This is the authors’ final peer reviewed version of the item published as:


Available from Deakin Research Online: http://hdl.handle.net/10536/DRO/DU:30006516

Reproduced with the kind permission of the copyright owner.

Copyright : 2006, AFAANZ
MEASURING MORAL JUDGEMENT AND THE IMPLICATIONS OF
COOPERATIVE EDUCATION AND RULE-BASED LEARNING

Steven Dellaportas *
La Trobe University

Barry J Cooper
RMIT University

Philomena Leung
Deakin University

*corresponding author
Steven Dellaportas
Department of Accounting and Management
School of Business
La Trobe University
Victoria, Australia 3086
Ph: 61 3 9479 1471
Fax: 61 3 9479 5971
E-mail: s.dellaportas@latrobe.edu.au.

Classification code: M41 Accounting.

Key words: Defining Issues Test, Welton instrument, dilemmas, cooperative education, rule-based learning.
MEASURING MORAL JUDGEMENT AND THE IMPLICATIONS OF
COOPERATIVE EDUCATION AND RULE-BASED LEARNING

ABSTRACT

The Defining Issues Test (DIT), developed by Rest (1986), measures a person’s level of moral development using hypothetical social dilemmas. While the DIT is useful for measuring moral development in social settings, it may not adequately capture an individual’s moral judgement abilities in solving work-related problems (Weber, 1990; Trevino 1992; Welton et al., 1994). In this study, the moral judgement levels of 97 accounting students were measured over a one year period using two separate test instruments, the DIT and a context-specific instrument developed by Welton et al. (1994). The test scores are significantly higher on the DIT than the Welton instrument (between the instruments and over time), suggesting that accounting students use higher levels of moral reasoning in resolving hypothetical-social dilemmas and lower levels of moral reasoning in resolving context-specific dilemmas. The difference in test scores was highest during cooperative education (work placement program), implying that the environment is a significant determinant on students’ test scores.
1. INTRODUCTION

The Defining Issues Test (DIT) is a self-administered questionnaire, developed by Rest (1986) that elicits a subject’s level of moral reasoning. The DIT provides a useful contribution to the understanding of an individual’s moral reasoning abilities based on dilemmas derived from social issues. However, one limitation of the DIT is that it may not accurately reflect an individual’s moral reasoning in a professional setting. The DIT is an accurate measure of moral reasoning in dealing with personal issues (Fraedrich et al., 1994) but a context-specific instrument may work better in understanding an individual’s moral reasoning in professional settings (Welton et al., 1994).

Welton et al. (1994) applied a self-designed test instrument comprising accounting-specific issues to measure the effect of ethics education on the moral reasoning levels of graduate accounting students. Test results reveal that students receiving ethics instruction demonstrated gains in moral reasoning within the decision context. However, Welton et al. (1994) did not simultaneously use the DIT and, therefore, were unable to provide comparative analysis between the scores derived from the DIT and the Welton instruments. Consequently, whether accounting students resolve context-specific dilemmas the same way they resolve social dilemmas is still unclear. The primary aim of this paper is to address this limitation and compare accounting students’ moral judgements using two separate test instruments, the DIT, which contains hypothetical dilemmas on broad life issues, and the context-specific instrument developed by Welton et al. (1994), which contains realistic dilemmas particular to business and accounting. The second aim of this paper is to explore the potential the effects of one year of formal
accounting education and one year of accounting related work experience (cooperative education) on the moral development of accounting students. A comparative analysis of the test scores between the DIT and Welton instruments will add evidence to the debate on whether accounting students reason at their moral capacity in resolving context-specific ethical dilemmas, and may explain in part how accountants respond to work-based dilemmas. Furthermore, the findings in this paper will determine the extent and immediacy, if any, of situational factors (scenarios in the Welton instrument and students’ cooperative education) on cognitive moral reasoning in ethical decision-making. The literature, particularly the popular press, provide numerous examples of unethical behaviour in the accounting profession. They range from individual tax fraud to serious breaches of professional and ethical conduct. However, rather than blame such offences on the ethics of the individual accountant, questionable acts may be context-specific. The results in this study may help us to better understand how people respond to problems in professional settings.

The remainder of this paper is structured as follows. The next section describes the theory of moral reasoning and development. The third section explains the process by which an individual’s moral reasoning and development is measured. The fourth section further examines the relevant literature and establishes the hypotheses. The fifth section displays the results and the final section discusses the findings followed by the conclusion.
2. THE THEORY OF MORAL REASONING AND DEVELOPMENT

Kohlberg’s (1969) theory of cognitive moral reasoning and development (CMD) centres on how one’s belief system dictates conflict resolution and problem solving in everyday life. Kohlberg’s theory posits that individuals have identifiable cognitive skills that are used to resolve ethical dilemmas. These skills, determined by the reasons given about why certain actions are perceived as morally just or preferred, were used by Kohlberg to provide a hierarchical continuum of six stages of cognitive moral development, with each successive stage representing a higher level of reasoning regarding the definition and nature of right and wrong. The rationale used by individuals to resolve ethical dilemmas, displays characteristics that enable the researcher to classify people according to a particular stage of moral development (Weber, 1990). Much has been written on Kohlberg’s stages of moral development (see Weber, 1990; Lovell, 1995; Fisher, 1997); it is not the purpose of this paper to elaborate upon them unduly. However, some description of Kohlberg’s stages of moral development is useful to understand the theoretical background and significance of this study.

