J Bus Ethics (2016) 133:141–163
DOI 10.1007/s10551-014-2349-7

Insights Regarding the Applicability of the Defining Issues Test to Advance Ethics Research with Accounting Students: A Meta-analytic Review

Anne L. Christensen · Jane Cote · Claire K. Latham

Received: 16 May 2013 / Accepted: 31 August 2014 / Published online: 10 September 2014

Abstract Numerous researchers have investigated accounting students' levels of moral reasoning, ethical choice and judgment employing the Defining Issues Test (DIT) and using its *P* score as an indicator of moral reasoning. Not surprisingly, a number of DIT studies report conflicting results. Moreover, despite widespread use of the DIT, there is concern that it may not adequately measure all facets of ethical judgment (cf. Bailey et al., Behav Res Account 22(2):1–26, 2010). Thus, we endeavor to provide insight not only into the contradictory results but also about the applicability of the DIT for studying accounting students. To do so, we collect published and unpublished DIT studies employing accounting students as subjects and use meta-analysis to aggregate findings across these studies to quantify their results, examining commonly employed variables. We show significant relationships between *P* scores and some variables (length of professional experience, choice of major, political ideology, gender, GPA and education level) but not others (age). Further, our findings demonstrate that the DIT provides added insights when exploring questions of ethical choice, and ethics instruction, particularly when the instruction is embedded in an accounting course. Finally, we find that the level of DIT *P* scores reported in the studies relates to whether the study was published. We discuss the implications of our findings for future research.

Keywords Accounting · Defining Issues Test · Ethics research · Ethics education · Meta-analysis · Ethics assessment

A. L. Christensen
Jake Jabs College of Business and Entrepreneurship, Montana State University, P.O. Box 173040, Bozeman, MT 59717-3040, USA
e-mail: annec@montana.edu

J. Cote (✉) · C. K. Latham
Carson College of Business, Washington State University, 14204 NE Salmon Creek Avenue, Vancouver, WA 98686-9600, USA
e-mail: janecote@vancouver.wsu.edu

C. K. Latham
e-mail: clatham@vancouver.wsu.edu

Introduction

For more than 25 years, accounting researchers have studied accounting students to explore levels of moral reasoning, judgment and ethical choice. Many of these studies have investigated how ethics instruction affects the cognitive moral development of accounting students. Others have examined students' political ideology, gender, grade point average (GPA), age, and education level in relationship to moral reasoning or ethical judgment. The Defining Issues Test (DIT) (Rest 1979, 1986) measures moral reasoning development and has been a predominant variable in much of this research. It has been touted as "the measure of choice in higher-education assessments of moral judgment development" (Bailey et al. 2010, p. 8). Given its extensive use in accounting research, we explore quantitatively, using meta-analytic techniques, the extent to which these diverse studies provide common insights into the ethical reasoning of accounting students.

Understanding the value the DIT has in assessing the effectiveness of ethics education is critical, arguably now more than ever before, for both education and the accounting profession. Students entering the accounting profession are confronted with increasingly complex financial environments, presenting more complicated ethical challenges. Business education is expected to prepare

students for these challenges. The Association to Advance Collegiate Schools of Business Accounting Accreditation Standards now requires accounting degree programs to establish expectations for ethical behavior by students as well as provide effective learning experiences in ethical understanding and reasoning (AACSB International 2009). The Pathways Commission also underscores the critical need for education to successfully "build skills in ethical decision making and responsible judgment" (American Institute of Certified Public Accountants and American Accounting Association 2012, p. 133). From a legislative perspective, several states now require CPA candidates to have a strong foundation in ethics instruction with identifiable accounting ethics coursework (e.g., California, New York) and/or a Board-approved stand-alone 3-semester hour ethics course, including the topics of ethical reasoning, integrity, objectivity and independence, taken at a recognized college or university (e.g., Maryland, Texas).

We conduct meta-analyses to provide a means of quantitatively integrating the results of multiple accounting studies that employ the DIT, thus resulting in greater power and an increased likelihood of determining whether relationships are statistically significant. Our findings demonstrate that individual constructs such as length of professional experience, choice of major (accounting versus non-accounting), political ideology, gender, GPA, and education level are differentiators of moral reasoning ability as measured by the DIT. As well, the DIT provides added insights when exploring questions of ethical choice, and ethics instruction, especially when instruction is embedded in an accounting course. Finally, we find that the level of DIT *P* scores reported in the studies relate to whether the study was published.

The remainder of the article proceeds in four sections. In the next section, we briefly describe the DIT, using the Rest (1986) four-stage model of ethical decision-making and we present the research questions we investigate in light of previous studies using the DIT. We follow with a discussion of our methodology, including background on meta-analysis as well as methods used in compiling and coding the data. Third, we present our findings. Finally, we discuss our results, providing suggestions for future research.

Literature Review

DIT

The DIT was developed by Rest (1979, 1986) as an objective measure of moral reasoning. Rest based the scale on the six stages of cognitive moral development founded by Kohlberg (1969). Kish-Gephart et al. (2010) describe cognitive moral development (CMD) theory as explaining the progression of individuals' ability to reason concerning moral rightness or wrongness over time. Rest (1979) posits that the first two stages of cognitive moral development are pre-conventional reasoning in which individuals largely focus on themselves and make decisions based on the likelihood of punishment or rewards. The next two stages of CMD are considered conventional reasoning. Individuals who reason at this level are concerned with society's norms and rules for appropriate behavior. The final two stages of CMD are regarded as post-conventional reasoning and individuals' decisions are determined by their beliefs about what is just or fair. The ethics literature has generally supported a positive relationship between levels of moral reasoning and ethical judgment, yet less clear evidence exists concerning the relationship between moral reasoning and ethical action (Bailey et al. 2010; Kish-Gephart et al. 2010; O'Fallon and Butterfield 2005). The DIT captures *a capacity to make ethical judgments* rather than the ethical decision itself. Bailey et al. (2010) suggest that there exists a lack of clarity concerning what the DIT actually measures and, as such, it fails to reflect that ethical judgment is only part of the larger picture of ethical behavior.

These results may stem from many accounting research studies including only one or two of the four stages of Rest's model (1986) of ethical behavior. Rest's first stage, moral sensitivity, is where individuals must recognize the moral dimensions of an issue within a specific context. The next stage requires individuals to employ moral reasoning, which the DIT strives to measure. The third stage, moral motivation, is when individuals organize their priorities for taking ethical actions when facing competing demands. The fourth stage, moral character, requires individuals to determine whether to act consistently with their motivation when they are facing a particular situation. Thus, studying moral reasoning without including moral sensitivity, moral motivation and moral character may result in the lack of clarity noted by Bailey et al. (2010).

The *P* score is the most prevalent DIT indicator used in accounting student ethics research (cf., Bailey et al. 2010). It represents the relative importance individuals give to the highest two stages of moral development, stage 5 and 6, described as post- conventional moral reasoning. The *P* score, in particular, has been used to examine whether moral reasoning is linked to ethical student behavior and whether ethical reasoning increases with ethics instruction, as well as the influence of various personality and environmental characteristics. Some accounting ethics studies show ethical reasoning as measured by the DIT is linked to ethical student behavior or decisions (Abdolmohammadi and Baker 2008; Radtke 1999), but others find no significant association (Massey 1997). Research regarding the effectiveness of various forms of ethics instruction also has been mixed (Bloodgood et al. 2010; O'Fallon and

Butterfield 2005). Meta-analysis enables us to aggregate the studies that investigate the same variables to determine the relationship between the DIT measures of moral reasoning with variables commonly explored in accounting research.

Research Questions

Whereas dozens of accounting research studies use the DIT, our purpose is to examine the commonly employed variables across accounting student studies to determine the strength and direction of relationships between *P* scores and variables such as ethical choice, ethics instruction, political ideology, demographic factors and whether the *P* scores of accounting students significantly differ from those of other populations. As these are the variables most frequently expected to impact ethical attitudes and choice, this approach will lead to findings that can inform future research. Using the DIT *P* score as the measure of moral reasoning, we describe nine variables, comprising our nine research questions that represent nine key DIT associations investigated in prior ethics research.

We begin by exploring whether the population of accounting students differs, that is, does the ethical reasoning of accounting students vary from other groups. We first compare accounting students to accounting professionals and then compare accounting majors to other college majors. Table 1 provides a summary of all research questions linked to the studies examined.

Accounting Majors and Professionals

Do the ethical reasoning levels of accounting students differ from those of individuals who are working in the accounting profession? Senior accounting students are soon-to-be professionals and therefore the student subject pool is of interest (cf. Massey 2002). If there is no significant difference in moral reasoning between accounting college seniors and practicing professionals, then this aspect of entry to the profession may be a smooth transition. If significant differences exist between experienced accountants and senior accounting students, then effective ethics instruction becomes crucial to close the gap, provided ethics instruction can actually increase the moral reasoning of college students (Milne 2001). King and Mayhew (2002) posit that moral reasoning ability increases as individuals are confronted with moral dilemmas, which suggests that accounting professionals are more likely to reason at higher levels than students. However, controversy exists over whether accountants' moral reasoning actually decreases as they reach higher levels within their firms (Ponemon 1992; Bernardi and Arnold 2004; Scofield et al. 2004). Therefore, we examine the collective impact of the relationship between accounting experience and moral reasoning with the following research question.

RQ1 Do practicing accountants have different levels of moral reasoning, as measured by the DIT *P* score, than accounting college seniors?

Accounting versus Non-accounting Majors

Is there a difference in ethical reasoning levels between accounting majors and non-accounting majors while in college? Research on moral reasoning development and students pursuing majors in different disciplines has not produced consistent results (Maeda et al. 2009). Pascarella and Terenzini (2005) suggest that moral reasoning growth is likely to occur at a faster rate in a liberal arts environment, while moral reasoning growth among students in more technical fields is likely to be slower. By contrast, Snodgrass and Behling (1996) find no significant differences in moral reasoning levels of business students and non-business students majoring in arts and humanities, social sciences, natural sciences or with undeclared majors. In sum, as prior research related to major has presented mixed results, we include college major in our meta-analysis to explore the collective findings, as articulated in our second research question.

RQ2 Do DIT *P* scores differ between accounting majors and other business majors as well as non-business majors?

Ethical Choice

Whether an individual acts ethically is, arguably, the most critical research question. The term, ethical choice, is used here to encompass ethical behavior, ethical intentions, and ethical decisions in response to vignettes, hypothetical dilemmas or observed behavior. This definition is consistent with a meta-analysis study reviewing unethical choice in organizations (Kish-Gephart et al. 2010). Prior research demonstrates a gap between moral judgment and moral action. Therefore, several studies explore the ethical choices made by participants to assess the relationship between moral attitudes and decision making (Radtke 1999).

