

Measuring Ethical Sensitivity and Evaluation

Tara J. Shawver
John T. Sennetti

ABSTRACT. Measures of student ethical sensitivity and their increases help to answer questions such as whether accounting ethics should be taught at all. We investigate different sensitivity measures and alternatives to the well-established Defining Issues Test (DIT-2, Rest, J. R. et al. [1999, *Postconventional Moral Thinking: A Neo-Kohlbergian Approach* (Lawrence Erlbaum Associates, Mahwah, NJ)], frequently used to measure the effects of undergraduate accounting ethics education. Because the DIT measures cognitive development, which increases with age, the DIT scores for younger accounting students are typically lower, have limited range, and are not likely to vary sufficiently with corresponding choices in ethical dilemmas. Since the DIT measures only the moral judgment component of ethical decision-making, we consider the multidimensional ethical scale (MES) to allow respondents to provide explanations for their moral and other judgments. The MES has been used to measure attitudes related to justice, utility, contractualism, egoism, and relativism. Unfortunately, the MES is not comparable in one-dimension to the DIT, and unlike the DIT, the MES has no theoretical or objective base. Therefore, we construct a comparable one-dimensional relative measure, a Composite MES Score, obtained from previous research on practicing accountants. We compare the reliability of this measure to the DIT in explaining the ethical choices of 54 specially chosen, somewhat homogeneous students, whose ages range from 18 to 19, and who are taking a second semester freshman accounting course at a private, religion-affiliated university. These particular students are relatively untrained in the formal use of questionable accounting choices. These students are less likely to recognize the dilemmas of the MES and are also less likely to demonstrate sufficient variation in their DIT scores, traditionally low for freshmen students. As freshmen, they are recent graduates of high school and more likely guided by other ethical influences including friends, family, or contractual obligations (some of the MES constructs) rather than higher cognitive development. This study confirms suspicions. We find the DIT

scores do not vary sufficiently to explain the moral reasoning of freshmen. For eight dilemmas and 24 choices we find the DIT score correlates with only three choices, whereas the MES regression models have at least one significant construct for 23 out of 24 ethical choices. The Composite MES Score (a relative measure) also explains 23 out of 24 choices and is statistically related to the DIT in only one of the choices. Unlike the DIT, the Composite MES permits pretest and retesting with different dilemmas to evaluate changes in ethical sensitivity. These results argue for relative rather than absolute measures of sensitivity and guides beyond cognitive development (the DIT-score) to explain undergraduate student sensitivity.

KEY WORDS: accounting ethics, Defining Issues Test (DIT), Multi-dimensional Ethics Scale (MES), ethics training

Introduction

The Education Committee of the National Association of State Boards of Accountancy (NASBA) has proposed to increase the content of the 150 semester hour requirement to include 3 h of ethics in accounting and 3 h of ethics in business. The scandals of KPMG and the well-known failed audits of Enron, WorldCom, Adelphia, Tyco, Xerox, Health-South, and more recently, many mortgage companies, remind us that accounting programs still need to teach ethical conduct.

The added-value of ethical training is often measured using tests of ethical awareness, such as the Defining Issues Test (DIT, Rest, 1979; DIT-2, Rest et al., 1999; see studies by Armstrong, 1993; Hiltebeitel and Jones, 1991; Ponemon, 1993; St. Pierre et al., 1990; Shaub, 1994; and more). The DIT-2 score is obtained from asking twelve questions on

each of the five ethically challenged dilemmas, such as whether it is fair to steal from the wealthy to feed a starving family. Higher DIT scores imply higher moral reasoning, defined by the higher Neo-Kohlbergian stages of moral development.

Studies using the DIT often show mixed results and there are at least seven limitations of the DIT. It uses a set of (1) fixed and limited dilemmas, (2) which are not associated directly with business, and creates (3) a one-dimensional score not validly useful as (4) a valid pretest/post test (repeated) measurement from the same respondent; (5) it measures only at most the first two of the required four activities (sensitivity, judgment, motivation, and character) that are needed to predict behavior (Rest et al., 1999); (6) it is known to be subject to gender, geographic, religious, and discipline-related biases or disadvantages (Fisher and Sweeney, 1998; Rest and Narvaez, 1994). Also, (7) because the DIT measures cognitive moral growth from children to adults and since it increases with age (Rest, 1979), beginning college students under the age of 20 would not be expected to create much variation in their DIT scores. This reduced variation inhibits statistical correlation between the range of DIT scores and the range of ethical evaluations scores, affecting the ability of the DIT to show an effect of an ethical intervention, such as an accounting ethics course. To some of these criticisms, Rest et al. (1999, p. 101) respond that the DIT was intended to measure at most only two of the four moral components, and it does produce a P-score and N2 score that have different (but similar) purposes; in addition, the variance of younger, less educated persons should have more variance, not less (Rest, 1979, p. 110).

One solution to improving any DIT measure is to ask respondents their justification for their choices made on ethical dilemmas, as does the Multidimensional Ethics Scale (MES, Reidenbach and Robin, 1990). The MES is a questionnaire that asks respondents to reveal their reasons for deciding on dilemmas. In doing so, the MES reveals the respondent's ethical orientation according to the five generally held philosophical constructs (moral equity, relativism, egoism, utilitarianism, and contractualism). [A relativistic respondent may feel that an action is ethical because it is "culturally acceptable" to me and "my family."] The DIT is mostly limited to questions of justice, utility, and rights-based

morality, whereas the MES expands these to cover what Rest et al. (1999, p. 164) refer to as the next two, social and personal domains.

Therefore, we investigate MES measures of ethical sensitivity as alternatives to the well-established DIT. We compare the reliability of these measures to the DIT in explaining the ethical choices of 54 students of a second semester freshman accounting course at a private university. These particular students are generally untrained in the formal use of questionable accounting choices and fraud, the related dilemmas of the MES, and they are less likely to demonstrate large variation in their DIT scores. Being recent graduates of high school, they are more likely influenced by more informal ethical influences (friends and family, some of the MES constructs). We find for eight dilemmas and 24 choices the DIT score explains only three choices, whereas the Composite MES Score explains 23 out of 24 choices and is not statistically related to the DIT for any of the choices. These results argue for relative, rather than absolute measures [such as the traditional DIT-2] for assessing the change in undergraduate student sensitivity levels.

This paper is organized as follows: first, the hypothesis development considers the strengths and weaknesses of the DIT and MES; this is followed by the methodology used to study these ethical evaluations of the students. The study ends with corresponding recommendations and conclusions.

Hypotheses development

Defining Issues Test (DIT-2)

The advantages of the DIT

The typical measure of moral awareness and reasoning (cognition) is the DIT-2 (Rest et al. 1999, pp. 644–659). The DIT-2 score (Rest et al., 1999), first developed by Rest (DIT, 1979), is based on an updated view of Kohlberg's (1969) six-stages of moral development taken from Piaget's study of children (Piaget, 1965). As a child's physical development progresses from that of a child to an adult, cognitive development progresses through higher stages. Cognitive development measurement starts with responses made under pleasure and pain rewards and continues to responses motivated by

individually held principles of conscience and the greater good for society.