Kohlberg’s theory consists of three levels and within each level there are two distinct stages of moral development. The pre-conventional level, which embraces stages 1 and 2, reflects a level of moral reasoning that is exclusively self-centred. A person reasoning at these early stages complies with rules and social expectations only when personal consequences can be avoided (stage 1) and a personal benefit is to be derived (stage 2). At the conventional level, the focus is on maintaining relationships and the notion of living within a community assumes increasing significance. In this level, relationships
begin with acting to please one’s peers and superiors (stage 3), to a course of action defined by compliance to categorical rules such as laws and regulations (stage 4). The post-conventional, or principled level, reflects a growing moral autonomy as defined by self-determined, but not selfish moral reasoning. The focus shifts from strict compliance with formal rules, to embracing notions of a social contract (stage 5) and universal principles founded on deontological ethical theories of justice, duties and equal human rights (stage 6). Overall, the individual’s moral judgement grows less and less dependent on outside influences with each successive stage of moral reasoning.

Empirical evidence in accounting ethics research have discovered associations between moral development and independence judgements (Ponemon and Gabhart, 1990); the propensity of internal auditor’s to blow the whistle (Arnold and Ponemon, 1991); the ability to detect fraud (Bernardi, 1994); the ability to resist management pressure (Sweeney and Roberts, 1997; Windsor and Ashkanasy, 1995); and political ideology (Fisher and Sweeney, 1998). However, comparative analyses have regularly shown that the test scores for sample groups of accountants and accounting students are lower than the scores for other comparative groups, including college graduates and other professional groups (Gaa, 1994; Ponemon and Gabhart, 1994). Accountants often demonstrate an orientation towards conventional reasoning that emphasises the application of laws and rules. Lampe and Finn (1992) assert that the internalisation of GAAP has instilled a stage four reasoning that is predominant in accountants’ moral reasoning.
3. MEASURING COGNITIVE MORAL DEVELOPMENT

While Kohlberg’s primary interest was to devise a theoretical system to represent the logic of moral thinking, Rest (1986) developed the DIT, a valid, objective, reliable measurement instrument of CMD based on the six stages of moral development, as defined by Kohlberg (1969). The DIT is a self-administered questionnaire that elicits the subject’s level of moral development. To test moral judgement, subjects are asked to read six hypothetical moral dilemmas and then select from a series of “issues” or questions called “items for consideration” that the subject believes is critical to the resolution of the ethical dilemma. The DIT is based on the premise that people who are at different points of moral development interpret moral dilemmas differently, define the critical issues of the dilemma differently, and have their own intuitions about what is right and fair in a situation. Therefore, differences in the way dilemmas are recognised and the relative importance given to “items for consideration”, are taken as indications of the subject’s underlying cognitive capacity.

The DIT asks the individual to rate and rank the importance of “items” that represent different stages of moral reasoning. The ranking of the four most important issues enables the researcher to determine a developmental score of CMD known as the P-score. The P-score reflects the number of times that an item associated with stage 5 or 6 is chosen as an important item in determining a course of action in resolving a moral dilemma (Rest, 1986). In effect, the P-score represents the probability of a post-conventional (stages 5 and 6) response to a moral problem. The P-score is a continuous variable ranging from zero to one and expressed as a percentage of 100. A distinct advantage of a continuous
variable such as the P-score, is the ability to identify changes in moral reasoning by measuring and comparing differences in test scores before and after treatment effects. A change in the P-score implies real moral development.

The Welton instrument, which parallels the DIT in design, comprises four business and accounting dilemmas: (1) the ‘Reimbursement’ dilemma considers whether a finance officer should approve an expense reimbursement claim by his superior, knowing that the claim has been exaggerated; (2) the ‘Bankruptcy’ dilemma considers whether a banker should approve a loan application knowing that the applicants were former bankrupts; (3) ‘The Bank Audit’ dilemma considers whether a bank should increase its loan loss reserve at a time when it could hinder efforts to raise equity capital through a planned share issue; (4) ‘The Opinion’ dilemma considers whether an auditor should issue a qualified opinion due to weaknesses in the client’s computerised internal control structure implemented by the same audit firm.¹