Kish-Gephart et al. (2010) found an inverse relationship between moral reasoning, as measured by the DIT and *un*ethical choice in the workplace, suggesting that individuals with higher capacities for moral reasoning are uncomfortable making unethical choices. In turn, the findings suggest that as levels of moral reasoning increase, individuals make more ethical decisions and behave more ethically. However, the research findings are mixed. Bay (2002) cites prior research evidence indicating that for non-accounting subjects, only about half the studies found

Table 1 Summary of studies by research question

Studies		RQ1	RQ2	RQ3	RQ4	RQ5	RQ6	RQ7	RQ8	RQ9
Author(s) and year of publication	Journal/Dissertation/ Working Paper	Practicing Accountants	Major	Ethical Choice	Ethics Instruction	Political Ideology	Gender	GPA	Age	Class Level
Abdolmohammadi (2005)	Research on Professional Responsibility and Ethics in Accounting				× Pre/post		×		×	×
Abdolmohammadi and Baker (2008)	CPA Journal			×			×			
Abdolmohammadi et al. (2009)	Journal of Accounting Education						×	×	×	
Armstrong (1984)	Dissertation (unpublished)	×					×		×	×
Armstrong (1993)	Journal of Accounting Education				× Pre/ post, course		×		×	×
Bancroft (2002)	Dissertation (unpublished)			×	× Course				×	
Bernardi (1995)	Accounting Educators' Journal						×			×
Bernardi et al. (2002)	Research on Accounting Ethics				× Course		×	×		
Bonawitz (2002)	Dissertation (unpublished)		×		× Course					×
Brown-Liburd and Porco (2011)	Issues in Accounting Education						×	×		
Buell (2009)	Dissertation (unpublished)				× Course		×		×	×
Burks (2006)	Dissertation (unpublished)		×							×
Christensen et al. (2010)	Advances in Accounting Education			×			×			×
Cote et al. 2013	Advances in Behavioral Accounting Research			×			×	×	×	×
Cruz (2003)	Dissertation (unpublished)			×						×
Delaney (2005)	Dissertation (unpublished)				× Pre/post					
Dellaportas (2006)	Journal of Business Ethics				× Pre/post			×	×	×
Dellaportas et al. (2008)	Accounting & Finance									×
Douglas and Schwartz (1998)	Research on Accounting Ethics		×				×			×
Earley and Kelly (2004)	Issues in Accounting Education				× Pre/post					×
Eynon et al. (1996)	Journal of Accounting Education					×	×			×
Fisher and Ott (1996)	Research on Accounting Ethics							×		×
Fisher and Sweeney (1998)	Journal of Business Ethics					×				
Fisher and Sweeney (2002)	Advances in Accounting Behavioral Research					×				×
Hajjar (1998)	Dissertation (unpublished)		×	×						
Hickman (2008)	Dissertation (unpublished)		×		× Course				×	×
Ho (1997)	Research on Accounting Ethics	×								×
Icerman et al. (1991)	Accounting Educators' Journal		×							×
Jeffrey (1993)	Issues in Accounting Education		×							×

Table 1 continued

Studies		RQ1	RQ2	RQ3	RQ4	RQ5	RQ6	RQ7	RQ8	RQ9
Author(s) and year of publication	Journal/Dissertation/ Working Paper	Practicing Accountants	Major	Ethical Choice	Ethics Instruction	Political Ideology	Gender	GPA	Age	Class Level
Jeffrey and Ravenscroft (2000)	Research on Accounting Ethics		x							x
Klimek and Wenell (2011)	Academy of Educational Leadership Journal				× Course	×	×		×	
Lampe and Finn (1992)	Auditing-a Journal of Practice & Theory	×								×
Massey (1997)	Dissertation (unpublished)	×		×						
Massey (2002)	Research on Accounting Ethics									×
Massey and Thorne (2006)	Behavioral Research in Accounting	×								×
Ponemon (1993)	Journal of Accounting Education				× Pre/post		×			×
Ponemon and Glazer (1990)	Issues in Accounting Education	×								×
Radtke (1999)	Advances in Accounting Behavioral Research			×						×
Rogers and Smith (2001)	Teaching Business Ethics			×						
Shaub (1994)	Journal of Accounting Education	×			× Course		×	×	×	×
St. Pierre et al. (1990)	The Accounting Educators' Journal		×		× Course		×			×
Sweeney and Fisher (1998) (includes 2nd experiment)	Behavioral Research in Accounting					×				
Thorne (1999)	Journal of Accounting Education						×		×	×
Thorne (2000)	Behavioral Research in Accounting									×
Thorne (2001)	Journal of Accounting Education						×		×	
Venezia (2005)	The Journal of American Academy of Business				× Course		×		×	
Wilhelm (2010)	Delta Pi Epsilon Journal				× Pre/post					×

significant relationships between the DIT and ethical choice.

Therefore, we explore the association between ethical choice and cognitive moral development as measured by the DIT with the following research question.

RQ3 Do ethical choices differ for accounting students with different levels of moral reasoning, as measured by the DIT *P* score?

Ethics Instruction

The efficacy of ethics instruction is difficult to assess. With technical accounting topics, an examination can be administered to assess proficiency. An ethics examination can measure students' proficiency with philosophical ethical dimensions and the frameworks for making ethical decisions. However, it is more challenging to assess whether students' attitudes and moral reasoning ability have improved with ethics instruction. One approach several studies adopt has been to measure how DIT scores for individuals are affected in the presence of educational interventions.

Two meta-analyses, using non-accounting student subjects, find conditions where ethics instruction improves moral reasoning DIT scores. Schlaefli et al. (1985) found that instruction did positively impact DIT scores with students from junior high to adult. The impact was modest with younger subjects and more significant with adults.

Waples et al. (2009) narrowed their investigation to business ethics instruction and also found a positive association with instruction and DIT scores when the instruction was centered on moral reasoning abilities. These results are consistent with King and Mayhew's (2002) suggestion, drawn from their extensive literature review of DIT studies in higher education that college experiences that require students to wrestle with ethical issues are likely to increase students' ethical reasoning. However, Milne (2001) cites studies (Miller and Miller 1976; Baxter and Rarick 1987) indicating that values are set early in life and it may be too late to change the values of college-age students.

Some studies measure whether accounting students' ethical reasoning increases over the course of one term of ethics instruction using a pre and post-test design. Other accounting student studies measure the relationship between simply taking an ethics course and DIT *P* scores. Thus, we conduct two analyses, one with pre and post-test studies and one with studies of those completing an ethics course (versus those not completing an ethics course) to investigate the association of ethics instruction with level of DIT *P* scores in our fourth research question.

RQ4 Does moral reasoning, as measured by the DIT *P* score, differ when students participate in ethics instruction?

Political Ideology

College student political identification with conservative or liberal beliefs frequently is studied in conjunction with moral reasoning (Maeda et al. 2009). Rest et al. (1999) provide an extensive discussion of the relationship between political identification and moral reasoning. Although a number of studies find liberals generally score higher on the DIT than do conservatives, the explanations for these results are controversial (Emler et al. 1983; Bernardi et al. 2004). Bailey et al. (2005) and Thoma et al. (1999) posit political ideology is a distinct construct from moral reasoning. However, Fisher and Sweeney (2002) hold that the variance in moral reasoning largely stems from the variance in political ideology. The strength and level of significance of the relationship of political ideology to moral reasoning is thus a matter of dispute. The findings in studies involving accounting students have also varied. Thus, through meta-analysis, we investigate this relationship, which is articulated in research question 5.

RQ5 Does political ideology relate to the level of moral reasoning, as measured by the DIT *P* score?

Gender

Prior research in business ethics has highlighted gender differences in moral reasoning abilities and ethical decision-making (McCabe et al. 2006). Tenbrunsel and Smith-Crowne (2008, p. 566), in reviewing over 30 years of research related to ethical decision-making, find conflicting results, note that some studies find females make more ethical judgments than do males, few studies find males make more ethical judgments than do females, and many studies find no significant differences in the ethical judgments of females and males. They suggest that gender differences are likely to be small and the result of either male/female socialization differences (Gilligan 1982) or differences in situational factors (Mischel 1968). Similarly, Kish-Gephart et al. (2010) found that women make slightly more ethical choices than men, and Borkowski and Ugras (1998), in their meta-analysis using business students, find a similar relationship. In accounting, Bernardi and Bean (2008) analyze multiple studies of accounting students and find that female accounting majors score 10.4 % higher on the DIT than their male counterparts. In contrast to these results, Rest (1986) proposes there is no theoretical basis for finding gender differences, while Gilligan (1977) argues that females and males simply reason differently when faced with ethical issues. Thus, we investigate the relationship between gender and moral reasoning with the following research question.

RQ6 Does gender relate to the level of moral reasoning, as measured by the DIT *P* score?

GPA

Public accounting firms use GPA as a key measure of an accounting major's potential success in future employment. In the psychology research domain, Spickelmier (1983) found measures of academic success to be positively related to DIT *P* scores suggesting academically successful students may internalize moral reasoning development more effectively than their fellow students. Johnson et al. (1993) investigate the relationship between business writing and moral judgment and find that DIT *P* scores are strongly associated with GPA. Rest (1979) similarly finds that academic achievement correlates highly with moral reasoning; *r* commonly ranges from 0.2 to 0.5. Accordingly, we include GPA in our meta-analysis to assess its relationship to the moral reasoning of accounting students.

RQ7 Do students with higher GPAs exhibit different levels of moral reasoning, as measured by the DIT *P* score, than do students with lower GPAs?

Age

There is intuitive appeal to the expectation that with age comes more advanced moral reasoning ability. Rest (1986) posits and Kohlberg (1984) concurs that maturation in

moral reasoning occurs as individuals accumulate greater life experiences. Thus, we should witness advancement in moral reasoning not within a semester or possibly even an academic year, but over a longer span of time. This theory has motivated age as a common variable of interest in studies of moral reasoning. Significant linkages have been noted in some studies (Kelley et al. 1990; Kohlberg 1969; Trevino 1992) but not others (Ford and Richardson 1994).

Of two meta-analyses examining ethical choice in the workplace and among business students (Kish-Gephart et al. 2010 and Borkowski and Ugras 1998, respectively), only Borkowski and Ugras (1998) found age to be significant. While Kish-Gephart et al. (2010) found a small, but significant negative relationship between age and unethical choice, when level of moral reasoning was controlled, age was no longer statistically significantly associated with unethical choice. As age is a prominent variable in ethics research involving accounting students, we include a research question regarding its impact on DIT *P* scores.

RQ8 Does moral reasoning as measured by the DIT *P* score differ for students of different ages?

University Class Level

One might assume that advancement through a university education provides students with a more nuanced moral lens leading to the ability for advanced moral reasoning. However, the findings are not conclusive. Maeda et al. (2009) compared the moral reasoning ability of students in a wide range of majors at different education levels in different studies and find a small, but significant relationship. Mean *P* scores increase with university education level. By contrast, Kish-Gephart et al. (2010) do not find a significant relationship between university class standing and ethical choice. Yet, Bernardi and Bean (2008) conduct a meta-analysis of 13 studies with students enrolled in accounting courses (but not necessarily accounting majors) and do find that the DIT scores increase as university class standing advances.[1]

Accordingly, we evaluate whether moral reasoning ability differs in studies employing accounting students at one university class level from those in studies employing students at a different university class level, investigating the ninth research question.

RQ9 Does moral reasoning, as measured by the DIT *P* score, differ for students at different university class levels?

Method

Conducting a meta-analysis requires the collection published and unpublished research that uses common variables to study the constructs of interest (Borenstein et al. 2009). Using the data provided in the studies listed in Table 1, we statistically aggregate the findings to assess the relative strength of the relationships between DIT *P* scores and variables such as ethical behavior, ethical instruction, and demographics. Here we provide a brief background on our data collection and meta-analysis methods.

Data Collection

Data for a meta-analysis is the output from prior research. Therefore our data collection method is to search for the population of research that uses the DIT and employs accounting students as subjects. Accounting students were defined as those taking accounting courses beyond the introductory level. We only include studies using accounting students in introductory courses (freshmen and sophomore level courses), when the students majoring in accounting DIT scores are separately reported from other enrolled students. We include in this search studies that use professional accountants as subjects not only for completeness, but also because some studies use both as subjects.

Bailey et al. (2010) serves as the starting point for our data collection. The authors review prior research utilizing the DIT and accounting students or accountants and present an extensive listing of prior research. The references in each study listed in Tables 1 and 2 of Bailey et al. (2010) are collected to determine if there are additional studies meeting our screening criteria of research utilizing the DIT and accounting students. Bailey et al.'s (2010) Table 1 focuses on studies that investigate accounting students, our subject of interest, whereas Table 2 focuses on studies that explored accountants. As noted above, we still review the references in Bailey et al.'s Table 2 as some studies investigate both populations and we wish to ensure

[1] Bernardi and Bean's (2008) paper includes 13 published studies with subjects that are students enrolled in accounting courses. We include 47 unpublished and published studies using the DIT instrument to investigate various variables in relation to accounting majors. In contrast to Bernardi and Bean (2008), our analyses only include students clearly identified as accounting majors in studies using students enrolled in accounting principles courses. If separate scores for accounting majors are not provided, we do not include the study in our analyses. Thus, our samples differ from those included in Bernardi and Bean (2008). In addition to investigating college level and gender, we analyze studies comparing the DIT *P* scores of accounting majors to accounting professionals, accounting majors to other college students pursuing other majors, and studies investigating the relationship of accounting majors' DIT *P* scores with variables such as ethics instruction, ethical choice, political ideology, grade point average and age.

completeness. We additionally screen each reference in the Bernardi and Bean (2008) study of accounting students' DIT scores for our criteria. Further, we collect literature reviews and meta-analysis studies in the ethics domain but outside accounting or conducted outside the United States (e.g., Kish-Gephart et al. 2010; McPhail 2006; Pierce 2006; O'Fallon and Butterfield 2005). The studies in the reference listings in each of these studies are screened based on our criteria.

