhe, Lee, & Rajadhyaksha, 2010; Griffin, Putman, Moser, & Kilgore, 2005; Jennings, 2004).

Merrit (2003) stressed that fixing this problem must occur “where careers begin – with management education” (p. 105). Critics and scholars have argued that the curriculum of business schools has added to corporate scandal, but has done little to address the issues (Bennis & O’Toole, 2005; Ghoshal, 2005; Mitroff, 2004). Mitroff (2004) blames business schools and their faculty who are “guilty of having provided an environment where the Enrons and Andersens (sic) of the world could take root and flourish...we delude ourselves seriously if we think
we played no part whatsoever“ (p. 185). Ghoshal (2005) agrees, offering the assessment that business schools have pushed students away from acting morally responsible. By presenting amoral ideologies, business schools have allowed its students (and future business leaders) to rationalize their thoughts and actions. This has created a bifurcated existence where business students and leaders have one set of values for their family, social, spiritual, and civic lives and a separate set of values for their professional lives.

These scholars state that business educators need to acknowledge their responsibility and assume blame for not doing a better job with the values of these unethical corporate leaders (Ghoshal, 2005; Mitroff, 2004). The popular press has joined in their call for more ethical training in business schools (for example, Alboher, 2008; Baer, 2009; Brown, 2009; Danko, 2009; Gentile, 2009; Holland, 2009; Jacobs, 2009; Klein, 2008; Krehmeyer, 2007; Lavelle, 2006; Navarro, 2008; Schmidt, 2008; Sternberg, 2009). This study explores one attempt to correct the situation.

**Focus of the Study**

This study investigated the issue of whether business ethics can be taught, and assessed the optimal conditions for increases in the moral development of freshmen business students enrolled in a business ethics course. It addresses the
question directly, by taking advantage of a natural experiment in which freshmen were randomly distributed across small and medium-sized sections of a business ethics course taught by the same instructors. The study involved a pre- and post-test to measure growth in principled moral reasoning across the first semester of college among first-year business students who were enrolled in small and medium sized versions of the same business ethics class. The instrument used to measure this growth was the Defining Issues Test (DIT2).

The Defining Issues Test (DIT) is the most widely used instrument within research on moral development (Thoma, 2002; Trevino, 1992). Since the 1974 inception of the DIT (Rest, 1974) there have been over 500 studies utilizing the DIT (Thoma, 2002; King & Mayhew, 2002) which have included over 500,000 participants (Rest, Bebeau, Narvaez, & Thoma 1999b; Walker, 2002).

King and Mayhew (2002), in a review of post-1980 research on the moral judgment of college students, found 172 studies that used the DIT with U.S. college and university undergraduates. A subsequent review winnowed the list to 157 studies (King & Mayhew, 2004). This review researched the relationship between students and collegiate factors that affected moral reasoning. Factors included student characteristics (age, ethnicity, gender, SES, aptitude/intelligence, political identification, religious identification, and other
characteristics) and collegiate contexts (institutional type, academic disciplines, curricular experiences, and co-curricular experiences). Not surprisingly, several studies have investigated the effect of an ethics course on the moral judgment of college students (Armstrong, 1993; Bonowitz, 2002; Boss, 1994; Ponemon, 1993). Results have often largely demonstrated significantly higher DIT scores for students who participated in an ethics course or intervention. However, “studies that more carefully examine the conditions for growth are needed” (King & Mayhew, 2004). One potentially important condition for growth is class size, and whether small classes are better in increasing the moral judgment of college students (King & Mayhew, 2004).

Class size is the oldest and most widely studied topic within the research on teaching and learning (McKeachie, 1990). However, despite being the earliest research question of college teaching, studies on class size have been inconclusive and unclear (Follman, 2009). The vast majority of research on class size is at the elementary school level, with secondary school studies next, and college studies lacking and inconsistent. While college teachers and students have consistently favored small classes (McKeachie, 1980), most studies with the criterion of achievement tests demonstrate no advantage of small classes over large ones (Laughlin, 1976; McKeachie, 1963; Vincent, 1969; William et al., 1985).
This combination of “unclear” (Glass & Smith, 1979, p. 2) and inconclusive research demonstrates the need for much more work in this area.

Class size is a hotly debated topic. As the costs of higher education continue to rise, administrators seek ways to reduce them. Increasing class size is often considered as a way to create economies of scale (Hancock, 1996; Kokkenlenberg, Dillon, & Christy, 2008). “But putting more students into each class goes against strong conventional wisdom that class size affects learning quality” (Hancock, 1996), and often is resisted by faculty as a prudent option (Kokkenlenberg, Dillon, & Christy, 2005). This topic will investigate the debate of class size, especially with respect to how it affects moral reasoning of students.

Given the need for ethics in the business world, the dearth of research on curricular conditions for effective ethics instruction, the importance of fiscally responsible practices, and an incomplete understanding of class size effect; this study will attempt to measure the effect of class size on the cognitive moral development of first year college students enrolled in a business ethics course. Do larger class sizes outweigh the possible future costs due to ethical lapses of students who are tomorrow’s business leaders? Do smaller classes sufficiently increase moral judgment, despite the university’s costs in administering smaller classes? These are the questions push this study.
**Research Questions**

To understand the impact of class size on the moral reasoning of freshmen business ethics students, this study investigates the following questions:

1. Do students enrolled in a one-semester business ethics class have significant gains in moral reasoning, measured over time from the pretest to the posttest?

2. Do students enrolled in smaller sections of a business ethics course have greater gains in moral reasoning scores than students enrolled in larger sections of the same course, taught by the same instructor?

3. What is the relative effect of class size in cognitive moral development when accounting for the variables of gender, academic aptitude, instructor, class time, class size and possible interactions among these variables?

**Theoretical Framework**

Moral development has been a distinct purpose of American higher education since its foundations (Rudolph, 1977, 1991; Rueben, 1996). Despite the many changes throughout the past 400 years of higher education in the United States, the liberal arts tradition of educating the whole person is still quite prevalent in today’s ethos (for example, Astin, 1997; Chickering & Reisser, 1993;
Heath, 1968; Pascarella & Terenzini, 1991, 2005; Sanford, 1967). According to a recent report by the Association of American Colleges and Universities (2002), college students should be “responsible learners” with a sense of social responsibility and strong ethical judgment.

Fortunately, there exists a large collection of research on the effects of college on moral development. “By far the dominant theoretical framework that guides this inquiry has been that of Lawrence Kohlberg” (Pascarella and Terenzini, 1991, p. 336). Kohlberg built upon Piaget’s models used to investigate moral development (observing children playing games and interviews based on short scenarios). Piaget (1960) theorized that a child reasoned in phases that “broadly speaking, follow one another without, however, constituting definite stages” (p. 195). These phases included obedience to adults and cooperation with peers (Krebs and Denton, 2005).

Kohlberg’s dissertation revised Piaget’s work in two main ways. First, he replaced Piaget’s short observations and short scenarios with nine moral dilemmas and structured interviews. Kohlberg read these hypothetical dilemmas to 84 adolescent boys and coded their responses. From these interviews, Kohlberg became convinced of Piaget’s theory on the moral development of children (Evans, Forney, Guido-DeBrito, 1998). He then further defined Piaget’s
earlier stages, and identified three more advanced stages for a total of six.

Kohlberg’s six stages fit into one of three levels - pre-conventional, conventional, or post-conventional (principled) – moral thinking. The following table illustrates Kohlberg’s levels and stages of moral development (Kohlberg, 1981)

**Table 1.1**

*Kohlberg’s Six Stages of Moral Reasoning*

<table>
<thead>
<tr>
<th>Level</th>
<th>Stage</th>
<th>Priorities</th>
<th>Moral Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preconventional</td>
<td>1</td>
<td>Avoid harm</td>
<td>Punishment and Obedience</td>
</tr>
<tr>
<td>Focus on Self</td>
<td>2</td>
<td>Self-interest</td>
<td>Instrumental Exchange</td>
</tr>
<tr>
<td>Conventional</td>
<td>3</td>
<td>Expectations of Others</td>
<td>Interpersonal Concordance</td>
</tr>
<tr>
<td>Focus on Others</td>
<td>4</td>
<td>Duties/Rights</td>
<td>Law and Order</td>
</tr>
<tr>
<td>Postconventional</td>
<td>5</td>
<td>Nonrelative obligations first</td>
<td>Social Contract</td>
</tr>
<tr>
<td>Focus on Principles</td>
<td>6</td>
<td>Self-chosen principles</td>
<td>Universally applied code of rational ideal</td>
</tr>
</tbody>
</table>

Kohlberg believed that individuals moved through stages structurally, sequentially, and hierarchically (Walker, 1988). In the structural criterion, individuals would apply the same moral reasoning regardless of the perspective. A stage one thinker would attempt to avoid punishment in every scenario. The individual will always move through the stages in a sequential format. Different
individuals will move through stages at different rates, but they cannot jump stages. This is because the stages are hierarchical. Each stage represents a more highly developed stage than the previous one. Kohlberg described this invariant hierarchical staging as a “hard” stage model of moral development (Kohlberg, 1983).

Kohlberg’s theory, while widely used, is not without its critics (Kohlberg, 1983; Krebs & Denton, 2005; Rest et al., 1999a). Criticisms include the lack of the perspective of cultural relativism (Schweder, 1982; Simpson, 1974), a structuralistic bias (Gibbs, 1979; Sullivan, 1977), and a lack of orientation to the ethic of care (Gilligan, 1982).

Rest built upon a foundation of Kohlberg’s theory with the problems associated with his work to form a neo-Kohlbergian approach to moral reasoning (Rest et al., 1999a; Rest et al., 2002). Rest’s six stages of moral development (1979) are directly borrowed from Kohlberg’s model (Evans et al., 1998). Rest rejected Kohlberg’s hard stage model in favor of a soft stage one. In Rest’s model, an individual’s reasoning is content-driven rather than structurally-formed (Rest et al., 1999b). Rest therefore refers to schema rather than stages. He abandoned the staircase model for a more variant one. In Rest’s model, an individual can use different schema to reason. No longer could individuals be categorized within a
single stage (Thoma, 2002). Instead, Rest “use(d) a model of development that conceptualizes change in terms of shifting distributions instead of the staircase stage model - defining developmental stages not in terms of operations but more in terms of the content of schemas” (Rest et al., 1999a, p. 182). Rest and his colleagues used data from the DIT to measure these schemas. Kohlberg was interested in the individual’s qualitative responses to structured interviews, which would be used to categorize the individual within a stage. The DIT is a quantitative instrument which assigns individuals a number. The number is reported as the “P score,” which is the percentage in which an individual uses post-conventional reasoning. In conclusion, this study will assess what kind of classroom interventions are most effective at increasing moral development, operationalized as increases in principled moral reasoning. Therefore, Kohlberg’s seminal work on the cognitive moral development of adolescents will be seen through a post-Kohlbergian lens throughout this study.

Significance of the Study

This study investigates the moral reasoning of freshmen business students. Business is increasingly is the most popular major for undergraduates. Students are drawn to the professional opportunities that a business degree
affords. However, public confidence in the business is low. A recent Gallup Poll showed that only 17 percent of Americans have a great deal of confidence in big business. This compares unfavorably to 82 percent who have a great deal of confidence in the military and 51 percent who have a great deal of confidence in the presidency (Saad, 2009). This confidence is certainly affected by financial crises that have cost Americans trillions of dollars. These crises are commonly cited as matters of greed and corruption by business leaders, and business schools have been asked to help remedy this situation.

This study is significant in several ways. First, it addresses the demand for better pedagogy in business ethics courses. It investigates the moral reasoning of undergraduate business students, as a measure for increasing moral development of future business leaders. It examines classroom conditions for optimal growth of moral reasoning. There is a need to use empirical studies to look at the interactions of measures within ethical interventions. Empirical findings of moral reasoning often look at measures individually, but lack in their application of the interactions of measures. For example, research indicates that measures of gender are more significant with post conventional thinking than measures of race or ethnicity (King & Mayhew, 2004). However, the research is
lacking in studies that use the interactions between the measures. In other words, how does gender interact with race when examining moral reasoning?

A missing measure within the research on moral reasoning is class size. The condition of class size is important when considering the cost efficiencies that higher education institutions pay close attention to. Increasing class size is one way to keep costs controlled, but not enough is known about potential downsides to increased class sizes. Anecdotally, students and professors seem against raising the size of classes, yet there is not enough research on class size at the college level to demonstrate clear findings. This study plans to examine conditions for the growth of moral reasoning of undergraduate business students, paying particular attention to the condition of class size, and the interactions of several important measures such as gender, instructor, class time, ethnicity, aptitude (SAT scores), college grades, and political orientation.

Design of the Study

To explore the questions and hypotheses, this study uses a quasi-experimental group design. This type of design is strongest when it includes more than one group, a common measured outcome, and random assignment (Gribbons & Herman, 1997). This study meets all of these qualifications.
The study utilized the data collected during the 2009 Fall semester at Boston College. The Defining Issues Test (DIT) was given to 400 freshmen business students enrolled across 18 sections of a business ethics course. The students were assigned to instructors, and then further assigned to sections. These sections were categorized as small or medium-sized. The two class sizes became the two treatment groups. Each group completed the DIT2 as both pretest and posttest, with the intervention occurring in between the pretest and posttest. The research design notation is as follows:

\[
\begin{align*}
R_s & \quad O & \quad X & \quad O \\
R_m & \quad O & \quad X & \quad O
\end{align*}
\]

“Rs” is the group of students randomly assigned to small classes. “Rm” is the group of students randomly assigned to medium-sized classes. “O” is the DIT2 which is used as both the pretest and posttest. X is the intervention of ethics instruction within the course.

The pre-tests of each were administered to course members in the beginning of the 2009 Fall semester. Post-tests were conducted in December. Results were analyzed through T-tests and a multiple regression of demographic variables on DIT post conventional reasoning scores to determine class size effect.
upon the moral reasoning of students enrolled in sections of a business ethics class. More on the design of the study will be found in Chapter Three.

Limitations of the Study

The possibility of threats to internal and external validity may exist. Factors that may affect internal validity include testing, maturation, and attrition (Campbell & Stanley, 1966). Testing threats may be caused by the participants’ pre-testing, which may cause them to “learn enough from the pretest to improve performance on the posttest” (Mertler, 2002, p. 334). The DIT2 does not have clearly “better” answers so this threat will be limited. The effects of maturation could threaten internal validity. Potentially, growth within the moral domain may be due to factors outside of this class. A comparison group would be helpful in minimizing this threat. Also, students may drop out of both or either group, causing internal validity to be threatened by attrition (Campbell & Stanley, 1966). Records were kept to help track attrition or mortality. Attendance was required of the course, and no student missed more than three classes. Only two students, from an original population of over 500, withdrew from this course; both due to family emergencies. This low mortality rate ensures the threat of mortality was limited.
Ecological validity (Bracht & Glass, 1968) requires the setting of the groups to be similar. This study examined all students who participated in the Fall 2009 Portico course. This sample was composed of several smaller groups, each which underwent a standard curriculum, yet with a different dynamic. Therefore, “some groups of participants, especially when involved in innovations, develop a group spirit that motivates high achievement” (Charles & Mertler, 2002, p. 335). This “Hawthorne Effect” may influence performance, as participants might have responded more favorably to the instrument since they were being studied. However, the DIT2 instrument offers a variety of responses, with no clear “right” or “wrong” answers. Furthermore, internal consistency checks occur throughout the DIT2 to ensure that responses are consistent throughout the instrument.

Another limitation involved the ability to draw conclusions about a population from a sample. Purposeful sampling is preferred over convenience sampling, as it allows researchers to use their judgment in selecting a sample that best represents the population. “The major disadvantage of purposive sample is the researcher’s judgment may be in error – he or she may not be correct in estimating the representativeness of a sample or their expertise regarding the information needed” (Fraenkel & Wallen, 1996, p. 103). However, there are a
number of studies utilizing the DIT which involve purposeful sampling (Porco, 2003). For this study, the sample is restricted to undergraduate business students in the first year of their college study. Moreover, the sample is participating in a first semester introductory business course. Therefore, “caution must be used when generalizing to other populations, even other undergraduate business students” (Adkins, 2009, p. 60).

Furthermore, this study observed one institution and one curriculum of the institution over a short time period. Therefore, possible institutional effects (self-selection of students who attend Boston College, with a certain mission and ethos) cannot isolate this treatment from other experiences. This is also a one semester study, with no longitudinal follow-up. Finally, the intended outcome of the curriculum is principled moral action, and principled moral reasoning is only part of moral action (Rest et al., 1999b). In other words, knowing the right thing to do may not always lead to doing the right thing.

Definition of Terms

Many who have grappled with issues of morality and ethics have agreed that the meanings of these terms are varied. Research on class size has also been inconsistent and ambiguous (Follman, 2009). Therefore, there is a need for clear
definitions of some of the key terms of this study. The following italicized terms appear throughout this study, which will be further discussed in Chapter Two.

Class size is the number of students assigned to and enrolled in a specific class under the instruction of a specific teacher (Follman, 2009). A major issue of class size studies at the college and university level is the lack of a magic number. There is no consensus for what defines the ideal class size. The annual rankings done by the U.S. News and World Report use a weighted sum of categories, and rank schools based on the sum (out of 100). One such category, “Faculty Resources,” comprises 20 percent of the final score. Class size weighs in as 30 percent of the “Faculty Resources” category. Therefore, class size accounts for six out of 100 possible points. Schools with a greater proportion of classes under 20 students receive higher scores. (http://www.usnews.com/articles/education/best-colleges/2008/08/21/how-we-calculate-the-rankings.html?PageNr=2, retrieved on 5/5/2009). There is no justification, however, for why 20 is chosen as the ideal class size number. The Glass and Smith meta-analysis of class size demonstrates that the ideal class size is 15 – 18 (1979). Therefore, I will use that range in classifying small classes. According to the Tennessee study (Project STAR), a difference of eight students is a significant one (Finn & Achilles, 1999). Therefore, I will define medium sized classes as having between 23 and 26 students enrolled.
For the purpose of this study, *moral reasoning* and *ethical reasoning* will be used interchangeably. Both can be defined as the degree in which an individual uses a principled thought process in thinking about moral dilemmas (Kohlberg, 1981). Kohlberg theorized that there are pre-conventional, conventional, and post-conventional stages of moral reasoning. The theory of these stages is known as *cognitive moral development* (CMD).

The Defining Issues Test is interested in measuring *post-conventional moral reasoning*. Post conventional reasoning is the highest level in Kohlberg’s theory of moral development. At this level, the individual recognizes alternative moral courses, explores the options, and then decides on a personal moral code. The *post-conventional* level is also referred to as the *principled* level of moral reasoning, and both will be used interchangeably throughout this study.

*Overview of the Study*

This study explores the class size effect on the moral judgment of first year college students. Chapter one provided an introduction to the study including the topic of the study, the research questions, the study’s theoretical rationale and significance to the higher education literature, an overview of the research design, predicted limitations of the study, and definitions of key terms. Chapter
two will review the literature by integrating the theoretical framework of
cognitive moral development with business ethics education and research on
class size effect. Chapter three will explain the research design methodology
utilized in light of quantitative quasi-experimental studies. Chapter four will
present the data and findings of the research. Chapter five will summarize these
findings, discuss implications for higher education practice, describe the
limitations of this study, and suggest future possibilities for research.
CHAPTER TWO
REVIEW OF THE LITERATURE

This chapter reviews the literature related to moral development theory and business ethics. The first section is an historical review of moral education in American higher education. This provides a backdrop to the current state of moral development of college students. The second section presents major theories of moral development, with an emphasis on Kohlberg’s model of cognitive moral development (CMD). Also included in this section is a description of neo-Kohlbergian theories, especially the work of James Rest and his colleagues, as well as the work of Krebs and Denton. The third section investigates the Defining Issues Test, which has been the most often used tool in assessing CMD of college students. The fourth section focuses on business ethics: its history, place in the curriculum, pedagogical methods, criticisms, and studies related to business ethics interventions (in regard to how they impact CMD). Fifth, class size, a particular condition of business ethics instruction, is presented as a missing variable in the research of CMD of college students. The final section focuses on the strengths and gaps in these literatures and in their intersections as a set.
Moral Education in American Higher Education

Moral development has been a distinct purpose of American higher education since its foundations (Rudolph, 1991; Rueben, 1996). This history began with the colonial college, which sought to prepare young men for professions in the clergy (Brubacher & Rudy, 2004; Geiger, 2005). Preparation for the clergy soon expanded to other vocations, such as medicine and law. While rooted in the classical curriculum of the tridium, the goal of the colonial college was to form men of high moral standards, develop the whole person, and create the gentlemen scholar (Brubacher & Rudy, 2004). The final course was a seminar in moral philosophy, often taught by the university president (Brubacher & Rudy, 2004; Rudolph, 1977). The colonial college experience was, ultimately, “the unity of values training, professional education, and the liberal arts” (Pamental, 1987, p. 3).

The training of moral values extended beyond the classroom. Courses in moral philosophy were supplemented by a “collegiate way of life” (Brubacher & Rudy, 2004, p. 41). Modeled after Oxford and Cambridge, the early American college was a residential experience for the elite and future leaders of the country. It “(made) men out of boys” (Rudolph, 1962, p. 140) through its unheated dormitories, religious revivals, extracurricular activities, and dedicated
faculty acting *en loco parentis*. “Thus, the academic curriculum and the entire campus environment clearly viewed the formation of student character as a central mission” (Pascarella & Terenzini, 2005, p. 335).

As American higher education continued into the 19th century, it became more diverse. The research university, influenced heavily by the German system, became the prevalent model. Professors, often trained in Germany, did not seek to educate the whole person (Altbach, 2005; Brubacher & Rudy, 2004). Rather, they specialized in disciplines. This model also provided a professional education, but adherence to a strict subject matter had replaced the holistic education of the earlier colleges (Altbach, 2005; Pamental, 1987). In many sections of academia, well-roundedness and moral fortitude were replaced with technical ability and scientific research.

Concurrently, societal needs complemented the German influence to create interest in the study of agriculture, business, and engineering. Paradigmatic events such as westward expansion and the industrial revolution precipitated new knowledge, which differed from the colonial college model of transmitting old knowledge (Brubacher & Rudy, 2004). Soon after, new curricula were developed at institutions such as RPI, Union College, and West Point (Rudolph, 1977, 1990). These shifts – increased specialization, the creation of new
knowledge, and the emergence of disciplines – altered the unity of the colonial curriculum. The well-rounded moral philosopher became relegated to the philosophy department. Instead of teaching a capstone moral philosophy course to all students, that “course became merely one of several courses in a separate department, avoided by many because it did not relate to their future careers” (Pamental, 1987, p. 13). This distinction is still quite current with today’s college students.

These shifts – towards proliferation of new knowledge through research and specialization – were not without sacrifice. The collegiate way of life had become a bifurcated life. The twentieth century, however, saw a movement towards a more holistic educational philosophy. According to the Student Personnel Point of View (1937)

This philosophy imposes upon the educational institutions the obligation to consider the student as a whole – his intellectual capacity and achievement, his emotional make up, his physical condition, his societal relationships, his vocational aptitudes and skills, his moral and religious values, his economic resources, and his aesthetic appreciations. It puts emphasis, in brief, upon the development of the student as a person rather than upon his intellectual training alone (p. 39).
So, despite the many changes throughout the past 374 years of higher education in the United States, the liberal arts tradition of educating the whole person is still quite prevalent in today’s ethos (for example, Astin, 1997; Chickering & Reissner, 1993; Heath, 1968; Pascarella & Terenzini, 2005; Stanford, 1967).

According to a recent report by the Association of American Colleges and Universities (2002), college students should be “responsible learners” with a sense of social responsibility and strong ethical judgment. Today, the colonial college model is reflected in the modern liberal arts education as most colleges include moral development in their mission (King & Mayhew, 2002; 2004). A key question, therefore, in assessing the impact of higher education upon its students, is Does college affect moral development?

Theories of Moral Development

Kohlberg

Fortunately, there exists a large collection of research on the effects of college on moral development. “By far the dominant theoretical framework that guides this inquiry has been that of Lawrence Kohlberg” (Pascarella & Terenzini, 1991, p. 336). Kohlberg’s interest in morality began with his experiences during World War II, during which he served time in a concentration camp due to his
involvement in helping Jewish immigrants. Kohlberg’s personal observations of the horrific injustices of the Holocaust caused him to reject moral relativism in search of a more universal morality (King & Mayhew, 2004; Kohlberg, 1991).

Kohlberg’s sustained interest in the study of moral development stemmed from his dissertation work at the University of Chicago (Evans, Forney, & Guido-DeBrito, 1998). His seminal study examined the moral reasoning of adolescent boys. Kohlberg built upon Piaget’s methods used to investigate moral development observing children playing games and interviews based on short scenarios. Piaget (1960) theorized that children reasoned in phases that “broadly speaking, follow one another without, however, constituting definite stages” (p. 195). These phases included obedience to adults and cooperation with peers (Krebs & Denton, 2005).