4. HYPOTHESES DEVELOPMENT

4.1 Hypothetical-social and context-specific dilemmas

Realistic case dilemmas are more likely to elicit representative reasoning processes because they are seen to generate interest and therefore involvement by the respondent (Fredrickson, 1986). “Story pull” occurs when an individual’s familiarity and experience with a particular dilemma, “pulls” or “elicits” at a particular stage response (Rest, 1986). The extent of story pull depends upon the subject’s ability to associate with the character in the dilemma or the nature of the conflict, possibly due to the occurrence of a similar
personal experience (Elm and Weber, 1994). Weber (1991) argues that remoteness is an important influence on moral reasoning. Tolleson et al. (1996) added an element of remoteness to the “Heinz and Drug” dilemma in the original DIT, by changing the key actors from a husband and wife relationship to strangers. The authors concluded that remoteness had a significant impact on moral reasoning and behaviour. As strangers, respondents were more detached or less willing to become involved in a dying person’s plight. Similarly, Kenny and Eining (1996) rely on attribution theory, which posits that individuals behave differently when they are actors, as opposed to observers in the ethical dilemma. They discovered that accounting students apply a different standard of ethical behaviour in relation to others than they hold for themselves.

The dilemmas contained in the DIT deal in broad life issues, which may not fairly represent professional environments (Weber, 1990; Elm and Nichols, 1993; Elm and Weber, 1994; Ponemon, 1990, 1993; Trevino, 1992; Welton et al., 1994). The dilemmas comprise a wide variety of moral issues ranging from discontinuing a newspaper for its disturbing social influence to stealing a drug in order to save the life of one’s spouse. However, the DIT does not contain dilemmas of a business or professional nature, which may cause subjects to abandon their professional role in favour of their role as a member of society. For example, Weber (1990) investigated the moral reasoning abilities of managers using three dilemmas contained in the DIT. Two of the dilemmas were altered

---

1 The Welton instrument was selected for this study because it was the only context-specific instrument published at the time of data collection for this study.
to reflect dilemmas that are typically faced in a corporate context. Managers subsequently reasoned at lower levels on these two dilemmas compared with the original and broader unaltered social dilemma. Elm and Nichols (1993) caution against using the DIT in assessing managers’ moral reasoning when addressing business-related ethical dilemmas. An alternative test instrument containing dilemmas of a business nature provides a more systematic assessment of how decisions about ethical dilemmas are dealt with at work (Wyld et al., 1994).

In the field of accounting, Ponemon (1993) argues that the resolution of hypothetical dilemmas of a social nature may be too crude a measure to capture accounting students’ moral reasoning. Ponemon suggests that an instrument based on accounting and auditing dilemmas provides a more context specific measure of moral reasoning. The solution by some researchers has been to develop a context-specific test instrument containing dilemmas particular to the profession (Welton et al., 1994; Fisher, 1997; Thorne, 2000). Context-specific instruments are based on the DIT and thus contain the same psychometric properties. The difference in test instruments, compared with the DIT, is in the nature of the dilemmas and related “items for consideration”. Welton et al. (1994) suggest that accounting-specific dilemmas provide a better measure of moral reasoning from a professional perspective, because they are more likely to elicit stage responses that reflect the level of reasoning used in resolving accounting dilemmas. The primary

---

2 The DIT comprises six dilemmas: 1. Heinz and the Drug – examines whether Heinz should steal a drug that might save the life of his wife who is dying from cancer; 2. Student Take-Over – examines university students’ freedom of speech and their right to protest; 3. Escaped Prisoner – examines whether a man should pay for a past crime after living 8 years of a virtuous existence that contributed to the well-being of the local community; 4. The Doctor’s Dilemma – examines the issue of euthanasia; 5. Webster – examines discrimination against minorities; and 6. Newspaper – examines freedom of speech as it relates to the press.
objective of this study is to determine whether accounting students use different levels of moral reasoning in resolving realistic professional dilemmas (Welton instrument) compared with hypothetical-social dilemmas (DIT). Welton et al. (1994) and others (see Trevino, 1986; Weber, 1990) contend that moral judgements on realistic professional dilemmas are not expected to score as high as moral judgements on hypothetical-social situations. In fact, the test scores on instruments comprising business (Weber, 1990) and accounting (Thorne, 2000, 2001) dilemmas were lower than the test scores based on the DIT. Ponemon (1990) contends that the DIT is based on conflicts in human life, which are on a higher plane than client needs. Therefore, test scores elicited by the DIT should at least be equal or higher than the test scores elicited from a context-specific instrument. The first hypothesis is thus stated:

**H1** The mean DIT P-score will be significantly higher than the mean Welton P-score for the accounting students being examined.