In taking the DIT, a subject responds to several standardized vignettes (dilemmas). The uniform ranking of the most important rationales for decisions made in each situation eventually determines the DIT score. Higher DIT-2 scores imply higher levels of moral reasoning. The advantages of the DIT-2 to other tests are that it is a structured, self-administered test that can be objectively scored. The DIT is well known and accepted among many academic disciplines (Rest et al., 1999).

Many studies of students and practicing accountants, e.g., Ponemon and Gabhart (1993), show the relationship of higher DIT scores (higher moral reasoning) to more frequent ethical choices. Duncan and Knoblett (2000, p. 28) find “a corporate accountant’s ethical development (a higher DIT score)” associates with “the accountant’s decision to engage in managing earnings.” So, generally speaking, a higher DIT score on a dilemma will suggest a stronger ethical choice, especially with accounting practitioners, but there are exceptions which lead to the “U”-shaped curve for the DIT and opposite-than-anticipated results. A respondent with a higher DIT score may believe an illegal action is just (e.g., copying software makes it ‘available to the public’) resulting in a higher DIT score for lower moral choices. This possible inverted “U”-shaped curve in DIT scores was first reported by Ponemon (1988) and later by Ponemon and Gabhart (1993). However, this U-shape does not seem to appear in studies of practicing accountants (Abdomohammadi and Arian, 2007; Bernardi and Arnold, 2004; Scofield et al., 2004). Hence, it would appear that generally,

H1: Students with higher moral reasoning are more likely to evaluate certain questionable actions as unethical.

The disadvantages of using the DIT score

The use of the DIT to measure student behavior has many disadvantages besides a possible “U-shape” scoring problem. First, Rest (1979) identifies four stages as the sequential steps an individual must take to make a decision: interpretation, judgment of the morality of the situation, choice of action, and implementing the choice of action. Admittedly (Rest

et al., 1999), the DIT-2 measures at most, the first two. Also, the DIT has been shown to be subject to gender, geographic, religious, and discipline-related biases (Fisher and Sweeney, 1998; Rest and Narvaez, 1994).

Accounting and business majors exhibit DIT scores that are well below the norms for college students in non-accounting academic disciplines (Armstrong, 1987; Ponemon and Gabhart (1993); Rest, 1986; Shaub, 1994). Accounting students generally score lower than business students (Thornton, 2000). Shaub (1994) points to Kohlberg’s Cognitive Moral Development theory where conventional thinkers identify with, or internalize, rules and expectations, particularly the expectations of authorities such as those emphasized within the accounting curricula and the business professions. Given that accounting and business programs are typically rule based, it can be expected that in response to ambiguous, non-rule-based DIT ethical dilemmas, accounting and business students would have different ethical sensitivities and judgments than students from other academic disciplines.

Four other limitations of the DIT include its (1) fixed and limited dilemmas or vignettes, (2) each of which is not associated directly with business. Since the vignettes are fixed, (3) valid pretest and post-test (repeated) measurements from the same respondent on the same vignettes are limited, as students may remember the same vignettes. Also, (4) the DIT creates only a one-dimensional score without explanation for behavior. Finally, because the DIT measures cognitive moral growth, a fixed age group, such as beginning college students, would not be expected to have much variation in their DIT scores and this would inhibit statistical correlation tools which attempt to show a relation between the range of DIT scores and the range of ethical evaluations scores, reducing the value of the DIT for measuring the effects of ethical interventions such as ethics courses.

The Multidimensional Ethics Scale (MES)

Advantages and disadvantages of using the MES questionnaire

Well-known alternatives to the DIT are few. One solution to predicting behavior may lie in using more than one measurement. [Rest et al. (1999, Chap. 5) suggest adding a “utilizer moderator” measurement to the DIT score, since the DIT score by itself is

insufficient to predict behavior.] Prior research on ethical awareness using accounting students as subjects (e.g. Cohen et al., 1996) suggests multi-dimensional measures such as the Multi-Dimensional Ethics Scale (MES, Reidenbach and Robin, 1988) found in Appendices A and B. Respondents first evaluate questionable vignettes (actions) in Appendix A and then make choices in answering questions 13–15 of Appendix B. Respondents then provide their own (relative) reasons for their choices by answering questions 1–12 found in Appendix B. These twelve questions represent five generally recognized accepted philosophical areas (constructs) as to why one may “do the right thing.”

Respondent actions identified as “ethical” in dilemmas (question 15) could be explained because the action (1) is just (has moral equity), and is consistent (2) with the respondent’s personal understanding of the world (is relative), (3) with the respondent’s long-term interests (is egoistic), (4) with the respondent’s understanding of universal human costs and benefits (has utility value), and (5) with the respondent’s implied or explicit contractual obligations (is performed for deontological reasons).

An action then identified as “ethical” (score of 1) in question 15, Appendix B, is one where the respondent also likely believes the action is correspondingly just (a score of 7 for questions 1–3). This perceived ethical action score of 1 may also likely correspond to scores (near 7) in the answers to the relativistic questions 4 to 6 because it is acceptable to “my friends and family.” This perceived ethical action score of 1 may or may not correspond to scores (near 7) for (3) egoistic questions 7–8 because the action promotes an individual’s long-term interests. This perceived ethical action may also likely correspond to scores (near 7) in the answers to the utilitarian questions 9 to 10, since the action is done for the greatest good. The action may also likely correspond to scores (near 7) in the answers to the contractual questions 11–12 because it satisfies implied or explicit contracts, as in the appropriate use of generally accepted accounting principles.

Cohen et al. (1998) uses the eight vignettes (Appendix A) developed by Cohen et al. (1993) in a study evaluation made by 645 undergraduates. Sennetti et al. (2004) and Shawver et al. (2006) use these same vignettes for studies of accountants working for companies offering initial public offer-

ings (IPOs) and find their MES measures explain more choices than just the measure of cognitive development, the DIT-2. Shawver and Clements (2007) and Shawver and Shawver (2008) use the MES and selected vignettes from the original eight that Cohen et al. developed to explain choices made by accounting students and accounting professionals, respectively. Cruz et al. (2000, p. 226) use three different and questionable tax dilemmas to explain “not only what the respondent (the CPA tax practitioner) believes regarding the conflict of serving the client versus serving the IRS, but why he or she believes it.” Unlike the one score, fixed dilemmas of the DIT, different (but correlated) dilemmas can be substituted in the MES questionnaire for the purpose of creating pretests and post-tests needed to measure the intervention (ethics course) effect, and in this way the MES may overcome the pretest–post-test problem of the DIT using the same dilemmas.