Kohlberg’s dissertation revised Piaget’s work in two main ways. First, he replaced Piaget’s short observations and short scenarios with nine moral dilemmas and structured interviews, known as the Moral Judgment Interview (Kohlberg, 1981). The most famous example is the Heinz dilemma, in which Kohlberg (1981) states:

A woman was near death from a special kind of cancer. There was one drug that the doctors thought might save her. It was a form of radium that
a druggist in the same town had recently discovered. The drug was expensive to make, but the druggist was charging ten times what the drug cost him to produce. He paid $200 for the radium and charged $2,000 for a small dose of the drug. The sick woman's husband, Heinz, went to everyone he knew to borrow the money, but he could only get together about $1,000 which is half of what it cost. He told the druggist that his wife was dying and asked him to sell it cheaper or let him pay later. But the druggist said: "No, I discovered the drug and I'm going to make money from it." So Heinz got desperate and broke into the man's store to steal the drug for his wife.

Should Heinz have broken into the store to steal the drug for his wife?

Why or why not? (p. 12)

Kohlberg read these hypothetical dilemmas to 84 adolescent boys and coded their responses. From these interviews, Kohlberg became convinced of Piaget’s theory on the moral development of children (Evans, et al., 1998). He then further defined Piaget’s earlier stages, identified three more advanced stages, and presented a “more fine grained analysis” (Porco, 2003, p. 40). Kohlberg’s six stages are outlined in Table 2.1 (Kohlberg, 1981).
Table 2.1 - Kohlberg’s Six Stages of Moral Reasoning

<table>
<thead>
<tr>
<th>Level</th>
<th>Stage</th>
<th>Priorities</th>
<th>Moral Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preconventional:</td>
<td>1</td>
<td>Avoid harm</td>
<td>Punishment and Obedience</td>
</tr>
<tr>
<td>Focus on Self</td>
<td>2</td>
<td>Self-interest</td>
<td>Instrumental Exchange</td>
</tr>
<tr>
<td>Conventional:</td>
<td>3</td>
<td>Expectations of Others</td>
<td>Interpersonal Concordance</td>
</tr>
<tr>
<td>Focus on Others</td>
<td>4</td>
<td>Duties/Rights</td>
<td>Law and Order</td>
</tr>
<tr>
<td>Postconventional:</td>
<td>5</td>
<td>Nonrelative obligations first</td>
<td>Social Contract</td>
</tr>
<tr>
<td>Focus on Principles</td>
<td>6</td>
<td>Self-chosen principles</td>
<td>Universally applied code of rational ideal</td>
</tr>
</tbody>
</table>

Each of the stages fit into one of three levels - pre-conventional, conventional, or post-conventional (principled) – moral thinking. The first level is that generally seen in young children. In the first stage of this level, what is right is defined by the rules of an authority figure (e.g., parent or teacher). People obey what is right to avoid punishment. In the second stage of this level people follow rules if it is in their own best interest. Though they begin to see the needs of others, their own hedonism is paramount. (Evans, et al., 1998).

“The second level of moral thinking is that generally found in society, hence the name conventional” (Barger, 2000). The first stage of this level (stage 3)
is characterized by fulfilling social roles in a dutiful way ("good-boy/good-girl").

In the second stage of this level (stage 4), individuals act in accordance to the laws and duties of society. Here, what is right is still driven by external forces rather than principled perspectives.

Kohlberg thought that very few people attained the third level of moral reasoning (Barger, 2000). “At this stage, the rightness of laws and social systems are evaluated on the basis of the extent to which they promote fundamental human rights and values” (Evans, et al., 1998, p. 175). Individuals at this stage are driven by social contracts, which are necessary to uphold for the sake of all in society. For Kohlberg, a social contract moves beyond the stage 4 rule utilitarianism (everyone must follow rules, otherwise, chaos ensues) in order to create a blueprint for society. It is a development from a “law maintaining” society to a “law creating” one (Kohlberg, 1981, p. 153). The last stage (stage 6) is based on abstract reasoning and grounded in justice, following the philosophical work of Kant and Rawls. Judgments at this stage are based on universally-generalizable principles that are categorically imperative. Contracts and laws do not guide decisions. Instead an action is the end in and of itself, because it is the right thing to do. While Kohlberg believed in the theoretical validity of this stage, he found little evidence that many people were in stage six. He cited Martin
Luther King, Jr. and Gandhi among the small sample of individuals who had reached this stage.

Kohlberg believed that individuals moved through stages structurally, sequentially, and hierarchically (Walker, 1988). In the structural criterion, individuals would apply the same moral reasoning regardless of the perspective. A stage one thinker would attempt to avoid punishment in every scenario. The individual will always move through the stages in a sequential format. Different individuals will move through stages at different rates, but they cannot jump stages. This is because the stages are hierarchical. Each stage represents a more highly developed, more complex, stage than the previous one. Kohlberg described this invariant hierarchical staging as a “hard” stage model of moral development (Kohlberg, 1983).

**Critiques of Kohlberg**

Kohlberg’s theory, while widely used, has been widely criticized as well (Kohlberg, 1983; Rest et al., 1999a). Criticisms include the lack of the perspective of cultural relativism (Schweder, 1982; Simpson, 1974) and a lack of orientation to the ethic of care (Gilligan, 1982; Murphy & Gilligan, 1979). Both criticisms are rooted in the research sample of Kohlberg’s study, and the resulting emphasis on justice.
As stated earlier, Kohlberg's initial work involved interviewing adolescent boys in the United States. Simpson (1974) argues that the Kohlberg's stages, therefore, may not be (and actually are not) culturally universal. She objects to Kohlberg's theory of universal and sequential stages based on empirical and philosophical claims. Empirically speaking, Simpson cites cross-cultural studies where reasoning beyond stage four does not exist, and other cultures which reverse the developmental sequences. Philosophically, Kohlberg's western and Eurocentric study cannot be universal “because it is the product of a researcher or theorist who has a particular cultural identity and background” (Kohlberg, et al., 1983, p. 110). She feels that Kohlberg's worked is biased and utilizes a measuring stick that implies western culture as “morally superior” (Simpson, 1974, p. 91) to others.

A second, and related criticism originates from the work of Carol Gilligan. Again, because Kohlberg's work focused on the moral reasoning of boys, Gilligan argues that it is androcentric and does not adequately explore the concerns of women (Gilligan, 1982). Gilligan, who was a student of Kohlberg's, conducted studies in which women also participated. Her college study "explored identity and moral development in the early adult years" (Gilligan, 1982, p. 2), and included both women and men participants. Her abortion
decision study included 29 young women and “considered the relation between experience and thought and the role of conflict in development” (p. 3). Gilligan’s research found that patterns of development between the men and women in her and Kohlberg’s studies were similar (Colby, Gibbs, Lieberman, & Kohlberg, 1983; Walker, 1988). However, rather than focusing on the value of justice, which “fails to capture the distinctly female voice on moral matters” (Crain, 1985, p. 135). Gilligan observed that women’s decisions focused on interpersonal relationships and an ethic of care in which abstract principles of justice are subordinated to the maximizing of benefit for everyone involved in a particular concrete situation. Gilligan introduces a possibility that moral development may occur along two lines – one that focuses on justice while the other focuses on interpersonal relationships and care. Perhaps the goal of adult moral development is to draw together these lines? Or, in Gilligan’s own words, “the dialogue between fairness and care not only provides a better understanding of relations between the sexes but also gives rise to a more comprehensive portrayal of adult work and family relationships” (1982, p. 174).

**Neo Kohlbergian Theories**

A more recent update of Kohlberg’s work (Krebs & Denton, 2005) attempted to find more practical implications for Kohlberg’s theory of
development. Krebs and Denton (2005) found that Kohlberg’s model left people unprepared to make everyday decisions. In order to assess this hypothesis, they updated Kohlberg’s examples as follows:

Kohlberg’s dilemmas were modified by substituting for Heinz a homosexual man whose boyfriend was sick from AIDS and by asking participants to imagine themselves as the protagonists in the dilemmas. New dilemmas involved decisions about whether to help a victim in an emergency, whether to keep a promise to drive friends home after becoming intoxicated, whether to engage in prostitution, whether to disclose damaging information during the sale of a business, and whether to support free trade when it went against business interests (Krebs & Denton, 2005, p. 632).

Krebs and Denton offer some clear and practical methods of updating Kohlberg. By utilizing “real” scenarios rather than “hypothetical” Kohlbergian scenarios, they introduce concepts of interaction and emotion. This is an important distinction from Kohlberg and responds to the earlier criticisms of non-relativistic reasoning that lacks in an emphasis of care. The outcome of their work is a set of eleven propositions which “account for the goals people use moral judgments to achieve” (Krebs & Denton, 2005, p. 640). Ultimately, while
Kohlberg focuses on moral reasoning as an end in itself, Krebs and Denton see moral reasoning as a means to another end – that of moral action. More on the correlation between moral reasoning and moral behavior will follow later in this chapter.

*James Rest and the Defining Issues Test*

James Rest produced what is perhaps the greatest expansion of Kohlberg’s theory. Like Kohlberg, his interest in moral development originated from his own childhood experiences. Rest grew up in the Deep South during the peak of the civil rights movement (Thoma, 2002). He was the son of a progressive minister, whose conservative congregation was at odds with his parents’ view of civil rights. Rest wrestled with this tension, trying to resolve its conflicts through academic pursuits. He studied philosophy, then theology, then psychology, without satisfaction (Thoma, 2002). Finally, Rest was exposed to Kohlberg’s theory of moral development and began to make sense of the conflicting viewpoints. Rest was particularly drawn to Kohlberg’s emphasis on macro-morality (Thoma, 2002). In the case of his father’s congregation during the civil rights movement, the congregation was characterized by traditions and conventions and a fear of change. Rest’s parents, meanwhile, emphasized
challenging the conventional order for a more just one. By using the lens of Kohlberg’s theory of moral development, Rest began to more clearly understand the tensions between the traditional Congregationalists and the proponents of the civil rights movement (Thoma, 2002).

At the end of his time studying with Kohlberg at Harvard, Rest became concerned that the gap between theory and data was too big. Rest was committed to Kohlberg’s theory of morality as cognitive and developmental, yet he was concerned with the empirical methods for measuring data. Based on his dissertation work, Rest sought to rank and rate the responses from Kohlberg’s Moral Judgment Interviews (MJI). In fact, Rest was able to construct an instrument that measured cognitive growth, without the empirical concerns of the MJI, and with a much more efficient method of data collection.

Upon completing his dissertation, Rest moved to the University of Minnesota, where his work was used to develop the Defining Issues Test (DIT) to measure the moral reasoning of individuals (Rest, 1994). It is a paper and pencil test derived from Kohlberg’s theories of moral development (Kohlberg, 1984). Rest borrowed from Kohlberg’s MJI, with its open-ended responses to moral dilemmas (Bebeau & Thoma, 2003). Rather than coding free responses to hypothetical dilemmas given in an interview format, the DIT presents vignettes
for a subject to rate and rank. The responses represent the degree in which “moral schema” of thinking are activated. Rest understood schema to be structures of thinking that are activated by content or stimulus (Rest et al., 1999a). He and his colleagues used the terminology in order to differentiate DIT research from Kohlberg’s stage theory. The schema are closely related to Kohlberg’s stages of moral development (Bebeau & Thoma, 2003), yet different in the following ways (Rest, Narvaez, Thoma & Bebeau, 2000):

1. Hard stages versus soft schema. Rest’s schema are seen as “shifting distributions rather than a staircase” (p. 384).

2. Schema are more concrete and specific than stages.

3. Cognitive operations vs. content of operations. While Kohlberg stressed operations of moral thinking, Rest focused on content.

4. Universality. “Kohlberg postulated universality as a characteristic of stages whereas we regard cross-cultural similarity as an empirical question” (p. 385).

5. Articulation vs. tacit knowledge. Rest used multiple choice responses in order to test that reasoning was part of an individual’s discernment but not based on their ability to articulate open-ended responses.
In his book, *Moral Development: Advances in Research Theory*, Rest (1986) describes the DIT as follows:

The DIT is based on the premise that people at different points of development interpret moral dilemmas differently, define the critical issues of the dilemmas differently, and have different intuitions about what is right and fair in a situation. Differences in the way that dilemmas are defined therefore are taken as indications of their underlying tendencies to organize social experience. These underlying structures of meaning are not necessarily apparent to a subject as articulate rule systems of verbalizable philosophies – rather, they may work “behind the scenes” and may seem to a subject as just commonsensical and intuitively obvious (p. 196).

*The Four Phases of the DIT*

At the University of Minnesota, Rest’s research program developed into the Center for the Study of Ethical Development. The work of the center was designed around the DIT, and can be viewed in four phases (Thoma, 2002).

Phase 1 is concerned with the development of the DIT, specifically with assessing and validating the measure of moral reasoning. Phase 1 culminated in
1979 with the validation of the P score (P for principled). It represents the proportion of responses on the DIT that reflect principled (P) moral reasoning (King & Mayhew, 2002). The P score is computed by adding a respondent’s Stage 5 and 6 responses, and then weighting these ranks in order to calculate a score. The scores range from 0 – 95 and represent the percent of an individual’s preference for post-conventional thinking (Bebeau & Thoma, 2003).

Phase 2 shifts from assessment of the DIT instrument to “concerns related to the conditions and detriments of change in moral judgments” (Thoma, 2002, p. 228). By 1981 the DIT was established as a reliable and valid instrument and a Center for collecting and scoring data had been created. These conditions created an explosion of research using the DIT. By 1986, which marks the end of Phase 2, over 500 DIT studies had been conducted. Rest’s book, Moral Development: advances in theory and research (Rest, 1986), summarized these findings and gave way to Phase 3.

Phase 3 is tied to the development of Rest’s own theory of moral development, the Four Component Model of morality. For Rest, Kohlberg’s issue of moral judgment was but one of four parts of what is needed for moral action. Rest, et al. (1999a, p. 101), outline the four processes as follows:
1. Moral sensitivity (interpreting the situation, role taking how various actions would affect the parties concerned, imagining cause-effect chains of events, and being aware that there is a moral problem when it exists)

2. Moral judgment (judging which action would be most justifiable in a moral sense – purportedly DIT research has something to say about this component)

3. Moral motivation (the degree of commitment to taking the moral course of action, valuing moral values over other values, and taking personal responsibility for moral outcomes)

4. Moral character (persistent in a moral task, having courage, overcoming fatigue and temptations, and implementing subroutines that serve a moral goal). (p. 101)

Moral judgment (component 2) is often the intended outcome for classroom ethics interventions, and is the construct measured by the DIT.

However, moral behavior is the ultimate outcome of any moral judgment (Rest, 1984). In other words, “action is nearer to being the litmus test for bona fide morality” (Haan, 1985, p. 53). For Rest, morality is not solely a cognitive process.

The four component model is concerned with the interactions between cognition
and effects (Bebeau, Rest, & Narvaez, 1999; Narvaez & Rest, 1995; Rest, 1986) rather than “the traditional tripartite classification of moral functioning into cognitive, affective, and behavioral domains” (Walker, 2002, p. 354). Nonetheless, Pascarella and Terenzini offer the hypothesis of an indirect effect between moral judgment and moral behavior. Their distillation of research on moral development shows a systematic link between moral reasoning and “a range of principled behaviors, including resisting cheating, social activism, keeping contractual promises, and helping those in need” (Pascarella & Terenzini, 1991, p. 367). Rest and his colleagues have also demonstrated correlations between DIT scores and moral behavior, though the correlations are most powerful when all four components are integrated. This presents interesting conditions for moral action; however, this study focuses on the second component of moral judgment.

Phase 4 is linked to the focus on moral development and moral education within the professions. Furthermore, Phase 4 represents “the process of taking stock of theoretical and empirical underpinnings of the measure as influenced by the accumulation of data generated by the DIT” (Thoma, 2002, p. 228). By this phase, the DIT had accumulated norming samples of more than 45,000 individuals. This sample size presented the opportunity “to revisit some of the indexing and procedural issues that had been tabled in the 1970s because the
available samples would not support the necessary statistical analysis” (Thoma, 2002, p. 238). From this analysis, a new instrument, the DIT2, and a new scoring index, the N2 score, were born.

Compared to the DIT, the DIT2 has the advantage of being shorter, more contemporary, and clearer. In addition, a recent study encourages researchers to substitute the DIT2 for the DIT due to gains in validity, a high correlation of the DIT with the DIT2 \( r = .79 \). The Cronbach’s alpha – an index of the internal consistency of a test based on the extent to which test-takers who answer a test item one way respond to other items in the same way (Gall, 1998) – of the DIT2 is .90 (Rest, Narvaez, Thoma, & Bebeau, 1999b). This compares favorably to the internal consistency of the original DIT, which had a Cronbach’s alpha of .76 (Rest, 1984). For these reasons, the DIT2 has become the preferred option for measuring moral development and is utilized in this study.

The N2 index is a relatively new version (hence N2) of the DIT index score. The score combines the percentage of post-conventional thinking with the rejection of simplistic thinking, both of which are desirable (Bebeau & Thoma, 2003). The score is weighted and adjusted to have the same mean and standard deviation as the P score so comparisons can be made. The N2 index permits more subjects to pass subject reliability checks, which allows sample retention (Rest &
Narvaez, 1998). In a comparison of studies of N2 to P scores, Rest et al. (1997) concluded that the N2 index is “generally better than the P index (p. 505), largely due to the rejection of simplistic thinking and the retention of more data. The DIT and its index scores will be more clearly examined in Chapter 3.

**Empirical Findings of the Defining Issues Test**

The Defining Issues Test (DIT) is the most widely used instrument within research on moral development (Thoma, 2002; Trevino, 1992). Since the 1974 inception of the DIT (Rest, 1974) there have been over 500 studies utilizing the DIT (King & Mayhew, 2002; Thoma, 2002) which have included over 500,000 participants (Rest, et al., 1999b; Walker, 2002).

Whether intentional or not, formal education increases the moral reasoning of its students (King & Mayhew, 2004). For example, DIT P scores by educational levels are as follows: Junior high students = 21.9; senior high students = 31.8; college students = 42.3; graduate students = 53.3 (Rest, 1994).
Table 2.2 – DIT P Score Means and Standard Deviations

<table>
<thead>
<tr>
<th>Age/Education</th>
<th>Mean P Score</th>
<th>Standard Deviation</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior High</td>
<td>21.9</td>
<td>8.5</td>
<td>1,322</td>
</tr>
<tr>
<td>High School</td>
<td>31.8</td>
<td>13.5</td>
<td>581</td>
</tr>
<tr>
<td>College</td>
<td>42.3</td>
<td>13.2</td>
<td>2,479</td>
</tr>
<tr>
<td>Graduate School</td>
<td>53.3</td>
<td>10.9</td>
<td>183</td>
</tr>
</tbody>
</table>

Though this may, at a glance, seem to be a matter of age and maturation, research has shown that formal education is “the most consistent and powerful correlates of moral judgment development” (Rest & Thoma, 1985, p. 709). Rest’s (1979) secondary analysis of over 4,500 subjects discovered that formal education accounted for between 39 to 48 percent of the variance of DIT scores. Rest and Thoma (1985) also performed a secondary analysis on 6,000 subjects and found that formal education 53 percent of the variance of DIT scores. This compared to gender, which contributed to only .2 percent of the variance. Rest and Thoma performed a longitudinal study with a subset of these subjects to compare DIT increases in people who attend college with DIT scores of those who do not attend. Rest & Thoma divided their subjects into “low education” (2 years or less
of college) and “high education” (three or more years of college) groups. They found quite divergent paths between the two groups, as seen in table four.

### Table 2.3 – DIT P Scores for Low Education and High Education Group

<table>
<thead>
<tr>
<th></th>
<th>Low Education Group</th>
<th>High Education Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score at HS Graduation</td>
<td>33</td>
<td>37</td>
</tr>
<tr>
<td>Score Six Years Later</td>
<td>34.5</td>
<td>51</td>
</tr>
</tbody>
</table>

Rest & Thoma used ANCOVA to show that the DIT scores increased significant in the high education group, above and beyond what can be attributed for by the initial score at high school graduation.

The relationship of formal education to moral reasoning has been the central topic of at least 44 other studies, with more than 85 percent showing a significant correlation. Of the 44 studies, 38 have reported a significant relationship between formal education and moral reasoning, while six (Bonawitz, 2002; Cohen, 1982; Cummings, Dyas, Maddux, & Kochman, 2001); Galotti, 1989; Quarry, 1997; Shaub, 1994) have demonstrated no significant relationship. These studies verify the assertion that participation in college significantly increases DIT scores, compared to what is attributed to age alone, as
formal education is “by far the most powerful demographic correlate of DIT P scores, typically accounting for 30 percent to 50 percent of the variance in large, heterogeneous samples” (Rest, et al., 1999b, p. 70).

*Student Characteristics*

Other often-researched characteristics of DIT subjects include ethnicity, gender, socio economic status (SES), aptitude (intelligence), political identification, and religious identification (King & Mayhew, 2004). Findings of a review by King & Mayhew (2004) show that gender, aptitude, and political identification are more likely to affect moral reasoning than ethnicity, SES, or religious identification. The following explores studies that focus on these characteristics.

Because gender has been an area of interest in DIT studies, Thoma & Rest (1986) conducted a meta-analysis of 56 DIT studies with over 6,000 subjects. Thoma concluded, as stated above, that education was the strongest indicator of increased moral reasoning, 250 times more powerful than gender (1985). Nonetheless, he reports that women have higher scores than men. This is consistent with a majority of studies on gender differences in moral reasoning. King & Mayhew’s (2004) review of 40 studies that focus on the relation of gender
to moral reasoning show that 22 of the studies report higher scores for women than men, 16 studies report no differences between women and men on the DIT, and only two studies report that men score higher. These findings show that, unlike Kohlberg’s MJI findings, there are no gender biases with the DIT (Brabeck, 1983).

Several studies have investigated the effects of SES on moral reasoning, without consistent findings. Ways in which SES has been operationalized include by measuring parent’s educational level, parent’s occupation, and family income. These studies (Biggs & Barnett, 1981; Finger, Borduin, and Baumstark, 1992; Gfellner, 1986; Gongre, 1981; Mentkowski & Strait, 1983, Rest, 1979; Stepp, 2002) show no relationship between SES and DIT scores.

The variables of aptitude and intelligence has been measured in a variety of ways. Aptitude has been measured using SAT scores (Quarry, 1997; Sanders, Lubinski, and Benbow, 1995), ACT scores (Gongre, 1981; Hendel, 1991; Stepp, 2002), and PSAT scores (Hendel, 1991). Intelligence has been measured using high school GPA (Green, 1981; Mentkowski & Strait, 1983; Quarry, 1997; Stepp, 2002) and by performance on written assignments (Johnson, Insley, Motwani, and Zbib, 1993). Aptitude has demonstrated a relationship between test scores and DIT scores in all but one study on SAT scores (Quarry, 1997). The majority of
studies have shown that higher aptitude test scores relate to higher DIT scores. Measures of intelligence are less predictable, as high school grades do not relate to DIT scores in any significant way.

Other demographic characteristics include political identification and religious identification. As a measure of social justice, political identification has been an area of interest to DIT research (e.g., Elmer, Renwick, and Malone, 1983; Elmer, Palmer-Canton, and St. James, 1997; Narvaez, Getz, Rest, and Thoma, 1999; Thoma, 1994). DIT participants are asked to self-report their political orientation in response to the question, how would you characterize yourself? Likert scale responses include very liberal, liberal, neither liberal nor conservative, conservative, and very conservative. These studies report a relationship between liberal political orientations and post conventional reasoning. Similarly, research on religious identification has shown that students who self-report a more liberal religious ideology also exhibit more post conventional reasoning than students who self-report as religiously conservative (Clouse, 1985; Foster & LaForce, 1999; Murk & Addelman, 1992; Stepp, 2002). Most studies on religious measures, however, show no differential effects on DIT scores and moral reasoning (King & Mayhew, 2004). These measures include religious affiliation, church attendance, and religious training.
Borkowski & Urgas (1998) performed a meta-analysis of empirical studies from 1985 to 1994, specifically reviewing the ethical development of business students. They reviewed 56 studies of almost 20,000 business students, investigating measures of gender, age, and undergraduate major. Their finding were mixed, and were consistent with previous research of moral reasoning. The only measure which showed a relationship with ethical development was gender. Of the 47 studies that investigated gender, 49 percent yielded a significant difference between female and male students, 34 percent were non-significant, and 17 percent yielded mixed findings. The other measures, age and college major, were inconclusive in their findings (Borkowski & Urgas, 1998).