In an effort to locate the highest number of studies meeting our necessary criteria, we search online databases using relevant search terms. General databases searched include EBSCOhost (which incorporates Academic Search Complete, Business Source Complete, and EconLit), ERIC, Social Science Citation Index, Sociological Abstracts, Emerald Library Journals, and PsychInfo. The American Accounting Association (AAA) Digital Library, which includes 15 AAA journals, yielded further relevant studies. We use a variety of keywords and phrases to capture use of the various forms of the DIT in studies of accounting student and accountants, as detailed in prior reviews (see, e.g., Bernardi and Bean 2008), including *defining issues test, DIT, DIT2, moral action, moral behavior, moral intention, moral judgment, moral reasoning, ethic*,* and *ethical decision-making* also adopting "and" *accounting students, account*, audit*, tax* and manage**. We also use keywords for unethical constructs including *unethical* and *cheating*. Because *Research on Professional Responsibility and Ethics in Accounting* (formerly *Research on Ethics in Accounting*) is not available in any online database we use, we manually searched all volumes of this journal (1995–2013) for studies meeting our criteria.

Finally, unpublished studies, including dissertations, working papers, and conference-presented papers provide relevant data for this analysis. The "file drawer problem," where studies with non-significant results are not published can bias the meta-analysis findings (Borkowski and Ugras 1998; Waples et al. 2009). Therefore, it is typical to include non-published studies such as dissertations and conference proceedings. We search for unpublished papers in ProQuest Dissertations & Theses, and the Accounting Research Network (SSRN). The 2008–2013 American Accounting Association conference proceedings are collected including concurrent sessions, research interaction forums, emerging and innovative research, and new scholars concurrent sessions where applicable. AAA conferences screened are Annual Meeting, ABO, ATA, Auditing, FIA, MA and the Annual Ethics Symposium.[2]

Our search process produces over 3,100 "hits." Approximately 2,400 of these papers are immediately eliminated as the various search terms picked up on individual words and are not associated with the DIT. We reduce the number an additional 525 by removing studies that did not employ accounting students as subjects. Eliminating duplicates and working papers that are eventually published decreases the number of studies to 86. Using this screened list of studies, we conduct a final search with Google Scholar and uncover no additional studies that we did not discover in our previous searches. Where authors do not include DIT version, standard deviations, *P* scores, or complete descriptions of the participants, we contact authors via email and phone for clarification. In some cases we are unable to obtain the needed statistical data to include particular studies in our meta-analyses. Our resulting data set includes 47 studies meeting the criteria that they collect DIT *P* scores, use accounting students as participants, and provide sufficient statistical data to be included in one or more analyses.[3] These studies are presented in Table 1 along with information detailing which variables are included in each study (e.g., ethical choice, ethics instruction, political ideology, etc.).

Meta-analysis

Meta-analysis provides a summary of empirical findings on a particular topic by statistically combining the quantitative results of multiple studies that investigate the same or similar factors and determining the strength of those relationships (Borenstein et al. 2009; Littell et al. 2008). Thus, meta-analysis provides a way of standardizing quantitative findings among a group of studies. As such, with meta-analysis, we control for between-study variation and gain a higher statistical power than with a single study (Card 2012).

A common measure used in meta-analysis is referred to as an effect size (Lipsey and Wilson 2001). In meta-analysis, a summary effect size, which represents the weighted mean of the individual studies' effect sizes, is computed (Borenstein et al. 2009). In addition, with meta-analysis, it is possible to make comparisons on the basis of categorical variables (e.g., level of education, which is the subject of research question 9) by calculating weighted means using techniques similar to ANOVA (Card 2012).

[2] Proceedings were reviewed in this window to capture new studies that may not yet have made it through the journal review process. Our various search procedures, as described, indicated accounting student DIT studies included in proceedings prior to this period had made it to publication.

[3] As shown in Appendix, of the 86 studies, 17 are eliminated due to the inability to determine the number of accounting majors (e.g., the study included other business majors) or the study employed MBAs, 11 are removed as the study used a modified DIT or an instrument similar to a DIT, seven did not include the necessary data and four use the same subjects or data as a study already included in our meta-analysis.

For research questions 1 through 8 (research question 9 is discussed subsequently), we use the statistical results from studies reporting DIT *P* scores whose subjects are primarily accounting majors and contain the same variables (e.g., ethics instruction or age). For each relationship we analyze, there are a minimum of three studies reporting independent effect sizes (Rosenthal 1987). More studies and larger samples within those studies increases the statistical power in meta-analysis (Borenstein et al. 2009)

We code each study's reported results in terms of student education level (e.g., graduate, senior, junior, or lower division), and publication status. We separately code education levels that are mixed such as juniors and seniors when statistical results are not provided on the basis of specific levels. In coding the variables related to our research questions, we adjust the coding as needed for consistency such that variables like liberal or conservative for political ideology and male or female for gender are coded in the same direction.

For each study, our analysis requires sample sizes, mean *P* scores, standard deviations and effect size. However, not all studies report an effect size, which can be calculated on the basis of correlations, mean differences or proportions (Borenstein et al. 2009); rather, some studies report t-statistics, F-statistics, χ^2 statistics or means for two groups. In those cases, we convert reported statistics into study correlations (r) using common conversions (cf. Borenstein 2009; Card 2012).[4] In turn, we use Comprehensive Meta Analysis© software to compute meta-analysis effect sizes from the individual studies' correlations. We use a random effects model.[5] Thus, in our study, and following Hunter and Schmidt (2004), the meta-analysis effect sizes represent estimated population correlations (rho, ρ) for DIT *P* scores and each variable of interest such as ethics instruction or age. Since sample size and variance impact effect sizes (correlations), we also present sample size and variance in our analysis. Additionally, we present the 95 % confidence interval for the meta-analysis effect sizes (correlations).

We categorize the reported meta-analysis effect sizes (correlations) in terms of Cohen's (1988) benchmarks—namely, effect sizes (correlations) of $r \approx 0.10$ are small, $r \approx 0.30$ are medium, and $r \approx 0.50$ are large. For the confidence intervals, we distinguish between non-significant and significant results. If the confidence interval includes zero, the relationship is statistically insignificant. As effect sizes (correlations) become smaller, the relationship between variables is less likely to be statistically significant.

Because our meta-analysis effect sizes are computed from a variety of studies, it is also important to assess whether the variance of effect sizes differ significantly from zero and the likelihood that there is in fact more than one effect size (correlation) in the population. To do so, we utilize the Q value statistic (Lipsey and Wilson 2001, p. 116) to provide an estimate of heterogeneity (Card 2012). In cases where the sample size is adequate and the Q value is statistically insignificant (does not exceed the χ^2 value given the degrees of freedom and chosen level of statistical significance), we conclude that heterogeneity is not substantial and our observed effect size accurately represents the one effect size in the population (Card 2012).[6] The interpretation of the Q value does not depend upon the statistical significance of the effect size.

[4] Among the conversions we used were the following (cf. Borenstein 2009; Card 2012):

Studies reporting t statistics: $r = \text{SQRT}(t^2/(t^2 + \text{df}))$.
Studies reporting F statistics: $r = \text{SQRT}(F/(F + \text{df}))$.
Studies reporting χ^2 statistics: $r = \text{SQRT}(\chi^2/N)$.
Studies reporting means for 2 groups: Step 1: $d = (M_1 - M_2)/\text{sd}_{\text{pooled}}$, where M_1 and M_2 are the means for groups 1 and 2, respectively and $\text{sd}_{\text{pooled}}$ is the pooled sample standard deviation. Step 2: $r = d/\text{SQRT}(d^2 + a)$, Note $a = (n_1 + n_2)^2/(n_1 n_2)$, where n_1 and n_2 are the sample sizes for groups 1 and 2, respectively).

[5] There are two models for computing meta-analysis effect sizes (correlations): the fixed-effects model and the random effects model. Under the fixed-effects model, the assumption is there is homogeneity among individual studies in the meta-analysis and that the studies share one "true" effect size or, in other words, there is one "true" population correlation. The fixed-effects model is appropriate where the studies share the same design and similar subjects. The fixed effect model should be used when all factors that could influence the effect size are included in all studies in the analysis (Borenstein et al. 2009).

Under the random-effects model, it is assumed that there is heterogeneity among individual studies and thus, the true effect may vary somewhat from study to study resulting in no one true rho (population correlation ρ). In some DIT studies included in our analyses, some subjects differ with respect to factors such as GPA and age, which could influence the effect size for different studies (Borenstein et al. 2009). In addition, some studies that measure the relationship between DIT *P* scores and a factor such as gender include other factors such as GPA or age and some do not include the factors GPA and age. Hence, we employ the random-effects meta-analysis model in all our analyses, which estimates the mean of a distribution of effects and allows for potential between-study differences (Erez et al. 1996; Borenstein et al. 2009). We report results based on the more conservative random effects model because of the variability in design and the subjects of the studies we analyze (Card 2012).

[6] Although a Q value that is statistically insignificant indicates that there is just one effect size in the population, it does not mean that a fixed-effects model should be used in an analysis of studies that may include different factors such as GPA or age in some studies, but not others as these different factors could influence the effect size (Borenstein et al. 2009). Further, random effects models allow generalization of conclusions to the population, whereas fixed effects models limit conclusions to the studies included in the analysis (Card 2012). The Q value is calculated in the same manner for both fixed-effects and random effects models and it measures the heterogeneity of effect sizes for the studies included in the analysis.

When the meta-analysis effect size is statistically significant and the sample size is adequate, a Q value that is statistically insignificant, suggests the differences observed in the individual studies included in the analysis is likely due to sampling fluctuation. If the Q value statistic is significant (exceeds the χ^2 value given the degrees of freedom and chosen level of statistical significance), it indicates the effect sizes (correlations) observed in the studies differ by more than that which is expected solely from subject-level sampling. Thus, if the meta-analysis effect size is significant and the Q value is significant, there may be other variables or moderators influencing the observed effect size and thus, there may be more than one population effect size (Card 2012).

Since the Q value indicates only the presence, but not the degree, of heterogeneity, it is important to assess the degree of heterogeneity when the Q value is significant. To do so, we report the I^2 index as a complement to the Q value when the Q value is statistically significant (Huedo-Medina et al. 2006). Higgins et al. (2003) suggest when $I^2 \approx 25$, there is a small amount of heterogeneity, when $I^2 \approx 50$ there is a medium amount of heterogeneity, and when $I^2 \approx 75$ or more there is a large amount of heterogeneity.

For research question 9, we compare mean P scores by level of university education (graduate through lower division). This is an across-studies comparison and allows us to calculate a group weighted mean effect size for studies reporting the mean P scores for seniors and to calculate a group weighted mean effect size for studies reporting the mean P scores for graduate students and determine whether the difference in means by education level is statistically significant (Konstantopoulos and Hedges 2009).[7] The comparison of these group weighted mean effect size scores is similar to ANOVA (Card 2012). We make additional comparisons for studies reporting mean P scores for other levels of university education.

Finally, we perform sensitivity analysis by comparing weighted mean P scores for published and unpublished studies. Borenstein et al. (2009) emphasize the importance of including unpublished studies in meta-analyses as unpublished studies often report smaller effect sizes; hence, effect sizes may be overstated, if unpublished studies are excluded. Thus, sensitivity analysis aids in explaining whether significant heterogeneity or between study differences in effect sizes are the result of characteristics of the sample, e.g., published or unpublished studies.

Results

Of the 47 studies we include in one or more analyses, four studies did not include the needed statistical information for calculating the overall mean P scores for all accounting majors.[8] We use 50 samples from 43 studies (i.e., several studies include more than one sample) and determine the mean P scores for all accounting students ($N = 4{,}499$) is 35.815 (variance = 0.612) with a range of 25.94–49.5. We then examine how P scores may relate to the specific variables of interest specified in our research questions.

Accounting Majors and Professionals

Analyzing RQ1, which asks whether the ethical reasoning levels of accounting students differ from individuals who are in the accounting profession, we examine seven studies (College Seniors $N = 370$; Practicing Accountants $N = 610$) that report DIT P scores for both senior accounting students and practicing accountants. Table 2 indicates that there is a small positive association between professional experience and DIT P scores (population estimate of weighted average correlation $\rho = 0.118$) and that heterogeneity is moderately significant (Q value = 11.909, $p = 0.064$, $I^2 = 49.6$).This marginally significant Q value and accompanying I^2 value suggests the correlation effect sizes are not all estimates of just one population value, but could represent multiple population values (Card 2012). These seven studies include accountants with differing levels of experience, which may influence the correlation effect size among the studies. Ponemon and Glazer (1990) measure only accountants with at least 5 years' experience compared to Massey and Thorne (2006) who measure the moral reasoning of accountants with an average of 1.3 years' experience. Ponemon (1992) notes that the ethical reasoning of managers and partners is lower than that of staff accountants. We eliminate the four samples with the lowest levels of accounting experience and retain those samples of accountants with 5 or more years of experience (managers from Lampe and Finn 1992; Ponemon and Glazer 1990; Armstrong 1984) and repeat the analysis. The effect size for the second analysis is slightly smaller ($\rho = 0.108$;

[7] Table 1 presents a total of 33 studies for RQ9 that incorporate data for at least one level of university education. Some of these studies report mean P scores for lower division students, some for juniors, some for seniors, and some for graduate students. Few of these studies report mean P scores for more than one level of university education. For example, we have only four studies that report mean P scores for both lower division and senior students. However, by conducting our analysis across studies, we are able to include seven studies that report lower division mean P scores and 24 studies that report senior level mean P scores, thus providing substantially more power than if we used only the four studies that report scores for both lower division and senior students.