Disadvantages of the MES

The MES elicits the reasons one may use for explaining ethical or unethical behavior. Higher scores for each MES construct do not necessarily represent corresponding proportional increases in moral cognition, as intended with the DIT, correlated with the stages Kohlberg’s Stages of Moral Development (Rest, 1979). Since the MES has no corresponding objective scale, it may fail “to provide a useful psychometric construct for measuring the ethical thinking process of professional accountants” and “does not provide an insight into the fundamental question what makes accountants more or less ethical” (Jones and Poneman, 1993, p. 411). Therefore, the five MES scores can at best present the respondent’s “explanation” (correlation) for the respondent’s choices.

H2: The MES constructs of moral equity, relativism, egoism, utilitarianism, or contractualism explain ethical (and unethical) choices by accounting students.

Since the MES is not one-dimensional, and therefore cannot be compared in one dimension to the DIT, it “does not offer researchers a substitute instrument (Weber 1996, p. 517).” Therefore, we suggest one previously used on practicing accountants by Sennetti et al. (2004) and Shawver et al. (2006).

The Composite MES

For the MES in Appendix B, line 15, if the action is chosen as “ethical,” then the reasons for this choice (‘just,’ ‘fair,’ etc.) should be scored nearer to 7, and the composite sum of answers to questions 1 through 12 should be high, and if the action is chosen as “unethical,” the sum should be lower. Sennetti et al. (2004) and Shawver et al. (2006) name this sum, the Composite MES. This Composite MES has no corresponding absolute, objective meaning as does the DIT, but it can be used as a relative comparison to measure (explain) a student’s improvement in sensitivity in the respondents’ perceived concept of justice, rightful obligation, and so forth.

H3: The Composite MES explains the ethical (and unethical) evaluations of accounting students.

The Composite MES is also one way to compare the MES to the DIT. Higher ethical sensitivity measured by a higher DIT score should correspond to lower MES scores, each near “1” (unjust, unfair, etc.), and therefore should correspond to lower Composite MES Scores. Hence, we would anticipate a relationship between the Composite MES and the DIT, the measure of moral reasoning:

H4: The Composite MES is related to moral reasoning.

Methodology

We study specially chosen somewhat homogeneous students, aged 18–19, at a private, religion-affiliated university, taking the freshman second semester introductory accounting course (Introduction to Managerial Planning and Control, 120 students in 7 sections). Having had some exposure to the first course, Introduction to Financial Accounting, these students are aware of, but are not necessarily trained to recognize questionable accounting choices. Each participant completed three separate questionnaires including the Multidimensional Ethics Scale (MES), the Defining Issues Test (DIT-2), and a control questionnaire to moderate responses bias taken from the Balanced Inventory of Desirable Responding (BIDR), the Impression Management (IM) scale.

The Multi-Dimensional Ethics Scale (MES) questionnaire

Each participant received a set of eight dilemmas or vignettes (Appendix A), and for each vignette action, responded to the sixteen lines of the MES questionnaire (Appendix B). Vignettes 3 and 7 are accounting related, and 5, 6, and 8 present earnings management related, specifically. Vignette 3 is a direct violation of the 1977 Foreign Corrupt Practices Act, a well-known accounting rule that also requires firms to establish accounting controls over their financial statements. Vignette 7, expensing personal gifts, is one familiar to professional (tax) accountants. Vignette 5 suggests an early shipment of goods, a standard fraudulent practice used to overstate income. A bank that extends an unfair loan, Vignette 6, fraudulently increases bank income (similar to that conducted in recent mortgage-related fraud). To unfairly reduce the bad debt, Vignette 8, is to similarly reduce expenses and increase income.

Those that are not accounting issues, Vignettes 1, 2, and 4, indirectly identify with the respondent’s short-term self-interest. Vignette 1 addresses the unfair layoff of a younger worker. Vignette 2 challenges product safety (but to the employee’s advantage). Vignette 4 considers the common problem of software copyright violation, mentioned earlier.

Scoring of the MES

The respondents decide if each vignette suggests (a) behavior that is ethical or unethical (ranked 1 to 7, question 15 in Appendix B), (b) an action they would choose to perform (question 13, Appendix B), and (c) an action their peers would choose to perform (question 14, Appendix B). [Israeli (1988, p. 263) points out “that the best predictor of a respondent’s behavior is their beliefs and perceptions concerning their peers’ behavior.”] Respondents next give the reasons (answers to questions 1–12, Appendix B) for their choices, (a)–(c). Appendix B also permits respondents to explain why they believe the action is ethical.

The DIT

Participants’ completed DIT questionnaires that were scored objectively by the Center for the Study of Ethical Development at the University of

Minnesota. The DIT scoring sheets provide a 5-digit grid for subject identification and instructions for coding are provided in the DIT manual. In keeping with a promise to protect participant identity, a four-digit number was used to match the DIT forms with the experimental materials without receiving any identifying information from the participants.

The scoring of the DIT

The most commonly reported score for the DIT is the “P” (principled reasoning) score. “P” scores are intended to measure the percentage of Kohlberg’s Stage Five or Stage Six cognitive moral development levels. “P” scores represent the simple sum of scores converted to a percentage for items pertaining to principled moral considerations and range from 0 to 95 (Rest, 1979). The higher the participant’s “P” scores, the higher the percentage of Kohlberg’s post-conventional moral reasoning level.

Specifically, each participant received the DIT-2’s five-version ethical dilemma questionnaire and a set of standard responses conforming to Kohlberg’s six-stage model. Respondents are asked to read each ethical scenario and to indicate for each scenario whether it raises a questionable issue, and if so, how important the respondent considers the issue to be. Respondents are asked to determine the importance of each item by a rating scale of five (great importance) to one (no importance). In addition, the questionnaire asked respondents to rank the 12 ethical dilemmas in terms of the four most important. Further, after reading each dilemma story, respondents are asked to state their preference for what action to take in the story on a seven-point scale (1 = strongly favor some action; 7 = strongly oppose the action). From these responses descriptive statistics and valid “P” scores are determined and computed by the Center for the Study of Ethical Development located within the University of Minnesota.

The Balanced Inventory of Desirable Responding (BIDR)-controlling for response bias

Research finds females more likely to present a social desirability (reporting) bias (Paulhus, 1991). Since studies of respondents are always subject to response bias, we filter all responses through a bias test called the IM scale, which is a subset of the Balanced Inventory

of Desirable Responding scale (BIDR, the short version 6, Paulhus, 1991, pp. 17–51) that measures self-deceptive positivity and impression management. Responses to twelve statements of impression management, such as “I never cover up my mistakes,” are given on a 7-point likert scale ranging from “not true” to “very true” to evaluate whether a person is deliberately underreporting socially undesirable acts while over reporting socially desirable acts. Paulhus (1991) found typical scores for males to average 2.93 with a standard deviation of 2.8 and for females to average 3.21 with a standard deviation of 2.8 on a 7-point scale. Sennetti et al. (2004) find in their study of IPO accountants a correlation of the IM to the DIT P-score, but not to the Composite MES score.