In conclusion, there is a large body of research concerning student characteristics and moral reasoning. Educational level is by far the most relational measure, as it is 250 times more powerful than the next measure of gender (Thoma & Rest, 1985). Other measures that show relationship with DIT scores include aptitude, intelligence (to a lesser extent), and political orientation. Age, ethnicity, SES, and religious orientation are less likely to affect moral reasoning. Research focused solely on business students is consistent with the larger body of research (Borkowski & Ugras, 1998).
More research is needed to show the interaction effects of these measures. Perhaps due to the small sample sizes of subgroups, too few studies have investigated these interactions (e.g., gender by political orientation). This study attempts to show interactions of instructor, class time, academic aptitude (SAT scores), first year college grades, and gender.

Curricular Interventions

The DIT has also been widely used to investigate the effect of curricular interventions on moral reasoning (Pascarella & Terenzini, 1991). In particular, ethics courses have often utilized the DIT to assess moral reasoning as an outcome of the course. Intervention studies are similar to longitudinal studies, as they test and re-test individuals who therefore act as controls (Rest et al., 1999a). However, they are typically shorter than longitudinal studies, as they last less than one year, and sometimes much less. This study investigates the curricular intervention of a business ethics course.

Business Ethics

The birth of formal business education in the United States can be traced to 1881. That year, a Philadelphia businessman donated $100,000 to the
University of Pennsylvania to establish the School of Finance and Economy. Soon after, the school was named the Wharton School, to honor its donor, Joseph Wharton (Pamental, 1987). Wharton’s aim was

To provide for young men special means of training and correct instruction in the knowledge and in the arts of modern finance and economy, both public and private, in order that, being well informed and free from delusions on these important subjects, they may either serve the community skillfully, as well as faithfully, in office of trust, or, remaining in private life, they may prudently manage their own affairs and aid in maintaining sound financial morality; in short, to establish means for imparting a liberal education in all matters concerning finance and economy. (Herrick, 1904, p. 255).

Despite being a new type of education, the aim of the early business school shared more in common with the colonial college than the German-influenced research universities established in the same era as Wharton. The early curriculum consisted of languages (Latin, German, and French), English (composition and literature), philosophy, social sciences (economics and history), natural sciences (mineralogy, geology, physics, astronomy), and law (Pamental, 1987). Another feature shared with the colonial college was the capstone course
in moral philosophy. The first study of business education, by Charles Herrick during the turn of the 20th century, supported the emphasis of morality in business. In Herrick’s (1904) view, “whenever schools of commerce are established, whatever their local problems, their supreme aim should be the production of those who have the ballast of integrity of purpose, whose ships of life shall be ever on the even keels of strict morality…(business education) will be perilous if it does not remain moral” (p. 64).

By 1916, seventeen leading American colleges and universities gathered to form The Association to Advance Collegiate Schools of Business (AACSB) (www.aacsb.edu retrieved on 4/26/10). The new organization began to accreditate schools in 1919 when it adopted its first standards for business degree programs. Expansion of business schools has been exponential. Currently, AACSB membership includes more than 1,100 universities in 70 countries.
Table 2.4 – Number of Business programs in the U.S., by year

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>1881</td>
<td>1</td>
</tr>
<tr>
<td>1913</td>
<td>29</td>
</tr>
<tr>
<td>1924</td>
<td>183</td>
</tr>
<tr>
<td>1956</td>
<td>581</td>
</tr>
<tr>
<td>2010</td>
<td>1100</td>
</tr>
</tbody>
</table>

(Pamental, 1989; AACSB, 2010)

The number of students studying business likewise grew exponentially. Currently, it is the most popular program of studies in the United States with 327,531 undergraduate degrees in business conferred in 2006-2007.
Table 2.5 – Number of bachelor degrees conferred in business, from 1923 to 2007

<table>
<thead>
<tr>
<th>Year</th>
<th>Degrees conferred</th>
</tr>
</thead>
<tbody>
<tr>
<td>1924</td>
<td>5,091</td>
</tr>
<tr>
<td>1940</td>
<td>18,549</td>
</tr>
<tr>
<td>1948</td>
<td>37,328</td>
</tr>
<tr>
<td>1958</td>
<td>50,090</td>
</tr>
<tr>
<td>2007</td>
<td>327,531</td>
</tr>
</tbody>
</table>

(Pamental, 1989; AACSB, 2010)

However, this growth outpaced the ability to teach business effectively. A study by the Ford Foundation in 1950 stated that the “simple fact of the matter is that academic standards are too low in most of the business schools in the United States” (Gordon & Howell, 1959, p. 136). This study, and a similar one commissioned by the Carnegie Foundation (Pierson, 1959) reported that business schools offered and required too many courses, and this diluted quality. These studies motivated an emphasis on the scholarship of business, and the growth of business as a field of study.
The focus on research and quantitative rigor led to a decline in soft fields such as business ethics until the 1960s. The civil rights movement and social justice of the 1960s raised consciousness in society and education, including business. Price fixing scandals at General Electric and Westinghouse offered fodder for business courses (Freeman, Stewart, & Moriarty, 2009). These scandals (and others like Watergate) mobilized business ethics as its own field of study (DeGeorge, 1987). In the mid-1970s, a group of scholars, mostly at Catholic universities, began to study business from an ethical point of view. The University of Kansas hosted the discipline’s first conference in 1974 (Freeman et al., 2009). This group went on to form the Society for Business Ethics, which has created two academic journals: The Journal of Business Ethics in 1982 and Business Ethics Quarterly in 1991 (Freeman et al., 2009). In 1976 the AACSB began to encourage all member schools to include ethics in the curriculum (Desplaces D, Melchar D, Beauvais L, Bosco S., 2007). By 1979, the AACSB required schools to show evidence of curricular efforts to cover business ethics as part of the accreditation process (Dean & Beggs, 2006).

Despite the encouragement of the AACSB, schools have lagged in their response. A 1982 survey of AACSB institutions showed that approximately 40 percent of the 655 institutions who responded offered a course on business ethics.
and/or the social responsibility of financial corporations (Hosmer, 1985).

Therefore 60 percent of member institutions did not offer any business ethics courses in 1982. A more recent survey of 50 top-ranked business schools reported that nine out of 36 undergraduate business schools required business ethics (Wilhelm, 2005). Recent crises and scandals in business have increased these numbers, yet widespread support for business ethics in the curriculum is slow to arrive.

*Impediments to the Study of Business Ethics*

Support for business ethics in the curriculum is gaining, yet there are several conditions that impede this progress. The first, and most prevalent, is the objectivist philosophy of business as a profit-making enterprise. This philosophy is often attributed to the economist Milton Friedman, whose famous article title speaks for itself: “The Social Responsibility of Business is to Increase its Profits” (Friedman, 1970). Friedman suggests that there is no such thing as business ethics and it would be a waste of time to teach it in business education. Individuals in the business arena should be bound to the same moral and legal codes as people in other sectors, and spaces in business curricula should be reserved for important and relevant topics such as finance and accounting. In
Friedman’s view, courses which focus on ethics are not realistic or appropriate in the business environment (McDonald & Donleavy, 1995).

Though Friedman’s position is not as strongly held as it once was, ethics education in business is still not widespread. A more contemporary opinion is that it should be an “optional extra,” if anything at all (Mathieson & Tyler, 2008). The authors (Mathieson & Tyler, 2008) suggest three common objections to ethics instruction in business programs: (1) Values are already formed before students arrive at the university, (2) Ethics instruction is indoctrination, (3) Anything goes in the business world so it is unrealistic to educate students based on a false hope.

Despite the philosophical and anecdotal criticisms of business ethics, empirical finding support its place in the curriculum. The following section will present arguments to the above criticisms of business ethics, providing empirical support.

Mathieson and Tyler’s first objection is that values are formed at a younger age, therefore business schools nor companies cannot change an individual’s beliefs or behaviors. However, psychological research demonstrates that individuals continue to change throughout their lives. While early-life experiences can be formational and influential, a person can have a different set
of values as they age. Values are shaped by several factors, including cognitive
development, social development, and environment, all of which do change
beyond childhood.

Several studies have demonstrated the impact of cognitive development,
which changes throughout the lifespan, upon moral development (Baxter
Magolda, 2003; Erikson, 1968; Kegan, 1984; King & Kitchener, 1994; Love, 2002;
Perry, 1981). Guthrie (1997) studied why reflective thinkers are better able to
avoid effects such as peer pressure in order to make more reasoned judgments.
Pascarella and Terenzini (1991, 2005) cite several studies which show that
cognitive and moral development increase due to college attendance,
independent of other influences.

The modeling of others, beyond parents or childhood heroes, can also be
inspirational agents for change. Moberg (2000) showed that individuals who are
wrestling with a dilemma can be inspired to act justly by witnessing others who
have also struggled with the same dilemma. Haidt (2000) and Keltner and Haidt
(2003) found that seeing a moral behavior can cause one to improve oneself
morally, and be inspired to think of new and better ideas of what is ethically
possible.
Environments and organizations also influence ethical behaviors. For example, Trevino, Weaver, Gibson, and Toffler (1999) reviewed 2,882 ethics surveys and discerned that companies with ethics and compliance programs significantly lowered unethical and illegal practices, increased collegiality, simplified instances of reporting bad news to supervisors, enhanced morale and commitment to the organization, and positively influenced decision making. McCabe, Trevino, and Butterfield (2001), in a study of undergraduate students, found that peer influences and institutional practices (e.g., honor codes) influenced (positively or negatively depending on the environment) the willingness to cheat. These studies show that environment can influence ethical thinking and behavior.

A second objection to business ethics in the curriculum is the belief that ethics are culturally relative, and teaching ethics can be paramount to indoctrination (Mathieson & Tyler, 2008). However, this ignores several current assumptions. First, certain behaviors are universal. For example, Schwartz (2005) found that following seven business practices are commonly held: trustworthiness, respect, responsibility, fairness, caring, and citizenship. Dautheribes, Kernes, and Kinnier (2000) performed an ethnographic content analysis of moral and religious texts throughout history. They found several
universal beliefs that spanned across cultures and religions, including justice, serving a higher purpose, and respect for others. The golden rule – to treat others as you want to be treated – existed in several texts throughout history and across religions. Second, ethics instruction should be about creating more critical thinkers, which in turn may lead to the resistance of indoctrination (Williams & Dewett, 2005). Third, professors already transmit values in the classroom. A professor who prefers quantitative research methods over qualitative ones is stating her bias with that preference. Ethics should be viewed as a subject that one should aspire to teach from an objective perspective, but should also be aware of their biases. This can be no different than with other subjects.

A third objection to teaching business ethics is that “in the business world, anything goes” (Mathieson & Tyler, 2008, p. 6). This sentiment is closely aligned with the objectivist notion of self-interest. More recently popularized by the likes of Milton Freidman, Ayn Rand, Ronald Reagan and Alan Greenspan, self-interest has its roots in the very individualistic nature of America’s origins. This dominance of valuing self-interest above all has laid way to greed, dishonesty, and relativism. Others use notions of utilitarianism as a way to measure value and worth. Often, ethical matters such as rights and justice are sacrificed for the good of the company and its shareholders. Therefore, business and capitalism
require thinking about profits above all else. If greater good comes from the profits, so be it. But, it is argued, ethics should not override profits.

However, this is not a rational argument against teaching of business ethics. In fact, reviews of financial firms find that there is a link between ethical practices and financial performance. Two separate studies found that companies with ethics codes performed better on measures of market value added and return on capital (Verschoor, 1999; Webley & More, 2003). Reputation is reported by CEOs to be a company’s most important tangible asset (MacMillan & Joshi, 1997). A company creates its reputation through a brand name image. Ethical behavior is an important part of this brand name reputation (Brickley, Smith, & Zimmerman, 2002).

Employees prefer working for ethical companies. Seventy nine percent of employees reported that a firm’s “ethical concerns” were a key reason for why they stayed with their current firms (Josephson Institute, 2004). Other studies have shown that employees prefer working for companies with clear and unambiguous ethical expectations (Gellerman, 1989).

Individuals should pursue business ethics as a natural step towards their humanity. The “anything goes” relativism and cost-benefits utilitarianism is seated in more pre-conventional or conventional ways of thinking. Business
ethics can help students move to higher orders of post-conventional thinking. Once there, they will be able to critically address the structures of capitalism so as not to succumb to its potential for ethical breaches. The post-conventional business person can be one who does well while doing good.

A final, and perhaps most concrete, argument against the “anything goes” mentality is that unethical practices are often severely penalized. For example, legislation in the wake of the Enron scandal created the Sarbanes-Oxley Act. This act punished those who destroyed certain kinds of evidence (20 years in prison), executed fraudulent securities (25 years in prison), or misreported financial information (10 years in prison, and fined up to $1,000,000). More recently, analysts predict that the Securities and Exchange Commission may fine Goldman Sachs more than $1 billion for defrauding investors in a collateralized debt obligation linked to home loans (Westbrook & Scheer, 2010). Considering these costs, it pays to act ethically.

Business Ethics Pedagogy – Best Practices

Despite the arguments in favor of emphasizing business ethics, there seems to be no consensus for how best to teach it (Wilhelm, 2008). A recent longitudinal review of moral reasoning of undergraduate students enrolled in a
business foundations course at a Midwestern university was conducted over five semesters (Wilhelm, 2008). This review recommended several conditions for optimizing moral development in business foundations courses with an ethical component. They included:

1. Teaching foundations in philosophical ethics (e.g., deontology, teleology, virtue theory)

2. Presenting readings that “introduce a step-by-step ethical decision-making framework that is logical, concise, and utilizes the evaluative components from the ethical theories” (Wilhelm, 2008, p. 27).

3. Utilizing case studies (between two and six) in order to practice the ethical frameworks.

4. Requiring graded assignments that include writing and quizzes.

5. Allowing ample time for reflection which extends throughout the semester.

6. Dedicating “considerable effort and attention to case analyses employing an ethical decision-making framework” (Wilhelm, 2008, p. 28) is required of the instructor. However, Wilhelm found that
professors do not need to be specialists in ethics in order for their students to have positive changes in their moral reasoning ability.

Though these conditions are not present in all business ethics courses, they do present pedagogical techniques which attempt to meet intended goals of ethics instruction. The three main goals include building ethical awareness, promoting moral development, and enhancing ethical decision-making (Williams & Dewett, 2005), which align well with both the Portico intervention and Rest’s four component model. This dissertation utilizes the Defining Issues Test as a way of measuring moral reasoning. Several studies, as previously noted, have also utilized the DIT to assess ethics interventions on moral reasoning (see earlier section). However, a missing body of research is the interactions between measures. Understanding these interactions can create better conditions for optimum growth.

Conditions for Moral Development/Class Size

One missing measure in moral reasoning is the impact of class size. Class size is the oldest and most widely studied topic within the research on teaching and learning (McKeachie, 1990). Discussion on class size can be found in the Babylonian Talmud of the sixth century, which explicated the ideal class size for Bible study. “The number of pupils assigned to each teacher is 25. If there are 50,
we appoint two teachers. If there are 40, we appoint an assistant, at the expense of the town” (Epstein, 1976, p. 3). The first empirical study dates back to 1895 (Holland, 1954). From that initial study, by Rice in 1895, to the 1950s there were over 70 more studies on class size, but only 13 on the college level (Monroe, 1950; Von Borgersrode, 1941).

These studies peaked in the 1920s, largely due to the seminal work of Edmonson and Mulder (1924). Their research investigated class size at the University of Minnesota, where they compared the achievements of students in different-sized classes. Edmundson and Mulder found the performance of the two groups to be equivalent (McKeachie, 1990), with students preferring the smaller classes.

These investigations resulted in a large-scale study at the University of Minnesota. “The Committee of Research at the University of Minnesota” conducted 59 experiments in a variety of subjects. Forty-six of the experiments found student performance, measured by classroom exams, to be better in the large class setting (Hudelson, 1928).

After peaking in the 1920’s, single variable experiments on class size continued to be conducted until World War II. At that point, over 1,700,000 World War II veterans had taken intelligence tests (McKeachie, 1990, p. 189). This
proliferation of empirical knowledge, and the resulting confidence of the psychology profession, created opportunities for further research in areas such as college teaching and learning. Vanguards of the Post World War II era included Macomber and Siegel (1957) and Glass and Smith (1979).

More recently, despite being the earliest research question of college teaching, studies on class size have been inconclusive and unclear (Follman, 2009). The vast majority of research on class size is at the elementary school level, with secondary school studies next, and college studies lacking and inconsistent. While college teachers and students have consistently favored small classes (McKeachie, 1980), most studies with the criterion of achievement tests demonstrate no advantage of small classes over large ones (Laughlin, 1976; McKeachie, 1963; Vincent, 1969; William et al., 1985). This combination of “unclear” (Glass & Smith, 1979, p. 2) and contradictory research demonstrates the need for much more work in this area. Furthermore, budgetary concerns require cost-efficient planning. Increasing class size is a partial solution for university administrators, but is there a cost to student learning?

Given the need for ethics in the business world, the dearth of research on curricular conditions for effective ethics instruction, and an incomplete understanding of class size effect, this study will attempt to measure the effect of
class size on the cognitive moral development of first year college students enrolled in a business ethics course.

**Conclusion**

However, “studies that more carefully examine the conditions for growth are needed” (King & Mayhew, 2004). A much needed study should ask whether small classes are better in increasing the moral judgment of college students (King & Mayhew, 2004).

A thorough review of the literature on class size and moral development uncovered no studies that have investigated the impact of class size on the moral development of college students. In fact, only two studies that utilized the DIT even mentioned class size at all (Abdolmohammadi, 2005; Ponemon, 1993). Neither of these studies presented class size as a main effect of the study.

Given the cost of attending higher education, institutions are interested in maximizing economic efficiencies. A popular method for keep costs in check is the increase of class sizes. For example, compare the potential costs for one section of 200 students with 10 sections of 20 students. Or, more appropriate to the realities of this study’s sample, compare the following scenarios presented in table 2.6.
Table 2.6 – A comparison of estimated costs between small and medium-sized sections of a course

<table>
<thead>
<tr>
<th></th>
<th>Small Class</th>
<th>Medium-Sized Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of Students</td>
<td>505</td>
<td>505</td>
</tr>
<tr>
<td>Students per Section</td>
<td>19</td>
<td>26</td>
</tr>
<tr>
<td>Sections</td>
<td>27</td>
<td>20</td>
</tr>
<tr>
<td>Cost per Section (estimated)</td>
<td>$5,200</td>
<td>$5,200</td>
</tr>
<tr>
<td>Total Cost (estimated)</td>
<td>$104,000</td>
<td>$140,400</td>
</tr>
</tbody>
</table>

Considering that this course is but one of many, one can quickly see how increasing class sizes can affect economic efficiencies. But is there a hidden cost to these cost saving measures? Do students learn, both cognitively and affectively, as well in different sized classes? What class size conditions (and other conditions) are most effective? This study attempts to investigate these issues, as it examines the impact of class size on the cognitive moral development of freshmen business ethics students. The next chapter reviews these issues while outlining the research questions, the design of the study, a fuller examination of the research instrument, the sampling method, a description of the sample and populations, the data collection process, methods of data analysis, formats for reporting the data, and frameworks for discussing the findings.
CHAPTER III

RESEARCH METHODOLOGY

The purpose of this study is to examine the impact of class size on the changes in moral reasoning of freshmen business ethics students. This chapter reviews the research question and provides a rationale for the quasi-experimental quantitative design of the study. Reasons also are given for the use of the “Defining Issues Test 2” (DIT2) as the research instrument. Sampling method, the description of the sample and population, and overall data gathering procedures are also reviewed. This chapter concludes with a discussion of the method of data analysis, formats for reporting the data, and frameworks for discussing the findings.

Research Questions

To understand the impact of class size on the moral reasoning of freshmen business ethics students, this study investigates the following questions:

1. Do students enrolled in a one-semester business ethics class have significant gains in moral reasoning, measured over time from the pretest to the posttest?
2. Do students enrolled in smaller sections of a business ethics course have greater gains in moral reasoning scores than students enrolled in larger sections of the same course, taught by the same instructor?

3. What is the relative effect of class size in cognitive moral development when accounting for the variables of gender, academic aptitude, instructor, class time, class size and possible interactions among these variables?

**Research Design**

A true experimental design attempts to answer, *if a program/treatment/experiment is given, then a certain outcome will occur*. Experimental designs use equivalent groups that are randomly selected. When this is not available, quasi-experimental designs may be favorable. While quasi-experimental designs look like experimental designs, they lack the random assignment component. This threatens the internal validity of a study more than when using an experimental study. Nonetheless, this study uses a quasi-experimental group design. This type of design is most appropriate when it includes more than one group, a common measured outcome, and random assignment (Gribbons & Herman, 1997). This study meets all of these qualifications.
The null hypotheses for this study are:

H0₁. There is no statistically significant difference between moral reasoning scores at the beginning and end of a one semester business ethics course.

H0₂. There is no statistically significant difference of the moral reasoning scores between students in small sections and medium sized sections of a business ethics class.

H0₃. Independent variables of instructor, class time, measures of academic ability (SAT scores), gender, and class time are unrelated to N2 score alone or in interaction of class size.

To test the questions and hypotheses, first year business students were surveyed before and after their mandatory freshmen ethic class in the fall of 2009. The students were assigned to one of three instructors, and then further assigned to one of eighteen sections of the business ethics course. These sections are categorized as small (18 to 19 students) or medium-sized (25 to 27 students). Each group completed the DIT2 as both pretest and posttest, with the intervention occurring in between the pretest and posttest. The research design notation is as follows:

\[
\begin{align*}
R_s & : O \quad X \quad O \\
R_m & : O \quad X \quad O
\end{align*}
\]
“Rs” is the group of students randomly assigned to small classes. “Rm” is the group of students randomly assigned to medium-sized classes. “O” is the DIT2 which is used as both the pretest and posttest. X is the intervention of ethics instruction within the course.

Treatment

The treatment for this study is called “Portico.” Portico is a required course which all Boston College business students take in the fall of their freshman year. The goal of the course is two-fold. First, it hopes to introduce students to the study of business, in a broad sense. Second, it offers students the opportunity to “develop a nuanced method for recognizing and responding to the ethical challenges” of the world in which they live (see Appendix C, “Portico Syllabus”).

The basic framework for the class is a funnel in which students move from macro to micro/personal issues. During the first half of the course, students begin with a wide and historically informed consideration of global, national, and regional issues and end with a discussion of industry, organizational, and functional issues. During the second half of the course, the ‘funnel’ narrows and
students consider more personal issues, including ethics, leadership, and personal/professional development.

The choice of readings and assignments is designed to reinforce the interconnections across the levels of the funnel. In the first six weeks, readings focus on economics, globalization and business topics. Assignments are individual and team-based, and include industry analyses, group projects, and essays on globalization and economic development. The second part of the semester shifts to readings in philosophy. The class is taught using a blend of Socratic method and case study, in which students are required to reflect and apply theoretical ethical frameworks to contemporary examples and business cases. Assessment tasks are varied, but include much writing which focuses on personal experiences and require reflective judgments.

The three main goals of business ethics courses are building ethical awareness, promoting moral development, and enhancing ethical decision-making (Williams & Dewett, 2005). These goals mirror the goals of Portico, and are measurable with the DIT2. The goals are also supported through sound pedagogical practices. A longitudinal review of moral reasoning in undergraduates enrolled in a business foundation course at a Midwestern university reported six conditions for optimizing moral development in business
ethics courses (Wilhelm, 2008). They included teaching foundations in philosophy; developing frameworks based on those philosophical foundations; using case studies to practice working within the frameworks; requiring multi-modal assignments; and allowing ample time for reflection. The Portico curriculum was developed with these goals and conditions in mind, and is well-suited for measurement using the DIT2.

Sample

The population for this study consists of all students enrolled in Portico. The university has an enrollment of 504 freshmen in its business program, all of whom are required to enroll in one of the 23 sections of the course. Of the 23 sections, 18 of them are taught by a core group of three full-time instructors. The other five sections are taught by four different adjunct instructors.

The sample consists of 400 students who enrolled in the 18 sections taught by the three core instructors. Five class sections were eliminated from this study so that there is less variance within the variable of “instructor.”

The 18 sections were divided into nine “small” (N=18 students) sections (treatment 1) and nine “medium-sized” (n=26 students) sections (treatment 2). These two course sizes serve as the two treatments which are examined in this
study. Research shows that the ideal class size is between 15 and 18 (Glass & Smith, 1979) and a difference of eight students is a significant one (Finn, 1997). Therefore, these are the sizes of small and medium-sized classes in this study. The differently sized classes are evenly divided among the three instructors teaching the 18 sections (see Table 3.2).

Participation for this study was voluntary and consisted of completing the Defining Issues Test, 2nd version (DIT2) at the beginning and end of the course’s intervention. Only those who completed the survey at both pretest and posttest and who met the reliability criteria were included in the analysis of this study. Reliability criteria are met according to the participant’s responses to test items. Several “checks” exist within the scoring procedure of the DIT2. These checks (called the New Check total score) recognize issues with participant reliability (Rest, et al., 1999). Four issues the New Check total score addresses include random participant responses, too many missing responses, participants who choose answers based on wording instead of meaning, and participants who select the same answer on too many items (Williams, 2005). Each respondent of the DIT2 has a running tally of these issues called the New Check total score. When a score greater than 200 occurs, the participant is purged from the sample (Bebeau & Thoma, 2003).
Overall, the population and sample resembled each other when viewed by gender, SAT scores, and college grade point average. T Tests of the population and sample determined that there was not a statistically significant difference between the population and sample (see Table 3.1).