[8] One study did not provide the mean DIT P score for accounting students and three studies did not provide the standard deviation. However, these four studies did provide the needed statistics for one or more of our other analyses.

Table 2 Meta-analytic weighted average correlation effect sizes of factors with DIT *P* scores (adjusted scores)

Variable	RQ	*k*	*N*	Random effects correlation			Heterogeneity	
				Est. (size)	Variance	95 % CI	*Q* value (*p* value)	I^2 (amount)
Practicing Accountants: College Seniors/Professional Accountants	1	7	370/610	0.118(S)	0.002	0.020 to 0.213*	11.909 (0.064*)	49.616 (M)
Practicing Accountants: College Seniors/Experienced Accountants	1	3	180/212	0.108(S)	0.005	−0.034 to 0.246	3.525 (0.172)	
Practicing Accountants: College Seniors/Staff level Accountants	1	4	190/398	0.0.011(S)	0.000	−0.094 to 0.116	2.377 (0.498)	
Major: Accounting/Non Accounting Business Major Senior	2	7	459/941	0.101(S)	0.001	0.043 to 0.159*	7.100 (0.312)	
Major: Accounting/Non Accounting Business Major Lower Division	2	3	102/218	0.158(S)	0.003	0.048 to 0.264*	0.883 (0.643)	
Major: Accounting/Non Business Major Senior	2	4	245/203	0.137(S)	0.002	0.045 to 0.228*	1.859 (0.602)	
Major: Accounting All Levels/Business Major All Levels	2	12	640/1,305	0.111(S)	0.001	0.066 to 0.155*	8.866 (0.634)	
Ethical Choice	3	9	602	0.210(S)	0.002	0.131 to 0.287*	3.996 (0.857)	
Pre and Post Ethical Instruction	4	8	271	0.156(S)	0.004	0.032 to 0.274*	2.769 (0.906)	
Ethics Course/No Ethics Course	4	10	453/462	0.001(S)	0.003	−0.103 to 0.106	20.423 (0.015*)	55.932 (M)
Political Ideology	5	6	461	−0.312(M)	0.004	−0.420 to −0.196*	8.782 (0.118)	
Gender: Female/Male	6	22	1,372/1,106	−0.161(S)	0.001	−0.222 to −0.098*	47.382 (0.001*)	55.679 (M)
GPA	7	7	1,078	0.154(S)	0.004	0.024 to 0.279*	23.683 (0.001*)	74.666 (L)
Age	8	14	1,342	0.066(S)	0.004	−0.058 to 0.187	59.860 (0.000*)	78.217 (L)

RQ Research Question Number

k Number of separate samples reported among studies (some studies, such as Fisher and Sweeney 1998, reported results based on two samples, e.g. Experiment 1 and Experiment 2)

N number of subjects in combined samples

Est. (size) estimated correlation for effect size; correlations of ≈0.10/0.30/0.50 are Small (S)/Medium (M)/Large (L) (see Cohen 1988)

Variance variance associated with estimated effect size correlation

C.I. the 95 % confidence interval (C.I.) for the random effects correlation estimate; the estimated effect size is statistically significant if the C.I. does not include zero and is not significant if it includes zero. We are 95 % confident that the true effect size is within the range of the confidence interval. * Statistically significant effect size

Q value (p value) a weighted measure of between-studies heterogeneity (similar to ANOVA's partition of variance between groups); the *p* value indicates the likelihood that the effect sizes in the studies differ by more than what is expected solely from subject-level sampling; * statistically significant

I^2 *(amount)* the proportion of total variation resulting from heterogeneity; I^2 values of ≈25/50/75 correspond to Small (S)/Medium (M)/Large (L) amounts of heterogeneity (see Huedo-Medina et al. 2006). The ability to detect different amounts of heterogeneity is affected by the number of studies included in the analysis. When the *Q* value is not statistically significant, we conclude that heterogeneity is not substantial and do not focus on the I^2 index (see Card 2012)

Q value = 3.525, $p = 0.172$, College Seniors $N = 180$; Practicing Accountants = 212) and statistically not significant. We then use the four samples with staff level accountants and senior accounting students in another analysis and find no statistically significant correlation effect ($\rho = 0.011$; Q value = 2.377, $p = 0.498$, College Seniors $N = 190$; Staff Accountants $N = 398$). The lack of significance in the ρ's and Q values for both subsamples when the combined sample is significant suggests a larger number of samples of more experienced and staff level accountants as well as samples of students is needed to determine the true relationship between accounting experience and DIT P scores.

Accounting versus Non-accounting Majors

RQ2 then explores students' major to assess whether accounting students differ in moral reasoning abilities from their counterparts in other business disciplines or compared to university students of any major. In seven studies comparing senior accounting students ($N = 459$) to senior business (non-accounting) students ($N = 941$) the weighted average correlation effect size is $\rho = 0.101$. The analysis of three studies that clearly identify lower division students majoring in accounting ($N = 102$) and lower division non-accounting business students ($N = 218$) results in a weighted average correlation effect size of $\rho = 0.158$. Comparing senior accounting majors ($N = 245$) to senior non-business majors ($N = 203$) in four studies, the correlation effect size is $\rho = 0.137$. Finally, when 12 samples are used to compare accounting majors at all levels ($N = 640$) to business (non-accounting) majors at all levels ($N = 1305$), the correlation effect size is $\rho = 0.111$. In sum, we find a small, positive, statistically significant effect size for each comparison indicating accounting students' P scores are slightly higher than students in other majors. We also find no significant between-studies heterogeneity in any of the analyses, thus suggesting that the small but statistically significant observed ρ measuring effect sizes stem from college major rather than some other factor.

Ethical Choice

To address RQ3, we test the correlation between DIT P scores and ethical choice in nine accounting student studies with a collective 602 subjects (N) where we can calculate the needed correlations. We find an estimated weighted average effect size (correlation) of 0.210, as shown in Table 2. The correlation falls in the small to medium range and is significant, indicating that accounting students with higher levels of moral reasoning are more likely to make ethical choices. The heterogeneity test, Q value, is 3.996 and not significant, which indicates that the proportion of observed variance is not due to the heterogeneity of treatment effects (Higgins et al. 2003). However, because of the different forms of ethical choice measured (e.g., lying, plagiarism, etc.), a larger sample of studies may be needed to conclude there is an absence of substantial heterogeneity (Card 2012). In sum, we find a statistically significant relationship between the DIT P score and ethical choice.

Ethics Instruction

We explore RQ4 in two ways, incorporating studies that examine ethics instruction embedded in an accounting course; and, separately, incorporating studies that examine a stand-alone ethics course. We analyze eight samples in seven studies ($N = 271$) that report DIT P scores before and after ethics instruction embedded in a course with accounting students.[9] The statistical results presented in Table 2 indicate a small, but statistically significant positive relationship between post ethics embedded instruction and P scores, with an estimated population effect size (correlation) of $\rho = 0.156$. Further analysis reveals no significant between-studies variance (heterogeneity) ($Q = 2.769$; $p = 0.906$) that could be detected in the eight samples from seven studies included in our analyses.

The meta-analysis of ten samples that reported the relationship between P scores and accounting students who had taken an ethics course ($N = 453$) versus those who did not take an ethics course ($N = 462$) indicated no statistically significant relationship ($\rho = 0.001$). In addition, the between-studies variance was statistically significant (Q value = 20.423, $p = 0.015$) with a medium amount of the total variance stemming from heterogeneity ($I^2 = 55.932$). The significant difference in the studies' heterogeneity indicates there is likely a range of true effect sizes related to taking an ethics course and DIT P scores, rather than just one true population correlation effect size and the observed variance is likely due to factors other than just whether one has taken a stand-alone ethics course. These findings imply that accounting student ethics instruction embedded in an accounting course generally has a small but significant positive effect on P scores, but simply taking an ethics course does not.

Political Ideology

With RQ5, we analyze five studies (including Sweeney and Fisher 1998, which includes two distinct samples in their experiments, thus $k = 6$ samples, $N = 461$[10]) that

[9] One of these studies (Ponemon 1993) provides scores for two samples, one with seniors and one with graduate students.

[10] In five samples, 154 students were identified as liberal and 255 as conservative. One additional sample (Klimek and Wenell 2011) provides the correlation between Liberal/Conservatism and DIT P scores, but not the breakdown of liberal and conservatives in their

investigate the relationship between political ideology (conservative or liberal) and students' DIT *P* scores.[11] As presented in Table 2, students who identified themselves as conservative (coded 1) had significantly lower DIT *P* scores than students who identified themselves as liberal (coded 0) resulting in an estimated population correlation (ρ) of −0.312, a medium effect size. The *Q* value is 8.782 and not statistically significant ($p = 0.118$) indicating the between-studies heterogeneity is not significant. Hence, political ideology in these studies is highly related to the DIT *P* scores of accounting students and the lack of significant heterogeneity suggests that this finding is not likely the result of other factors.

Gender

For RQ6 we analyze 22 samples (Females $N = 1372$; Males $N = 1106$) that report the results of comparing female and male accounting students' DIT *P* scores.[12] We code female = 0 and male = 1. The correlation effect size is small to medium ($\rho = -0.161$) and statistically significant indicating female students generally have higher *P* scores than males. However, the heterogeneity between studies is significant (*Q* value = 47.382, $p = 0.001$) with a medium amount of between-studies observed variance ($I^2 = 55.679$). Thus, the statistical tests associated with evaluating heterogeneity suggest the significant differences in observed female and male *P* scores in these studies may be due to factors other than gender alone.

GPA

For RQ7, we analyze the seven studies ($N = 1078$) identified in Table 1 that measure GPA in relationship to students' DIT *P* scores. The estimate of the population correlation is 0.154 and statistically significant (see Table 2). However, heterogeneity between these studies is large (*Q* value = 23.683, $p = 0.001$; $I^2 = 74.666$). Hence, although DIT *P* scores is significantly related to students' GPA, the large heterogeneity indicates there are other unexplored factors affecting this result.

Age

Fourteen studies include sufficient statistical data to analyze the relationship between student age and DIT *P* scores (our RQ8). The relationship between age and DIT *P* scores, while positive, is not significant. The point estimate of the population correlation effect size ($\rho = 0.066$) is quite small and the heterogeneity between studies is quite large (*Q* value = 59.86, $p = 0.000$; $I^2 = 78.217$) thus providing a negative answer to RQ8. Level of moral reasoning does not significantly increase as accounting students age. However, the large heterogeneity between studies suggests that the observed correlation effect sizes stem from some factor or factors other than age.

University Class Level

After examining the direct relationships of DIT *P* scores and accounting experience, ethical choice, ethics instruction, political ideology and demographics in terms of weighted average correlation effect size, we determine *P* score effect sizes based on different levels of education to determine if mean *P* scores differ by university level (RQ 9). These analyses determine the weighted average mean *P* scores for different university levels and whether the observed differences are statistically significant. Some studies report mean *P* scores for graduate students, some for seniors, some for juniors, and some for lower division students. Thus, unlike our analyses for RQ 1 to RQ 8 where each study in each analysis contained the same variables (e.g., ethics course/no ethics course, females/males), for RQ 9 different studies provide samples of students at different university levels. In these comparisons, unlike the tests for RQ1–RQ8, the *Q* values and I^2s that test for heterogeneity indicate whether there are significant differences in the population mean *P* scores for students at different university levels.

First, we analyze all studies that report either lower division accounting students' DIT *P* scores (7 studies, $N = 362$) or senior accounting students' DIT *P* scores (24 studies, $N = 1{,}558$) and determine the weighted average mean effect size for each level. The results appear in Table 3 and indicate that lower division students' mean *P* scores are significantly lower than senior students' *P* scores (32.944 vs. 37.127). Further analysis indicates that a substantial amount of heterogeneity between studies can be explained by whether the accounting student is a lower division or senior student (*Q* value 6.422 and $p = 0.011$; $I^2 = 86.189$)

We also analyze studies that include lower division students as well as those including juniors (6 studies, $N = 379$) and find, as reported in Table 3, no statistically significant difference in *P* scores (32.944 vs. 34.804). Additional analysis indicates that the observed between

Footnote 10 continued
sample of 52 students. These students' political ideology was evaluated on a 5 point scale ranging from 1 very liberal to 5 very conservative. A few studies, but not all, provide separate statistics for political moderates so moderates are excluded from our analysis.