Results

Table I – The respondents are homogeneous in their DIT scores

Of 120 students taking the second semester accounting course, 97 agreed to participate in this study. Of these, 77 presented usable, DIT-validated questionnaires. We further selected only 54 of those students whose age was less than 20 years in order to create a more homogeneous group. Demographic information relating to these students is summarized in Table I; 33 are males and 21 are females. Calculated from the DIT, their “political view” response, scored from liberal to conservative scored as 1–4, averages 2.69 (about in the middle) and is perhaps less conservative than expected for these students attending a religion-affiliated university.

The mean IM score is 2.94 for males and 3.38 for females. This score is comparable to those reported by Paulhus (1991) for undergraduates: norms for males as 2.93 and for females 3.21 suggesting that these 54 students have not responded in a socially biased way to provide a more favorable reporting image.

Finally, the student mean DIT-2 (P-Score) is 26.26, shown in Table I. [Not shown are the median and mode, both equal to 26.] These scores compare favorably to those of older studies made with the DIT (Rest and Narvaez, 1994). “Institutionalized Delinquents” [as would be expected under Kohlberg Cognitive Moral Development Stages] then scored the lowest average, 18.4, and “Senior High School

TABLE I
Descriptive statistics of the 54 students

Gender	<i>n</i>	
Male	33	
Female	21	
Total	54	

	Mean	Standard deviation
Political view	2.69	1.01
P-score	26.26	14.31
IM score females	3.38	1.60
IM score males	2.94	1.69

Students” scored 31.8. This freshman average score, 26.26, although lower than 31.8, is reasonably consistent with research that suggests lower than expected scores are found for religious, one-year post high school college students who choose business as a major (e.g., Sweeney and Fisher, 1998).

Most important, we find these freshmen DIT scores consistent with expectations and homogeneous, as

shown by the coefficient of variation (cv) computed from the ratio of the standard error (of the mean) to the mean. The cv is then computed as ($[14.31/\text{square root of } 54]/26.26$ or) 0.074 (7.4%). This low cv (at 10% or less) suggests data that are more homogeneous (Cochran, 1963, p. 54; see also Medicare Carriers Manual, Part 3, Sampling Guidelines, 2000, p. 24).¹ Freshmen DIT scores are typically low and do not vary much, as shown here. This influences the tests of correlation that use the DIT to explain ethical evaluations.

H1 is not supported by Table II

H1: Students with higher moral reasoning are more likely to evaluate certain questionable actions as unethical.

Table II shows each student’s DIT score correlated with his/her answers to all three (column) ethical evaluations (questions 13–15) given for each dilemma.² The evaluations on the MES are numbered

TABLE II

The Pearson correlation coefficients created from the DIT score and the ethical evaluations made for each vignette

Vignette	(15) Ethical	(13) Undertake action	(14) Peers undertake action	Sig. (1-tailed)
V1 layoff	−0.130	−0.015	0.114	0
<i>p</i> -value	0.174	0.458	0.205	
V2 product safety	−1.810	−0.047	−0.082	1 (ethical)
<i>p</i> -value	0.096**	0.368	0.278	
V3 bribe	0.015	0.129	0.040	0
<i>p</i> -value	0.457	0.176	0.386	
V4 software	−0.043	0.122	0.051	0
<i>p</i> -value	0.378	0.189	0.358	
V5 early shipment	−0.076	−0.059	−0.136	0
<i>p</i> -value	0.293	0.336	0.163	
V6 loan	−0.269	−0.134	−0.179	2 (ethical, peers)
<i>p</i> -value	0.025*	0.167	0.097**	
V7 gifts	−0.067	0.001	−0.053	0
<i>p</i> -value	0.315	0.496	0.351	
V8 bad debt	0.086	0.114	−0.048	0
<i>p</i> -value	0.267	0.206	0.365	

*Significance determined from *p*-values <0.05, **marginally significant for *p*-values <0.10.

Total of statistically significant correlations – 3.

Total of possible significant correlations – 24.

as (13) “The probability that I would undertake this action,” (14) “The probability that my peers would undertake this action,” and (15) “The action described is ethical.” Under column one [(15) Ethical], there are only two of eight questionable actions that are statistically significant, even at the level of 10%: V2 (Product safety) and V6 (Unfair loan). The DIT fails to explain 21 of the 24 ethical evaluations of Table II; H1 fails.

Under the binomial sign test, H1 fails again. For this test we assume that the DIT explains 50% of the actions. The chance of observing 21 or more out of 24 as unexplained is 1%, less than the critical (*p*-value) of 5%. Once again, H1 fails.

H2 is supported by Table III

H2: The MES constructs of moral equity, relativism, egoism, utilitarianism, or contractualism, explain ethical evaluations of business students.

First evaluation: a given action (vignette) described is ethical (question 15)

The first part of Table III provides the mean scores of the question 15, “Is the action ethical,” for each vignette. None of the averages, which range from 3.81, for the dilemma of a foreign bribe, to 5.48, for the question of product safety, are statistically close to 1 (ethical) or to 7 (unethical).

From Table III it can be argued that all these means are statistically different from 1 or 7, using either the pooled standard deviation of this table, or conservatively the largest standard deviation, 2.06. The standard error of the mean then becomes 2.06/ (square root of 54) or 0.28. Then 1 plus two standard errors give the upper 95% limit, 1.56, much below the lowest mean, 3.81. Therefore, on average, none of the dilemmas are considered “ethical,” as they are too far above 1.0. The same applies to the score for “unethical,” 7 which has a lower limit of 7–0.56 or 6.44, much higher than the highest mean, 5.48. On average the students did not find any one of the actions ethical or unethical. But there are students

TABLE III
Means, Adj. *R*², coefficients and *p*-values

Dependent variable: is it ethical									
Vignette	Ethical mean	Std dev.	Adj. <i>R</i> ²	J	R	E	U	C	Sig. (2-tailed)
V1 layoff	4.17	1.76	0.427	−4.238 0.000*	−1.151 0.256	−1.102 0.276	−1.403 0.167	−1.040 0.303	J
V2 product safety	5.48	1.58	0.295	−1.828 0.074**	−0.819 0.417	0.303 0.763	−0.313 0.756	−1.826 0.074**	JC
V3 bribe	3.81	1.53	0.417	−2.785 0.008*	0.273 0.786	1.872 0.067**	−0.779 0.440	−1.796 0.078**	JEC
V4 software	4.02	2.06	0.793	−4.827 0.000*	−0.472 0.639	1.853 0.070**	−3.057 0.004*	−4.888 0.000*	JUEC
V5 early shipment	4.59	1.75	0.641	−4.306 0.000*	−2.255 0.029*	0.355 0.724	−0.133 0.895	−1.148 0.257	JR
V6 loan	4.74	1.43	0.640	−3.238 0.002*	−2.807 0.007*	−0.426 0.672	0.381 0.705	−3.124 0.003*	JRC
V7 gifts	4.96	1.67	0.612	−4.149 0.000*	−0.228 0.821	1.176 0.246	−1.576 0.122	−1.265 0.212	J
V8 bad debt	5.17	1.65	0.771	−7.221 0.000*	0.698 0.488	0.203 0.840	−0.962 0.341	−1.218 0.229	J

For is it ethical, response closer to 1 indicates an ethical action, closer to 7 indicates unethical actions. For MES items, *Significance determined from *p*-values <0.05, **marginally significant for *p*-values <0.10. Justice (J), Relativism (R), Egoism (E), Utilitarianism (U), Contractualism (C).

who do find some actions unethical as shown in the second part of Table III.