4.2 Length of formal education

The second objective of this paper is to determine whether additional explanatory variables (formal education and cooperative education) have an effect on students’ test scores. Cognitive development research shows that without significant intervention, or an appropriate environment, the majority of the adult population will never surpass the conventional level (stages 3 and 4) of Kohlberg’s model (Kohlberg, 1971). One type of intervention conducive to moral development is formal education. Research has repeatedly shown that moral development is highly correlated with education levels.
(Blasi, 1980; Rest 1986; Rest 1988). Increasing DIT P-scores in post secondary education suggests that formal education is a major factor transforming moral judgment structures. As the Welton instrument is a measure of moral development with the same psychometric properties as the DIT, advances in the moral development of accounting students should be reflected in the test scores in both the DIT and Welton instruments. The second hypothesis is thus stated:

**H2** The DIT and Welton P-scores at the end of the year for the accounting students being examined will be significantly higher than the P scores of the respective instruments at the beginning of the year.

### 4.3 Cooperative Education

The degree program from which the students were selected to participate in this study is a four-year program comprising two years of academic study, one year of cooperative education, followed by the final year of academic study. Cooperative education is an industry placement program in which students work in industry and public accounting firms to gain first-hand work experience and apply the theory learned in the first two years of the degree program. Cooperative education students are full-time employees and subject to the same terms and conditions as other employees of the organisation in which they work. The required status of cooperative education makes the degree program unique to the Australian education environment.

Elm and Nicholls (1993) suggest that the organisational context in which one works can contribute to moral development and that its influence could be in the direction of either
higher or lower levels of moral reasoning skills, depending on the individual’s self
monitoring propensity and the actual experiences of the organisation. Similarly, Jones
and Hiltebeitel (1995) contend that moral attitudes change as new accountants assimilate
the culture of the profession and their behaviour can be influenced by the conduct of
superiors and accepted industry practices. Socialisation is therefore a process of
adopting the “official position” rather than giving individual judgement (Bebeau et al.,
1985).

The early work of Arlow and Ulrich (1980) noted the potential impact of the
socialisation-effect on moral judgements. They found that ethics education improves the
ethical awareness of business students, but the effect is not persistent. A follow-up study
four years later found that test scores reverted to pre-test levels. Arlow and Ulrich (1985)
suggest that the decline in test scores is accounted for by the socialisation process
students undergo as they move into the world of business. Welton et al. (1994) similarly
discovered a positive movement in test scores for a group of students exposed to ethics
instruction, but the effects of ethics intervention were transitory, with test scores
returning to pre-test levels after a six-month interval (La Grone et al. 1996). Cooperative
education exposes accounting students to a maximum of one-year of practical
accounting related work experience. Considering the duration that students are exposed
to the work environment and the temporary status of their employment, the socialisation
process is not expected to have a significant effect on test scores for the sample of
accounting students in this study. Thus the following hypothesis is stated in the null
form:
H3: Cooperative education does not have a significant impact on the DIT and Welton P-scores for the accounting students being examined.

The hypotheses will be analysed by comparing the test scores between the instruments and the pre- and post-test scores within each instrument. This analysis has multiple purposes. First, to determine whether accounting students’ resolve context-specific dilemmas in the same way they resolve hypothetical-social dilemmas (H1). Second, to determine whether the moral reasoning levels of accounting students improve after one year of formal accounting education, and if so, whether their ability to resolve context-specific dilemmas improves to the same extent as their ability to resolve hypothetical social dilemmas (H2). Finally, to examine the potential influence of the organisational work environment on students’ moral reasoning and development (H3).

5. DATA COLLECTION

The long version of the DIT comprises six dilemmas and, according to Rest (1986), it can take up to 45 minutes to complete, while the short version comprising three dilemmas can take up to 35 minutes to complete. Thus, the short version is not necessarily more time efficient. It was discovered during the pilot phase of this study that many students required a minimum of one hour to complete the six-dilemma version of the DIT. To overcome the problems with administering two test instruments that could up to two hours to complete, the short version (three dilemmas) of the DIT was
combined with the short version (three dilemmas) of the Welton instrument to form the equivalent of one six-dilemma test, consistent with Rest’s (1979) original DIT. Limiting the time required to complete the test instrument will minimise the number of instruments that will fail the internal consistency checks, due to time and fatigue.

The short version (three dilemmas) has similar characteristics to the long form (six dilemmas), but the major drawback is that fewer dilemmas result in less reliability. However, Rest (1986) has shown that the three dilemmas consisting of Heinz, Prisoner and Newspaper, have the highest correlation of P-scores, 0.93 of any three-dilemma set from the full six-dilemma set. The short version of the DIT comprising the dilemmas as recommended by Rest (1986) is utilised in this study. Welton et al. (1994) developed an accounting-specific test instrument containing four dilemmas. After careful consideration of all four dilemmas, one dilemma was omitted to make it consistent with the short (three-dilemma) version of the DIT. This dilemma entitled “The Opinion” was omitted, because the critical issues in this dilemma rely on an understanding of auditing concepts that may be unfamiliar to entry-level accounting students.