[11] Bernardi et al. (2004) measure the political ideology and DIT *P* scores of business and other students enrolled in lower division accounting principles classes. However, separate scores for students majoring in accounting are not provided, so we do not include this study in our analysis.

[12] In some studies investigating ethical instruction pre and post-test *P* scores for females and males are provided. We include only the post-instruction scores in our analysis.

Table 3 Meta-analysis of class level and weighted average mean DIT *P* scores (adjusted scores)

Variable	RQ	*k*	*N*	Mean effects			Heterogeneity	
				Point Est. (size)	Standard error	CI	*Q* value (*p* value)	I^2 (amount)
Lower Division vs. Seniors								
Lower Division	9	7	362	32.944	1.350	30.229–35.59		
Seniors	9	24	1,558	37.127	0.946	35.273–38.98		
Total between overall	9	31	1,920	35.750	0.774	34.232–37.27	6.422* (0.011)	86.189 (L)
Ju Lower Division vs. Juniors								
Lower Division	9	7	362	32.944	1.350	30.299–35.59		
Juniors	9	6	379	34.804	1.315	32.227–37.38		
Total between overall	9	13	741	33.898	0.942	32.052–35.74	0.974 (0.324)	
Juniors vs. Seniors								
Juniors	9	6	379	34.804	1.315	32.227–37.38		
Seniors	9	24	1,558	37.127	0.946	35.273–38.98		
Total between overall	9	30	1,937	36.335	0.768	34.830–37.84	2.057 (0.151)	
Gr *Seniors vs. Graduate Students*								
Seniors	9	24	1,558	37.127	0.946	35.273–38.98		
Graduate Students	9	7	569	36.467	2.088	32.374–40.56		
Total between overall	9	31	2,127	37.014	0.861	35.326–38.70	0.083 (0.774)	
Lower Division vs. Graduate Students								
Lower Division	9	7	362	32.944	1.350	30.229–35.59		
Graduate Students	9	7	569	36.467	2.088	32.374–40.56		
Total between overall	9	14	931	33.982	1.134	31.760–36.20	2.008 (0.157)	
Entire Sample		50	4,499	35.815	0.612	34.615–37.02		

RQ Research Question Number

k Number of separate samples reported among studies (some studies, such as Fisher and Sweeney 1998, reported results based on two samples, e.g. Experiment 1 and Experiment 2)

N number of subjects in combined samples

Point Est. (size) uncorrected weighted average mean point estimate

Variance variance associated with estimated weighted average mean

C.I. the 95 % confidence interval (C.I.) for the mean effect size

Q value (p value) a weighted measure of between-studies heterogeneity (similar to ANOVA's partition of variance between groups); the *p* value indicates the likelihood that the effect sizes in the studies differ by more than what is expected solely from subject-level sampling. * indicates statistically significant difference

I^2 *(amount)* the proportion of total variation resulting from heterogeneity; I^2 values of ≈25/50/75 correspond to Small (S)/Medium (M)/Large (L) amounts of heterogeneity (see Huedo-Medina et al. 2006).The ability to detect different amounts of heterogeneity is affected by the number of studies included in the analysis. When the *Q* value is not statistically significant, we conclude that heterogeneity is not substantial and do not focus on the I^2 index (see Card 2012)

studies heterogeneity is not related to whether the student is lower division or a junior.

Our analysis of studies using juniors or seniors reveals no statistically significant differences (*Q* value = 0.974 and $p = 0.324$). The estimated population mean effect size for juniors' *P* scores is 34.804 compared to 37.127 for seniors (*Q* value = 2.057, $p = 0.151$).

Similarly, when we compare seniors to graduate students (7 samples, $N = 569$), the mean effect size of seniors' *P* scores is not significantly different from graduate students' *P* scores (37.127 vs. 36.467, *Q* value 0.083 and $p = 0.774$). We also compare lower division students to graduate students and find the mean effect size of lower division students' *P* scores is not significantly different from graduate students' *P* scores (32.944 vs. 36.467, *Q* value 2.008 and $p = 0.157$). Thus, accounting student moral reasoning as measured by DIT *P* scores consistently increases with level of education through the senior year, however, not all the observed increases are statistically significant. While the mean *P* score for accounting seniors is slightly higher than mean scores for accounting graduate students, the difference is not significant.[13] In addition, although accounting graduate students' have higher mean

[13] The trend in our results (using both published and unpublished studies) of accounting majors at higher levels of college scoring higher on moral reasoning as measured by the DIT is consistent with results obtained by Bernardi and Bean (2008) using published studies of students enrolled in accounting courses (both accounting and other business majors).

Table 4 Summary of results

Variable	RQ	Effects		Heterogeneity	
		Size	Significance	*p* value	Amount
Practicing Accountants: College Seniors/Professional Accountants	1	S	*	**0.064**	M
Practicing Accountants: College Seniors/Experienced Accountants	1	S	NS	0.172	NA
Major: Accounting/Non Accounting Business Major Senior	2	S	*	0.312	NA
Major: Accounting/Non Accounting Business Major Lower Division	2	S	*	0.643	NA
Major: Accounting/Non Business Major Senior	2	S	*	0.602	NA
Major: Accounting All Levels/Business Major All Levels	2	S	*	0.634	NA
Ethical Choice	3	S	*	0.857	NA
Pre and Post Ethical Instruction	4	S	*	0.906	NA
Ethics Course/No Ethics Course	4	S	NS	**0.015**	M
Political Ideology	5	M	*	0.118	NA
Gender: Female/Male	6	S	*	**0.001**	M
GPA	7	S	*	**0.001**	L
Age	8	S	NS	**0.000**	L
University Level: Lower Division/Senior	9	Seniors > Lower Division	*	**0.011**	L
University Level: Lower Division/Junior	9	Juniors > Lower Division	NS	0.324	NA
University Level: Junior/Senior	9	Seniors > Juniors	NS	0.151	NA
University Level: Senior/Grad Student	9	Seniors > Grad Student	NS	0.774	NA
University Level: Lower Division/Grad Student	9	Grad Student > Lower Division	NS	0.157	NA

RQ Research Question Number

Size For RQ 1–8, correlations of ≈0.10/0.30/0.50 are Small (S)/Medium (M)/Large (L) (see Cohen 1988); for RQ 9, the weighted average mean point estimates for the university levels are compared (e.g., the senior level average mean point estimate exceeds the lower division point estimate)

Significance significant = *; not significant = NS

p value the likelihood that the effect sizes in the studies differ by more than what is expected solely from subject-level sampling (*p* value <0.10 in bold)

Amount the proportion of total variation resulting from heterogeneity; I^2 values of ≈25/50/75 correspond to Small (S)/Medium (M)/Large (L) amounts of heterogeneity (see Huedo-Medina et al. 2006).The ability to detect different amounts of heterogeneity is affected by the number of studies included in the analysis. When the Q value is not statistically significant, we conclude that heterogeneity is not substantial and do not focus on the I^2 index (see Card 2012). These amounts are shown as NA (not applicable) on the table

scores than lower division accounting students, again, the difference is not significant.

Sensitivity Analyses

To determine if mean DIT *P* scores differ between published and unpublished studies we conduct an additional meta-analysis. In comparing 36 samples ($N = 3{,}632$) from published studies to 11 samples ($N = 601$) from unpublished studies we note a significant difference in mean *P* scores; 36.339 for published samples and 32.02 for unpublished samples. The between-studies heterogeneity is significant (Q value = 9.977, $p = 0.002$) and explains a large amount of the total observed variance ($I^2 = 87.285$). Thus, it appears there are differences in observed mean *P* scores between published and unpublished studies.

We provide a summary of our results in Table 4.

Summary and Discussion

Results Summary

Our findings demonstrate the context and relevant ranges where the DIT can aid the assessment of accounting students' moral reasoning abilities. Specifically, with respect to RQ1, accounting experience remains positively correlated with DIT *P* scores, but is not statistically significant when accountants with 5 or more years' experience are compared to college accounting seniors or staff level accounting professionals are compared to college accounting seniors. This finding adds to the conversation surrounding the controversial inverted-U effect (Bailey et al. 2010). While Ponemon (1992) builds the case for selection socialization and decreased ethical reasoning in the senior ranks of accounting firms, Scofield et al. (2004) use large samples of accounting professionals and find little support for this proposition.

RQ2 explored whether accounting students' DIT *P* scores differed from other business and non-business students. Our findings point to small differences; accounting students have slightly higher moral reasoning abilities than do non-accounting students.

While ethical choice has been examined using a variety of choice scenarios, collectively, the findings indicate that higher DIT *P* scores are associated with more ethical choices (RQ3). Ethical choice is one of the ultimate outcome goals derived from ethics education and therefore this finding reassures future researchers seeking to use the DIT as a metric to explain and predict ethical choice.

With some state boards of accountancy mandating a stand-alone accounting ethics course as prerequisite for CPA certification, RQ4 tests the assumptions related to form of ethics instruction. There are conflicting results in the literature. For example, Dellaportas (2006) and Armstrong (1993) find accounting students' ethical reasoning as measured by DIT P scores increases with ethics instruction, while Delaney (2005) and Ponemon (1993), similar to Milne (2001), do not. Our findings do indicate the success that ethics instruction can have at the university level—but only with instruction embedded within an accounting course. The evidence does not support the efficacy of a stand-alone ethics course when efficacy is measured by the DIT.

Five prior studies were included in our analysis of political ideology (RQ5). Collectively, we find that those students identifying themselves as politically conservative had significantly lower DIT *P* scores than those self-identifying as liberal. While this finding is moderately robust, we also recognize the need to interpret it cautiously given the methodological debate surrounding three of the five studies in our sample. Bernardi et al. (2004) suggest that the results reported in two Fisher and Sweeney (1998, 2002) studies and the Sweeney and Fisher (1998) study may have resulted from scaling issues in the measurement of political attitudes. Further, Bailey et al. (2005) report that in two large samples ($n = 719$ and $n = 253$) of accounting professionals, political attitudes accounted for less than 10 percent of the variance in DIT *P* scores.

As referenced in Bernardi and Bean (2008), initial expectations in moral reasoning research predicted trivial if any differences in the DIT score between males and females. However, Bernardi and Bean (2008) in their meta-analysis of students enrolled in accounting courses find that females score higher on the DIT than do males. With our larger sample of studies, we confirm in our analysis of RQ6 that gender is a significant variable explaining the difference in DIT scores with accounting students.

GPA (RQ7) is found to be positively associated with the DIT *P* score, indicating the advantage for collecting this variable to explain variation in findings in future research. However, Age (RQ8) is not significantly associated with moral reasoning ability and therefore is less useful as an explanatory variable. In contrast to age, we note that moral reasoning as measured by DIT *P* scores generally increases with education level (RQ9). However, although the increases from lower division students to seniors are statistically significant, the increases from lower division students to juniors and juniors to seniors are not statistically significant. The mean *P* score of graduate accounting students is slightly, but not significantly lower than that of senior accounting students and the mean *P* score of graduate accounting students is not significantly higher than that of lower division accounting students.

We also find that published studies do report significantly higher mean DIT *P* scores than do unpublished studies. Thus, as Borenstein et al. (2009) suggest, including unpublished studies in meta-analysis results in smaller weighted average correlations and mean effect sizes than otherwise would be calculated.[14]

Discussion, Implications, and Future Research

As Bernardi and Bean (2008) point out, the use of statistical meta-analysis methods provides a rigorous alternative to descriptive literature reviews (Glass 1976). Findings can also highlight conflicting results, thus providing guidance for the design of future research.