The reasons a given action (vignette) described is unethical: the MES constructs

In the second part of Table III, following Cohen et al. (1993), we present the outcomes of three regression models, each one with the dependent variable as one of the three evaluations, questions 13–15 of Appendix B, and each one having the independent variables, the means of the responses to the questions of for each of the five constructs: justice (J), relativism (R), egoism (E), utilitarianism (U), and contractualism (C).

Each vignette-dilemma, V, has at least one statistically significant predictor (construct) variable, justice (J). For vignette V1, the unfair layoff, even though the ethical mean score, 4.17, is not statistically close to unethical (7), the coefficient of justice (J), -4.238, is statistically significant for the model with adjusted R^2 , 0.0427. This implies there are those respondents who feel this action is both ethical

(responses closer to 1) and just (responses closer to 7), and respondents who feel this action is both unethical (responses closer to 7) and unjust (responses closer to 1). The average, 4.17, suggests more think it unethical, but not enough to be statistically significant.

But the justice construct explains the ethical choices (has a significant coefficient) in all eight vignettes. Contractualism is found in four vignettes, utilitarianism in one, relativism in two, and egoism is found to be statistically significant in two vignettes. Therefore, we see in Table III that since at least one construct of the MES explains every one of the eight ethical choices, H2 is supported, at least for the question, “The action described is ethical,” question 15.

H2 is supported by Tables IV and V

Tables IV and V describe similar results for the other ethical evaluations (questions 13 and 14) that relate

TABLE IV
Means, Adj. R^2 , coefficients and p -values

Dependent variable: would you do it									
Vignette	Do it mean	Std. dev.	Adj. R^2	J	R	E	U	C	Sig. (2-tailed)
V1 layoff	4.30	1.63	0.463	-3.783 0.000*	-1.169 0.248	-0.454 0.652	-2.495 0.016*	-1.493 0.142	JU
V2 product safety	5.15	1.80	0.368	-1.913 0.062*	-2.338 0.024*	0.890 0.378	0.595 0.555	-0.659 0.513	JR
V3 bribe	3.43	1.61	0.390	-1.883 0.066**	-0.419 0.677	-0.483 0.631	-2.335 0.024*	-2.044 0.046*	JUC
V4 software	2.87	1.69	0.388	-0.362 0.719	-0.758 0.452	0.504 0.616	-4.252 0.000*	-1.842 0.072**	UC
V5 early shipment	4.11	1.97	0.649	-6.615 0.000*	-1.974 0.054**	0.335 0.739	1.182 0.243	1.169 0.248	JR
V6 loan	4.09	1.69	0.534	-2.590 0.013*	-1.932 0.059**	-2.133 0.038**	-3.225 0.002*	-1.153 0.255	JREU
V7 gifts	4.37	1.98	0.437	-2.880 0.006*	-1.015 0.315	1.983 0.053**	-2.755 0.008*	0.775 0.442	JEU
V8 bad debt	4.91	1.69	0.672	-4.473 0.000*	0.414 0.681	0.442 0.661	-1.531 0.132	-2.226 0.031*	JC

For would you do it, response closer to 1 indicates would complete action, closer to 7 indicates would not complete action.

For MES items, *Significance determined from p -values <0.05, **marginally significant for p -values <0.10. Justice (J), Relativism (R), Egoism (E), Utilitarianism (U), Contractualism (C).

to undertaking the action or choosing the action that peers would undertake. For Table IV under the question (13), the pirating of software (V4) now has the lowest mean, 2.87, and seems more likely to be chosen as undertaken (since 2.87 is the mean nearest to 1, seemingly more likely undertaken). Still, this mean is above the largest possible two standard deviation limit 1.56, and therefore statistically different from 1, “undertaken.” On the average pirating software (V4) is not to be undertaken, although since 2.87 is closer to 1, it is more likely than it is not to be undertaken. If a student chooses to pirate software as an action to be undertaken, it may be chosen for two reasons, one for greater utility (U), with statistically significant coefficient, -4.252, and one for its implied contract (C), with statistically significant coefficient, -1.842. Therefore, we see in Table IV that since at least one construct of the MES explains every one of the eight “to be undertaken” choices. H2 is supported.

For Table V, “Would your peers undertake this action?” the pirating software vignette (V4) again has

the lowest mean, 2.78, the one nearest to 1. But this mean is still outside the maximum two standard deviation limit, 1.56, and therefore is statistically significant from 1, in this case, peers would not take this action. On the average pirating software (V4) is not to be undertaken by peers, although since 2.78 is closer to 1, it is more likely than it is not to be undertaken. If a student chooses to pirate software as an action to be undertaken by peers, it may be chosen for only one of the two reasons given earlier, for greater utility (U), with statistically significant coefficient, -3.968. All but one of the vignettes in Table V, V1, unfair layoff, have at least one statistically significant construct. Therefore, in Tables III–V for 23 out of 24 actions there is at least one construct that is statistically significant. H2 is supported.

H3 is supported by Table VI

H3: The Composite MES explains the ethical evaluations of business students.

TABLE V
Means, Adj. R², coefficients and p-values

Dependent variable: would your peers do it									
Vignette	Peers mean	Std dev.	Adj. R ²	J	R	E	U	C	Sig. (2-tailed)
V1 layoff	4.30	1.38	(0.027)	0.559	-1.523	0.172	-0.342	-1.153	None
				0.579	0.134	0.864	0.734	0.255	
V2 product safety	4.04	1.69	0.267	-0.062	-3.454	-1.429	0.047	0.872	R
				0.951	0.001*	0.159	0.962	0.387	
V3 bribe	3.22	1.53	0.127	-1.192	0.046	-0.742	-2.142	-0.725	U
				0.239	0.964	0.462	0.037*	0.472	
V4 software	2.78	1.70	0.299	-0.686	0.133	-0.368	-3.968	-0.231	U
				0.496	0.895	0.714	0.000*	0.818	
V5 early shipment	3.43	1.87	0.381	-3.174	-2.533	-0.257	0.623	1.509	JR
				0.003*	0.015*	0.798	0.536	0.138	
V6 loan	3.69	1.44	0.435	-2.274	-1.295	-2.736	-3.877	0.757	JEU
				0.027*	0.202	0.009*	0.000*	0.453	
V7 gifts	3.76	1.62	0.150	-2.650	0.364	-0.142	0.079	0.083	J
				0.011*	0.717	0.888	0.938	0.934	
V8 bad debt	4.26	1.52	0.525	-2.050	-1.327	-0.474	-0.255	-1.890	JC
				0.046*	0.191	0.638	0.799	0.065*	

For peers, response closer to 1 indicates would complete action, closer to 7 indicates would not complete action. For MES items, *Significance determined from p-values <0.05, **marginally significant for p-values <0.10. Justice (J), Relativism (R), Egoism (E), Utilitarianism (U), Contractualism (C).