<table>
<thead>
<tr>
<th>TABLE 3.1</th>
<th>Student gender, academic aptitude, and first-year grades, by population and sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Population (n=504)</td>
</tr>
<tr>
<td>GENDER</td>
<td></td>
</tr>
<tr>
<td>Female (%)</td>
<td>n=138 (27%)</td>
</tr>
<tr>
<td>Male (%)</td>
<td>n=366 (73%)</td>
</tr>
<tr>
<td>SAT SCORES</td>
<td></td>
</tr>
<tr>
<td>Score</td>
<td>1335 (113.36)</td>
</tr>
<tr>
<td>GPA</td>
<td>3.30 (0.44)</td>
</tr>
</tbody>
</table>

Generally speaking, statistical analyses require that n > 30 (Ott, 1988) in order to insure statistical power. For this study, the sample size for each group is determined by $n = \frac{2\sigma^2t^2}{D^2}$ (Borg & Gall, 1989). The standard deviation ($\sigma^2$) used is 13.2 which is the norm for college students (Rest, 2000). t is 2.042 which is the test value of 30 cases with significance at the .05 level (Ott, 1988). D is 8.1, which is the expected change in DIT “N2” scores (Rest, 2000). This demonstrates a minimum sample size of 22, however, statistical analysis requires a sample size
of at least 30. The sample for each group in this study (see table 4) is greater than 30.

<table>
<thead>
<tr>
<th></th>
<th>Instructor 1</th>
<th>Instructor 2</th>
<th>Instructor 3</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Classes</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>168</td>
</tr>
<tr>
<td>Medium Classes</td>
<td>78</td>
<td>77</td>
<td>77</td>
<td>232</td>
</tr>
<tr>
<td>TOTAL</td>
<td>134</td>
<td>133</td>
<td>133</td>
<td>400</td>
</tr>
</tbody>
</table>

**Instrument**

This study uses the Defining Issues Test, 2nd version, created in 1999 to update the Defining Issues Test (DIT). James Rest designed the original DIT at the University of Minnesota in 1979 to measure the moral reasoning of individuals (Rest, 1994). The DIT is a paper and pencil test derived from Kohlberg’s theories of moral development (Kohlberg, 1984). Rest borrowed from Kohlberg’s Moral Judgment Interview (MJI), which used open-ended responses to moral dilemmas (Bebeau & Thoma, 2003). Rather than coding free responses to hypothetical dilemmas given in an interview format, the DIT presents vignettes for a subject to rate and rank. The responses represent the degree in which “schemas” of thinking are activated. The schemas are closely related to
Kohlberg’s stages of moral development (Bebeau & Thoma, 2003), yet different in the following ways (Rest, Narvaez, Thoma & Bebeau, 2000):

1. Hard stages versus soft schema. Rest’s schema are seen as “shifting distributions rather than a staircase” (p. 384).

2. Schemas are more concrete and specific than stages.

3. Cognitive operations vs. content of operations. While Kohlberg stressed operations of moral thinking, Rest focused on content.

4. Universality. “Kohlberg postulated universality as a characteristic of stages whereas we regard cross-cultural similarity as an empirical question” (p. 385).

5. Articulation vs. tacit knowledge. Rest used multiple choice responses in order to test that reasoning was part of an individual’s discernment but not based on their ability to articulate open-ended responses.

In his book, Moral Development: Advances in Research Theory, Rest (1986) describes the DIT as follows:

The DIT is based on the premise that people at different points of development interpret moral dilemmas differently, define the critical issues of the dilemmas differently, and have different intuitions about what is right and fair in a situation. Differences in the way that dilemmas
are defined therefore are taken as indications of their underlying
tendencies to organize social experience. These underlying structures of
meaning are not necessarily apparent to a subject as articulate rule
systems of verbalizable philosophies – rather, they may work “behind the
scenes” and may seem to a subject as just commonsensical and intuitively
obvious (p. 196).

The responses in the DIT are carefully selected to reflect the schema of
moral reasoning. After reading a vignette that includes a moral dilemma,
subjects must choose an action from the story. For example, in the famous
“Heinz” dilemma, a man must choose whether or not to steal medicine for his
dying wife. Next, a respondent considers to what extent twelve issues are
important, rating the responses from 1 to 5 (“Great,” “much,” “some,” “little,” or
“no” importance, respectively). Finally, a respondent ranks the four (out of
twelve) most important issues. Rest and his colleagues (Rest, Narvaez, Thoma,
and Bebeau, 1999a) explain how the DIT works as follows:

The DIT is a device for activating moral schemas (to the extent that a
person has developed them) and for assessing them in terms of
importance judgments. . . . As the subject encounters an item that both
makes sense and taps into his or her preferred schema, that item is rated and ranked as highly important. Alternatively, when the subject encounters an item that either doesn’t make sense or seems simplistic and unconvincing, the item receives a low rating and is passed over for the next item. . . . In the DIT, we are interested in knowing which schemas the subject brings to the task (are already in his or her head). Presumably, those are the schemas that structure and guide the subject’s thinking in decision making beyond the test situation (p. 6).

Though Rest and his colleagues often considered updating the DIT in various ways, they opted for stability and left it unchanged (Rest et al., 1999a). However, after 25 years of use, the vignettes became outdated. The Defining Issues Test Version 2 (DIT2) is a revision of the DIT. Compared to the DIT, the DIT2 has updated stories, is a shorter test, has clearer instructions, retains more subjects, and has more validity (Rest et al., 1999b).

The DIT2 has five contemporary vignettes and is scored in the same manner as the DIT. The five moral dilemmas are followed by the selection of an action, twelve issues to rate when choosing an action, and the selection and ranking of the top four issues in choosing an action. Individual responses are
aggregated into scores. The two most widely researched scores are the P score and the N2 score (Bebeau & Thoma, 2003).

**Scoring the Instrument: The P Score**

The P score is the original measurement of the DIT which represents the proportion of responses that reflect principled or postconventional (P) moral reasoning (King and Mayhew, 2002). The P score is computed by adding a respondent’s postconventional responses, and then weighting these ranks in order to calculate a score. (Bebeau & Thoma, 2003; Rest et al., 1997). For example, the DIT2 has five dilemmas. Each dilemma has 12 items (see appendix A). The participant’s task is to rate each item in terms of how important it is, then to rank their four most important items. The P score is calculated based only on the ranking data. If a participant ranks a postconventional item within their top four (generally, four of the twelve items per dilemma are postconventional), then the score is weighted as seen in Table 3.3.
TABLE 3.3
Weights for calculating the P score in the DIT2

<table>
<thead>
<tr>
<th>Rank</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Total possible score</td>
<td>10</td>
</tr>
</tbody>
</table>

Since there is a maximum of ten points for each dilemma, the maximum base score is 50 (ten points times five dilemmas equals 50). This score is then converted to a percentage which becomes the P score. For example, a participant with a total weighted score of 40 would convert to a P score of 80 (percent, because 40 out of a possible 50 points is equal to 80 percent). The range of P scores is 0 to 95 (note: not 100 because some dilemmas do not have four possible principled items in them).

The P score, despite constant research efforts to find a better measure, survived as the most valid and reliable measure for the first 20 years of the DIT. It consistently offered better trends than other measures, was easy to compute, and simple to interpret (Rest et al., 1997). However, it was not without its detractors. The main criticism of the P score is that much of the data are thrown away. Since the P score is computed using the rankings of the DIT, all of the...
ratings data is discarded. Another criticism concerns missing data. If a participant leaves a ranking item blank (i.e., ranks two or three items in a dilemma rather than four), then the score is recalculated based on the responses that are provided. For example, the maximum weighted score may become a 44 rather than a 50, and that score is then converted into a percentage. However, critics argue that omitted items can be the same as not prioritizing a P item, but the P score does not account for this.

**Scoring the Instrument: The N2 Score**

The N2 Index is a newer version (hence N2) of the index score. The score, which was introduced in 1997, combines the percentage of post-conventional thinking with the rejection of simplistic thinking, both of which are desirable (Bebeau & Thoma, 2003). While the P score only uses the ranking data in computing a score, the N2 uses both ranking and rating data. Therefore, an N2 score is calculated using two parts: the prioritization of principled items and the degree in which lower stage items are rated lower than the higher stage items (Rest et al., 1997).

The first part is calculated almost identically to the P score, except that no adjustment is made for missing items. The second part of the N2 score is based
on rating, not ranking, data which solves the problem of discarded data. The researchers use the idea of “discrimination” (Rest et al., 1997), which is a calculation of the difference averages given to lower stage (stages 2 and 3, which are preconventional or conventional) items and higher stage (stages 5 and 6, which are postconventional) items. To calculate the score, the average ratings of lower stage responses are subtracted from the average ratings of higher stage responses. The ratings are standardized by dividing the difference by the difference of the participant’s standard deviation of higher and lower stage ratings. These two parts – using ranking and rating data – are then computed into one score for each participant by adding the ranking score to the rating score when weighted by three. The researchers “weight the discrimination component by three because this component has about 1/3 the standard deviation of the P scores; therefore weighting equalizes the two parts of the N2 score” (Rest et al., 1997, p. 501). N2 scores are adjusted by this weighting, so that they are in the same range of scores as P scores (0 to 95) (Bebeau & Thoma, 2003).

In conclusion, the N2 index permits more subjects to pass subject reliability checks, which allows sample retention (Rest et al., 1999a). In a comparison of studies of N2 to P scores, Rest et al. (1997) concluded that the N2 index is better than the P index, largely due to the rejection of simplistic thinking.
and the retention of more data. The new checks within the N2 scoring method include guards against random responding, tolerate more missing data, check against participants who choose responses based on tricky or lofty words rather than the meaning of the item, and require participants to be discriminating in their responses. When scored, each participant has a running tally of checks if their responses fall into one of the above four problem areas. Participants with New Check scores of greater than 200 are purged from the sample.

Overall, these new checks allow for more measures being tested and less discarded data. When analyzing DIT results, Rest et al. found that the N2 score purges less participants than the P score. In fact, the N2 participant reliability checks allow for over 96 percent of responses to be used. When scoring the same sample with the P score method, only 77 percent of responses passed the previous checks used for P scores (Rest et al., 1999a). Therefore, the N2 index is used in this study.

**Validity**

The DIT has been assessed in terms of seven criteria set forth by Rest and his colleagues in order to operationalize construct validity for the instrument (Rest, Narvaez, Thoma, & Bebeau, 2000, p. 390). These seven criteria must be met
in order for a test of moral development to be valid (Rest et al., 1999a; Rest et al., 2000). First, scores must differentiate across educational levels. In fact, a meta-analysis of a large composite of samples show that 30% to 50% of the variance of DIT scores is attributable to level of education. Second, gains occur over time. A review of a dozen studies of freshmen to senior college students (n = 755) show effect sizes of .80, “making gains in DIT scores one of the most dramatic effects of college” (Bebeau & Thoma, 2003). Third, the DIT is a cognitive developmental measure, which is significantly related to moral comprehension (r = .60s) (Bebeau & Thoma, 2003). Fourth, the DIT is affected by educational interventions. A meta-analytical review of 55 interventions examined DIT scores between experimental and control groups (Schlaefli, Rest, & Thoma, 1985). The comparison of dilemma discussion interventions with control groups shows that the intervention groups have moderate gains in moral reasoning scores, while the comparison groups show little to no gains. Fifth, the DIT is connected to desirable decision-making in professional settings. Sixth, the DIT is correlated to political attitudes. When DIT scores are coupled with political attitudes, the correlate closely (r = .40 to .65) and predict “up to two-thirds of the variance of controversial public policy issues (such as abortion, religion in the public school, women’s roles, rights of the accused, rights of homosexuals, free speech issues)”
(Bebeau & Thoma, 2003). Seventh, reliability is adequate. Cronbach alpha is in the upper .70s/low .80s. In addition, several innovations within the DIT, such as the creation of the DIT2, have been made to increase the power of the validity of the measurement (Rest et al., 2000).

Compared to the DIT, the DIT2 has the advantage of being shorter, more contemporary, and clearer. In addition, a study of 200 participants across four educational levels who took both the DIT and DIT2 compared those results (Rest et al., 1999b). The findings led the researchers to encourage the substitution of the DIT2 for the DIT due to a high correlation of the DIT with the DIT2 ($r = .79$). The Cronbach’s alpha – an index of the internal consistency of a test based on the extent to which test-takers who answer a test item one way respond to other items in the same way (Gall, Borg, and Gall, 1996) – of the DIT2 is .90 (Rest et al., 1999b). This compares favorably to the internal consistency of the original DIT, which had a Cronbach’s alpha of .76 (Rest, 1988). For these reasons, the DIT2 has become the preferred option for measuring moral development and is utilized in this study.
### Variables

The dependent variable in this study is the students’ change in N2 scores over time (from pretest to posttest). N2 scores are a score of the DIT2 which measures the cognitive moral reasoning of respondents (as further described in the following section). The independent variables for this study are instructor, class size, class time, measures of academic ability (SAT scores and grade point average), gender, and political viewpoints. Table 3.4 presents a full list of variables.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Variable type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2 Difference</td>
<td>Dependent</td>
<td>The measured change in N2 scores over time (pretest to posttest)</td>
</tr>
<tr>
<td>Instructor</td>
<td>Independent</td>
<td>The three instructors were coded 1, 2, or 3</td>
</tr>
<tr>
<td>Class Size</td>
<td>Independent</td>
<td>Class size was either small or medium, coded 1 or 2 respectively</td>
</tr>
<tr>
<td>Class Time</td>
<td>Independent</td>
<td>Class times were coded as early morning (code =1), late morning (2), early afternoon (3), or late afternoon (4)</td>
</tr>
<tr>
<td>SAT Scores</td>
<td>Independent</td>
<td>SAT scores in the top third of each were coded 1, SAT scores in the middle third were coded 2, and SAT scores in the bottom third were coded 3</td>
</tr>
<tr>
<td>GPA</td>
<td>Independent</td>
<td>GPAs in the top third of each were coded 1, GPAs in the middle third were coded 2, and GPAs in the bottom third were coded 3</td>
</tr>
<tr>
<td>Gender</td>
<td>Independent</td>
<td>Gender was coded using the convention of the center which scored the data: 1 = male, 2 = female</td>
</tr>
<tr>
<td>Conlib</td>
<td>Independent</td>
<td>Political viewpoints was assessed using a five-point Likert scale from 1 (very liberal) to 5 (very conservative)</td>
</tr>
</tbody>
</table>
Data Collection Procedures

This study conducts a secondary analysis of an existing dataset collected by the Office of the Dean for Undergraduates. The original collection used the DIT2 as both pretest and posttest. The DIT2 was administered at Boston College as both a pre-test and post-test. The pre-test was administered in class the week of October 26, 2009. This was the week prior to the ethical interventions of the course. The post-test was administered the week of December 9, 2009, which was the last week of the semester. A human subjects exemption form was submitted and approved for collection of data.

There were 18 sections of the course in which the DIT2 was administered. The primary investigator visited each of these classes to administer the instrument, and participation was voluntary. Eighty-six percent of the sample of 400 students completed the pretest (n=342). Materials (informed consent form, #2 pencils, DIT2 instructions, DIT2 answer sheet) were distributed, then, the informed consent form was reviewed before the test was administered according to its instructions. The DIT2 took approximately 30-35 minutes to complete. The informed consent form was attached to the answer sheet. Then, the primary investigator assigned a five digit identification number to each answer sheet. At that point, the primary investigator had three sets of data. First were the answer
sheets with unique and non-descript identification numbers. Second were the informed consent forms. Third was a spreadsheet with student names and the corresponding five digit identification number. These three sets were saved in different and secure locations. The answer sheets were secured in a file cabinet and eventually mailed to the "Center for the Study of Ethical Development" at the University of Alabama for scoring. The informed consent forms were filed in a locked cabinet of the researcher’s desk. The spreadsheet was saved on a password protected electronic storage device.

In December 2009, the post-test was administered to the students who completed the pretest (342). Participation was voluntary, and 89 percent completed the posttest (n=303). The students and primary investigator repeated the test-administrations steps. The spreadsheet was used to assign a matching five digit identification number to the answer sheets. The answer sheets were immediately mailed to "The Center for the Study of Ethical Development." The informed consent forms and spreadsheet were securely stored as stated above.

Data Analysis Procedures

"The Center for the Study of Ethical Development" scored the forms and sent an electronic file to the researcher. The data were imported into an SPSS file,
merged with demographic data, and then used to analyze data. The analysis was two-part. First, descriptive findings were reported. This included a comparison to national norms, as well as an investigation of the descriptive findings of N2 scores, reported for each independent variable. The second part of the analysis assessed the significance of the finding. A paired samples T Test was performed to see if there was a statistically significant difference between the pre-test and post-test N2 scores. Then, two-way ANOVA was used to show differences in N2 scores for the factors of class size and instructors. In this case, two-way repeated measures tests were conducted, to assess change over time, per case. Finally, N2 scores were analyzed in light of any descriptive findings in order to understand the statistical significance of the independent variables such as class time.

**Limitations**

The possibility of threats to internal and external validity exists. Factors that may affect internal validity include testing, maturation, and attrition (Campbell & Stanley, 1966). Testing threats may be caused by the participants’ pre-testing, which may cause them to “learn enough from the pretest to improve performance on the posttest” (Charles & Mertler, 2002, p. 334). The DIT2 does not have clearly “better” answers so this threat was limited. The effects of
maturation could threaten internal validity. Potentially, growth within the moral domain may be due to factors outside of this class. A comparison group would have been helpful in minimizing this threat. However, since the entire class was required to take Portico, there was no available comparison group. To combat this limitation, the study results were compared to a national group in order to test the significance of possible gains in moral reasoning scores. This is further discussed in Chapter four.

Another limitation is that students may drop out of both or either group, causing internal validity to be threatened by attrition (Campbell & Stanley, 1966). Records were kept to help track attrition or mortality. Attendance was required of the course, and no student missed more than three classes. Only two students, from an original population of over 500, withdrew from this course; both due to family emergencies.

Ecological validity (Bracht & Glass, 1968) requires the setting of the groups to be similar. I studied all students who participated in the fall 2009 Portico course. This sample was composed of several smaller groups, each which underwent a standard curriculum, yet with a potentially different dynamic. Therefore, “some groups of participants, especially when involved in innovations, develop a group spirit that motivates high achievement” (Charles &
Mertler, 2002, p. 335). This may have influenced performance, as participants might have responded more favorably to the instrument since they were being studied. However, the DIT2 instrument offers a variety of responses, with no clear “right” or “wrong” answers. Furthermore, internal consistency checks occur throughout the DIT2 to ensure that responses are consistent throughout the instrument.

Another limitation involves the ability to draw conclusions about a population from a sample. For this study, the sample is restricted to undergraduate business students in the first year of their college study. Moreover, the sample is participating in a first semester introductory business course. Therefore, “caution must be used when generalizing to other populations, even other undergraduate business students” (Adkins, 2009, p. 60).

Furthermore, this study observes one institution and one curriculum of the institution over a short time period. Therefore, possible institutional effects (self-selection of students who attend Boston College, with a certain mission and ethos) cannot isolate this treatment from other experiences. This is also a one semester study, with no longitudinal follow-up. Finally, the intended outcome of the curriculum is principled moral action, but principled moral reasoning is only part of moral action (Rest, et al., 1999b). In other words, knowing the right thing
to do may not always lead to doing the right thing. Three separate reviews of the literature demonstrate that measures of moral judgments and actions are related, but only weakly (Arnold, 1989; Blasi, 1980; Thoma & Rest, 1986). So, if moral action is the intended outcome, then studies which specifically measure moral action must occur. Further research on this topic is merited, and should be a long-term goal of the community.
CHAPTER FOUR

PRESENTATION OF THE FINDINGS

As stated earlier, the purpose of this study was to measure the changes in moral reasoning of freshmen business ethics students. The Neo-Kohlbergian theories of moral reasoning as described by James Rest framed the research and informed the research questions. In order to measure changes in moral reasoning, the Defining Issues Test, version 2 (DIT2) was used. Chapter three described the research methodologies followed in investigating this quasi-experiment.

This chapter presents the descriptive data, statistical findings, and multivariate results; and is divided into five sections. Section one prepares the reader for the analysis by describing the population, samples, and variables. Section two presents a descriptive analysis of the respondents. The subsequent three sections present results of the hypothesis testing for each of the three research questions.

Section One: Population, Sample, and Variables

This quasi-experimental study was conducted during the fall 2009 semester at Boston College. It assessed the moral reasoning of freshmen business
students enrolled in an ethics class (Portico, which is described in Chapter three). The population for this study consisted of all students enrolled in Portico. The university had an enrollment of 504 freshmen in its 2009 business program, all of whom were required to enroll in one of the 23 sections of the course. Of the 23 sections, 18 of them were taught by a core group of three full-time instructors. The other five sections were taught by four different adjunct instructors.

The study population consisted of all 400 students who enrolled in the 18 sections taught by the three core instructors. The additional five class sections were eliminated from the overall population so that there was less variance within the variable of “instructor.” Participation for this study was voluntary and consisted of completing a survey (the DIT2) at the beginning and end of the course’s intervention. Only those who completed the survey at both pre-test and post-test and who met the reliability criteria were included in the analysis of this study. A further description of the sample is as follows.

Eighty-six percent of the study population of 400 students completed the pre-test (n=342). The post-test was administered to the students who completed the pre-test (n=342). Participation was again voluntary, and 89 percent completed the post-test (n=303). The DIT2 contains reliability checks (called the New Check total score) that recognize issues with participant reliability (Rest, et al., 1999).
These checks were run on the 303 matching sets of data. Four issues the New Check total score addresses include:

the problem of participants responding to the question randomly, the problem of too many missing responses, the problem of participants selecting items based on wording instead of meaning, and the problem of participants selecting the same answer for all or many items (Williams, 2005, pp. 76-77).

Each respondent of the DIT2 has a running tally of these issues called the New Check total score. When a score greater than 200 occurs, the participant is purged from the sample (Bebeau & Thoma, 2003). Thirty-five participants’ scores were purged from this study due to having a New Checks total score over 200.

Consequently, usable pairs of surveys (pre-test and post-test) were obtained from 268 participants. Table 4.1 compares the gender, SAT scores, GPA, and pre-test N2 scores of the total population to the study population and sample. An analysis of these groups shows no statistically significant differences between the three groups.
TABLE 4.1
Student gender, academic aptitude, and grades, by group*

<table>
<thead>
<tr>
<th></th>
<th>Total Population (n=504)</th>
<th>Study Population (n=400)</th>
<th>Sample (n=268)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENDER</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female (%)</td>
<td>n=138 (27%)</td>
<td>n=109 (27%)</td>
<td>n=78 (29%)</td>
</tr>
<tr>
<td>Male (%)</td>
<td>n=366 (73%)</td>
<td>n=291 (73%)</td>
<td>n=191 (71%)</td>
</tr>
<tr>
<td>SAT SCORES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score</td>
<td>1335 (113.36)</td>
<td>1332 (116.24)</td>
<td>1338 (111.48)</td>
</tr>
<tr>
<td>GPA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>3.30 (0.44)</td>
<td>3.28 (0.45)</td>
<td>3.33 (0.44)</td>
</tr>
<tr>
<td>PRE-TEST</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 Score</td>
<td>NA</td>
<td>35.15 (13.47)</td>
<td>35.44 (13.15)</td>
</tr>
</tbody>
</table>

*Total population includes all first year students in the course; study population = all students in eligible course sections; and sample are those who completed pre and post-test with reliable responses.

Internal reliability of the sample of participants passing the built in checks was performed by analysis of the N2 scores for each of the five stories making up the DIT2. As stated in Chapter three, reliability for the DIT2 when measured by Cronbach’s alpha is in the upper .70s to low .80s. This was based upon a sample that contained a range of educational levels from junior high school to graduate school (Bebeau & Thoma, 2003). The DIT2 guide advises that if a sample does not contain the entire range of educational levels, the reliability is likely to be lower. The sample in this study had a Cronbach’s alpha of .74, in line with what was expected given that all participants are at the same educational level. It also compares favorably to the internal consistency of the original DIT, which had a Cronbach’s alpha of .76 (Rest, 1988).
DIT2 data for the 268 participants were merged into one file with demographic data from Boston College’s student information systems (BCSIS). N2 scores from the DIT2 were combined with six variables from the BCSIS for this study. The following describes the data and their coding procedures.