Students from all year levels of the degree program were invited to participate in this study (see Table 1 for an outline of the sample description). The sample comprises 97 full-time volunteer accounting students ranging in age from 18 to 22 (mean 21.04 years) with adequate gender representation (55 females and 42 males). The resulting questionnaires were organised into four groups based on the students’ year of enrolment with each group comprising students from different year levels. The number of students
enrolled in each group is: Year 1, 18; Year 2, 20; Year 3, 12; Year 4, 47. Students completed the test instrument three times throughout a one year period; the beginning of the year (pre-test), the middle of the year (mid-test), and at the end of the year (post-test). During this process, students were asked to mark the test instrument with a self determined identifier so that the test instruments could be matched for the purpose of paired sample testing. According to Rest (1979), the effects of dilemma familiarity from re-testing the same subjects with the same dilemmas are negligible. Test-retest studies with an interval of one to three weeks, show that the effect on mean P-scores is insignificant (Davison and Robbins, 1978). However, changes in P-scores, if any, beyond the three week period are related to the time between the tests and presumably “real moral development”, rather than simply retaking the test. Therefore, Rest (1986) recommends using the same dilemmas.

**Place Table 1 about here**

The DIT contains reliability and consistency checks that prevent subjects from artificially inflating their test scores. A completed instrument that fails either of these checks is rendered unusable and withdrawn from the sample. The Welton instrument is based on the same psychometric properties as the DIT and, therefore, contains the same reliability and consistency checks that apply to the DIT. Completing the DIT and Welton instruments three times throughout the year (3 tests x 97 students) resulted in 291 Welton instruments and 291 DITs. The loss rate (test instruments invalidated due to the reliability and consistency checks: 39 DIT and 16 Welton) combined with the attrition
rate (did not complete: 36 DIT and 36 Welton), resulted in a total of 239 useable Welton instruments and 216 useable DITs (see Table 2). The useable DIT and Welton instruments were then matched using the identifying mark inscribed by the students. The matching process resulted in a further eight DITs and 21 Welton instruments being removed from the sample because they could not be matched with a corresponding test instrument. This process resulted in a total of 208 paired samples.

Place Table 2 about here

6. FINDINGS

The mean DIT P-score for the entire sample of accounting students in this study (208 matched pairs) is 3.37 percentage points higher than the mean Welton P-score (see Table 3). The repeated sample $t$-test reveals that the difference in P-scores is statistically significant, $t(207) = 2.40$, $p \leq .05$, one-tailed, suggesting that accounting students score higher on the DIT than the Welton instrument. Furthermore, the Pearson correlation indicates that the relation between the DIT and Welton P-scores is weak, $r=0.31$, $n=208$, $p>.01$. Two interpretations may result from this finding. First, the Welton instrument is not an accurate measure of moral reasoning compared to the DIT, or second, subjects use different levels of moral reasoning in analysing accounting-specific dilemmas compared with hypothetical-social dilemmas. The authors of the Welton instrument report a Chronbach alpha of 0.62 suggesting high internal validity particularly when it is compared with similar studies that report lower Chronbach alphas. Additionally, Mary Beth Armstrong, a noted researcher in the field of moral cognitive development research,
independently validated the wording of the instrument and the scoring scheme (Welton et al., 1994, p.40). Assuming the Welton instrument is an accurate measure of moral reasoning in a professional context, the findings in this study support the first hypothesis and it is concluded that accounting students appear to apply different ethical considerations in resolving context-specific dilemmas (such as the dilemmas contained in the Welton instrument) and hypothetical-social dilemmas (such as the dilemmas contained in the DIT).

The second hypothesis is analysed by comparing the pre- and post- test scores for the DIT and Welton instruments. The means, standard deviations and \( t \)-statistics are reported in Table 3. The DIT P-scores increase progressively throughout the year with a significant increase in post-test scores (post-test – pre-test: 6.09 percentage points), \( t(136) = 1.88, p<.05 \), one tailed. However, the Welton P-scores remain static throughout the year and the difference in the pre- and post-test scores (0.19 percentage points) is not significant, \( t(136) = 0.07, p>.05 \), one tailed. This finding is contrary to Welton et al. (1994) who found a positive movement in test scores. The increasing difference between the DIT and Welton P-scores, which is statistically significant at the end of the year (post-test), occurs because the DIT P-score increases with time but the Welton P-score does not. The findings in this study suggest that students are capable of improving their moral capacity as measured by the DIT P-score but are not demonstrating the same improvement in the Welton P-scores. Thus, the second hypothesis is partially rejected and it is concluded that only the DIT P-scores increase after one year of accounting education but not the Welton P-scores.
The third hypothesis is tested by clustering the test instruments into four groups based on the students’ year of enrolment. The purpose of this analysis is to determine if the P-scores of any one group, particularly those undertaking cooperative education, displayed a significant difference between the test instruments. The means, standard deviations, and the results of the repeated sample *t*-tests for each group are presented in Table 4 below. The mean DIT P-scores recorded in this study are higher than the Welton P-scores in all four groups but the size of the difference is greatest during the second and third years of the degree program. A comparison of the P-scores between the groups reveal that the difference in P-scores is not significant for years 1 and 4, however, the difference in P-scores is significant in years 2 and 3. It appears that there was a significant change in the P-scores of the period during which accounting students were either preparing for or undertaking cooperative education, with the largest difference occurring during the actual work placement period.