Accounting student studies were selected for our investigation because the curriculum for accounting students across universities has many commonalities and some states are requiring ethics education. Bernardi and Bean (2008) support using accounting students as subjects based on Rest's (1986) assertions that moral reasoning levels stabilize after formal education ends. Our findings can be insightful for future research with other student

[14] Rest (1990) reports that the *P* score of the 3 Story version of the DIT has a correlation of 0.93 with the *P* score of the 6 Story DIT. Rest et al. (1999, p. 651) also indicate that the DIT 6 Story and DIT2 are correlated with each other at a level similar to that shown in prior studies of test–retest correlations of the DIT 6 Story (Rest 1979, p. 239). Hence, we combine studies using all three versions of the DIT. However, Bernardi and Bean (2008) indicate that the scales for *P* scores differ between the 3 Story and 6 Story versions of the DIT with a maximum *P* score of 90 on the 3 Story and 95 on the 6 Story. They suggest similar scaling issues may arise with respect to the DIT2. To determine if our results might be influenced by DIT version we conduct two tests in which we converted DIT 3 Story *P* scores to the DIT 6 Story scale and excluded studies using the DIT2. We then repeated our analyses for pre and post ethics instruction and our comparison of lower division students with seniors and found our results did not significantly change from our results using unadjusted *P* scores in analyses combining all three versions of the DIT. Thus, consistent with the strong correlations among *P* scores in all DIT versions reported by Rest (1990) and Rest et al. (1999) our results are not significantly affected by the scaling issues noted by Bernardi and Bean (2008).

populations and have consistencies with prior findings where business and accounting students were studied.

Our meta-analysis explores nine variables commonly used in research with accounting students to explain differences in DIT scores. To our knowledge, ours is the most comprehensive study of its kind, in terms of the number of constructs examined and the use of highly sensitive statistical indicators. Our results provide insight into the mixed findings from accounting student studies employing the DIT that are not evident from each individual study alone. The meta-analyses conducted here provide statistical power to identify meaningful implications for future research. Not only have we identified whether commonly employed variables are statistically associated with DIT *P* scores when examined across multiple studies, we have measured the association of each variable in relation to accounting students' DIT *P* scores.

Factors Associated with the DIT

The results of our analyses indicate the common demographic variables where the DIT is relevant. While changing demographic characteristics for study participants is not possible, our findings do point to the need to use these variables as control or moderating variables. Even those with small but significant effect sizes should be included in future research designs to improve the predictive value of the results.

Our findings motivate future research in several complementary areas. First, the inverted-U effect, where senior accounting professionals exhibit decreased levels of moral reasoning compared to staff accountants and senior accounting students can be explored further. Whereas our findings suggest there may be some decrease in moral reasoning among senior level accountants, our study did not include studies without accounting students. Thus, large scale studies focusing exclusively on professionals, such as Scofield et al. (2004) were not analyzed, which would likely affect our findings. Further, there might be a confounding variable, such as gender (Bay and Greenberg 2001) or an environmental factor that influences this relationship. Likewise, the lessons from this research can be instructive for all business domains.

Second, accounting majors are shown to have higher DIT *P* scores than do non-accounting majors. What are the factors that lead to this effect? Could there be a self-selection bias where those with higher moral reasoning ability choose to study accounting? Or could there be a socialization effect, where the curricular emphasis on professional standards and values inculcates a positive moral reasoning perspective, leading to higher DIT scores? Identifying other majors where professional values and standards are emphasized in the curriculum (e.g., nursing) can serve as a comparison with accounting students to assess the reasons why they score higher than non-accounting students on the DIT.

Emphasizing the complicated interplay among factors affecting moral reasoning, we evaluated and found heterogeneity among weighted average correlation effect sizes of *P* scores and factors such as accounting experience, gender, and GPA. We provide statistically confirming evidence that in some cases multiple factors do provide greater insights into moral reasoning ability. Incorrect conclusions may be reached, if only the correlation between gender and moral reasoning is considered without measuring the heterogeneity, which suggests other factors are needed to explain this observed relationship. Our results reinforce those in the Herington and Weaven (2008)—namely, consistency in DIT moral reasoning scores improves when cluster analysis is used to examine the factors that influence the scores of different groups of people such as males and females.

Borkowski and Ugras (1998) conducted a meta-analysis with studies of American business students to explore the relationships between demographic characteristics and ethical behavior/attitudes. They found that gender was a significant variable with females demonstrating more ethical development than males. Bernardi and Bean's (2008) meta-analysis of published studies of students enrolled in accounting courses found that females scored significantly higher than males. Using both published and unpublished studies of accounting majors, we also find a small but significant effect for gender with females having higher DIT *P* scores than males. In addition, our results suggest that gender may be interacting with other factors to produce the observed results.

In contrast to our findings, where we only included studies that used the DIT, Borkowski and Ugras (1998) found that age was also a discriminating factor in ethical development. Older students were more ethical than were younger students. Herington and Weaven (2008) found moral reasoning increased with age among business students, but only when they eliminated the small number of students age 50 and above in their sample as those individuals had lower levels of moral reasoning than the other students. Kohlberg's (1984) theory posits that maturity positively impacts moral behavior. While we did not find age to be significant, we did find that between lower division students and seniors there was a significant difference in the DIT scores, but not between lower division and graduate students. This leads to a question for future research regarding the impact of education on moral reasoning which could be explored with traditional and non-traditionally aged students to assess whether age confounds the education effect.

Ethical Choice

Although the DIT is a convenient and reliable measure, frequently used in research studies of ethical choice, our findings reveal that the DIT does not capture all dimensions of ethical decision making processes. Rather, its applicability must be carefully tailored to the research question and method. The DIT *P* score of respondents was used as a predictor or explanatory variable in 25 of the 47 studies in our sample. The remaining 22 studies used DIT *P* scores as a metric to compare moral reasoning ability between different demographic groups. Where the DIT is employed to predict or explain a decision outcome, process or instruction method, it is necessary to match the task to the appropriate moral reasoning stage within the Rest paradigm. The DIT measures moral reasoning ability, which is most closely associated with the second component in the Rest 4-component model. If a particular task is designed to explore moral sensitivity (stage 1), moral motivation (stage 3) or moral character (stage 4), the DIT may not be the appropriate predictor of task behavior. Some studies in our sample used the DIT where the task was identification (stage 1), rather than reasoning (stage 2). For instance, Christensen et al. (2010) found no relationship between the DIT and their decision task. However, their task was one of moral decision identification, rather than a moral reasoning task. The majority of the students did not recognize the task as one that presented an ethical choice, therefore, they had not initiated moral reasoning. As such, the DIT provided no explanatory evidence to support the students' behavior.

Chan and Leung (2006) underscore the need to use the DIT for tasks with a moral reasoning element. Chan and Leung (2006) measured ethical sensitivity, the first stage in Rest's model and assessed whether it was correlated with DIT scores. They found no relationship between a student's ethical sensitivity and their score on the DIT. They point out that while some students may have high moral reasoning ability, if they are unable to recognize the event as one with a moral choice, they may not behave ethically and vice versa. Their findings reinforce Rest's assertions that morality is a complex multi-dimensional construct with separable components. Where the DIT is used, the research questions and method need to clearly be focused on the moral reasoning component. Future research exploring DIT effectiveness might examine its efficacy within a moral reasoning task as compared to tasks at higher stages such as moral motivation and character. Clearly linking the task to the Rest paradigm would be one approach that could further delineate the boundaries where the DIT is effective. Bailey et al. (2010) categorize future research opportunities in the context of Rest's four stages, providing a starting point to implement the implications of our findings.

Ethics Instruction

Our findings indicate certain types of ethics instruction are more effective than others. For example, we find students who complete a stand-alone ethics course do not exhibit higher levels of moral reasoning than students who do not complete an ethics course. In contrast and similar to Waples et al. (2009) we found that ethics instruction embedded in a course is more effective at building moral reasoning skills. Waples et al. (2009) looked at ethics instruction in the university and professional domains. They also found ethics instruction to be more effective with older students. While our studies showed no age effect, our participant populations are university students, a group that are typically within a narrow age band.

We find DIT *P* scores can provide some explanatory information where ethics instruction is embedded into an accounting course. Although statistically significant the weighted average correlation and mean effect sizes are relatively small. However, each variable alone is rarely a sufficient indicator for studies seeking to measure how interventions change students' moral reasoning abilities. This implies that DIT *P* scores may be influenced by a complex set of factors making it difficult to capture meaningful changes in moral reasoning abilities when ethics interventions are undertaken. As one factor, Massey and Van Hise (2009) suggest different pedagogies, such as context-specific active learning techniques, provide more impact in accounting ethics instruction. Using both quantitative and qualitative assessments, they provide support for the positive impact of student-led sessions, prompting questions specific to the profession, interviewing, reflective assignments and case analyses.

In addition to the demographic variables commonly captured in studies of the effectiveness of ethics instruction, it may be important to capture additional variables. In a study of college students' moral reasoning, Maeda et al. (2009) found that DIT scores of nonnative English speakers were significantly lower than those of native English speakers. While the difference was not large, it was significant. Thus, for example, the cumulative effects of individual student characteristics may disguise the effectiveness or lack of effectiveness of education efforts to increase students' moral reasoning abilities.

Published and Unpublished Studies

While many literature reviews and some meta-analyses include only published studies (Brierley and Cowton 2000; Day et al. 2002), omitting unpublished studies results in upwardly biased results. Numerous studies of publications in the social sciences and medicine document a bias towards publishing only studies with statistically significant results

(Pigott 2009; Begg 1994; Lipsey and Wilson 1993; Rosenthal 1979). When only published studies are included in meta-analysis there is a high likelihood of overstating results. We follow the approach of Kish-Gephart et al. (2010) and Borkowski and Ugras (1998) and include both published and unpublished studies in our analyses. We then conduct sensitivity tests and do find that the *P* scores are statistically significantly higher in published than unpublished studies. Thus, meta-analysis researchers concerned with accurately measuring effect sizes should consider including unpublished works in their studies.

Possible Limitations

The results reported in this study have potential limitations. Every effort was made to collect the population of studies using the DIT *P* score and accounting majors as subjects. However, our search may not have fully revealed all potential data or there may be unpublished studies of which we are not aware. We also excluded some studies that included students enrolled in accounting principles classes, where accounting majors were not specifically identified. Although there are likely some non-accounting majors in upper division accounting courses the number is usually small relative to accounting principles courses where the number is usually large. We also include only studies available in the English language. In addition, conducting meta-analysis requires specific statistical data for calculating effect sizes. Not all studies reported this information and some attempts to contact authors for clarification were unsuccessful. Therefore, of the eligible studies for our meta-analysis, some could not be included. Maeda et al. (2009) also note DIT *P* scores of students at state and liberal arts universities differ from those at certain categories of religious universities. We did not control for educational context in our analyses. It is unknown how these limitations may affect our findings.

In summary, our findings can provide guidance for future research to develop more efficient and effective research designs. The DIT is a convenient and reliable measure that when used appropriately can provide insight into the process of moral reasoning.

Acknowledgments The authors would like to express their appreciation for helpful comments from reviewers and discussants at the 2012 Western American Accounting Association Annual Meeting and the 2012 American Accounting Association Annual Meeting. The guidance provided by Professor Ioannis Kareklas has also been insightful.

Appendix: Studies Excluded

Unable to determine breakdown of undergraduate accounting majors (includes introductory accounting courses incorporating all business majors)

Abdolmohammadi, M., and J. Sultan. (2002). Ethical reasoning and the use of insider information in stock trading. Journal of Business Ethics, 37(2), 165–173.

Abdolmohammadi, M. J., D. L. Gabhart, and Reeves, M. (1997). Ethical cognition of business students individually and in groups. Journal of Business Ethics, 16(16), 1717–1725.

Bailey, C. D. (2011). Does the DIT measure ethical judgment ability or political position? Journal of Social Psychology, 151(3), 314–330.

Bay, D. D., and R. R. Greenberg (2001). The relationship of the DIT and behavior: A replication. Issues in Accounting Education, 16(3), 367–380.

Bernardi, R. A., D. F. Bean, and D. W. Massey. (2004). The influence of political ideology on DIT scores: Fact or artifact. Research on Professional Responsibility and Ethics in Accounting, 9, 21–48.

Bernardi, R. A., R. L. Metzger, R. Bruno, M. Hoogkamp, L. E. Reyes, and G. H. Barnaby. (2004). Examining the decision process of students' cheating behavior: An empirical study. Journal of Business Ethics, 50(4), 397–414.

Chan, C. W., C. S.Troutman, and D. O'Bryan. (2000). An expanded model of taxpayer compliance: Empirical evidence from the United States and Hong Kong. Journal of International Accounting, Auditing & Taxation, 9(2), 83–103.

Falk, H., B. Lynn, M. Mestelman, and M. Shehata. (1999). Auditor independence, self-interested behavior and ethics: Some experimental evidence. Journal of Accounting and Public Policy 18(Winter), 395–429.

Fisher, D. G. (1997). Assessing taxpayer moral reasoning: The development of an objective measure. Research on Accounting Ethics, 3, 141–171.

Shawver, T., and J. Sennetti. (2009). Measuring ethical sensitivity and evaluation. Journal of Business Ethics, 88(4), 663–678.