The three dependent variables in the analysis were pre-test, post-test and differences (from pre to post) of N2 scores. A series of categorical independent variables were used in the analysis for selected background characteristics and course experiences. These included gender, class size, instructor, class time, SAT scores, and students’ GPAs. Table 4.2 presents a description of the variables.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Variable type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2 Pre-test</td>
<td>Dependent</td>
<td>The measured N2 pre-test scores</td>
</tr>
<tr>
<td>N2 Post-test</td>
<td>Dependent</td>
<td>The measured N2 post-test scores</td>
</tr>
<tr>
<td>N2 Difference</td>
<td>Dependent</td>
<td>The measured change in N2 scores over time (pre-test to post-test)</td>
</tr>
<tr>
<td>Instructor</td>
<td>Independent</td>
<td>The three instructors were coded 1, 2, or 3</td>
</tr>
<tr>
<td>Class Size</td>
<td>Independent</td>
<td>Class size was either small or medium, coded 1 or 2 respectively</td>
</tr>
<tr>
<td>Class Time</td>
<td>Independent</td>
<td>Class times were coded as early morning (code =1), late morning (2), early afternoon (3), or late afternoon (4)</td>
</tr>
<tr>
<td>SAT Scores</td>
<td>Independent</td>
<td>SAT scores in the top third of each were coded 1, SAT scores in the middle third were coded 2, and SAT scores in the bottom third were coded 3</td>
</tr>
<tr>
<td>GPA</td>
<td>Independent</td>
<td>GPAs in the top third of each were coded 1, GPAs in the middle third were coded 2, and GPAs in the bottom third were coded 3</td>
</tr>
<tr>
<td>Gender</td>
<td>Independent</td>
<td>Gender was coded using the convention of the Center: 1 = male, 2 = female</td>
</tr>
</tbody>
</table>
Descriptive Analysis of the Participants

The sample for this study includes a total of 268 students, all freshmen at Boston College. The average SAT score of the students is 1338, and the average grade point average is 3.33. The sample includes 78 females and 190 males. One hundred and fifty two students enrolled in the medium sized classes and 116 students are in the small classes. Further comparisons show the demographics by class size and instructor in tables 4.3 and 4.4.

<table>
<thead>
<tr>
<th>TABLE 4.3</th>
<th>Gender, SAT scores, and Freshmen GPA of Portico students, by class size</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENDER</td>
<td>Total (n=268)</td>
</tr>
<tr>
<td>Female (%)</td>
<td>n=78 (29%)</td>
</tr>
<tr>
<td>Male (%)</td>
<td>n=190 (71%)</td>
</tr>
<tr>
<td>SAT SCORES</td>
<td>Score</td>
</tr>
<tr>
<td>Score</td>
<td>1338 (111.48)</td>
</tr>
<tr>
<td>GPA</td>
<td>GPA</td>
</tr>
<tr>
<td>GPA</td>
<td>3.33 (0.44)</td>
</tr>
</tbody>
</table>

When students are grouped by instructor, the gender ratios, SAT scores, and GPAs are again similar to each other. SATs range from 1331 to 1353 and GPAs ranged from 3.30 to 3.36. Gender ratios and size of the groups were more varied, yet still not statistically significantly different (see Table 4.4).
TABLE 4.4
Gender, SAT scores, and Freshmen GPA of Portico students, by instructor

<table>
<thead>
<tr>
<th>GENDER</th>
<th>Total (n=268)</th>
<th>Instructor 1 (n=105)</th>
<th>Instructor 2 (n=84)</th>
<th>Instructor 3 (n=79)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female (%)</td>
<td>n=78 (29%)</td>
<td>22 (21%)</td>
<td>31 (37%)</td>
<td>25 (31%)</td>
</tr>
<tr>
<td>Male (%)</td>
<td>n=190 (71%)</td>
<td>83 (79%)</td>
<td>53 (63%)</td>
<td>54 (69%)</td>
</tr>
<tr>
<td>SAT SCORES</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score</td>
<td>1338</td>
<td>1333</td>
<td>1331</td>
<td>1353</td>
</tr>
<tr>
<td>(Standard Deviation)</td>
<td>(111.48)</td>
<td>(122.73)</td>
<td>(101.75)</td>
<td>(105.65)</td>
</tr>
<tr>
<td>GPA</td>
<td>3.33</td>
<td>3.35</td>
<td>3.30</td>
<td>3.36</td>
</tr>
<tr>
<td>(Standard Deviation)</td>
<td>(0.44)</td>
<td>(0.47)</td>
<td>(0.39)</td>
<td>(0.42)</td>
</tr>
</tbody>
</table>

Significance testing revealed no statistically significant differences between the total (n=268) and the subgroups of class size and instructor. A z-test for two proportions was run to compare gender ratios between groups. Results showed there was no statistically significant difference by gender at either the p ≤ .05 or p ≤ .01 levels. A student’s T test was run to compare the sample SAT and GPA scores with the SAT and GPA scores by class size and instructor. Again, results showed there was no statistically significant difference by SAT scores or GPA at either the p ≤ .05 or p ≤ .01 levels.

N2 scores

The N2 score represents the degree to which an individual uses higher order moral reasoning. It is the proportion of items selected on the Defining Issues Test – Version 2 (DIT2) that appeal to moral ideals and/or theoretical
frameworks for resolving complex moral issues. This study utilizes three
different N2 scores – pre-test scores, post-test scores, and the difference between
the pre-test and post-test. The following sections present the N2 score results of
Portico students; first as they compare to the national sample, and then by the
independent variables used in the study.

Norms for DIT2 N2 Scores

The DIT2 was developed by the Center for the Study of Ethical
Development in 1998. In the first five years of the DIT2, there were 176 datasets
scored by the Center, which has maintained a database of the findings of these
studies. This database is able to ensure consistent scoring and reporting of scores
between the studies, allows for a sufficient sample size, and is able to sort
responses into demographic categories (such as educational level) (Bebeau &
Thoma, 2003).

Bebeau, Maeda, and Tichy-Reese (2003) conducted a meta-analysis of the
176 studies, in order to generate normative information for N2 scores. There
were 10,870 responses in the analysis, and the findings are reported in Table 4.4a.
TABLE 4.4a - DIT2 Means and Standard Deviations for N2 Scores by Educational Level

<table>
<thead>
<tr>
<th>Year in College</th>
<th>Mean</th>
<th>N2 SCORE Standard Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freshman</td>
<td>31.05</td>
<td>14.42</td>
<td>2096</td>
</tr>
<tr>
<td>Sophomore</td>
<td>31.24</td>
<td>14.94</td>
<td>1028</td>
</tr>
<tr>
<td>Junior</td>
<td>32.65</td>
<td>16.04</td>
<td>1,333</td>
</tr>
<tr>
<td>Senior</td>
<td>36.85</td>
<td>15.53</td>
<td>2,441</td>
</tr>
</tbody>
</table>

According to Bebeau and Thoma (2003), "this study enables researchers to compare N2 scores (the better index of moral judgment) with age and education means from a larger data set" (p. 41). The normative data show that the average N2 score of college freshmen is 31.05, which increases to a score of 31.24 by the sophomore year. The Portico group had significantly higher scores, as seen in Table 4.5.

TABLE 4.5 - N2 scores for Portico students, compared to the national sample

<table>
<thead>
<tr>
<th></th>
<th>N2 Pre-test For Portico Students</th>
<th>N2 Post-test for Portico Students</th>
<th>N2 Scores for National Sample of Freshmen</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>268</td>
<td>268</td>
<td>2096</td>
</tr>
<tr>
<td>Mean</td>
<td>35.37</td>
<td>40.16</td>
<td>31.05</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>13.42</td>
<td>15.43</td>
<td>14.42</td>
</tr>
<tr>
<td>N2 t Test</td>
<td>4.65** (df = 2362)</td>
<td>9.66** (df = 2362)</td>
<td></td>
</tr>
</tbody>
</table>

** = A comparison of N2 scores between Portico students and the national sample of freshmen are statistically significantly different at the .01 level.
Independent sample T tests are used to test for significance between non-matching groups. Since the national sample and the Portico sample are not matching groups, an independent sample T test compared the N2 scores of the two groups. Compared to the national sample (Bebeau, Maeda, & Tichy-Reese, 2003), both the pre-test and the post-test N2 scores of the Portico group are statistically significantly different than the normative N2 scores of college freshmen. In fact, N2 scores of Portico freshmen most closely compare to the normative data scores of college seniors and graduate students.

A more in-depth descriptive analysis of N2 scores follows, and looks at N2 score means by gender, instructor, class size, class times, SAT scores, and grade point averages.

*N2 Scores by Gender*

When looking at the means of N2 scores by gender, females started and finished with higher scores than males, and also had a greater difference in N2 scores (see Table 4.6). When comparing N2 scores between males and females by using an independent samples t test, there was a statistically significant difference in both the pre-test and post-test results between men and women, but
the overall change from pre-test to post-test (the N2 difference score) is not statistically significantly different between males and females.

| TABLE 4.6 |
| N2 Score Means and Standard Deviations for Portico students, by gender |

<table>
<thead>
<tr>
<th></th>
<th>Female (n = 78)</th>
<th>Male (n = 190)</th>
<th>Gender t test (df = 266)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2 Pre-test</td>
<td>38.68 (14.14)</td>
<td>34.01 (12.90)</td>
<td>-2.62 **</td>
</tr>
<tr>
<td>N2 Post-test</td>
<td>44.98 (14.48)</td>
<td>38.19 (15.40)</td>
<td>-3.34 **</td>
</tr>
<tr>
<td>N2 Difference</td>
<td>6.30 (12.69)</td>
<td>4.18 (12.76)</td>
<td>-1.24</td>
</tr>
</tbody>
</table>

** = Significant at .01 level

N2 Scores by Instructor

Instructors all had N2 gains in their classes from pre-test to post-test (see Figure 4.1), and their scores were relatively close to one another.
Pre-test scores were all within 1/10th of a standard deviation from the mean, as were post-test scores and N2 differences. So, despite an increase in N2 scores within groups, none of these scores (see Table 4.7) were statistically significantly different from one another between groups.
TABLE 4.7
N2 Score Means and Standard Deviations for Portico students, by instructor

<table>
<thead>
<tr>
<th></th>
<th>Instructor 1 (n = 105)</th>
<th>Instructor 2 (n = 84)</th>
<th>Instructor 3 (n = 79)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2 pre-test</td>
<td>35.29 (14.56)</td>
<td>34.06 (13.17)</td>
<td>36.86 (12.04)</td>
</tr>
<tr>
<td>N2 post-test</td>
<td>40.73 (15.75)</td>
<td>38.56 (15.66)</td>
<td>41.12 (14.79)</td>
</tr>
<tr>
<td>N2 difference</td>
<td>5.44 (11.97)</td>
<td>4.50 (14.65)</td>
<td>4.26 (11.68)</td>
</tr>
</tbody>
</table>

**N2 Scores by Class Size**

When looking at the N2 scores by class size, both small and medium classes had gains in N2 scores. Medium-sized classes started and ended with higher N2 scores, but smaller classes had greater gains in N2 scores, as seen in Table 4.8, and further depicted in Figure 4.2.
Table 4.8
N2 Score Means and Standard Deviations for Portico students, by class size

<table>
<thead>
<tr>
<th></th>
<th>Small Classes (n = 116)</th>
<th>Medium Classes (n = 152)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2 pre-test (SD)</td>
<td>32.58 (13.03)</td>
<td>37.49 (13.36)</td>
</tr>
<tr>
<td>N2 post-test (SD)</td>
<td>38.22 (14.84)</td>
<td>41.65 (15.74)</td>
</tr>
<tr>
<td>N2 difference (SD)</td>
<td>5.64 (12.14)</td>
<td>4.15 (13.20)</td>
</tr>
</tbody>
</table>

Note: One way ANOVA of scores by instructor were not statistically significantly different.

Figure 4.2 shows that both groups had increases in N2 Scores from pre-test to post-test. And, while the medium sized classes had higher N2 scores, the N2 Score difference was greater in the small classes. However, one way ANOVA of scores by instructor were not statistically significant different.
All class times, whether taught in the morning or afternoon, experienced gains in N2 scores. The greatest gains were in 8am and 9am classes, which saw an N2 difference of 6.80 from pre-test to post-test. Later morning classes also had N2 score gains of over 6, while both afternoon time groups had N2 score gains around 3 (see Table 4.9). Differences in pre-test, post-test, and N2 difference were statistically significant within and between class times and will be discussed more deeply later in this chapter.
TABLE 4.9
N2 Score Means and Standard Deviations for Portico students, by class time

<table>
<thead>
<tr>
<th>Class Time</th>
<th>8am, 9am (n=69)</th>
<th>10am,11am, 12pm (n=64)</th>
<th>1pm, 2pm (n=78)</th>
<th>3pm, 4pm (n=57)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2 pre-test (SD)</td>
<td>37.77 (14.58)</td>
<td>30.67 (13.88)</td>
<td>36.70 (11.75)</td>
<td>35.90 (12.59)</td>
</tr>
<tr>
<td>N2 post-test (SD)</td>
<td>44.57 (15.61)</td>
<td>36.90 (14.38)</td>
<td>40.07 (15.94)</td>
<td>38.62 (14.77)</td>
</tr>
<tr>
<td>N2 difference (SD)</td>
<td>6.80 (10.89)</td>
<td>6.23 (13.18)</td>
<td>3.37 (14.01)</td>
<td>2.72 (12.30)</td>
</tr>
</tbody>
</table>

N2 Scores by SAT scores and GPA

When looking at N2 scores by SAT scores and GPA scores, both have similar trends. The higher the SAT and GPA scores, the higher the N2 scores. In both cases, the highest tercile scores have the highest N2 pre-test and post-test scores. However, for SAT scores, the top and bottom terciles have the greatest N2 score gains. Results for SATs and GPAs are found in tables 4.10 and 4.11, respectively.
For GPAs, the middle tercile has the greatest gains, which is the inverse of SAT scores.

**TABLE 4.10**

N2 Score Means and Standard Deviations for Portico students, by SAT terciles

<table>
<thead>
<tr>
<th>SAT Group</th>
<th>Top Third (98)</th>
<th>Middle Third (85)</th>
<th>Bottom Third (84)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2 pre-test (SD)</td>
<td>38.27 (13.85)</td>
<td>35.66 (11.93)</td>
<td>31.34 (13.24)</td>
</tr>
<tr>
<td>N2 post-test (SD)</td>
<td>44.34 (14.82)</td>
<td>38.66 (14.46)</td>
<td>36.59 (16.03)</td>
</tr>
<tr>
<td>N2 difference (SD)</td>
<td>6.07 (11.12)</td>
<td>3.00 (12.00)</td>
<td>5.24 (15.05)</td>
</tr>
</tbody>
</table>

**TABLE 4.11**

N2 Score Means and Standard Deviations for Portico students, by GPA terciles

<table>
<thead>
<tr>
<th>GPA Group</th>
<th>Top Third (101)</th>
<th>Middle Third (81)</th>
<th>Bottom Third (82)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2 pre-test (SD)</td>
<td>37.97 (13.43)</td>
<td>34.23 (11.88)</td>
<td>32.68 (14.21)</td>
</tr>
<tr>
<td>N2 post-test (SD)</td>
<td>41.99 (15.32)</td>
<td>40.41 (14.20)</td>
<td>36.89 (16.34)</td>
</tr>
<tr>
<td>N2 difference (SD)</td>
<td>4.02 (11.53)</td>
<td>6.17 (12.54)</td>
<td>4.21 (14.48)</td>
</tr>
</tbody>
</table>
In conclusion, a review of the descriptive statistics helps to uncover several interesting trends. To start, the sample – freshmen at Boston College – had moral reasoning scores (N2 scores) at the level of a national sample of college seniors (n=2,441) (Bebeau et al., 2003). And, even though their pre-test scores started at a higher level than the national norms of college freshmen, they still gained 1/3 of a standard deviation when measuring moral reasoning scores. Compared to the national sample (Bebeau et al., 2003), the Portico participants went from the moral reasoning level of college seniors to graduate students.

Within the sample, it was interesting to see what made a difference and what did not. Taking this kind of a course made a difference for virtually everyone. Similar to past research, women started and finished with higher N2 scores, yet also had greater gains in N2 scores, than male students. In terms of instructor, all instructors noticed gains with their students. Instructors taught an equal number of small and medium-sized classes. Interestingly, the medium-sized classes had higher pre-test scores, as well as slightly higher post-test scores. However, the small classes had greater gains in N2 scores.

One of the more interesting findings was that earlier classes had higher gains than later classes. In fact, both groups of morning classes had N2 score increases of greater than 6, while afternoon class gains were around 3. The
earliest classes (8am and 9am) had the highest pre-test scores, post-test scores, and gains in N2 scores.

In terms of academic variables (SAT scores and GPAs), both had a similar trend. The higher the scores (SAT and GPA), the higher the pre-test N2 score. In both variables, the higher tercile also had the highest N2 score, the middle tercile had the next highest N2 score, and the lowest tercile had the lowest N2 score. However, for SAT scores, the greatest gains were found with the top and bottom tercile while with GPA it was the middle tercile that had the greatest N2 score gains. So, while GPA and SAT had the same trend with pre-test scores, their difference scores were the inverse of each other.

Section Three: An Analysis of the Research Questions

This section will probe further, attempting to answer the following research questions.

1. Do students enrolled in a one-semester business ethics class have statistically significant gains in moral reasoning, measured over time from the pre-test to the post-test?
2. Do students enrolled in smaller sections of a business ethics course have greater gains in moral reasoning scores than students enrolled in larger sections of the same course, taught by the same instructor?

3. What are the significant relationships between the differences in N2 scores and the independent variables of instructor, class size, measures of academic ability, gender, and class time?

The null hypotheses for this study are:

H01. There is no statistically significant difference between moral reasoning scores at the beginning and end of a one semester business ethics course.

H02. There is no statistically significant difference of the moral reasoning scores between students in small sections and medium sized sections of a business ethics class.

H03. Independent variables of instructor, class size, measures of academic ability, gender, and class time are unrelated to a difference in N2 scores.

Findings for Research Question One

The first research question investigated if students enrolled in a one-semester business ethics class have significant gains in moral reasoning, measured over time from the pre-test to the post-test. To examine this question, a
Paired Samples T Test was conducted on the N2 scores variable. This test is used to compare the means of two sets of data, and assumes a normal distribution. To test the distribution, a QQ plot test was conducted. The results of this test determined that all three N2 scores (pre-test, post-test, and difference) were normally distributed (see Figures 4.3 – 4.5 for the QQ plots to test normal distribution).

**Figure 4.3 - Normal Q-Q Plot of N2 Pre-test Scores**
Figure 4.4 - Normal Q-Q Plot of N2 Post-test Scores

Figure 4.5 - Normal Q-Q Plot of the Difference in N2 Scores, from pre-test to post-test
The Paired Samples T Test showed that there is a statistically significant difference in the N2 difference scores of Portico students, measured over time from the pre-test to the post-test. Students made statistically significant gains in moral reasoning from the beginning to the end of the Portico semester. The results were significant at the $p \leq .05$ and .01 levels, therefore the null hypothesis was rejected (see Table 4.12).

Table 4.12 – Paired Sample T Test of the Difference in N2 scores from Pre-test to Post-test

<table>
<thead>
<tr>
<th>Paired Differences</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2 Difference (post – pre)</td>
<td>4.80</td>
<td>12.75</td>
<td>.78</td>
<td>6.16</td>
<td>267</td>
<td>.000</td>
</tr>
</tbody>
</table>

Findings for Research Question Two

The second research question explored the impact of class size on the N2 scores. Do students enrolled in smaller sections of a business ethics course have greater gains in moral reasoning scores than students enrolled in larger sections of the same course, taught by the same instructor? To answer this question, Two-way ANOVA was conducted to see if the independent variables of class size
(medium/small) and instructor (1, 2, 3) had any effect upon the difference in N2 scores of Portico students.

Since the two class size groups that are being compared differ in sample size (n of students in small class = 116; n of students in medium sized classes = 152), the equality of means and analysis of variance assumption were important in this analysis. To assure that the equality of variance assumption was being met, Levene's Test of Equality of Error Variances was performed. This tests the null hypothesis that the error variance of the dependent variable is equal across groups. Results showed that the null hypothesis of equal variance could not be rejected as the observed significance level was .310, which exceeded 0.05 (Table 4.13). Therefore, the equality of means assumptions are satisfied (p ≥ 0.05).

<table>
<thead>
<tr>
<th>Table 4.13 - Levene's Test of Equality of Error Variances*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable: N2.difference</td>
</tr>
<tr>
<td><strong>F</strong></td>
</tr>
<tr>
<td>1.199</td>
</tr>
</tbody>
</table>

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Instructor + ClassSize + Instructor * ClassSize

Two-way ANOVA testing examines the effects of two independent variables – instructor and class size – as well as their interaction, concurrently.
Looking at these data, the N2 score difference is greatest between the two class size groups, and would be significant at the 0.10 level. However, for the purposes of this study, none of the differences are statistically significant. This means that, among the Portico group, there is no statistically significant difference in the difference of the N2 scores (from pre-test to post-test) when comparing class size, instructors, and the interaction of class size and instructor. Therefore, the null hypothesis is accepted for research question 2. Table 4.14 summarizes these findings.

Table 4.14 – Two-way ANOVA testing of N2 difference scores by instructor, class size, and their interaction effect

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig. (p ≥ 0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructor</td>
<td>58.30</td>
<td>2</td>
<td>29.15</td>
<td>0.20</td>
<td>0.817</td>
</tr>
<tr>
<td>ClassSize</td>
<td>415.98</td>
<td>1</td>
<td>415.98</td>
<td>2.89</td>
<td>0.090</td>
</tr>
<tr>
<td>Instructor * ClassSize</td>
<td>363.05</td>
<td>2</td>
<td>181.52</td>
<td>1.26</td>
<td>0.285</td>
</tr>
</tbody>
</table>

Findings for Research Question Three

The third research question explored the significant relationships between the dependent and independent variables. Specifically, what are the significant relationships between the differences in N2 scores and the independent
variables? To answer this question, a rank biserial correlation analysis was performed on the dependent and certain continuous independent variables. Non-dichotomous categorical independent variables were omitted from this analysis, as they did not adhere to assumptions of correlation testing. In order to include class time as a variable, it was re-coded into two categories. Early and later morning classes were re-coded into “morning classes” and early afternoon and late afternoon classes were re-coded into “afternoon classes.” Therefore, the class time variable became a dichotomous one in order to fit the requirements of the rank correlation testing. Table 4.15 reflects the correlation analysis. Those that were found to be significant were noted.

<table>
<thead>
<tr>
<th></th>
<th>N2 Difference</th>
<th>SAT Score</th>
<th>GPA</th>
<th>Class Size</th>
<th>Class Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2 Difference</td>
<td>1</td>
<td>.021</td>
<td>.051</td>
<td>-.093</td>
<td>-.126*</td>
</tr>
<tr>
<td>SAT Score</td>
<td>.021</td>
<td>1</td>
<td>.464**</td>
<td>.136*</td>
<td>-.020</td>
</tr>
<tr>
<td>GPA</td>
<td>.051</td>
<td>.464**</td>
<td>1</td>
<td>.089</td>
<td>-.066</td>
</tr>
<tr>
<td>Class Size</td>
<td>-.093</td>
<td>.136*</td>
<td>.089</td>
<td>1</td>
<td>.001</td>
</tr>
<tr>
<td>Class Time</td>
<td>-.176**</td>
<td>-.020</td>
<td>-.066</td>
<td>.001</td>
<td>1</td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

Class time was the only independent variable that had a statistically significant correlation with the dependent variable (N2 difference score).
Descriptive data also showed that there were discrepancies in the N2 difference scores based on class time. Therefore, further research was needed to investigate these findings.

A one-way ANOVA of N2 difference score by Class Time (early morning, late morning, early afternoon, late afternoon) was run to examine further these differences. First, due to unequal sizes of the four class time groups, assumption testing was conducted. To assure that the equality of variance assumption was being met, Levene’s Test of Equality of Error Variances was performed. This tests the null hypothesis that the error variance of the dependent variable is equal across groups. Results showed that the null hypothesis of equal variance could not be rejected as the observed significance level was .702, which exceeded 0.05 (see Table 4.16). Therefore, the equality of means assumptions are satisfied ($p \geq 0.05$).

| Table 4.16 - Test of Homogeneity of Variances for ANOVA of N2 Difference score by Class Time |
|-----------------------------------------------|----------|--------|-------------|
| N2.difference                                | Levene Statistic | df1 | df2 | Sig. |
|                                              | .471     | 3    | 262 | .702 |

One-way ANOVA testing analyzed the variance of the means of “N2 difference scores” and “class time.” Looking at these data, there is a statistically
significant difference in the difference of the N2 scores (from pre-test to post-test) compared to class time. Table 4.17 summarizes these findings.