### 7. COOPERATIVE EDUCATION AND RULE-BASED LEARNING

The increase in DIT P-scores after one year of accounting education is encouraging and consistent with empirical research, which suggests that P-scores increase with years in formal education. However, the Welton P-scores remained static throughout the year and
do not compare favourably with the DIT P-scores. This finding is consistent with other studies, which suggest that accountants and managers employ lower levels of moral reasoning in the resolution of context-specific dilemmas and higher levels of moral reasoning in the resolution of context-free dilemmas (Weber, 1990; Thorne, 2000, 2001). In accounting, studies using the DIT as the measure of analysis typically report that accountants and accounting students do not develop moral reasoning capacities commensurate with individuals having similar social, economic, or educational backgrounds (Gaa, 1995). In other words, accounting professionals lag behind their counterparts in other disciplines (Puxty et al., 1994). If accountants lag behind their counterparts based on DIT evidence, the results in this study have greater significance because the gap in test scores may be larger using a context-specific instrument. The question now facing accounting researchers is why accountants exhibit lower scores on the context-specific instrument than the DIT when the construct is arguably the same. This paper does not attempt to determine causal links between accounting education and the test scores, but it explores possible reasons for the findings.

7.1 Cooperative education

It was reported above that the Welton P-scores are significantly lower during the middle stages of the degree program, particularly cooperative education (Year 3) where the difference is largest. According to Thorne (2001), cooperative education students do not appear to resolve ethical dilemmas at their cognitive moral capacity. Thorne (2000, 2001) suggests that cooperative education students, like accountants, respond to social factors when formulating an ethical judgement and respond to self-interest in the
exercise of professional judgement. Reall et al. (1998) provide an interesting analysis upon which to elucidate the effect of the work environment on P-scores. In the spirit of a competitive game situation, students took off their “moral reasoning caps” and put on their “game faces”. Their game faces involved exhibiting moral reasoning levels that were well below their capability based on their DIT P-scores. Students used game rules to define appropriate boundaries of acceptable conduct and in effect used lower levels of moral reasoning than their moral capability when taking part in the spirit of a competitive game. In this study, the students’ DIT P-scores, a measure of their moral capacity, are significantly higher than their Welton P-scores; therefore, it may be inferred that accounting students do not always resolve moral dilemmas using their most principled reasoning. It appears short-term contextual factors, like the work environment, may adversely affect Welton P-scores.

Ponemon (1990) suggests that the self-selection or selection-socialisation process within the public accounting profession may be the cause of a net decrease in measured moral cognition with increasing levels of seniority in the firm hierarchy. Selection-socialisation occurs when senior management promote like-minded employees and in turn, employee accountants unwittingly adopt the “culture” of the firm in the hope of furthering their careers within the referent culture. Eventually, employee attitudes change and resemble the culture of the profession as they progress to senior ranks. In a study of auditor’s decision processes, Lampe and Finn (1992) found that the socialisation process appears to be present in the first two years of public accounting experience. With rapid turnover in the first two years of employment, entry level auditors self-select and employers counsel
out auditors with too high, or too low, a level of moral development. The socialisation-effect is not new, but where it was assumed to occur with time, the cooperative education students in this study suggest that the socialisation-effect is immediate — within one year.

The accounting students in this study who reason at the conventional level are more likely to accept the norms of the referent culture and maximise their opportunities for career advancement. This is particularly so in large accounting firms which have strong socialisation processes that are designed to recruit, reshape, and retain graduate accountants that fit into a hierarchical team-orientated structure that promotes conformity (Wilkinson et al. 2003). Therefore, lower Welton P-scores may be the result of an organisational culture that typifies reasoning at stages lower than the student’s own stage of development. The socialisation process combined with a desire for success — to obtain full-time career positions with the student’s respective employers at the completion of the degree program — may unwittingly lead students to abandon their personal values (Ludwig and Longenecker, 1993) as measured by the DIT, for a successful career. In general, it appears accounting students will operate at different levels of moral reasoning in different contextual environments. An accounting student may be capable of exercising considerable ethical maturity in the freethinking university environment, but may learn not to do so in a context-specific environment.
7.2 The potential effects of rule-based learning