TABLE VI

The Pearson correlation coefficients created from the Composite MES and evaluations made for each vignette

Vignette	Ethical (E)	Undertake it (U)	Peers undertake it (P)	Sig. (1-tailed)
V1 layoff	-0.667 0.000*	-0.682 0.000*	-0.127 0.179	EU
V2 product safety	-0.555 0.000*	-0.561 0.000*	-0.456 0.000*	EUP
V3 bribe	-0.507 0.000*	-0.635 0.000*	-0.398 0.001*	EUP
V4 software	-0.792 0.000*	-0.559 0.000*	-0.381 0.002*	EUP
V5 early shipment	-0.761 0.000*	-0.681 0.000*	-0.553 0.000*	EUP
V6 loan	-0.746 0.000*	-0.742 0.000*	-0.583 0.000*	EUP
V7 gifts	-0.718 0.000*	-0.581 0.000*	-0.389 0.002*	EUP
V8 bad debt	-0.820 0.000*	-0.793 0.000*	-0.736 0.000*	EUP

*Significance determined from p -values < 0.05 , **marginally significant for p -values < 0.10 .

Total of statistically significant correlations – 23.

Total of possible significant correlations – 24.

The responses to all questions 1–12 in Appendix B for each vignette, computed as a sum, the Composite MES, can be used to test H3. For the “unethical” evaluation of an action score (near the number 7) there should be a corresponding reason (questions 1–12) ranked near “1” in Appendix B. Therefore, a total of lower MES scores (each nearer to 1) would be anticipated for all actions evaluated as unethical (question 15). This relationship implies a negative or inverse correlation between the Composite MES and the ethical evaluation score.

Table VI presents these negative correlations coefficients for each dilemma. Unlike the DIT in Table II, the Composite MES, similar to the MES regressions of Table III, explains most of the evaluations, 23 of the 24 correlation coefficients are statistically significant [and this number 23 of 24 exceeds the binomial sign test of mere chance]. As similarly seen in Table V, only the “peer” evaluation

TABLE VII

The correlation between the DIT and Composite MES score

Vignette	Pearson correlation	Sig. (1-tailed)
V1 layoff	-0.022	0.438
V2 product safety	0.128	0.178
V3 bribe	0.028	0.421
V4 software	-0.079	0.284
V5 early shipment	0.140	0.157
V6 loan	0.160	0.124
V7 gifts	0.179	0.097**
V8 bad debt	-0.053	0.351

*Significance determined from p -values < 0.05 , **marginally significant for p -values < 0.10 .

correlation in one dilemma, vignette V1 unfair lay-off, is not significant. H3 is supported.

H4 is not supported by Table VII

H4: The Composite MES is related to moral reasoning.

This Composite MES can be compared to the DIT score. MES scores nearer to “1” (unfair, unjust, etc.) would be anticipated for each questionable action by persons of higher moral reasoning (DIT) scores. Hence, we would anticipate a negative correlation between the DIT-2 and the Composite MES for questionable actions, but we find no evidence of this in Table VII [as found in Shawver et al. (2006) for practicing accountants]. Although the some of signs of the correlation coefficients are the correct direction as anticipated for three vignettes, only one of the eight correlations (for vignette V7, gifts) can be considered statistically significant, and only with one tailed p -value of 0.097 when typically we would expect at least 0.05. H4 is not supported.

This lack of correlation between the Composite MES and the DIT can be explained. While the DIT measures some generally recognized degree of justice and the sensitivity to the feelings of others, the Composite MES only includes the respondent’s

feelings of what is just, which in fact may lead to injustice. Acting according to only what is just for friends and family is described as injustice by Socrates (Kaplan, 2005, Lecture 3, Plato – *The Republic*). Therefore, we carefully observe the Composite MES as a relative, not an absolute measure of ethical sensitivity.

Summary and conclusions

This study empirically confirms what many researchers have surmised. Table II shows the DIT may not sufficiently explain the moral reasoning of somewhat homogeneous freshman accounting students (H1). We find explanations in at most only three of the 24 comparisons, three ethical choices for eight dilemmas. Tables III–V show the MES regression models identify 23 out of the 24 ethical evaluations (H2). Further, Table VI also shows the Composite MES (the total MES score) also explains 23 out of the 24 evaluations made (H3), unlike the DIT. But the respondent who correctly recognizes the questionable action does not necessarily do so because of higher moral reasoning, since Table VII shows this Composite MES is not statistically related to the DIT (H4). As demonstrated by the constructs given by the MES, there are reasons other than those given by higher cognitive moral development for recognizing questionable actions. While these may not be the “right” reasons, they are consistent with Socrates’ expectations (Kaplan, 2005, Lecture 3, Plato – *The Republic*) for under-developed moral reasoning.

Finally, this research by creating a Composite MES suggests one resolution to the non-comparability issue raised in prior research by Weber (1996) and Flory et al. (1993). “The multivariate approach is a situation-specific, temporal, and process-oriented, while the moral development approach is general (not situation specific), enduring, and trait-oriented” (Flory et al., 1993, p. 418).

Limitations and recommendations

The limitations of response

Limitations of this study begin with the limitations of self-reporting questionnaires, the size of the

sample, the location and type of the respondents who attended a Northeast, private, religion-affiliated university. This study specifically limited its respondents to those who would likely not vary much in their DIT scores, reflect less cognitive moral development, and more than likely be influenced by other considerations, such as friends and family, some of the MES constructs. Under different conditions other respondents could provide different results.

Another limitation is that the change in the Composite MES Score assumes this change would be consistent over all constructs. For example, one respondent may sense that an action now, in a post-training class environment, is more unethical and therefore has less justice (the justice score decreases), but at the same time, this respondent feels a contradictory need to perform this unethical act according to some duty (and therefore increases the contractual obligation score). In this unlikely case the Composite Score may not change or change in the unanticipated direction.

It is also possible that the DIT has a similar limitation in that if respondents’ choices on the DIT increase the P-Score for one of the five situations but decreases in another, this may cause no change in the repeated DIT P-Score, but this is more likely for the DIT since the dilemmas are fixed and repeated in post-test evaluations. This could be one explanation as to why so few studies of accounting ethics courses have found changes (pretest and post-test) DIT P-scores as a result of ethics interventions and coursework. This is less likely for the Composite MES score since it is possible to vary the dilemmas for the MES.

Further, since each respondent gave three evaluations on the each of the eight dilemmas, it could be argued that there were only eight choices, not 24, since the questions, ‘ethical?’, ‘will you do it?’ and ‘will your peers do it?’ seem related. First, prior research rejects this (see Israeli, 1988). Israeli (1988) finds that the choice made ‘whether “peers would undertake” the action’ corresponds to evidence that the respondent (not the peer) would actually behave in this way.