Table 4.17 – One-way ANOVA of N2 Difference score by Class Time

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1181.133</td>
<td>3</td>
<td>393.711</td>
<td>2.786</td>
<td>.041</td>
</tr>
<tr>
<td>Within Groups</td>
<td>37025.703</td>
<td>262</td>
<td>141.319</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>38206.836</td>
<td>265</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Since there are statistically significant differences in the sampled means, post hoc Scheffe testing was conducted. This tests all pairs for differences between means and all possible combinations of means. These results showed no statistically significant differences in the data. So, although the overall variance was significant, there was no significance when analyzing class times in pairs (see Table 4.18).
Table 4.18 – Post-hoc Scheffe Testing of the Multiple Comparisons of N2 Difference Scores, by Class Times

<table>
<thead>
<tr>
<th>(I) ClassTimeCode</th>
<th>(J) ClassTimeCode</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early Morning</td>
<td>Late Morning</td>
<td>-.114690468</td>
<td>2.071539203</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Early Afternoon</td>
<td>4.218955883</td>
<td>1.970641076</td>
<td>.208</td>
</tr>
<tr>
<td></td>
<td>Late Afternoon</td>
<td>4.075250299</td>
<td>2.127768183</td>
<td>.302</td>
</tr>
<tr>
<td>Late Morning</td>
<td>Early Morning</td>
<td>.114690468</td>
<td>2.071539203</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Early Afternoon</td>
<td>4.333646351</td>
<td>2.019525734</td>
<td>.206</td>
</tr>
<tr>
<td></td>
<td>Late Afternoon</td>
<td>4.18940767</td>
<td>2.173121115</td>
<td>.296</td>
</tr>
<tr>
<td>Early Afternoon</td>
<td>Early Morning</td>
<td>-4.218955883</td>
<td>1.970641076</td>
<td>.208</td>
</tr>
<tr>
<td></td>
<td>Late Morning</td>
<td>-4.333646351</td>
<td>2.019525734</td>
<td>.206</td>
</tr>
<tr>
<td></td>
<td>Late Afternoon</td>
<td>-1.143705584</td>
<td>2.077163200</td>
<td>1.000</td>
</tr>
<tr>
<td>Late Afternoon</td>
<td>Early Morning</td>
<td>-4.075250299</td>
<td>2.127768183</td>
<td>.302</td>
</tr>
<tr>
<td></td>
<td>Late Morning</td>
<td>-4.18940767</td>
<td>2.173121115</td>
<td>.296</td>
</tr>
<tr>
<td></td>
<td>Early Afternoon</td>
<td>.143705584</td>
<td>2.077163200</td>
<td>1.000</td>
</tr>
</tbody>
</table>

A plot of the N2 difference scores by class time revealed that the greatest differences appear to be divided between morning and afternoon. The N2 difference scores of both afternoon groups were shown to be close. The N2 difference scores of the two morning groups were also parallel. Figure 4.6 illustrates these findings.
To see if the division of class times into two groups (morning and afternoon) proved to be significant, a new variable called “class time code2” was created. The early morning and late morning groups were now coded as “morning.” The early afternoon and late afternoon groups were now coded as “afternoon.” The one-way analysis previously conducted was re-run with this new variable. Test results showed that there was a strong statistically significant difference ($p \geq 0.01$) when comparing the N2 difference scores of morning and afternoon classes. See table 4.19 for the results.
Table 4.19 – T test of N2 Difference Scores, by “Morning” and “Afternoon” Class Times

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1180.023</td>
<td>1</td>
<td>1180.023</td>
<td>8.414</td>
<td>.004</td>
</tr>
<tr>
<td>Within Groups</td>
<td>37026.813</td>
<td>264</td>
<td>140.253</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>38206.836</td>
<td>265</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These differences suggest a strong relationship between class time and N2 scores. To further explore, a two-way ANOVA was run. This test added the variable of instructor to see if that had an effect upon the significant difference in N2 scores. First, Levene's Test of Equality of Error Variances tested the null hypothesis that the error variance of the dependent variable was equal across groups. Results showed that the null hypothesis of equal variance could not be rejected as the observed significance level was .551, which exceeded 0.05 (see 4.20). Therefore, the equality of means assumptions are satisfied (p ≥ 0.05).

Table 4.20 - Levene’s Test of Equality of Error Variances

<table>
<thead>
<tr>
<th>F</th>
<th>df1</th>
<th>df2</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>.800</td>
<td>5</td>
<td>260</td>
<td>.551</td>
</tr>
</tbody>
</table>

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + ClassTimeCode2 + Instructor + ClassTimeCode2 * Instructor
As with previous testing, results showed a statistically significant difference of the N2 difference scores when comparing morning and afternoon classes. However, differences between instructors and the interaction of instructors and class time were not statistically significant. Post hoc Scheffe testing further showed that there were no statistically significant relationships between pairs of instructors and class times (see table 4.21).

Also tested was the effect size of class times. Effect size measures the magnitude of the difference between groups. It complements significance testing by providing further descriptive analysis. Effect size moves beyond simply asking whether or not there is a difference by asking how large the difference is (Prentice & Miller, 1992). A comparison of the N2 scores of the morning and afternoon classes found that the effect size could be considered medium (Cohen’s d = 0.36). This can be interpreted to show that the mean N2 scores of the morning classes were at the 65th percentile of the afternoon class mean N2 scores and so these results can be considered both statistically and substantively meaningful. Furthermore, there was nonoverlap in 25 percent of the two distributions. Nonetheless, while class time was not a variable in the original research question, it did prove to be a powerful variable in the positive changes of moral reasoning of freshmen business students.
Table 4.21 – Two-way ANOVA of N2 Difference Scores by Class Time, Instructor, and Their Interaction

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>1405.049&lt;sup&gt;a&lt;/sup&gt;</td>
<td>8</td>
<td>175.631</td>
<td>1.226</td>
<td>.284</td>
</tr>
<tr>
<td>Intercept</td>
<td>4803.681</td>
<td>1</td>
<td>4803.681</td>
<td>33.546</td>
<td>.000</td>
</tr>
<tr>
<td>ClassTimeCode</td>
<td>1282.423</td>
<td>3</td>
<td>427.474</td>
<td>2.985</td>
<td>.032</td>
</tr>
<tr>
<td>Instructor</td>
<td>94.512</td>
<td>2</td>
<td>47.256</td>
<td>.330</td>
<td>.719</td>
</tr>
<tr>
<td>ClassTimeCode * Instructor</td>
<td>96.379</td>
<td>3</td>
<td>32.126</td>
<td>.224</td>
<td>.879</td>
</tr>
</tbody>
</table>

<sup>a</sup> R Squared = .037 (Adjusted R Squared = .007)

Summary of Findings

Results of significance testing found that students enrolled in Portico in the Fall of 2009 had statistically significant gains in moral reasoning during the semester. And, while students enrolled in smaller sections (n = 19) of the course had greater gains in moral reasoning scores than students enrolled in medium sized (n = 27) sections of the same course, those gains were not statistically significantly different. Finally, the independent variable of class time had the most statistically significant relationship with gains in N2 scores. Table 4.22 presents the findings.
Table 4.22 - The Null Hypotheses of This Study and Their Findings

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H01.</strong> There is no statistically significant difference between moral reasoning scores at the beginning and end of a one semester business ethics course.</td>
<td>Rejected</td>
</tr>
<tr>
<td><strong>H02.</strong> There is no statistically significant difference of the moral reasoning scores between students in small sections and medium sized sections of a business ethics class.</td>
<td>Accepted</td>
</tr>
<tr>
<td><strong>H03.</strong> Independent variables of instructor, class size, measures of academic ability, gender, and class time are unrelated to a difference in N2 scores.</td>
<td>Rejected</td>
</tr>
</tbody>
</table>

These findings indicate that there are positive associations between Portico and gains on students’ moral reasoning scores. Chapter five includes an interpretation of these findings, and the implications these findings have on curricular decision making in higher education.
CHAPTER FIVE
DISCUSSION OF THE FINDINGS

As stated earlier, the purpose of this study was to measure the changes in moral reasoning of freshmen business ethics students. The neo-Kohlbergian theories of moral reasoning as described by James Rest were the focus of the research and informed the research questions. In order to measure changes in moral reasoning, the Defining Issues Test, version 2 (DIT2) was used. Chapter three described the research methodologies followed in investigating this quasi-experiment. Chapter four presented the data, statistical analyses, and results.

This chapter presents a discussion of the findings and is divided into five sections. Section one briefly summarizes the key findings of Chapter four. Section two interprets the results, comparing the findings to existing studies in moral reasoning. The subsequent three sections present implications for practice, limitations of the study, and recommendations for future research.

Summary of the Findings

This quasi-experimental study was conducted during the fall 2009 semester at Boston College. It assessed the moral reasoning of freshmen business students enrolled in an ethics class (Portico, which is described in Chapter three).
The sample for this study includes a total of 268 students, all of whom completed the Defining Issues Test – Version Two (DIT2) at the beginning and end of the Portico class.

Descriptive analyses show that the sample had moral reasoning scores (N2 scores) at the level of a national sample of college seniors (n=2,441) (Bebeau, et al., 2003). And, even though their pre-test scores started at a higher level than the national norms of college freshmen, they still gained 1/3 of a standard deviation when measuring moral reasoning scores. Compared to the national sample (Bebeau, et al., 2003), the Portico participants went from the moral reasoning level of college seniors to graduate students.

Within the sample, it was interesting to see what made a difference and what did not. Similar to past research, women started and finished with higher N2 scores, yet also had greater gains in N2 scores, than male students. In terms of instructor, student N2 scores occurred across all of them. Instructors taught an equal number of small and medium-sized classes. The medium-sized classes had higher pre-test scores, as well as slightly higher post-test scores. However, the small classes had greater gains in N2 scores.

One of the more interesting findings was that earlier classes had higher N2 score gains than later classes. In fact, both groups of morning classes had N2
score increases of greater than 6 points, while afternoon class gains were around 3 points. The earliest classes (8am and 9am) had the highest pre-test scores, post-test scores, and gains in N2 scores.

In terms of academic variables (SAT scores and GPAs), both had a similar trend. The higher the scores (SAT and GPA), the higher the pre-test N2 score. In both variables, the higher tercile also had the highest N2 score, the middle tercile had the next highest N2 score, and the lowest tercile had the lowest N2 score. However, for SAT scores, the greatest gains were found with the top and bottom tercile while with GPA it was the middle tercile that had the greatest N2 score gains. So, while GPA and SAT had the same trend with pre-test scores, their difference scores were the inverse of each other.

Empirically, this study asked the following research questions:

1. Do students enrolled in a one-semester business ethics class have statistically significant gains in moral reasoning, measured over time from the pre-test to the post-test?

2. Do students enrolled in smaller sections of a business ethics course have greater gains in moral reasoning scores than students enrolled in larger sections of the same course, taught by the same instructor?
3. What are the significant relationships between the differences in N2 scores and the independent variables of instructor, class size, measures of academic ability, gender, and class time?

Results of significance testing found that students enrolled in Portico in the Fall of 2009 had statistically significant gains in moral reasoning during the semester. However, the moral reasoning gains of students enrolled in the smaller sections (n = 19) were not statistically significantly different than students enrolled in medium sized (n = 27) sections of the same course. Finally, the independent variable of class time had the most statistically significant relationship with gains in N2 scores.

*Interpretation of the Findings*

There are several interesting findings within this study. Some validate the body of research within the field, yet some move in a new direction. This section will focus on the major findings of the study, specifically in relation to national norms, class size, and class time.
Comparisons between Portico Students and the National Sample

Studies show that measures of aptitude and intelligence are positively correlated with DIT scores (King & Kitchener, 1994; King & Mayhew, 2004; Sanders, Lubinski, and Benbow, 1995). Boston College is a highly selective liberal arts university. The average SAT scores of students in this study was 1338, which is 1.5 standard deviations above the mean of all SAT takers (http://professionals.collegeboard.com, retrieved on 1/20/2011). Therefore, it is not surprising that the pre-test N2 scores of Boston College freshmen are statistically significantly higher than the N2 scores for the national sample of college freshmen.

In addition to aptitude and intelligence, Rest (1979) asserts that moral reasoning growth is an element of cognitive complexity. In fact, studies have shown that higher levels of moral reasoning are associated with more sophisticated logical reasoning strategies (King & Mayhew, 2004). Santilli and Hudson (1992) tested students who were enrolled in a communications course at two points in the term, and found that formal reasoning is a precursor to the development of moral reasoning. LaRue and Olejnik (1980) used tests of Piagetian formal operations to analyze the relationship between formal operational thinking and moral reasoning. They tested introductory psychology
students who were randomly assigned to three experimental groups, and found that priming students with formal operational thinking was influential to moral reasoning.

Therefore, it can be argued that students admitted to a highly selective liberal arts university would have significantly higher moral reasoning scores than the general population. This could be due to the academic aptitude, intelligence, and critical thinking that is required of students at highly selective universities, and explains the significant difference in N2 scores. Despite this head start, Boston College students still showed gains in moral reasoning during the first semester. These gains were statistically significant in two ways. First, they are significant within the group. The N2 scores of Boston College students rose 4.79 points from the pre-test to the post-test, which was a statistically significant difference. Second, there is a statistically significant difference in N2 score gains when comparing the Boston College group to the national sample. The national sample had small gains from the freshmen to the sophomore year (0.19), compared to the significant gains of Boston College freshmen, whose N2 scores increased 4.79 points during the first semester of college.

These gains can be explained in several ways. First and foremost, the Portico course is taught using pedagogies which support growth in moral
reasoning. For example, studies have suggested the following strategies to be used in course interventions. These include opportunities for students to participate in role-taking (Reimer et al., 1983; Mason & Gibbs, 1993), the discussion of moral dilemmas (Keefe & Ashley, 2001), reflection exercises (King & Kitchener, 1993), active learning (Boss, 1994; Gorman et al., 1994; Hudec, 2002), and assignments which create cognitive disequilibrium (Rest, 1986; Rholes et al., 1982). Further, courses must be explicit in promoting the use of higher order thinking through philosophy or moral psychology (Mayhew & King, 2008). Portico strives to implement these strategies through class discussions, reflective exercises, written assignments, case studies, debates, and group projects.

Second, the Boston College curricular environment is that of a liberal arts university. First year students are taking liberal arts courses, such as writing, literature, philosophy, theology, economics, and calculus, in addition to Portico. Institutional type is a variable used in much of the literature on moral reasoning (Burwell, Butman, and Van Wicklin, 1992; Good & Cartwright, 1998; McNeel, 1991, 1994; Pascarella & Terenzini, 2005; Shaver, 1985, 1987). Across these studies, the only institutional type with large effect sizes was found in liberal arts colleges and universities. Perhaps students who choose these types of
institutions are more inclined to discuss issues, both in and out of the classroom. Also, the critical thinking of liberal arts curricula can enhance moral judgment.

Third, the mission of Boston College supports theories of social justice. This is exhibited through the curriculum, and perhaps more so in the co-curriculum. Volunteer opportunities, both international and domestic, abound. Many students are drawn to Boston College due to this culture of service, while others are affected by it as they matriculate. Research on community service shows a strong association with participation and growth in moral reasoning (King & Mayhew, 2004). In fact, this culture of student volunteerism complements the peer learning mentioned previously. Both findings are consistent with Astin’s (1993) conclusions, that “the student’s peer group is the single most potent source of influence and growth and development during the undergraduate years” (p. 398).

Class Size

This study compared the changes in moral reasoning scores between students enrolled in small classes of 19 and medium-sized classes of 26. The hypothesis was that students in the smaller classes would have greater gains in moral reasoning scores. This was the case, as the pre-test to post-test gains of
students in the smaller classes was 5.64 points compared to gains of 4.15 points in the medium-sized classes. The differences in these gains, however, are not statistically significant. This means that the pedagogical principles of the course matter more than the size of the course, at least when considering class size up to 26 students. The seven additional students do not significantly affect the gains in moral reasoning scores.

What seem to affect the scores are other pedagogical strategies, such as those mentioned above. Both class types (small and medium-sized) explicitly include moral content in the curriculum through the study of ethical philosophies (see Appendix, “Portico Syllabus” for content).

Both class types also strive to create supportive and safe learning communities. There is room to have thought provoking conversations on moral dilemmas and issues, and to think deeply and critically. However, these student interactions occur within a community which is carefully fostered by the instructors.

A final strategy is to promote much group work. Regardless of class size, both spend much time in groups of four or five students. These occur both in and out of class, and allow for strategies such as role taking, active learning, and discussions of dilemmas to be experienced by all students.
The way that the instructors structure and support the learning environment “affects the quality of the spaces they create for discussing moral issues and this may be as important as the content of the topics themselves” (Mayhew & King, 2008) as well as the differences between the size of the class.

Class Time

The most surprising and unexpected finding of the study was related to class time. Class time was found to be the most statistically significant variable related to gains in moral reasoning scores. The earlier the class time, the more the N2 scores increased from pre-test to post-test. This was unexpected as students were assigned to class times rather than choosing them. Therefore, there were no selection preferences. In other words, the less ambitious students could not select later class times, and the more ambitious students could not select the earlier class times. It was therefore surprising to see that, controlling for other variables, earlier class times were statistically significantly related to greater gains in N2 scores.

These findings go against the current circadian rhythms literature from the fields of bio-psychology and neuroscience. Circadian rhythms measure the daily variations in human behavior and functioning that happen over a 24 hour
cycle. These measurements can be helpful in determining optimal functioning for human beings. This optimal functioning then sorts people into chronotypes. There are three main chronotypes, morning people, intermediate people, and evening people (Horne & Ostberg, 1976).

Although there are people of varying chronotypes within all ages, there are trends within each age group. Studies have found that younger children (up to age 13) tend to be morning people (Kim et al., 2002). A shift occurs at age 13, as optimal functioning moves from morning to evening, though there is some variability within the ages. Evening is also the optimal functioning time for college students. A study of 210 university students ages 19 to 22 (May et al., 1993) found that 94 percent of college students were “definitely evening” or “moderately evening” types. Of the six percent of students who were morning types, none were “definitely morning” types.

Further studies have shown that there is a relationship between cognitive functioning and optimal circadian rhythms. This is called the synchrony effect (Hasher, 1998), and theorizes that there is a benefit to matching cognitive performance with chronotype preferences (Schmidt, et al., 2007). Since college students are not “morning types,” one would assume poorer performance in morning classes. However, this assumption runs counter to the findings of this
The greatest gains in cognitive moral reasoning occurred in the earliest class times (8 and 9 a.m.), with the gains decreasing as the time of the class became later. This was an unexpected finding, considering that only six percent of college students are found to be morning types (May et al., 2003).

While these findings are unexpected at first, there are several factors which make them less surprising. First, it could be that the earlier class times force students to be more alert and attentive earlier in the day. They may be going to bed earlier, engaging in less risky behaviors the night before class, and forming healthier routines. Second, instructors may be more prepared for the earlier class times. Portico faculty teach multiple sections of the course, and may be freshest and teaching optimally at the earlier times. In fact, while student chronotypes match that of a night owl, the older instructors’ circadian rhythms are those of the early bird. For this reason, the instructors may be teaching with more energy, clarity, and alertness than they do later in the day, possibly affecting the learning outcomes of the students. Further studies, including a mixed methods approach, would be helpful in analyzing the optimal course times for Portico.
Implications for Policy and Practice

The initial topic of this study asked if ethics could be taught, and, if so, how? The findings help to answer this question, providing pointers for curricular practices. They also help to guide policy makers and planners, especially at Boston College, with practices concerning class size and class time.

According to this study, ethical reasoning can be taught. Results from the Portico program exhibited statistically significant gains in the moral reasoning scores of freshmen business students. A review of the literature illustrates the suggestions for implementing meaningful interventions when moral reasoning is the intended outcome. These suggestions are curricular and pedagogical.

Curricular measures include explicitly presenting content on ethical theories. The ethical theories discussed in Portico were ethical egoism, ethical and cultural relativism, utilitarianism, deontology (via Kant, Rawls, and modern unifying theories), and virtue theory (Aristotle and Plato).

The explicitly ethical content of the course was supported by pedagogical strategies that are shown to increase moral reasoning in intervention courses. These strategies include role-taking, the discussion of moral dilemmas, reflection, active learning, and cognitive disequilibrium. These practices are engrained in the Portico program. Role-taking occurs throughout the semester, mainly in the
form of structured debates. Students form small groups and address ethical
issues through the framework of an ethical theory. Students must take the role of
the assigned theory, regardless of their personal preferences. These debates also
enhance moral growth through the discussion of moral dilemmas. The debates
are centered around ethical dilemmas. Students move through a framework to
analyze the dilemmas. First, they must identify the facts of the dilemma (or case).
Second, they must identify the ethical issues within the dilemma. Third, they
must identify options (based on the facts and issues). Fourth, they must choose
the best option of the dilemma. Fifth, and last, they must explain their choice by
examining the strengths and weaknesses of the decision, as well as reflecting
upon the decision and how it resonates with a student’s value system.

Reflection is another necessary component of successful ethical
interventions. Reflection occurs throughout the Portico program, both in class
and in graded and ungraded assignments. For example, students must reflect
upon their decisions, be they hypothetical ones (through cases) or real ones that a
student makes in the dining halls, residence halls, and classrooms. Students are
also asked to reflect upon the research assignments they complete, and how
those assignment topics fit with their process of vocational discernment. Finally,
throughout the semester, students begin class with short written reflections on a
topic provided by the instructor. These reflections are based on the Ignatian paradigm, which are techniques of reflecting upon the events of the day.

Particular points of the reflection include highpoints and low points of the day, how one felt during these points, and what one hopes to take into the next day. The stated goal of these reflections is that students develop habits that they will practice beyond the classroom requirements.

Active learning is another condition required of ethical interventions. For Portico, active learning is developed through “cold calling,” small class size, and small group work. Cold calling is the practice of involving all students in the discussion, rather than relying on volunteers. In Portico, a community quickly develops in which cold calling becomes unnecessary, as members freely contribute to the discussion. Small class size also helps promote active learning. And, while the smaller classes had greater gains in moral reasoning scores compared to the medium-sized classes at the $p \leq .10$ level, these gains were not statistically significant at the testing level of this study. Perhaps this was so due to the practice of small groups. Portico students were often divided into smaller groups of three to five students. These groups were student led, and required active involvement of all members. Students occasionally submitted 360 degree feedback in order to assess their and their peers’ involvement in these groups.
These assessments asked that group members recorded their relative strengths and weaknesses within the group, and also required all members to be engaged in order to receive optimal credit for the work. By making the class smaller, and by making the students the authors of their own meaning-making, the learning in Portico was active and interactive.

A final condition for optimal moral reasoning growth is cognitive disequilibrium. This disequilibrium helps to move a student through stages of ethical development. The disequilibrium must be enough to push and challenge a student, but not too great as to shutdown the student. In Portico, cognitive disequilibrium is presented through the curricular content of the course and the assignments. The content of the course is designed to expose students to ethical theories in a developmental way. For example, the readings begin with egocentric approaches (ethical egoism) and progress to theories of virtue and justice (doing good for good’s sake, and acting for the good of all rather than the good of self). These theories mirror Kohlberg’s stages of moral development, and explain why there is significant growth throughout the course. While students may locate their moral reasoning within a theory, they become aware of more advanced stages of moral reasoning throughout the semester. This awareness helps students increase their moral development. Another method of presenting
disequilibrium is through an assignment that asks students to imagine
themselves on a mountain climbing expedition which is interrupted by a dying
pilgrim on his way down the mountain (McCoy, 1988). For many students, this is
a complex and visceral assignment in which the dilemma of the case requires
much ethical discernment. Ultimately, the students are challenged to feel the
human impact of their decisions, while weighing the facts of the case. The
disequilibrium caused by combining emotion and reason contribute to the
growth of moral reasoning within Portico students.

All of these curricular and pedagogical strategies of Portico suggest
effective practices in courses that seek to promote the growth of moral reasoning.
At Boston College, this growth is a mission of the university. In fact, the first
objective of the Boston College Mission Statement is “the rigorous intellectual
development and the religious, ethical and personal formation of its
undergraduate, graduate and professional students in order to prepare them for
citizenship, service and leadership in a global society” (Trustees of Boston
College, 1996). With this stated goal, other classes can look to the Portico
Program as an effective method of achieving the mission of Boston College.
Further, the success of Portico can serve to inspire other departments and courses
to offer and require an ethics course in the curriculum. Currently, the only
undergraduate school at Boston College to require ethics is the School of Management. Other schools at Boston College offer ethics courses, but none require it. In fact, only the International Studies major in the College of Arts and Sciences requires a course in ethics within its curriculum. Portico can serve as an example of innovation within the curriculum, as a means towards strengthening the mission of the Boston College education through an ethics course. It is the hope of this study that other schools and departments, and perhaps the university at-large will see the benefits of this type of course in supporting and enhancing the university’s mission.