A growing body of research suggests that accounting education may actually inhibit student’s progression to higher levels of moral reasoning and ethical awareness because of the emphasis on rule-based thinking (McPhail, 1999, 2001). Rule-based thinking is the inadvertent adoption of principles or rules established by another (Kidder, 1995). An increasing emphasis in accounting education on technical competence or rule-based learning and less concern with the broader questions of human values and morality has resulted in students being trained rather than educated (McNeel 1994, Blundell and Booth 1988). Accounting courses are technical and rule-based, they focus on professional examinations, they do not to deal with ‘values, ethics and integrity’, and they fail to develop students’ critical thinking skills (Albrecht and Sack (2000, pp. 51-56). The implication for accounting students is that they overlook ethical issues because they focus too much on technical issues (Bebeau et al., 1985). Teaching standards and regulations throughout the degree program arguably leads students to concentrate on rules and technical correctness without discursive analysis, limiting their understanding of the issues involved. Teaching that emphasises conformance to regulations rather than the underlying ethical issues could promote a rigid understanding of professional responsibilities that is devoid of professional judgement. Consequently, students are trained in “hard and fast rules” at the expense of critical thinking abilities resulting from an overemphasis on technical matters (Lampe 1996, Welton et al., 1994).

This rule-based approach to accounting education and practice reflecting the internalisation of GAAP may have inadvertently instilled a stage 4 (conformance to
laws) orientation in the students’ moral reasoning abilities. Rest’s (1986) P-score is a measure of the extent to which a subject consistently gives high importance to stage 5 and stage 6 “items”. Therefore, advances in moral development up to and including stage 4 orientations will not elevate students’ P-scores even if moral development has occurred. Only when accounting students are taught to acquire stage 5 and 6 moral reasoning abilities will subjects improve their test scores. Arguably, rule-based learning has become so ingrained in the accounting students’ mindset that they may be stuck at stage 4 and unable to develop to higher stages of moral reasoning. Thus the lower Welton P-scores in years 2 and 3 of the degree program have been influenced by accounting education or a work environment that emphasises rule-based thinking.

The DIT P-scores were significantly higher than the Welton P-scores in years 2 and 3 of the degree program, but were not dissimilar in years 1 and 4. Yet, rule-based learning exists in all stages of the degree program. The pre-test scores for entry level students (year 1) in this study indicate that there is no difference in the way students resolve social and accounting or business related dilemmas (see Table 3). The first indication of a difference in the DIT and Welton P-scores appear in the second year of the degree program. It appears that the internalisation of rule-based learning requires at least one year of post-secondary education before its effects are reflected in test scores. Alternatively, limited exposure (up to one year) to rule-based learning, does not affect post-test scores. Further cognitive development research could explore the longitudinal effects of accounting education on students’ moral development.
If the potential effect of cooperative education on P-scores are set aside, the increasing difference in test scores between year 2 and year 3 may be indicative of the continuing effects of rule-based learning. However, in year 4, the difference in DIT and Welton P-scores are not statistically different, in spite of the rule-based learning that continues in this stage of the degree program. The intervention that may counter the continuing effects of rule-based learning is the inclusion of a dedicated and required course in ethics in the final year of the degree program undertaken by the students being examined. The ethics course is deliberately designed to counter the criticisms of rule-based learning and sensitise students to ethics in accounting and conflict resolution methods. The similarity in test scores between the DIT and the Welton instrument in year 4, and the decline in the growing gap experienced in the earlier years of the program, may be the result of the direct curricula approach to ethics education. After studying a discrete course in ethics, students are sensitised to ethics in accounting and are more aware of the ethical issues in the dilemmas posed in the Welton instrument and their professional and ethical responsibilities. Future research should investigate the effects of ethics education in accounting programs and in particular, a discrete ethics course in accounting.

8. CONCLUSION

The evidence in this study suggests that context-specific-dilemmas, through story pull, familiarity, cooperative education and rule-based learning, exert sufficient influence that prevents students from demonstrating their moral maturity as measured by the DIT. The significant increase is post-test DIT P-scores indicates that accounting students can progress to higher levels of moral reasoning but remain capable of expressing moral decision-making rationale at lower stages as well. Therefore, accounting students who
have developed the ability to reason at the post-conventional stage may not always advocate a solution consistent with that stage.