Tolleson, T., B. D. Merino, and A. G. Mayper. (1996). Applying behavioural models as prescriptions for ethics in accounting practice and education. Research on Accounting Ethics, 2, 21–49.

Trivedi, V., M. Shehata, and B. Lynn. (2003). Impact of personal and situational factors on taxpayer compliance: An experimental analysis. Journal of Business Ethics, 47(3), 175–197.

West, T., S. Pickard Ravenscroft, and C. B. Shrader. (2004). Cheating and moral judgment in the college classroom: A natural experiment. Journal of Business Ethics, 54(2), 173–183.

Studies which specifically used MBAs

Maroney, J. J., and R. E. McDevitt. (2008). The effects of moral reasoning on financial reporting decisions in a

post Sarbanes–Oxley environment. Behavioral Research in Accounting, 20(2), 89–110.
Kaplan, S. E., J. C. McElroy, S. P. Ravenscroft, and C. B. Shrader. (2007). Moral judgment and causal attributions: Consequences of engaging in earnings management. Journal of Business Ethics, 74, 149–164.
Kaplan, S. E., K. J. Newberry, and P. M. J. Reckers. (1997). The effect of moral reasoning and educational communications on tax evasion intentions. The Journal of the American Taxation Association, 19(2), 38–54.
Shapiro, B., J. Schatzberg, G. Sevcik, L. Thorne, and S. Wallace. 2005. A re-examination of behavior in experimental audit markets: The effects of moral reasoning and economic incentives on auditor reporting and fees. Contemporary Accounting Research 22(1), 229–264.

Studies which used an instrument similar to DIT or modified the DIT

Chan, S. Y. S., and P. Leung. 2006. The effects of accounting students' ethical reasoning and personal factors on their ethical sensitivity. Managerial Auditing Journal 21(4), 436–457.
Doyle, E., J. Frecknall-Hughes, and B. Summers. (2009). Research methods in taxation ethics: Developing the DIT for a tax-specific scenario. Journal of Business Ethics, 88(1), 35–52.
Fleming, D. M., R. N. Romanus, and S. M. Lightner. (2009). The effect of professional context on accounting students' moral reasoning. Issues in Accounting Education, 24(1), 13–30.
Green, S., and J. Weber. (1997). Influencing ethical development: Exposing students to the AICPA code of conduct. Journal of Business Ethics, 16(8), 777–790.
Hobson, J. L., M. J. Mellon, and D. E. Stevens. (2011). Determinants of moral judgments regarding budgetary slack: An experimental examination of pay scheme and personal values. Behavioral Research in Accounting, 23(1), 87–107.
Karacaer, S., R. Gohar, M. Aygün, and C. Sayin. (2009). Effects of personal values on auditor's ethical decisions: A comparison of Pakistani and Turkish professional auditors. Journal of Business Ethics, 88(1), 53–64.
Leitsch, D. (2006). Using dimensions of moral intensity to predict ethical decision-making in accounting. Accounting Education: an international journal, 15(2), 135–149.
Liyanarachchi, G., and C. Newdick. (2009). The impact of moral reasoning and retaliation on whistle-blowing: New Zealand evidence. Journal of Business Ethics, 89(1), 37–57.
Mintchik, N. M., and T. A. Farmer. (2009). Associations between epistemological beliefs and moral reasoning: Evidence from accounting. Journal of Business Ethics, 84(2), 259–275.
Stanga, K. G., and R. A. Turpen. (1991). Ethical judgments on selected accounting issues: An empirical study. Journal of Business Ethics, 10(10), 739–747.
Welton, R. E., R. Lagrone, and J. R. Davis. (1994). Promoting the moral development of accounting graduate students: an instructional design and assessment. Accounting Education, 3(1), 35–50.

Statistical measures needed for meta-analysis not reported

Cooper, B. J., P. Leung, S. Dellaportas, B. Jackling, and G. Wong. (2008). Ethics education for accounting students: A toolkit approach. Accounting Education, 17(4), 405–430.
Hiltebeitel, K. M., and S. K. Jones. (1991). Initial evidence on the impact of integrating ethics into accounting education. Issues in Accounting Education, 6(2), 262–275.
Kite, D., and R. R. Radtke. (1997). The effect of moral reasoning levels and political ideology on environmental accounting education. Research on Accounting Ethics, 3, 173–189.
Koeplin, J. P. (1998). A comparison of cognitive moral development of accounting students at a Catholic University with secular university accounting students. Unpublished doctoral dissertation, University of North Texas, Denton, TX.
Lampe, J. C. (1996). The impact of ethics education in accounting curricula. Research on Accounting Ethics, 2, 187–220.
Pope, K. R. (2005). Measuring the ethical propensities of accounting students: Mach IV versus DIT. Journal of Academic Ethics, 3(2–4), 89–111.
Schatzberg, J. W., G. R. Sevcik, and B. P. Shapiro. (1996). Exploratory experimental evidence on independence impairment conditions: Aggregate and individual results. Behavioral Research in Accounting, 8, 173–195.

Used same subjects or data as other study in sample (other study shown in [])

Abdolmohammadi, M. J., and C. Baker. (2007). The relationship between moral reasoning and plagiarism in accounting courses: A replication study. Issues in Accounting Education, 22(1), 45–55. [Abdolmohammadi, M. J., and C. Baker. (2008). Moral reasoning and questionable behavior. CPA Journal 78 (11), 58–61]
Nkenke, G. R. (2010).The impact of moral reasoning on ethical perception, intention, and orientation of upper level accounting students. Unpublished doctoral dissertation, Minneapolis, MN. [Bancroft, P. C. (2002). An investigation of moral reasoning as a predictor of ethical

awareness, ethical intention, and ethical orientation, Unpublished doctoral dissertation, Nova Southeastern, Ft. Lauderdale, FL.].

Porco, B. M. (2003). Factors affecting the cognitive moral development of undergraduate accounting students: Ethics education, internships, volunteerism, and Beta Alpha Psi. Unpublished doctoral dissertation, Fordham University, New York, NY. [Brown-Liburd, H. L., and B. M. Porco. (2011). It's what's outside that counts: Do extracurricular experiences affect the cognitive moral development of undergraduate accounting students? Issues in Accounting Education 26 (2), 439–54].

Sweeney, J. T., and D. G. Fisher. (1999). Politics, faking, and self-presentation: How valid is the *P* score of the DIT. Research on Accounting Ethics, 5, 51–75. [Fisher, D. G., and J. T. Sweeney. (1998). The relationship between political attitudes and moral judgment: Examining the validity of the DIT. Journal of Business Ethics 17(8), 905–916].

References

Note: References marked with an asterisk (*) indicate studies included in the meta-analysis (see Table 1)

AACSB International. (2009). *Eligibility procedures and accreditation standards for accounting accreditation*. Accessed December 15, 2011, at http://www.aacsb.edu/accreditation/accounting/standards/aacsb_accounting_standards.pdf.

*Abdolmohammadi, M. J. (2005). Ethical training in graduate accounting courses: Effects of intervention and gender on students' ethical reasoning. *Research on Professional Responsibility and Ethics in Accounting, 10*, 37–62.

*Abdolmohammadi, M. J., & Baker, C. (2008). Moral reasoning and questionable behavior. *CPA Journal, 78*(11), 58–61.

*Abdolmohammadi, M. J., Fedorowicz, J., & Davis, O. (2009). Accountants' cognitive styles and ethical reasoning: A comparison across 15 years. *Journal of Accounting Education, 27* (4), 185–196.

American Institute of Certified Public Accountants and American Accounting Association. (2012, July). The Pathways Commission on Higher Education: Charting a national course for the next generation of accountants. Accessed February 18, 2013, at http://commons.aaahq.org/files/0b14318188/Pathways_Commission_Final_Report_Complete.pdf.

*Armstrong, M. B. (1993). Ethics and professionalism in accounting education: A sample course. *Journal of Accounting Education, 11*, 77–92.

*Armstrong, M. B. (1984). Internalization of the professional ethic by Certified Public Accountants: A multidimensional scaling approach. Unpublished doctoral dissertation, University of Southern California, Los Angeles.

Bailey, C. D., Phillips, T. J., & Scofield, S. B. (2005). Does "political bias" in the DIT or DIT-2 threaten validity in studies of CPAs? *Behavioral Research in Accounting, 17*, 23–42.

Bailey, C. D., Scott, I., & Thoma, S. J. (2010). Revitalizing accounting ethics research in the neo-Kohlbergian framework: Putting the DIT into perspective. *Behavioral Research in Accounting, 22*(2), 1–26.

*Bancroft, P. C. (2002). An investigation of moral reasoning as a predictor of ethical awareness, ethical intention, and ethical orientation. Unpublished doctoral dissertation, Nova Southeastern, Ft. Lauderdale, FL.

Baxter, G. D., & Rarick, C. A. (1987). Education and the moral development of managers: Kohlberg's stages of moral development and integrative education. *Journal of Business Ethics, 6*(3), 243–248.

Bay, D. (2002). A critical evaluation of the use of the DIT in accounting ethics research. *Critical Perspectives on Accounting, 13*(2), 159–177.

Bay, D., & Greenberg, R. (2001). The relation of the DIT and behavior: A replication. *Issues in Accounting Education, 16*(3), 367–380.

Begg, C. B. (1994). Publication bias. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 399–409). New York: Russell Sage Foundation.

*Bernardi, R. A. (1995). Ethics research in accounting: A note on using the Defining Issues Test. *Accounting Educators' Journal, 7*(2), 1–16.

Bernardi, R. A., & Arnold, D. F. (2004). Testing the "inverted-U" phenomenon in moral development on recently promoted senior managers and partners. *Contemporary Accounting Research, 21*(2), 353–367.

Bernardi, R. A., & Bean, D. F. (2008). Establishing a standardized sample for accounting students' DIT scores: A meta-analysis. *Research on Professional Responsibility and Ethics in Accounting, 12*, 1–21.

Bernardi, R. A., Bean, D. F., & Massey, D. W. (2004). The influence of political ideology on DIT scores: Fact or artifact? *Research on Professional Responsibility and Ethics in Accounting, 9*, 21–48.

*Bernardi, R. A., Downey, A., Massey, D. W., & Thorne, L. (2002). Critical thinking and the moral reasoning of intermediate accounting students. *Research on Accounting Ethics, 8*, 73–102.

Bloodgood, J., Turnley, W., & Mudrack, P. (2010). Ethics instruction and the perceived acceptability of cheating. *Journal of Business Ethics, 95*(1), 23–57.

*Bonawitz, M. F. (2002). Analysis and comparison of the moral development of students required to graduate with an ethics course. Unpublished doctoral dissertation, Florida International University, Miami.

Borenstein, M. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 231–236). New York: Russell Sage Foundation.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex: Wiley.

Borkowski, S. C., & Ugras, Y. J. (1998). Business students and ethics: A meta-analysis. *Journal of Business Ethics, 17*(11), 1117–1127.

Brierley, J. A., & Cowton, C. J. (2000). Putting meta-analysis to work: Accountants' organizational–professional conflict. *Journal of Business Ethics, 24*(4), 343–353.

*Brown-Liburd, H. L., & Porco, B. M. (2011). It's what's outside that counts: Do extracurricular experiences affect the cognitive moral development of undergraduate accounting students? *Issues in Accounting Education, 26*(2), 439–454.

*Buell, K. (2009). The relationship of ethics education to the moral development of accounting students. Unpublished doctoral dissertation, Nova Southeastern University, Ft. Lauderdale, FL.

*Burks, B. D. (2006). The impact of ethics education and religiosity on the cognitive moral development of senior accounting and business students in higher education. Unpublished doctoral dissertation, Nova Southeastern University, Ft. Lauderdale, FL.

Card, N. A. (2012). *Applied meta-analysis for social science research*. New York: The Guilford Press.

Chan, S., & Leung, P. (2006). The effects of accounting students' ethical reasoning and personal factors on their ethical sensitivity. *Managerial Auditing Journal, 21*(4), 436–457.

*Christensen, A. L., Cote, J., & Latham, C. K. (2010). Unauthorized electronic access: Students' ethics, attitudes and actions. *Advances in Accounting Education, 11*, 247–262.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

*Cote, J., Latham, C. K., & Sanders, D. (2013). Ethical financial reporting choice: The influence of individual characteristics. *Advances in Accounting Behavioral Research, 16*, 115–148.

*Cruz, A. (2003). The awareness of undergraduate and graduate accounting students to ethical issues in internal audit. Unpublished doctoral dissertation, Nova Southeastern University, Ft. Lauderdale, FL.