One barrier to innovating courses is budgetary constraints. This study helps inform financial decisions around courses as well. As previously mentioned, this study found no statistically significant differences in gains in moral reasoning scores when comparing class sizes. This study shows that, in this type of course, students have significant gains in their moral reasoning scores, regardless of class size. Therefore, all sections of Portico (and perhaps classes similar to Portico) can be increased to 27 students and still have statistically significant gains in moral reasoning. Given the size of the freshmen class, if all classes were 27 students, the university could offer less sections than if all classes were 19 students. In fact, the difference would be eight sections in
Based on the instructor rate for Portico, capping classes at 27 would allow for significant moral reasoning gains, while saving the university over $40,000 per year. If these practices were implemented throughout all similar undergraduate courses, the savings would increase significantly. Financial planners and managers continue to look towards economic efficiency in financially difficult times. Increasing class size by eight students can help to maintain moral growth while still be economically efficient.

Administrators can also look at the class time findings of this study. While the circadian rhythms and personal preferences of students may lean towards later classes, this study found that students enrolled in earlier sections of Portico had greater gains in their moral reasoning scores. These findings may inform administrators in planning class times. For example, classes which promote moral growth, require active learning, and tend to require class participation to be instrumental to learning should consider earlier class times. This may increase the opportunities for moral growth, but may have secondary effects as well.

If students are required to attend earlier classes, it may impact their behaviors the night before the classes. For example, high risk drinking remains an important concern on college campuses (Mayhew et al., 2008). In fact, a report on 38,857 undergraduates from 89 colleges in the United States found that half of
these students were considered binge drinkers (CORE, 2005). Binge drinking is associated with several dangerous outcomes, including memory loss, sexual assault, violence, injury, and possible death (CORE, 2005).

Because of the dangers of alcohol on college campuses, the federal government has recognized the reduction of this behavior as a top priority (U.S. Department of Health and Human Services, 2005). And, while college leaders recognize the importance of curbing these behaviors, many interventions do not work (Kuh, 2004). The required attendance in early classes may persuade students to engage in less drinking the night before their class. University officials should consider the relationship between healthy behaviors and early classes, and more research should be conducted to investigate this possible relationship.

Limitations of the Study

There are several limitations to this study. First, this study observes one institution and one curriculum of the institution over a short time period. Therefore, possible institutional effects (self-selection of students who attend Boston College, with a certain mission and ethos) cannot isolate this treatment from other experiences. Second, this study does not include a control group, as
all students receive the intervention. Without a control group, it is more difficult to measure the effect of the intervention on the gains in moral reasoning. Comparing the intervention group to a control group could help to show that gains found in the intervention group were due to the interventions, as opposed to an overall college effect. Third, this is a one semester study, with no longitudinal follow-up. While gains in moral reasoning scores did occur, it is unclear if these trends will continue, plateau, or decrease. Fourth, the intended outcome of the curriculum is principled moral action, but principled moral reasoning is only part of moral action (Rest, et al., 1999b). In other words, knowing the right thing to do may not always lead to doing the right thing. Three separate reviews of the literature demonstrate that measures of moral judgments and actions are related, but only weakly (Arnold, 1989; Blasi, 1980; Thoma & Rest, 1986). So, if moral action is the intended outcome, then studies which specifically measure moral action must occur. Further research on this topic is merited.

**Recommendations for Future Research**

As stated above, this study had certain limitations. The following recommendations for future research are as follows:
1. Subsequent studies would replicate the research methodologies and procedures of the current study. Replications of this study would allow for stronger inferences by conducting a meta-analysis of the numerous studies (Johnson, 2001). Further, the various studies could be analyzed comparatively as well. If replications of the study supported the findings of this study, more inferences, and perhaps causal ones, could be made. For example, the findings of this study demonstrate a relationship between class time and gains in moral reasoning. Subsequent replications of this study might support the reliability of this relationship.

2. A subsequent study would include a control group. As noted above, the lack of a control group makes it difficult to assess the exact impact of the ethics intervention. While it appears that gains in moral reasoning scores are significantly impacted by the intervention, a comparison to a control group would make for a stronger study design.

3. Future research would follow the study group longitudinally. According to King and Mayhew (2002), the most needed research on moral development are longitudinal studies. Most of the current research, as with this study, focuses on snapshots in time or short-term interventions.
Longitudinal studies would help to measure the mediating variables that are most needed for gains in moral reasoning.

4. The longitudinal follow-up would include a measure to determine the relationship between moral reasoning and moral action. Since moral action is the ultimate goal, do the gains in moral reasoning amount to increases in moral action?

5. The independent variables of this study included gender, class size, instructor, class time, SAT scores, and students’ GPAs. Future research would analyze other variables such as socio-economic status, ethnicity, political orientation, college activities, time spent volunteering, and other possible variables. As stated before, research on moral development in college show that gains in moral reasoning is not an isolated incident that occurs in any one class, setting, type of institution, or type of student (Pascarella & Terenzini, 2005). Instead, it is an interconnected system of developmental conditions that lead to this development. Controlling for more variables than those of this study could help discern certain trends or relationships in this development.

6. Future research would include a mixed-methods approach. Mixed methods could help add some meaning to the quantitative findings of this
study. For example, students in smaller classes had greater gains in moral reasoning scores than students in medium-sized classes. However, these gains were not statistically significantly greater ones, so the study hypothesis was rejected. A mixed-methods approach could further investigate these differences in order to make meaning of the two class sizes. Likewise, there were statistically significant differences in moral reasoning scores of students in morning classes than those in afternoon classes. Mixed methods may help uncover why these differences existed.

While this study did uncover some interesting findings, it would be complemented by following the recommendations stated above. The limitations of the study could be corrected, potentially strengthening the reliability of the findings. The main ways to do so would be through replication of the study, repeating the study while adding a control group, longitudinally following the study group, including other variables in the analysis, and conducting a mixed-methods approach. Finally, the future research would ultimately assess the moral action of the participants.
Conclusion

Universities have a unique opportunity to foster the moral growth of its undergraduates. Great changes in technology and globalization require ethical sensitivities that are needed from future leaders. The creation of Portico, an innovative and rigorous business ethics course at Boston College, is a step towards helping to develop undergraduate students for ethical careers in business. This is particularly important given the current climate, in which poor decisions are increasingly costly to society.

This study demonstrated significant gains in moral reasoning scores of Portico students. This information, coupled with existing and possible future research, may help create curricular conditions for promoting the moral reasoning needed for good moral action. The ultimate goal is to create virtuous business leaders who will do well and do good in an increasingly interconnected and complex world. Our global economy may very well depend upon it.
References


level of education, gender and political orientation, paper presented at the
Association for Practical and Professional Ethics, Charlotte, NC, 27
February–2 March.

perspective on research in moral education. Educational Researcher, 28(4), 18-
26.

Issues Test, Version 2 (DIT2) and the scoring service of the Center for the
Study of Ethical Development, Unpublished manuscript, University of
Minnesota.

Review, 83(5), 96.

December 12, 2009, from
http://granitegrok.com/blog/2009/02/understanding_the_currrent_financial_c
ri.html


Retrieved May 5, 2002, from


Holland, J.B. (1954). The image of instructor as it is related to class size. *Journal of Experimental Education, 23* (2), 171. F


Hudelson, E. (1928). *Class Size at the College Level*. Minneapolis: University of Minnesota Press.


Mayhew, M.J., Caldwell, R.J., & Hourigan, A.J. (2009). The influence of
curricular-based interventions within first-year “success” courses on student alcohol expectancies and engagement in high-risk drinking behaviors.

*NASPA Journal, 45*(1), 45-72.


McKeachie, W. J. (1980). Class size, large classes, and multiple sections.

*Academe, 66*, 24-27.


*Journal of Educational Psychology, 82*(2), 189.


Mertler, C. A. (2002). *Advanced and multivariate statistical methods*


Appendices
Appendix A – The Defining Issues Test 2
**Instructions**

This questionnaire is concerned with how you define the issues in a social problem. Several stories about social problems will be described. After each story, there will be a list of questions. The questions that follow each story represent different issues that might be raised by the problem. In other words, the questions/issues raise different ways of judging what is important in making a decision about the social problem. You will be asked to rate and rank the questions in terms of how important each one seems to you.

This questionnaire is in two parts: one part contains the INSTRUCTIONS (this part) and the stories presenting the social problems; the other part contains the questions (issues) and the ANSWER SHEET on which to write your responses.

Here is an example of the task:

**Presidential Election**

Imagine that you are about to vote for a candidate for the Presidency of the United States. Imagine that before you vote, you are given several questions, and asked which issue is the most important to you in making up your mind about which candidate to vote for. In this example, 5 items are given. On a rating scale of 1 to 5 (1=Great, 2=Much, 3=Some, 4=Little, 5=No) please rate the importance of the item (issue) by filling in with a pencil one of the bubbles on the answer sheet by each item.
Assume that you thought that item #1 (below) was of great importance, item #2 had some importance, item #3 had no importance, item #4 had much importance, and item #5 had much importance. Then you would fill in the bubbles on the answer sheet as shown below.

<table>
<thead>
<tr>
<th>GREAT</th>
<th>MUCH</th>
<th>SOME</th>
<th>LITTLE</th>
<th>NO</th>
</tr>
</thead>
</table>

Rate the following 12 issues in terms of importance (1-5)

1. Financially are you personally better off now than you were four years ago?
2. Does one candidate have a superior moral character?
3. Which candidate stands the tallest?
4. Which candidate would make the best world leader?
5. Which candidate has the best ideas for our country’s internal problems, like crime and health care?

Further, the questionnaire will ask you to rank the questions in terms of importance. In the space below, the numbers 1 through 12 represent the item number. From top to bottom, you are asked to fill in the bubble that represents the item in first importance (of those given you to choose from), then second most important, third most important, and fourth most important. Please indicate your top four choices. You might fill out this part, as follows:

**Rank which issue is the most important (item number).**

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

Note that some of the items may seem irrelevant to you (as in item #3) or not make sense to you—in that case, rate the item as “No” importance and do not rank the item. Note that in the stories that follow, there will be 12 items for each story, not five. Please make sure to consider all 12 items (questions) that are printed after each story.

In addition you will be asked to state your preference for what action to take in the story. After the story, you will be asked to indicate the action you favor on a three-point scale (1 = strongly favor some action, 2 = can’t decide, 3 = strongly oppose that action).

In short, read the story from this booklet, and then fill out your answers on the answer sheet. Please use a #2 pencil. If you change your mind about a response, erase the pencil mark cleanly and enter your new response.

[Notice the second part of this questionnaire, the Answer Sheet. The Identification Number at the top of the answer sheet may already be filled in when you receive your materials. If not, you will receive instructions about how to fill in the number. If you have questions about the procedure, please ask now.]

Please turn now to the Answer Sheet.
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.

[If at any time you would like to reread a story or the instructions, feel free to do so. Now turn to the Answer Sheet, go to the 12 issues and rate and rank them in terms of how important each issue seems to you.]

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.

[Now turn to the Answer Sheet, go to the 12 issues for this story, rate and rank them in terms of how important each issue seems to you.]
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 decision. The first Open Meeting was a disaster. Passionate speeches dominated the microphones and threatened violence. The meeting barely closed without fist-fights. Later in the week, school board members received threatening phone calls. Mr. Grant wonders if he ought to call off the next Open Meeting.

[Now turn to the Answer Sheet, go to the 12 issues for this story, rate and rank them in terms of how important each issue seems to you.]

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 pain-killer 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?

[Now turn to the Answer Sheet, go to the 12 issues for this story, rate and rank them in terms of how important each issue seems to you.]

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 the 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?

[Now turn to the Answer Sheet, go to the 12 issues for this story, rate and rank them in terms of how important each issue seems to you.]
DIT-2 Answer Sheet

University of Minnesota
Copyright, James Rest and Darcia Narvaez
All Rights Reserved, 1998

Please read story #1 in the INSTRUCTIONS booklet.
Famine -- (Story #1)
What should Mustaq Singh do? Do you favor the action of taking the food? (Mark one.)
① Should take the food  ② Can't decide  ③ Should not take the food

Rate the following 12 issues in terms of importance (1-5)
① ② ③ ④ ⑤
1. Is Mustaq Singh courageous enough to risk getting caught for stealing?
2. Isn't it only natural for a loving father to care so much for his family that he would steal?
3. Shouldn't the community's laws be upheld?
4. Does Mustaq Singh know a good recipe for preparing soup from tree bark?
5. Does the rich man have any legal right to store food when other people are starving?
6. Is the motive of Mustaq Singh to steal for himself or to steal for his family?
7. What values are going to be the basis for social cooperation?
8. Is the epitome of eating reconcilable with the culpability of stealing?
9. Does the rich man deserve to be robbed for being so greedy?
10. Isn't private property an institution to enable the rich to exploit the poor?
11. Would stealing bring about more total good for everybody concerned or wouldn't it?
12. Are laws getting in the way of the most basic claim of any member of a society?

Most Important Item ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩⑪ ⑫
Second most Important ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩⑪ ⑫
Third most important ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩⑪ ⑫
Fourth most important ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩⑪ ⑫

Now please return to the Instructions booklet for the next story.

Reporter -- (Story #2)
Do you favor the action of reporting the story? (Mark one.)
① Should report the story  ② Can't decide  ③ Should not report the story

Rate the following 12 issues in terms of importance (1-5)
① ② ③ ④ ⑤
1. Doesn't the public have a right to know all the facts about all the candidates for office?
2. Would publishing the story help Reporter Dayton's reputation for investigative reporting?
3. If Dayton doesn't publish the story wouldn't another reporter get the story anyway and get the credit for investigative reporting?
4. Since voting is such a joke anyway, does it make any difference what reporter Dayton does?
5. Hasn't Thompson shown in the past 20 years that he is a better person than his earlier days as a shop-lifter?
6. What would best serve society?
7. If the story is true, how can it be wrong to report it?
8. How could reporter Dayton be so cruel and heartless as to report the damaging story about candidate Thompson?
9. Does the right of habeas corpus apply in this case?
10. Would the election process be more fair with or without reporting the story?
11. Should reporter Dayton treat all candidates for office in the same way by reporting everything she learns about them, good and bad?
12. Isn't it a reporter's duty to report all the news regardless of the circumstances?

Rank which issue is the most important (item number).

Now please return to the Instructions booklet for the next story.
School Board -- (Story #3)

Do you favor calling off the next Open Meeting?

[ ] Should call off the next open meeting  [ ] Can't decide  [ ] Should have the next open meeting

Rate the following 12 issues in terms of importance (1-5)

1. Is Mr. Grant required by law to have Open Meetings on major school board decisions?
2. Would Mr. Grant be breaking his election campaign promises to the community by discontinuing the Open Meetings?
3. Would the community be even angrier with Mr. Grant if he stopped the Open Meetings?
4. Would the change in plans prevent scientific assessment?
5. If the school board is threatened, does the chairman have the legal authority to protect the Board by making decisions in closed meetings?
6. Would the community regard Mr. Grant as a coward if he stopped the open meetings?
7. Does Mr. Grant have another procedure in mind for ensuring that divergent views are heard?
8. Does Mr. Grant have the authority to expel troublemakers from the meetings or prevent them from making long speeches?
9. Are some people deliberately undermining the school board process by playing some sort of power game?
10. What effect would stopping the discussion have on the community's ability to handle controversial issues in the future?
11. Is the trouble coming from only a few hotheads, and is the community in general really fair-minded and democratic?
12. What is the likelihood that a good decision could be made without open discussion from the community?

Rank which issue is the most important (item number):

Most Important Item  [ ] [ ] [ ] [ ] [ ] Third Most Important  [ ] [ ] [ ] [ ] [ ]
Second Most Important  [ ] [ ] [ ] [ ] [ ] Fourth Most Important  [ ] [ ] [ ] [ ] [ ]

Now please return to the instructions booklet for the next story.

Cancer -- (Story #4)

Do you favor the action of giving Mrs. Bennett more medicine?

[ ] Should give Mrs. Bennett an increased dosage to make her live  [ ] Can't decide  [ ] Should not give her an increased dosage

Rate the following 12 issues in terms of importance (1-5)

1. Isn't the doctor obligated by the same laws as everybody else if giving an overdose would be the same as killing her?
2. Wouldn't society be better off without so many laws about what doctors can and cannot do?
3. If Mrs. Bennett dies, would the doctor be legally responsible for manslaughter?
4. Does the family of Mrs. Bennett agree that she should get more painkiller medicine?
5. Is the painkiller medicine an effective analgesic drug?
6. Does the state have the right to force continued existence on those who don't want to live?
7. Is helping to end another's life ever a responsible act of cooperation?
8. Would the doctor show more sympathy for Mrs. Bennett by giving the medicine or not?
9. Wouldn't the doctor feel guilty from giving Mrs. Bennett so much drug that she died?
10. Should only God decide when a person's life should end?
11. Shouldn't society protect everyone against being killed?
12. Where should society draw the line between protecting life and allowing someone to die if the person wants to?

Rank which issue is the most important (item number):

Most Important Item  [ ] [ ] [ ] [ ] [ ] Third Most Important  [ ] [ ] [ ] [ ] [ ]
Second Most Important  [ ] [ ] [ ] [ ] [ ] Fourth Most Important  [ ] [ ] [ ] [ ] [ ]

Now please return to the instructions booklet for the next story.
Demonstration -- (Story #5)
Do you favor the action of demonstrating in this way?

① Should continue demonstrating in these ways ② Can't decide ③ Should not continue demonstrating in these ways

Rate the following 12 issues in terms of importance (1-5)

① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩ ⑪ ⑫
1. Do the students have any right to take over property that doesn't belong to them?
2. Do the students realize that they might be arrested and fined, and even expelled from school?
3. Are the students serious about their cause or are they doing it just for fun?
4. If the university president is soft on students this time, will it lead to more disorder?
5. Will the public blame all students for the actions of a few student demonstrators?
6. Are the authorities to blame by giving in to the greed of the multinational oil companies?
7. Why should a few people like Presidents and business leaders have more power than ordinary people?
8. Does this student demonstration bring about more or less good in the long run to all people?
9. Can the students justify their civil disobedience?
10. Shouldn't the authorities be respected by students?
11. Is taking over a building consistent with principles of justice?
12. Isn't it everyone's duty to obey the law, whether one likes it or not?

Rank which issue is the most important (item number):

Most important item ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩ ⑪ ⑫
Second most important ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩ ⑪ ⑫
Third most important ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩ ⑪ ⑫
Fourth most important ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩ ⑪ ⑫

Please provide the following information about yourself:

1. Age in years:

2. Sex (mark one): ① Male ② Female

3. Level of Education (mark highest level of formal education attained, if you are currently working at that level [e.g., Freshman in college] or if you have completed that level [e.g., if you finished your Freshman year but have gone on no further].)

① Grade 1 to 8
② Grade 9
③ Grade 10, 11, 12
④ Vocational/technical school (without a bachelor's degree) (e.g., Auto mechanic, beauty school, real estate, secretary, 2-year nursing program)
⑤ Junior college (e.g., 2-year college, community college, Associate Arts degree)
⑥ Freshman in college in bachelor degree program.
⑦ Sophomore in college in bachelor degree program.
⑧ Junior in college in bachelor degree program.
⑨ Professional degree (Practitioner degree beyond bachelor's degree) (e.g., M.D., M.B.A., Bachelor of Divinity, D.D.S. in Dentistry, J.D. in law, Masters of Arts in teaching, Masters of Education [in teaching], Doctor of Psychology, Nursing degree along with 4-year Bachelor's degree)
⑩ Masters degree (in academic graduate school)
⑪ Doctoral degree (in academic graduate school, e.g., Ph.D. or Ed.D.)
⑫ Other Formal Education. (Please describe.)

4. In terms of your political views, how would you characterize yourself (mark one)?

① Very Liberal
② Somewhat Liberal
③ Neither Liberal nor Conservative
④ Somewhat Conservative
⑤ Very Conservative

5. Are you a citizen of the U.S.A.?

① Yes ② No

6. Is English your primary language?

① Yes ② No

Thank You.
Appendix B – IRB Approval
IRB Protocol Number: 10.103.01e

DATE: December 7, 2009

TO: Ethan Sullivan

FROM: Office of Research Protections

RE: Assessing the Cognitive Moral Development of First Year Business Students

Notice of Evaluation – [Exempt 45 CFR 46. 101(b)] 2

The Office for Research Protections (ORP) has evaluated the project named above. According to the information provided, you intend to investigate whether business ethics courses increase ethical decision-making in students. This is a minimal risk study.

This study has been granted an exemption from Boston College IRB review in accordance with 45 CFR 46.101 (b) 2. This designation is based on the assumption that the materials that you submitted to the ORP contain a complete and accurate description of all the ways in which human subjects are involved in your research.

This exemption is given with the following conditions:

1. You will conduct the project according to the plans and protocol you submitted;
2. No further contact with the ORP is necessary unless you make changes to your project or adverse events or injuries to subjects occur;
3. If you propose to make any changes in the project, you must submit the changes to the ORP for IRB review; you will not initiate any changes until they have been reviewed and approved by the IRB;
4. If any adverse events or injuries to subjects occur, you will report these immediately to the ORP.

The University appreciates your efforts to conduct research in compliance with the federal
regulations that have been established to ensure the protection of human subjects in research.

Date of Exemption: December 07, 2009

Sincerely,

Stephen Erickson
Director
Office for Research Protections
COC
IRB Exemption Form
From IRB Review

I. Study Title: Assessing the cognitive moral development of first year business students (IRB Protocol 10.103.01e)

II. Principal Investigator Information
A. Name of Principal Investigator: Ethan Sullivan
B. Are You? (Please check)
   - Faculty
   - Staff
   - Undergraduate Student
   - Graduate Student
   - Postdoctoral fellow
C. Mailing Address: Fulton 414A
D. Department: GSOM
E. E-mail address: sullivan@bc.edu
F. Primary Phone Number: 617-552-0459
G. Alternate Phone:
H. Faculty Advisor's Name: 1. Faculty Advisor's Phone:
I. Faculty Advisor's E-mail:

III. Funding
A. None (Go on to Section IV)
   - Do you plan to apply for funding in the future? Yes* No * Please explain:
   - University Funded: List source:
   - External*: List source and grant number:
   - Federal*: List agency, department and grant number:
   *Wait until you have been notified that your project will be funded before seeking IRB approval unless otherwise instructed by funding source. Submit documentation of funding status with application and a complete copy of the grant with your IRB application.
E. Is BC the primary awardee for the grant? Yes No * If No Please list Primary Awardee:
F. Are there subcontracts? Yes* No If Please list sub-contractors:

IV. Study Information
Research must be "minimal risk" in order to qualify for an Exemption. Minimal risk means that the probability and magnitude of harm or discomfort anticipated in the research are not greater than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests (45 CFR 46).
A. Risk Level: Does this research pose more than minimal risk to participants? Yes* No
   * Greater than minimal risk must be reviewed by the Full Board. Please complete an Initial IRB Application Form
B. Prisoners: Does this research involve interaction with Prisoners or prisoner's private information? Yes* No
   * All prisoner research must be reviewed by the Full Board. Please complete an Initial IRB Application Form
C. Public Data: Will the study utilize archived data, documents, records or biological specimens? Yes* No
   * Provide Source:
   * When were these data collected:
D. Exempt Categories (45 CFR 46.101(b)) Check Category that best describes the study:
   - (1) Research conducted in established or commonly accepted educational settings, involving normal educational practices, such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.
   - This applies only Normal educational research in regular educational settings
   - (2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:
     (i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects, and
     (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be

BC IRB Exemption Form
Submit Materials by E-mail: irb@bc.edu
Version 02, 08/10/2008

Page 1 of 3

194
damaging to the subjects' financial standing, employability, or reputation. This exemption does not apply to children or prisoners.

☐ Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior that is not exempt under paragraph (b)(2) of this section, if:
   (i) the human subjects are elected or appointed public officials or candidates for public office; or (ii) federal statute(s) require(s) without exception that the confidentiality of the personally identifiable information will be maintained throughout the research and thereafter.
   This applies only to elected officials, not officials appointed via a regular hiring process.

☐ Research involving the collection or study of existing data, records, or special pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.
   All data must exist when the application is submitted (if data will be used that is collected or will be collected for clinical purposes, complete the IRB Review Form).

☐ Research and demonstration projects which are conducted by or subject to the approval of department or agency heads, and which are designed to study, evaluate, or otherwise examine:
   (i) Public benefit or service programs; (ii) procedures for obtaining benefits or services under those programs; (iii) possible changes in or alternatives to those programs or procedures; or (iv) possible changes in methods or levels of payment for benefits or services under those programs.
   This applies only to research and demonstration projects under the Federal Social Security Act. This does NOT apply to state or local public service projects that are not pursuant to the Social Security Act.

☐ Taste and food quality evaluation and consumer acceptance studies; (i) if wholesome foods without additives are consumed or (ii) if a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural or environmental contaminant at or below the level found to be safe, by the Food and Drug Administration or approved by the Environmental Protection Agency or the Food Safety and Inspection Service of the U.S. Department of Agriculture.