Also, the variation in the MES model coefficients found in Tables III–V show statistically significant changes in choice and in total evaluation.

Recommendations

Because, as indicated here with freshmen, (1) the DIT scores of students tend to be lower and to cluster without sufficient variation to support statistically significant findings and (2) sources other than cognition influence student ethical sensitivities, alternative measures are needed. The Composite MES is only a relative score and may not truly measure the DIT version of justice. But because it is a total of many attitudes or beliefs (lines 1–12 in Appendix B), it is subject to greater variation with a larger range, and as shown here, more likely a tool to show statistically significant correlation with ethical evaluations. It can be used by educators and accounting firms to assess the relative outcomes of professional training programs. Unlike the DIT, which uses fixed vignettes, the MES and the Composite MES can be computed from different sets of similar vignettes to test for increases in sensitivity. Unlike the DIT, changes in the dilemmas are possible, and changes in the Composite MES can be easily recognized and measured, as well as changes in the MES philosophical areas. In this way, the Composite MES provides at least one alternative to the pre and post-test assessment of ethical training and interventions.

Notes

¹ There is no test for relative variation. Cochran just gives an example. The Medicare Sampling Guidelines used to investigate overpayments in 2000 used the coefficient of variation, the cv, as a guide to homogeneity of the sample data. If the cv (called tolerable error by Medicare) is over 12% for a sample, then a sample size under 100 is considered too small, as the data are not homogeneous enough.

² It can be argued that these are not three independent and different evaluations since they are obtained each from the same person. Our results, consistent with others, do find differences.

Appendix A: The eight Multidimensional Ethics Scale vignettes

(#1) A firm has been hard hit by recessionary times and the partners realize that they must scale back.

An analysis of productivity suggests that the person most likely to be terminated is a longtime employee with a history of absenteeism due to illness in the family.

Action: instead, the partner in charge lays off a younger, but very competent, recent hire.

(#2) A company has just introduced a highly successful new kitchen appliance. The sales manager, who is paid partly on a commission basis, discovers that there has been insufficient product testing to meet government guidelines. The tests so far indicate no likelihood of any safety problem.

Action: the sales manager continues to promote the product.

(#3) A manager of a company is eager to do more business abroad and has been requested to make an undisclosed cash payment to a local distributor in a foreign country. The payment is requested as a “goodwill gesture” that will allow the company to introduce its products in that foreign country. This practice is considered normal business procedure in that country, and no laws prohibit such a payment there.

Action: the manager verbally authorizes the payment.

(#4) The owner of a local small business, which is currently in financial difficulty, approaches a longtime friend to borrow and copy a proprietary database software package which will be of great help in generating future business. The software package retails for \$500.

Action: the friend loans the software package.

(#5) A manager realizes that the projected quarterly sales figures will not be met, and thus the manager will not receive a bonus. However, there is a customer order which if shipped before the customer needs it will ensure the quarterly bonus but will have no effect on the annual sales figures.

Action: the manager ships the order to ensure earning the quarterly sales bonus.

(#6) A promising start-up company applies for a loan at a bank. The credit manager at the bank is a friend of and frequently goes golfing with the Company’s owner. Because of this company’s short credit history, it does not meet the bank’s normal lending criteria.

Action: the credit manager recommends extending the loan.

(#7) A salesman, the father of two small children, has been promoted to a job in what he has to travel

away from home for the firm on regular basis. Because the trips are frequent and inconvenience his family life, he is contemplating charging some small personal expenses while traveling for the company. He has heard that this is common practice in the company.

Action: the salesman charges the company \$50 family gifts.

(#8) The CEO of a company requests the controller to reduce the estimate for bad debts in order to increase reported income, arguing that this is common practice in the industry when times are hard. Historically, the company made very conservative allowances for doubtful accounts, even in bad years. The CEO’s request would make it one of the least conservative in the industry.

Action: the controller makes the adjustment.

References

Abdomohammadi, A. and D. Ariail: 2007, *Selection-Socialization and the “Inverted-U” Phenomenon in Moral Reasoning of Accountants in Public and Industry Practice* (American Accounting Association 2007 Ethics Symposium, Chicago, IL).

Armstrong, M.: 1987, ‘Moral Development and Accounting Education’, *Journal of Accounting Education* **5**(Spring), 27–43. doi:[10.1016/0748-5751\(87\)90036-4](https://doi.org/10.1016/0748-5751(87)90036-4).

Armstrong, M.: 1993, ‘Ethics and Professionalism in Accounting Education: A Sample Course’, *Journal of Accounting Education* **11**, 77–92. doi:[10.1016/0748-5751\(93\)90019-F](https://doi.org/10.1016/0748-5751(93)90019-F).

Bernardi, R. A. and D. F. Arnold Sr.: 2004, ‘Testing the “Inverted U” Phenomenon in Moral Development on Recently Promoted Senior Managers and Partners’, *Contemporary Accounting Research* **21**(2), 353–367. doi:[10.1506/L5PE-4JXY-8GTB-CEQN](https://doi.org/10.1506/L5PE-4JXY-8GTB-CEQN).

Appendix B

- | | | | | | | | | |
|---|---------------------------------|---|---|---|---|---|---|---|
| 1. Unjust | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Just |
| 2. Unfair | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Fair |
| 3. Not morally Right | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Morally Right |
| 4. Not acceptable to my family | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Acceptable to my family |
| 5. Culturally Unacceptable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Culturally Acceptable |
| 6. Traditionally Unacceptable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Traditionally Acceptable |
| 7. Not self-promoting for me | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Self-promoting for me |
| 8. Not personally satisfying for me | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Personally satisfying me |
| 9. Produces the least utility | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Produces the greatest utility |
| 10. Minimizes benefits while maximizes harm | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Maximizes benefits while minimizes harm |
| 11. Violates an unwritten contract | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Does not violate an unwritten contract |
| 12. Violates an unspoken promise | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Does not violate an unspoken promise |
| 13. The probability that I would undertake the same action is: | High 1 2 3 4 5 6 7 Low | | | | | | | |
| 14. The probability that my peers would undertake the same action is: | High 1 2 3 4 5 6 7 Low | | | | | | | |
| 15. The action described above is: | Ethical 1 2 3 4 5 6 7 Unethical | | | | | | | |
| 16. Please specify why you feel this action is either ethical or unethical. | _____ | | | | | | | |