Day, D. V., Unckless, A. L., Schleicher, D. J., & Hiller, N. J. (2002). Self-monitoring personality at work: A meta-analytic investigation of construct validity. *Journal of Applied Psychology, 87*(2), 390–401.

*Delaney, J. S. (2005). The impact of ethics education on the moral reasoning ability of accounting students. Unpublished doctoral dissertation, St. Ambrose University, Davenport, IA.

*Dellaportas, S. (2006). Making a difference with a discrete course on accounting ethics. *Journal of Business Ethics, 65*(4), 391–404.

*Dellaportas, S., Cooper, B. J., & Leung, P. (2008). Measuring moral judgment and the implications of cooperative education and rule-based learning. *Accounting & Finance, 46*(1), 53–70.

*Douglas, P. D., & Schwartz, B. N. (1998). A longitudinal study of the effect of teaching ethics throughout the accounting curriculum. *Research on Accounting Ethics, 4*, 89–112.

*Earley, C. E., & Kelly, P. T. (2004). A note on ethics educational interventions in an undergraduate auditing course: Is there an "Enron effect"? *Issues in Accounting Education, 19*(1), 53–71.

Emler, N., Renwick, S., & Malone, B. (1983). The relationship between moral reasoning and political orientation. *Journal of Personality and Social Psychology, 45*(5), 1073–1080.

Erez, A., Bloom, M. C., & Wells, M. T. (1996). Using random rather than fixed effects models in meta-analysis: Implications for situational specificity and validity generalization. *Personnel Psychology, 49*(2), 275–306.

*Eynon, G., Hill, N. T., Stevens, K. T., & Clarke, P. (1996). An international comparison of ethical reasoning abilities: Accounting students from Ireland and the United States. *Journal of Accounting Education, 14*(4), 477–492.

*Fisher, D. G., & Ott, R. L. (1996). A study of the relationship between accounting students' moral reasoning and cognitive styles. *Research on Accounting Ethics, 2*, 20.

*Fisher, D. G., & Sweeney, J. T. (2002). Morality vs. ideology: Implications for accounting ethics research. *Advances in Accounting Behavioral Research, 5*, 141–160.

*Fisher, D. G., & Sweeney, J. T. (1998). The relationship between political attitudes and moral judgment: Examining the validity of the Defining Issues Test. *Journal of Business Ethics, 17*(8), 905–916.

Ford, R., & Richardson, W. (1994). Ethical decision making: A review of the empirical literature. *Journal of Business Ethics, 13*(3), 205–221.

Gilligan, C. (1977). In a different voice: Women's conception of the self and morality. *Harvard Educational Review, 49*, 431–446.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.

Glass, G. (1976). Primary, secondary, and meta-analysis research. *Educational Researcher, 5*(10), 3–8.

*Hajjar, P. (1998). Predictability of reaction to ethical business situations based on levels of moral reasoning: a study of students. Unpublished doctoral dissertation, Nova Southeastern University, Ft Lauderdale, FL.

Herington, C., & Weaven, S. (2008). Improving consistency for DIT results using cluster analysis. *Journal of Business Ethics, 80*(3), 499–514.

*Hickman, M. (2008). Religiosity and moral reasoning of accounting students. Unpublished doctoral dissertation, Nova Southeastern University, Ft. Lauderdale, FL.

Higgins, J. P. T., Thomson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327* (September 6), 557–560.

*Ho, S. (1997). Ethical development of accounting students and public accounting practitioners' in Taiwan. *Research on Accounting Ethics, 3*, 293–303.

Huedo-Medina, T., Sanchez-Meca, J., Marin-Martinez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychological Methods, 11*(2), 193–2006.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error bias in research findings*. Thousand Oaks, CA: Sage.

*Icerman, R. C., Karcher, J. N., & Kennelley, M. (1991). A baseline assessment of moral development: Accounting, other business and nonbusiness students. *Accounting Educators' Journal, 3*(2), 46–62.

*Jeffrey, C. (1993). Ethical development of accounting students, non-accounting business students, and liberal arts students. *Issues in Accounting Education, 8*(1), 86–96.

*Jeffrey, C. G., & Ravenscroft, S. P. (2000). Rule observance attitudes and ethical development of accounting and non-accounting business students. *Research on Accounting Ethics, 7*, 1–18.

Johnson, J. L., Insley, R., Motwani, J., & Zbib, I. (1993). Writing performance and moral reasoning in business education. *Journal of Business Ethics, 12*(5), 397–406.

Kelley, S. W., Ferrell, O. C., & Skinner, S. J. (1990). Ethical behavior among marketing researchers: An assessment of selective demographic characteristics. *Journal of Business Ethics, 9*(8), 681–688.

King, P. M., & Mayhew, M. J. (2002). Moral judgment development in higher education: Insights from the Defining Issues Test. *Journal of Moral Education, 31*(3), 247–270.

Kish-Gephart, J. J., Harrison, D. A., & Treviño, L. K. (2010). Bad apples, bad cases, and bad barrels: Meta-analytic evidence about sources of unethical decisions at work. *Journal of Applied Psychology, 95*(1), 1–31.

*Klimek, J., & Wenell, K. (2011). Ethics in accounting: An indispensable course? *Academy of Educational Leadership Journal, 15*(4), 107–118.

Kohlberg, L. (1969). Stage and sequence: The cognitive developmental approach to socialization. In D. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347–480). Chicago: Rand McNally.

Kohlberg, L. (1984). *Essays on moral development: The nature and validity of moral stages* (Vol. 2). San Francisco: Harper & Row.

Konstantopoulos, S., & Hedges, L. V. (2009). Analyzing effect sizes: Fixed-effects models. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 279–293). New York: Russell Sage Foundation.

*Lampe, J. C., & Finn, D. W. (1992). A model of auditors' ethical decision-processes. *Auditing: A Journal of Practice & Theory, 11*(Supplement), 33–59.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48*(12), 1181–1209.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Littell, J. H., Cocoran, J., & Pillai, V. (2008). *Systematic reviews and meta analysis*. New York: Oxford University Press.

Maeda, Y., Thoma, S. J., & Bebeau, M. J. (2009). Understanding the relationship of moral development and individual characteristics: The role of educational contexts. *Journal of Educational Psychology, 101*(1), 233–247.

*Massey, D. W. (2002). The importance of context in investigating auditors' moral abilities. *Research on Accounting Ethics, 8*, 195–247.

*Massey, D. W. (1997). An investigation into the assessment of auditors' professional moral abilities and their improvement through the use of task-properties feedback. Unpublished doctoral dissertation, University of Connecticut, Storrs, CT.

Massey, D. W., & Van Hise, J. (2009). Walking the walk: Integrating lessons from multiple perspectives in the development of an accounting ethics course. *Issues in Accounting Education, 24*(4), 481–510.

*Massey, D. W., & Thorne, L. (2006). The impact of task information feedback on ethical reasoning. *Behavioral Research in Accounting, 18*, 103–116.

McCabe, A. C., Ingram, R., & Data-on, M. C. (2006). The business of ethics and gender. *Journal of Business Ethics, 64*(2), 101–116.

McPhail, K. (2006). In The Institute of Chartered Accountants of Scotland (Ed.), *Ethics and the individual professional accountant: A literature review*. Edinburgh: TJ International.

Miller, M. S., & Miller, A. E. (1976). It's too late for ethics courses in business schools. *Business & Society Review, 76*(17), 39–42.

Milne, M. J. (2001). Commentary on: Some thoughts on social and environmental accounting education. *Accounting Education, 10*(4), 369–374.

Mischel, W. (1968). *Personality and assessment*. New York: Wiley.

O'Fallon, M. J., & Butterfield, K. D. (2005). A review of the empirical ethical decision-making literature: 1996–2003. *Journal of Business Ethics, 59*(4), 375–413.

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of research*. San Francisco: Jossey-Bass.

Pierce, A. (2006). In The Institute of Chartered Accountants of Scotland (Ed.), *Ethics and the professional accounting firm: A literature review*. Edinburgh: TJ International.

Pigott, T. D. (2009). Handling missing data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 400–416). New York: Russell Sage Foundation.

*Ponemon, L. (1993). Can ethics be taught in accounting? *Journal of Accounting Education, 11*(2), 185–209.

Ponemon, L. (1992). Ethical reasoning and selection socialization in accounting. *Accounting, Organizations and Society, 17*(3–4), 239–258.

*Ponemon, L., & Glazer, A. (1990). Accounting education and ethical development: The influence of liberal learning on students and alumni in accounting practice. *Issues in Accounting Education, 5*(2), 195–208.

*Radtke, R. (1999). An experimental evaluation of the influence of moral reasoning on escalation errors. *Advances in Accounting Behavioral Research, 2*, 101–125.

Rest, J. (1990). *DIT Manual*. Minneapolis, MN: University of Minnesota Press.

Rest, J. (1986). *Moral development: advances in research and theory*. New York: Praeger.

Rest, J. (1979). *Development in judging moral issues*. Minneapolis, MN: University of Minnesota Press.

Rest, J., Narvaez, D., Thoma, S. J., & Bebeau, M. J. (1999). DIT2: Devising and testing a revised instrument of moral judgment. *Journal of Educational Psychology, 91*(4), 644–659.

*Rogers, V., & Smith, A. (2001). Ethics, moral development, and accountants in training. *Teaching Business Ethics, 5*, 1–20.

Rosenthal, R. (1987). *Judgment studies design, analysis, and meta-analysis*. New York: Cambridge University Press.

Rosenthal, R. (1979). The 'file drawer problem' and tolerance for null results. *Psychological Bulletin, 86*(3), 638–641.

Schlaefli, A., Rest, J., & Thoma, S. (1985). Does moral education improve moral judgment? A meta-analysis of intervention studies using the Defining Issues Test. *Review of Educational Research, 55*(3), 319–352.

Scofield, S. B., Phillips, T. J., & Bailey, C. D. (2004). An empirical reanalysis of the selection–socialization hypothesis: A research note. *Accounting, Organizations and Society, 29*(5–6), 543–563.

*Shaub, M. K. (1994). An analysis of the association of traditional demographic variables with the moral reasoning of auditing students and auditors. *Journal of Accounting Education, 12*(1), 1–26.

Snodgrass, J., & Behling, R. (1996). Differences in moral reasoning between college and university business majors and non-business majors. *Business and Professional Ethics Journal, 15*(1), 79–84.

Spickelmier, J.L. (1983). College experience and moral judgment development. Unpublished doctoral dissertation, University of Minnesota, Minneapolis, MN.

*St. Pierre, E., Nelson, E., & Gabbin, A. (1990). A study of the ethical development of accounting majors in relation to other business and non-business disciplines. *The Accounting Educators' Journal, 3*(Summer), 23–35.

*Sweeney, J. T., & Fisher, D. G. (1998). An examination of the validity of a new measure of moral judgment. *Behavioral Research in Accounting, 10*, 138–158.

Tenbrunsel, A. E., & Smith-Crowne, K. (2008). Ethical decision making: Where we've been and where we are going. *Academy of Management Annals, 2*, 545–607.

Thoma, S., Narvaez, D., Rest, J., & Derryberry, P. (1999). Does moral judgment development reduce to political attitudes or verbal ability? Evidence using the Defining Issues Test. *Educational Psychology Review, 11*(4), 325–341.

*Thorne, L. (2001). Refocusing ethics education in accounting: An examination of accounting students' tendency to use their cognitive moral capability. *Journal of Accounting Education, 19*(2), 103–117.

*Thorne, L. (2000). The development of two measures to assess accountants' prescriptive and deliberative moral reasoning. *Behavioral Research in Accounting, 12*, 139–169.

*Thorne, L. (1999). An analysis of the association of demographic variables with the cognitive moral development of Canadian accounting students: An examination of the applicability of American-based findings to the Canadian context. *Journal of Accounting Education, 17*(2–3), 157–174.

Trevino, L. K. (1992). Moral reasoning and business ethics: Implications for research, education, and management. *Journal of Business Ethics, 11*(5–6), 445–459.

*Venezia, C. C. (2005). The ethical reasoning abilities of accounting students. *The Journal of American Academy of Business, Cambridge, 6*(1), 200–207.

Waples, E., Antes, A., Murphy, S., Connelly, S., & Mumford, M. (2009). A meta-analytic investigation of business ethics instruction. *Journal of Business Ethics, 87*(1), 133–151.

*Wilhelm, W. J. (2010). Ethical reasoning instruction in non-ethics business courses: A non-intrusive approach. *Delta Pi Epsilon Journal, 52*(3), 152–167.