E. Categories of Sensitive Information (generally not eligible for exemption)

1. Information relating to sexual attitudes, preferences, or practices.
2. Information relating to the use of alcohol, drugs or other addictive products.
3. Information pertaining to illegal conduct.
4. Information that if released could reasonably damage an individual's financial standing, employability, or reputation within the community.
5. Information that would normally be recorded in a patient's medical record and the disclosure of which could reasonably lead to social stigmatization or discrimination.
6. Information pertaining to an individual's psychological well-being or mental health.
7. Genetic information.

Does the study include collection of any sensitive information? ☐ Yes ☐ No

F. Special Subject Populations (generally not eligible for exemption)

1. Minors (under 18 years of age), Not applicable to educational research.
2. Fetuses or products of labor and delivery.
3. Pregnant women (in studies that may influence maternal health)
4. Prisoners.
5. Individuals with a diminished capacity to give informed consent.

Does the study include any special subject populations? ☐ Yes ☐ No
* Indicate population:

G. Research Summary:
Please attach a brief (1-2 page) Research Summary that includes the following items in a Research Summary, labeled and presented in this order:

A. Introduction
B. Specific Aims
C. Methods of Data Collection and Analysis (Qualitative and Quantitative)
D. Description of the subject population, research setting, subject recruitment procedures
E. Informed consent procedure (if consent needed)
F. Provisions for subject and data confidentiality
G. Statement of potential research risks to subjects (e.g., breach of confidentiality, treatment complications)
H. Statement of potential research benefits to subjects (Monetary compensation is not a benefit of participation)
I. Investigator experience (Attach a current copy of your C.V. unless one is on file with the IRB)

Attach any research instruments that will be used for the study (interviews, questionnaires, advertisements) If the study is designed to
devise instruments and test the instruments for validity, state this in the Research Summary. Provide a copy of the materials to the IRB once developed using an Amendment Form. Please submit copies of IRB Training Certificates for all personnel who will interact or collect data.
SIGNATURE OF PRINCIPAL INVESTIGATOR

The undersigned accept(s) responsibility for the study, including adherence to the ethical guidelines set forth in the Belmont Report, Declaration of Helsinki, the Nuremberg Code, the ethical principles of your discipline, the Common Rule and Boston College policies regarding protections of the rights and welfare of human participants participating in this study. In the case of student protocols, the faculty supervisor and the student share responsibility for adherence to policies.

Ethan Sullivan
Print Name of Principal Investigator
Signature of Principal Investigator
Date 10/13/09

SIGNATURE OF FACULTY RESEARCH SUPERVISOR - REQUIRED FOR STUDENT RESEARCH

By signing this form, the faculty research supervisor attests that (s)he has read the attached protocol submitted for IRB review, and agrees to provide appropriate education and supervision of the student investigator, above and share the above Principal Investigator responsibilities.

Print Name of Faculty Research Supervisor
Signature of Faculty Research Supervisor
Date

SIGNATURE OF DEPARTMENT CHAIR OR DEAN - REQUIRED FOR FACULTY RESEARCH

Your signature below affirms you have been informed of the research.

Richard Keeley
Print Name of Chair or Dean
Signature of Chair or Dean
Date 10/13/09

BC IRB Exemption Form
Version 02, 06/10/2006
Submit Materials by E-mail: info@bc.edu

Page 3 of 3
G. Research Summary
A. Introduction and Background
A history of scandals in the business world has been exacerbated by the recent economic crisis. Many of the leaders at the forefront of these scandals are educated at top business schools. Many business schools have responded by including ethics into the curriculum. Do these ethical interventions have an effect upon the ethical decision-making of their participants? A large body of research, utilizing studies involving the Defining Issues Test (DIT), demonstrates that ethical decision-making increases with rising educational levels. Furthermore, interventions have proven to improve ethical decision-making in experimental studies. This body of research uses Kohlberg’s theory of moral development as its framework. Post-Kohlbergian frameworks by James Rest et al., who developed the DIT, have greatly added to the research.

B. Specific Aims/Study Objectives
This study proposes to investigate whether a business ethics course increases the ethical decision-making of its students. Furthermore, the following variables will be controlled in order to see what affects ethical decision-making: gender, SAT scores, ethnicity, political views, instructor, and class size.

C. Materials, Methods and Analysis
This research will use the Defining Issues Test-2. The Defining Issues Test (DIT) is a model of moral development devised by James Rest in 1979. The University of Minnesota formally established the Center for the Study of Ethical Development as a vehicle for research around this test in 1982 and there have been over 500,000 participants since that inception.

The DIT uses a Likert-type scale to give quantitative rankings to five moral dilemmas, the data of which are then analyzed. In 1999 the test was revised in the DIT2 for brevity, clarity and more powerful validity criteria (source: "Center for the Study of Ethical Development" (Website). DIT2. http://www.centerforthestudyofethicaldevelopment.net/DIT2.htm. Retrieved 2009-02-04).

The DIT2 will be administered at Boston College as both a pre-test and post-test. The pre-test will be administered in class the week of October 26th. This is the week prior to the ethical interventions of the course. The post-test will be administered the week of December 9th, which is the last week of the semester.

There are 18 sections of the course in which the DIT2 will be administered. The primary investigator will visit each of these classes to administer the test. He will pass out materials (informed consent form, #2 pencils, DIT2 instructions, DIT2 answer sheet). Then, he will review the informed consent form before continuing to administer the test according to its instructions. The DIT2 takes approximately 30-35 minutes to complete. The pre-test and post-test answer sheets need to be matched so that the data can match. There is a room on the answer sheet for a five digit identification number. The informed consent form will be attached to the answer sheet. Then, the primary investigator will assign a five digit identification number to each answer sheet. At that point, the primary investigator will have three sets of data. First will be the answer sheets with unique and non-descript identification numbers. Second will be the informed consent forms. Third
will be a spreadsheet with student names and the corresponding five digit identification number. These three sets will be saved in different and secure sources. The answer sheets will be secured in a file cabinet and eventually mailed to the "Center for the Study of Ethical Development" for scoring. The informed consent forms will be filed in a locked cabinet of my desk. The spreadsheet will be saved on a password protected electronic storage device.

In December, the post-test will be administered. The students and primary investigator will repeat the test-administrations steps. The spreadsheet will be used to assign the matching five digit identification number to the answer sheets. The answer sheets will be immediately mailed to "The Center for the Study of Ethical Development.” The informed consent forms and spreadsheet will be securely stored as stated above. "The Center for the Study of Ethical Development" will score the forms then will send me a printed report as well as the data on disk. The report will be an overview of the descriptive statistics. The data will be imported into an SPSS file, which I will use to analyze them. My analysis will use a two-way ANOVA to see if there is a statistically significant difference between the pre-test and post-test. If my assumptions are correct, and there is a difference (i.e. gain in scores of moral development), then I will perform post-hoc Sheffe tests to find out what variables may impact the difference. I will see if my dependent variable (difference score) is impacted by my independent variables, which include gender, ethnicity, SAT scores, political viewpoints, instructor, and class size.

D. Research Population and Recruitment Methods

This study investigates the effect of an ethics intervention on the moral development of freshmen business students. Participants will include students registered for this course, which is a requirement for freshmen business majors. Since I am interested in instructor and class size effect, I will study only the students who are taught by the full-time instructors of the course. Therefore, the population is 400, as there are 400 students enrolled in this course and taught by the three full-time instructors. I have communicated with these instructors and have received their permission to visit their classes to administer the test. Power analysis has determined that this is an appropriate number given the amount of independent variables of the study. There is no recruitment strategy for this study, since the PI will visit classes to collect the data. Students are not informed of this study prior to the day of administration, can opt out of the study, and do not receive compensation for their participation.

E. Informed Consent Procedure

The Primary Investigator will perform the informed consent procedure. He will review it with all participants in class. The Primary Investigator completed his Ph.D. coursework in Higher Education Administration at Boston College. As part of the coursework he completed the IRB Training at Boston College. Participants can ask questions about the study and the informed consent procedure at any time during the administration. They will also be given the contact information of the primary investigator should they have any follow-up questions.

F. Confidentiality

The pre-test and post-test answer sheets need to be matched so that the data can match. There is a room on the answer sheet for a five digit identification number. The informed
The primary investigator will assign a five digit identification number to each answer sheet. At that point, the primary investigator will have three sets of data. First will be the answer sheets with unique and non-descript identification numbers. Second will be the informed consent forms. Third will be a spreadsheet with student names and the corresponding five digit identification number. These three sets will be saved in different and secure sources. The answer sheets will be secured in a file cabinet and eventually mailed to the "Center for the Study of Ethical Development" for scoring. The informed consent forms will be filed in a locked cabinet of my desk. The spreadsheet will be saved on a password protected electronic storage device.

In December, the post-test will be administered. The students and primary investigator will repeat the test-administrations steps. The spreadsheet will be used to assign the matching five digit identification number to the answer sheets. The answer sheets will be immediately mailed to "The Center for the Study of Ethical Development." The informed consent forms and spreadsheet will be securely stored as stated above.

G. Potential Risks

The DIT2 measures the cognitive moral development of the participants. The questions are hypothetical and dilemma-based. There are minimal risks associated with taking the DIT2. The informed consent will explain any possible risks (i.e., being discomforted with the hypothetical dilemmas) and will offer the primary investigator as a resource for any discomfort that may occur. In the unlikely event that a participant could have an adverse effect, faculty of the course have been trained by the University’s Counseling Center on how to notice and refer any signs of psychological issues.

H. Potential Benefits

As with risks, there are minimal to no benefits for the individuals. Scores will be coded so that individuals cannot be identified, therefore individuals will not have access to their scores. It is possible, but highly unlikely, that taking this test can improve their moral development. However, it is statistically likely that the intervention of the course will improve their moral development. Therefore, the benefits would be due to the course, not the research study. While the research study will not benefit the participants, its findings could benefit subsequent enrollees in the course.
Introduction
• You are being asked to be in a research study of the impact of Portico on your decision making.
• You were selected as a possible participant because you are registered for one of the sections that is being studied.
• We ask that you read this form and ask any questions that you may have before agreeing to be in the study.

Purpose of Study:
• The purpose of this study is to see if Portico increases the moral development of freshmen, and to see what variables or conditions are optimal for this development.
• The total number of subjects is expected to be 400.

Description of the Study Procedures:
If you agree to be in this study, we would ask you to do the following things: read and sign this form and complete the Defining Issues Test-2 survey that is attached. The DIT2 will take approximately 30 minutes to complete. You will then re-take the DIT2 at the end of the semester. After all surveys are completed, they will be mailed to the Center for the Study of Ethical Development at the University of Alabama. The Center will score the data and will mail me a report and the raw data, which I will use to run a variety of statistical analyses.

Risks/Discomforts of Being in the Study:
• There are no reasonable foreseeable(or expected) risks. There may be unknown risks.

Benefits of Being in the Study:
• The purpose of the study is to see if Portico increases the moral development of freshmen, and to see what variables or conditions are optimal for this development.
• There are no expected personal benefits from participation in this study. However, this study will benefit subsequent Carroll School members, and will help us to gauge the impact of your freshmen experience. Due to the anonymous nature of the study,
we will not have individual results to share, but we will have results for the composite.

Payments:
• You will not receive any payment or reimbursement for participating in this study.

Costs:
• There is no cost to you to participate in this research study.

Confidentiality:
• The records of this study will be kept private. In any sort of report we may publish, we will not include any information that will make it possible to identify a participant. Research records will be kept in a locked file.
• All electronic information will be coded and secured using a password protected file.
• Access to the records will be limited to the researchers; however, please note that the Institutional Review Board and internal Boston College auditors may review the research records.

Voluntary Participation/Withdrawal:
• Your participation is voluntary. If you choose not to participate, it will not affect your current or future relations with the University.
• You are free to withdraw at any time, for whatever reason.
• There is no penalty or loss of benefits for not taking part or for stopping your participation. The subject does not jeopardize grades nor risk loss of present or future faculty/school/University relationships.

Contacts and Questions:
• The researcher conducting this study is Ethan Sullivan. For questions or more information concerning this research you may contact him at ethan.sullivan@bc.edu.
• If you believe you may have suffered a research related injury, contact Ethan Sullivan at 617-552-0459 who will give you further instructions.
• If you have any questions about your rights as a research subject, you may contact: Director, Office for Human Research Participant Protection, Boston College at (617) 552-4778, or irb@bc.edu

Copy of Consent Form:
• You will be given a copy of this form to keep for your records and future reference.

Statement of Consent:
• I have read the contents of this consent form and have been encouraged to ask questions. I have received answers to my questions. I give my consent to participate in this study. I will receive a copy of this form.
Signatures/Dates

Study Participant (Print Your Name):

___________________________________________

Participant Signature: ________________________________ Date

____
Appendix C

2009 Portico Syllabus
COURSE DESCRIPTION AND RATIONALE

Portico offers the beginning student a unique opportunity to situate contemporary business in a context that is global and historical; learn about business through engagement with faculty and practitioners; develop a nuanced method for recognizing and responding to the ethical challenges of contemporary business; and raise questions about personal aspirations and the opportunities available in the world of work. During the first half of the course, we begin with a wide and historically informed consideration of global, national, and regional issues and end with a discussion of industry, organizational, and functional issues. During the second half of the course, we consider more personal issues, including ethics, leadership, and personal/professional development. The choice of readings and assignments is designed to reinforce the interconnections across the levels of content.

COURSE FORMATS

As the description suggests, Portico’s goals are ambitious and to realize them will require a variety of teaching and learning strategies.

- You will do much writing. Some of that writing will be straightforward and analytical. A significant number of assignments, however, will require that you think and write in a more personal, reflective manner.
- You will work individually and in teams.
- You’ll benefit from the attention of the instructor, student assistants, faculty from several departments, alumni and business practitioners.

TEXTS AND READINGS

To purchase (available at the BC Bookstore)

- Sullivan (ed.), Portico (readings from this publication are marked +++ in the calendar section)
- The Wall Street Journal (15 week subscription through the bookstore)

On e-reserve

- Several of our readings are on e-reserves through the BC Library. In the calendar section, readings marked with *** are available on e-reserve. The direct link to the e-reserve appears
in the calendar section below and on the Portico website (www.bc.edu/portico, click on “readings”).

BlackBoard Vista

Blackboard vista is an online teaching tool that is accessed via cms.bc.edu or your agora portal. Portico has two folders. One is the “general” Portico folder that includes links to readings and videos of the Monday evening content. The other folder is for our section, MH100.03, and is specific to us.

GRADING AND EVALUATION

The following assignments will be due throughout the semester (refer to the assignments column of the calendar for details):

- Summer essay: 5%
- Assignment #2: 10%
- Group project: 25%
- Self-assessment and four year trajectory: 10%
- Final Exam: 25%
- Quizzes and Class Participation 25%

COURSE POLICIES

Academic Integrity:
The Portico program holds the highest possible standards for academic honesty. Instances of cheating, plagiarism, collusion, or any type of dishonesty will be treated seriously. To help understand the concept of academic integrity, all Portico students will complete an on-line tutorial. The complete university policy is found here: http://www.bc.edu/offices/stserv/academic/resources/policy.html#integrity

Class Attendance:
Portico requires class participation and all member of the class are valued members of the community. Class attendance is of utmost importance. In the event of having to miss class for a significant event or emergency, you should contact your instructor. Unexcused absences will affect your final grade.

Ethics Requirement:
Portico fulfills the ethics requirement within the Carroll School. A Carroll School student who does not successfully complete Portico will not meet this requirement, thereby jeopardizing their ability to remain enrolled in the Carroll School. In certain situations, a student will be able to fulfill the requirement by taking MH011: Introduction to Ethics.

Special Accommodations:
If you have a disability and will be requesting accommodations for this course, please register with either Kathy Duggan (Kathleen.duggan@bc.edu) Associate Director, Academic Support Services, the Connors Family Learning Center (learning disabilities and ADHD) or Suzy Conway (suzy.conway@bc.edu), Assistant Dean for Students with Disabilities (all other disabilities). Advance notice and appropriate documentation are required for accommodations.
## Calendar

<table>
<thead>
<tr>
<th>Weekly Topics</th>
<th>Readings</th>
<th>Assignments</th>
<th>Evening Sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Week One</strong></td>
<td><strong>Learning about business through the cell phone.</strong></td>
<td><strong>Summer assignment due August 28th</strong></td>
<td><strong>Labor Day – No evening session</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Gerard Goggin, <em>Cell phone culture</em></strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong><a href="http://irm.bc.edu/reserves/mh100/keel/mh10006.pdf">http://irm.bc.edu/reserves/mh100/keel/mh10006.pdf</a></strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>The Economist: <em>Nomads at last: A special report on mobile telecoms</em>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong><a href="http://irm.bc.edu/reserves/mh100/sull/mh10020.pdf">http://irm.bc.edu/reserves/mh100/sull/mh10020.pdf</a></strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sara Corbett, “Can the cell phone help end global poverty?”***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong><a href="http://irm.bc.edu/reserves/mh100/sull/mh10017.pdf">http://irm.bc.edu/reserves/mh100/sull/mh10017.pdf</a></strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week Two</strong></td>
<td><strong>Value chains and industry analysis</strong></td>
<td><strong>Assignment #2 due Sept. 30th at 4pm</strong></td>
<td><strong>Professors and alumni from Accounting and Finance</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Michael Porter, “How Competitive Forces Shape Strategy” (pp. 5 – 25 in <em>Portico</em>)</strong>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Apple Inc., 2008” (pp. 27 – 58 in <em>Portico</em>)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week Three</strong></td>
<td><strong>Innovation and entrepreneurship</strong></td>
<td><strong>Special sessions with Dean Andy Boynton</strong></td>
<td><strong>Special sessions with Dean Andy Boynton</strong></td>
</tr>
<tr>
<td></td>
<td>The Economist: <em>Global Heroes: A special report on entrepreneurship</em>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong><a href="http://irm.bc.edu/reserves/mh100/sull/mh10021.pdf">http://irm.bc.edu/reserves/mh100/sull/mh10021.pdf</a></strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Learn about venture capital at <strong><a href="http://www.nvca.org/def.html">http://www.nvca.org/def.html</a></strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week Four</strong></td>
<td><strong>The cell phone, revisited</strong></td>
<td><strong>Professors and alumni from Marketing and Operations</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Monday: How-to/workshop class for projects and presentations</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Introductory chapters from Accounting and Finance texts (on e-reserve)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Accounting – “Financial Statements and Business Decisions” by Libby,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong><a href="http://irm.bc.edu/reserves/mh100/keel/mh10004.pdf">http://irm.bc.edu/reserves/mh100/keel/mh10004.pdf</a></strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Finance - “The Corporation and the Financial Manager,” by Brealey,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong><a href="http://irm.bc.edu/reserves/mh100/keel/mh10003.pdf">http://irm.bc.edu/reserves/mh100/keel/mh10003.pdf</a></strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week Five</strong></td>
<td><strong>Ethics, an introduction</strong></td>
<td><strong>Professors and alumni from Marketing and Operations</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Lynn Sharp Paine, “Ethics: A Basic Framework” (pp. 79 – 86 in <em>Portico</em>)</strong>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Richard Spinello, “A Critical Overview of Ethical Frameworks” (pp. 63 – 78 in <em>Portico</em>)</strong>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Introductory chapters from Marketing and Operations texts (on e-reserve)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Marketing – “What is Marketing,” by Lamb,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong><a href="http://irm.bc.edu/reserves/mh100/keel/mh10008.pdf">http://irm.bc.edu/reserves/mh100/keel/mh10008.pdf</a></strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Operations - “Introduction to Operations Management,” by Davis,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong><a href="http://irm.bc.edu/reserves/mh100/keel/mh10009.pdf">http://irm.bc.edu/reserves/mh100/keel/mh10009.pdf</a></strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week Six</strong></td>
<td><strong>Team Projects</strong></td>
<td><strong>Columbus Day – No evening session</strong></td>
<td><strong>Columbus Day – No evening session</strong></td>
</tr>
<tr>
<td><strong>Week Seven</strong></td>
<td><strong>Presentations</strong></td>
<td><strong>Group projects due Oct. 19th at 4pm</strong></td>
<td><strong>Professors and alumni from Info Systems and Organizational Studies</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Group Presentations (assigned by lottery)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Introductory chapters from Info Systems and Organizational Studies texts (on e-reserve)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Org Studies – “Why mastering organizational behavior is essential to your career,” by Andre,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong><a href="http://irm.bc.edu/reserves/mh100/keel/mh10007.pdf">http://irm.bc.edu/reserves/mh100/keel/mh10007.pdf</a></strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Week Eight</strong></td>
<td><strong>Ethical Egoism,</strong></td>
<td><strong>Optional: Winston Forum on Business</strong></td>
<td><strong>Optional: Winston Forum on Business</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Utilitarianism (pp. 255 - 262 in <em>Portico</em>)</strong>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week Nine</td>
<td>Deontology</td>
<td>Ethics (see recommended events below)</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
<td>----------------------------------------</td>
<td></td>
</tr>
</tbody>
</table>
| Corporate Citizenship | Handouts on ethical egoism and relativism | Deontological Theories, (pp. 163 - 215 in Portico)+++
James Waters, “Catch 20.5: Corporate morality as an organizational phenomenon” (on e-reserve)
http://irm.bc.edu/reserves/mh100/sull/mh10018.pdf | Corporate Social Responsibility with Cheryl Kiser, Managing Director Center for Corporate Citizenship |

<table>
<thead>
<tr>
<th>Week Ten</th>
<th>Ethics</th>
<th></th>
</tr>
</thead>
</table>
| | Deontological Theories, (pp. 163 - 215 in Portico)+++
James Waters, “Catch 20.5: Corporate morality as an organizational phenomenon” (on e-reserve)
http://irm.bc.edu/reserves/mh100/sull/mh10018.pdf | Plato, Gorgias (pp. 89 - 159 in Portico)+++ |
| | William Cronon, “Only Connect”
http://www.williamcronon.net/writing/Cronon_Only_Connect.pdf | Ethical decision-making |
| | David Fritzshe, WorldCom’s Creative Accounting (pp. 349 – 351 in Portico)+++ | An evening with Landen Williams |

<table>
<thead>
<tr>
<th>Week Eleven</th>
<th>Virtue Ethics and beyond</th>
<th></th>
</tr>
</thead>
</table>
| Nicomachean Ethics, Books Two and Three (on e-reserve)
Book 2 - http://irm.bc.edu/reserves/mh100/keel/mh10010.pdf
Book 3 - http://irm.bc.edu/reserves/mh100/keel/mh10011.pdf | Plato, Gorgias (pp. 89 - 159 in Portico)+++ |
| The Natural Law tradition and Business Ethics (pp. 219 - 252 in Portico)+++ | View Sophie Scholl on BC cable |
| Robert Spitzer, S.J., “Six Steps for Remedi
Contemporary Ethical Problems” (pp. 265 – 279 in Portico)+++ | Natural Law with Professor Richard Spinello |

<table>
<thead>
<tr>
<th>Week Twelve</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Thanksgiving Break – no class</td>
<td>Self-assessment and four year trajectory due Nov. 23rd by 4pm</td>
<td>Thanksgiving Break – no class</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Week Thirteen</th>
<th>Ethical Leadership</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Johnson and Johnson Case &amp; James Burke, A Career in American Business A and B (pp. 285 – 328 in Portico)+++</td>
<td>Burke Video Sadhu Video</td>
<td>Burke Video Sadhu Video</td>
</tr>
<tr>
<td>Parable of the Sadhu (pp. 329 – 335 in Portico)+++</td>
<td></td>
<td>tbd</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Week Fourteen</th>
<th>Ethical Leadership</th>
<th></th>
</tr>
</thead>
</table>
| Chris Lowney, Heroic Leadership (on e-reserve)
pp. 13 – 35 http://irm.bc.edu/reserves/mh100/keel/mh10001.pdf
pp.277-295 http://irm.bc.edu/reserves/mh100/keel/mh10002.pdf | Final Exam due Dec. 11th at 4pm | An evening with Chris Lowney |

---

**Recommended Events**

The Winston Center for Leadership and Ethics has scheduled several fall events that complement our course. I hope that you will be able to take advantage of them.

**October 22**

A Winston Center Collaboration with the Boston College Alumni Association featuring Liz McCartney ’94, 2008 CNN "Hero of the Year." Fulton Hall Honors Library, 4:30pm.
October 28
Winston Forum on Business Ethics featuring Mike Dupee '91, Vice President for Corporate Social Responsibility, Green Mountain Coffee. Murray Room, Yawkey Center, 4:30pm.

November 12
The Chambers Lecture Series featuring Father Greg Boyle, S.J., Founder and Executive Director of Homeboy Industries, Los Angeles, CA. Gasson 100, 4:30pm.

November 16
Clough Colloquium featuring Alex Counts, President and CEO, Grameen Foundation. Heights Room, Corcoran Commons, 4:30pm.