- Cochran, W.: 1963, *Sampling Techniques*, 2nd Edition (John Wiley & Sons, Inc., New York, NY), p. 54.
- Cohen, J. R., L. W. Pant and D. J. Sharp: 1993, 'Culture-Based Ethical Conflicts Confronting Multinational Accounting Firms', *Accounting Horizons* **7**(3), 1–13.
- Cohen, J. R., L. W. Pant and D. J. Sharp: 1996, 'Measuring the Ethical Awareness and Ethical Orientation of Canadian Auditors', *Behavioral Research in Accounting* **8**(Supp), 98–119.
- Cohen, J. R., L. W. Pant and D. J. Sharp: 1998, 'The Effect of Gender and Academic Discipline Diversity on the Ethical Evaluations, Ethical Intentions, and Ethical Orientation of Potential Public Accounting Recruits', *Accounting Horizons* **12**(3), 250–270.
- Cruz, C. A., W. E. Shafer and J. R. Strawser: 2000, 'A Multidimensional Analysis of Tax Practitioner's Ethical Judgments', *Journal of Business Ethics* **24**(3), 223–245. doi:10.1023/A:1006140809998.
- Duncan, J. R. and J. A. Knoblett: 2000, 'Investigating Behavioral Antecedents of Earnings Management: Moral Reasoning, Earnings Pressure, and Likelihood of Detection', *Research on Accounting Ethics* **6**, 27–66.
- Fisher, D. G. and J. T. Sweeney: 1998, 'The Relationship Between Political Attitudes and Moral Judgment: Examining the Validity of the Defining Issues Test', *Journal of Business Ethics* **17**, 905–916. doi:10.1023/A:1005737508336.
- Flory, S. M., T. J. Phillips, R. E. Reidenback and D. P. Robin: 1993, 'A Reply to a Comment on "A Multi-Dimensional Analysis of Selected Ethical Issues in Accounting"', *Accounting Review* **68**(2), 417–421.
- Hiltebeitel, K. and S. Jones: 1991, 'Initial Evidence on the Impact of Integrating Ethics into Accounting Education', *Issues in Accounting Education* **6**(2), 262–275.
- Israeli, D.: 1988, 'Ethical Beliefs and Behavior Among Managers: A Cross-Cultural Perspective', *Journal of Business Ethics* **7**, 263–271. doi:10.1007/BF00381831.
- Jones, S. K. and L. A. Ponemon: 1993, 'A Comment on "A Multi-Dimensional Analysis of Selected Ethical Issues in Accounting"', *Accounting Review* **68**(2), 411–416.
- Kaplan, J.: 2005, Political Theory: The Classical Texts and Their Continuing Relevance. *The Modern Scholar* by Recorded Books, LLC. <http://www.modernscholar.com>.
- Kohlberg, L.: 1969, 'Stage and Sequence: The Cognitive-Development Approach to Socialization', in D. Goslin (ed.), *Handbook of Socialization Theory, Research* (Rand McNally College Publishing Company, Chicago, IL).
- Medicare Carriers Manual, Part 3 (HCFA-Pub.14–3): 2000, 'Sampling Guidelines', <http://www.cms.hhs.gov>, pp. 1–38. Accessed 31 August 2000.
- Paulhus, D. L.: 1991, 'Measurement and Control of Response Bias', in J. P. Robinson, P. R. Shaver and L. S. Wrightsman (eds.), *Measures of Personality and Social Psychological Attitudes*, 1st Edition (Academic Press, San Diego, CA), pp. 17–59.
- Piaget, J.: 1965, *The Moral Judgment of a Child* (Free Press, New York, NY).
- Ponemon, L. A.: 1988, A Cognitive Developmental Approach to the Analysis of Certified Public Accountants' Ethical Judgments. Unpublished Doctoral Dissertation, Union College, Schenectady, NY.
- Ponemon, L. A.: 1993, 'Can Ethics Be Taught in Accounting?', *Journal of Accounting Education* **11**, 185–209. doi:10.1016/0748-5751(93)90002-Z.
- Ponemon, L. A. and D. Gabhart: 1993, 'Ethical Reasoning in Accounting and Auditing', Research Monograph Number 21 (CGA-Canada Research Foundation, Vancouver, BC).
- Reidenbach, R. and D. Robin: 1990, 'Toward the Development of a Multidimensional Scale for Improving Evaluations of Business Ethics', *Journal of Business Ethics* **9**, 639–653. doi:10.1007/BF00383391.
- Rest, J. R.: 1979, *Development in Judging Moral Issues* (University of Minneapolis Press, Minneapolis, MN).
- Rest, J. and D. Narvaez: 1994, *Moral Development in the Professions* (Lawrence Erlbaum Associates, Mahwah, NJ).
- Rest, J. R., D. Narvaez, M. J. Bebeau and S. J. Thoma: 1999, *Postconventional Moral Thinking: A Neo-Kohlbergian Approach* (Lawrence Erlbaum Associates, Mahwah, NJ).
- Scofield, S. B., T. J. Phillips Jr. and C. D. Bailey: 2004, 'An Empirical Reanalysis of the Selection-Socialization Hypothesis: A Research Note', *Accounting, Organizations and Society* **29**(5&6), 543–563. doi:10.1016/S0361-3682(03)00021-7.
- Sennetti, J., T. Shawver and P. Bancroft: 2004, 'The Moral and Cultural Reasoning of IPO Accountants, Those Assumed Likely to Manage Earnings: A Small Sample Study', *Research on Professional Responsibility and Ethics in Accounting* **9**, 101–128.
- Shaub, M.: 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), 1–25. doi:10.1016/0748-5751(94)90016-7.
- Shawver, T. and L. Clements: 2007, 'The Intention of Accounting Students to Whistleblow in Situations of Questionable Ethical Dilemmas', *Research on Professional Responsibility and Ethics in Accounting* **11**, 177–191.
- Shawver, T., J. Sennetti and P. Bancroft: 2006, 'Can the "Clan Effect" Reduce the Gender Sensitivity to Fraud? The Case of the IPO Environment', *Journal of Forensic Accounting* **VII**, 185–208.
- Shawver, T. J. and T. A. Shawver: 2008, 'Will Corporate Governance and Whistleblowing Provisions Improve

- Financial Responsibility?', *Research on Professional Responsibility and Ethics in Accounting* **12**, 123–148.
- St. Pierre, K., E. Nelson and A. Gabbin: 1990, 'A Study of the Ethical Development of Accounting Majors in Relation to Other Business and Non-Business Disciplines', *The Accounting Educators' Journal* **6**(Summer), 23–36.
- Sweeney, J. and D. Fisher: 1998, 'An Examination of the Validity of a New Measure of Moral Judgment', *Behavioral Research in Accounting* **10**, 138–158.
- Thornton, J. M.: 2000, 'Challenges to the Defining Issues Test: A New Perspective on Accountants' Moral Development', *Research on Accounting Ethics* **7**, 225–252.
- Weber, J.: 1996, 'Welcoming Another CMD Instrument – The MES: But don't Throw Out the MJI or the DIT Just Yet!', *Business Ethics Quarterly* **6**, 517–522.
doi:[10.2307/3857502](https://doi.org/10.2307/3857502).

Tara J. Shawver
King's College,
Wilkes-Barre, U.S.A.
E-mail: tarashawver@kings.edu

John T. Sennetti
Nova Southeastern University,
Fort Lauderdale (Davie), U.S.A.

Copyright of *Journal of Business Ethics* is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Copyright of Journal of Business Ethics is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.