If one goal of accounting education is to develop in students an awareness of the role and impact of accounting on third parties and society in general, then the findings in this study have important implications for the way accounting is taught. The technical focus of accounting education using rule-based learning should be de-emphasised and replaced with an appreciation of the role of accounting in the political, economic and social contexts. The direct curricula approach to ethics education may be one way to achieve this goal. The findings in this study have implications for the accounting profession as well as accounting education. A professional accountant is relied upon and trusted by the general public in respect of their accounting affairs. In return for this trust, accountants have a fundamental duty of care to protect the interest of their client, employer and the public. However, the findings suggest that accountants become baffled when ideals conflict and are likely to resort to lower stages of moral reasoning for resolving conflict. For users of accounting information, this means accountants may overlook their third party obligations in favour of self-interest. The strength of the socialisation process and the potential influence of accounting firm cultures that emphasise revenue and short-term earnings (Wyatt, 2004) means that we must now question whether the value systems of accounting professionals are strong enough to withstand client and economic pressures, which could compromise ethical judgement (Douglas et al., 1995).
The results of the study should be limited to interpretations of students’ moral reasoning, not their moral actions. The test instruments ask the subject what a person ought to do in a particular hypothetical situation, but not what the subject actually would do, or did do, in a particular situation. Therefore, differences in test scores imply significant differences in behavioural intent (Robin and Gordon, 1996). Similarly, interpreting the findings of this study should be restricted to accounting education and any reference on how the findings relate to practising accountants is speculative and should be viewed with caution (Fisher and Ott, 1996). Finally, the sample of accounting students in this study consists entirely of volunteers, which lends itself to a self-selection bias. The subjects were selected by design from a single university and, combined with a small sample size, the sample may not be representative of the general population of accounting students. Future researchers should expand the research design beyond their own students and include several schools with similar and different educational philosophies to assist in the generalisability of the results.
9. REFERENCES


Table 1 Sample description, including the spread over the 4 years of the degree program

<table>
<thead>
<tr>
<th>Description</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>All students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>18</td>
<td>20</td>
<td>12</td>
<td>47</td>
<td>97</td>
</tr>
<tr>
<td>Female</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>32</td>
<td>55</td>
</tr>
<tr>
<td>Male</td>
<td>10</td>
<td>13</td>
<td>4</td>
<td>15</td>
<td>42</td>
</tr>
<tr>
<td>Mean age</td>
<td>19.01</td>
<td>21.35</td>
<td>20.01</td>
<td>22.25</td>
<td>21.04</td>
</tr>
</tbody>
</table>

Table 2 Usable matched test instruments

<table>
<thead>
<tr>
<th></th>
<th>DIT</th>
<th>Welton</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size (97 x 3)</td>
<td>291</td>
<td>291</td>
</tr>
<tr>
<td>Loss rate (failed reliability and consistency checks)</td>
<td>39</td>
<td>16</td>
</tr>
<tr>
<td>Attrition (did not complete)</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>Matched using student inscribed identifier.</td>
<td>208</td>
<td>208</td>
</tr>
</tbody>
</table>
Table 3 Paired sample comparisons within and between the 208 usable matched test instruments

<table>
<thead>
<tr>
<th>Group</th>
<th>(n=)</th>
<th>Mean DIT P-score (std dev)</th>
<th>Mean Welton P-score (std dev)</th>
<th>Difference</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>All students</td>
<td>208</td>
<td>34.14 (18.97)</td>
<td>30.77 (15.01)</td>
<td>3.37</td>
<td>2.40</td>
<td>0.01 *</td>
</tr>
<tr>
<td>Pre-test</td>
<td>77</td>
<td>30.84 (17.01)</td>
<td>30.30 (14.43)</td>
<td>0.18</td>
<td>0.23</td>
<td>0.41</td>
</tr>
<tr>
<td>Mid-test</td>
<td>70</td>
<td>35.33 (19.00)</td>
<td>31.54 (15.17)</td>
<td>3.79</td>
<td>1.59</td>
<td>0.06</td>
</tr>
<tr>
<td>Post-test</td>
<td>60</td>
<td>36.93 (20.89)</td>
<td>30.49 (15.75)</td>
<td>6.44</td>
<td>2.52</td>
<td>0.01 *</td>
</tr>
<tr>
<td>Post-test – Pre-test</td>
<td>+6.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test – Pre-test</td>
<td>+0.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* significant at .05, one tail.

Table 4 Paired sample comparisons based on year levels of the 208 usable matched test instruments

<table>
<thead>
<tr>
<th>Group</th>
<th>(n=)</th>
<th>Mean DIT P-score (std dev)</th>
<th>Mean Welton P-score (std dev)</th>
<th>Difference</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 1</td>
<td>38</td>
<td>29.38 (16.03)</td>
<td>29.29 (15.42)</td>
<td>0.09</td>
<td>0.02</td>
<td>0.49</td>
</tr>
<tr>
<td>Year 2</td>
<td>44</td>
<td>35.90 (21.22)</td>
<td>30.07 (14.04)</td>
<td>5.83</td>
<td>1.88</td>
<td>0.03 *</td>
</tr>
<tr>
<td>Year 3</td>
<td>29</td>
<td>38.50 (16.96)</td>
<td>24.02 (12.09)</td>
<td>14.48</td>
<td>14.26</td>
<td>.0001 *</td>
</tr>
<tr>
<td>Year 4</td>
<td>97</td>
<td>33.90 (19.38)</td>
<td>33.69 (15.49)</td>
<td>0.21</td>
<td>0.10</td>
<td>0.45</td>
</tr>
</tbody>
</table>

* significant at .05, one-tail