Journal of Business Ethics (2008) 80:499–514
DOI 10.1007/s10551-007-9451-z

Improving Consistency for DIT Results Using Cluster Analysis

Carmel Herington
Scott Weaven

ABSTRACT. In this article, cluster analysis is used to explore the conflicting results reported when the Defining Issues Test is used to explain moral reasoning ability in business situations. Using a convenience sample, gender, age, work experience, and ethics training were examined to determine their impact on the level of moral reasoning ability as measured by the Defining Issues Test. Using the whole sample, a significant difference was found for average *P* scores reported for males and females, but no significant differences were found based on age, work experience, and ethics training. However, the sample fell into distinct clusters that identified distinct male and female groupings. While females naturally fell into two distinct high- and low-moral reasoning ability clusters, male clusters were dominated more by work experience and ethics training. Clearly there are other factors mitigating the level of moral reasoning ability for males which require further exploration. The findings suggest that while the *P* score provides an initial point of comparison, the real benefit to the test is in exploring what is different for males and females in terms of training needs, and the impact of work experience on the moral reasoning ability, and most importantly, how to make ethics training enticing. Recommendations for future research are also discussed.

KEY WORDS: moral reasoning ability, ethics training, Defining Issues Test, age, gender, cluster analysis, ethical decision-making

Carmel Herington is a senior lecturer in the Department of Marketing, Griffith University Gold Coast Campus, Australia. Her research interests include service quality and the provision of service quality through relationship building, internal relationships and their impact on other relationships of the firm, ethical business practices and marketing education. She has published in Qualitative Market Research, European Business Review, Academy of Marketing Science Review, Journal of Travel Research, Journal of Marketing Education, *amongst others.*

Scott Weaven is a lecturer in the Department of Marketing, Griffith University, Gold Coast Campus, Australia. His research interests include service quality in asymmetric exchange relationships, motivational incentives analysis and business model choice, gender issues in small business management and marketing education. He has published in such scholarly journals as the International Small Business Journal, Journal of Marketing Channels, Academy of Marketing Science Review, *and* Assessment and Evaluation in Higher Education.

Ethics remains an important topic for business (Forte, 2004; Izzo, 2000; Kaynama et al., 1996). There is an identified need for improvement in business decision-making generally (Allen et al., 2005; O'Fallon and Butterfield, 2005). This includes discussion and debate about ethics training and the degree to which it is successful, the ethical decision-making of business people, and the reasons for, and manner in which ethical business decisions may be made. As such there remains a need to be able to investigate how an individual makes their decisions to allow for such debates and discussions to revolve around factual information.

The ethical decision-making process has enjoyed considerable examination and scrutiny (e.g., Hunt and Vitell, 1986; Rest, 1986a; Trevino, 1986), and it is generally agreed that the initial recognition of an ethical dilemma, and the decision as to how it should be resolved, relies on an individual using their own reasoning and judgment. This reasoning process has been termed moral reasoning ability (Kohlberg, 1976).

The purpose of an instrument called the Defining Issues Test (hereafter DIT) is to measure moral reasoning ability. The DIT has been of particular value within the business context and has been used to measure the moral reasoning thought processes of business students (e.g., King and Mayhew, 2002), business executives (Elm and Nicholls, 1993),

accountants (Armstrong, 1987; Ponemon, 1992), and marketing professionals (Castleberry et al., 1993; Goolsby and Hunt, 1992).

The DIT is well-established as a suitable means for measuring moral reasoning ability within the business context (Gibbs and Widaman, 1982; Goolsby and Hunt, 1992; Rest, 1979, 1986a). There is substantial support and voluminous reported usage of the DIT test. For example, Rest and Narvaez (1994) reported over a decade ago that there were then over 1000 published DIT studies. Most importantly, there is also support for a link between moral reasoning ability and ethical behavior (e.g., Abdolmohammadi and Sultan, 2002).

The aim of the DIT is to determine a score of each respondent's level of moral reasoning ability. This is called the *P* score. Using a detailed scoring guide (Rest, 1993), the *P* score is easily calculated by the individual researcher. It is also considered the most valuable score (Rest, 1993).

Interestingly, though, despite the established validity and reliability of the measure, there still remain enough contradictory results reported for debate to still exist over the impact of individual demographic variables such as gender and age on *P* score. As these are typically collected demographic characteristics, it would be expected that some kind of generalized conclusions regarding age and gender could have been made by now. While there has been more consistency with gender results, with females commonly scoring higher-average *P* scores than males (e.g., Abdolmohammadi et al., 1997; Forte, 2005; Self and Baldwin, 1998; Thoma, 1986; White, 1999), some differing results have been reported in relation to age (e.g., Elm and Nichols, 1993; Goolsby and Hunt, 1992). However, although many have speculated on possible reasons for these findings, no research has presented possible explanations for these conflicting results.

Although the DIT has enjoyed extensive use in the business ethics context (as evidenced by over 40 DIT related studies reported in the JBE alone over the 10-year period 1996–2005), its relevance to business has been questioned (e.g., Marnburg, 2001; Mudrack, 2003). Discrepancies in results examining the relationship between *P* scores and social variables such as work experience and ethics training have been reported and this has called into question the applicability of the DIT in the business setting. There have been instances reported where ethics training has and has not worked and work experience has and has not had an impact on the level of moral reasoning ability (Forte, 2004).

Analysis using the *P* score has been restricted to use of correlation, and forms of comparison utilizing t-tests, ANOVA, and MANOVA. Multiple regression has also been utilized to a lesser extent. This means that analyses have revolved around examination of single variables or characteristics in broad groups, rather than exploring for any patterns of individuals scoring at different *P* score levels. With the aim of providing greater explanation of differences in results and provide greater consistency with reported usage of the DIT, this research utilizes cluster analysis. *P* scores and pertinent demographic characteristics are explored in relation to identifying different groups of moral reasoners. Any identified differences will be examined to develop a more detailed taxonomy for moral reasoning. Being able to provide more consistent results among groups related to demographics will extend our current understanding of business ethics. In addition to exploring the sample as a whole, we also determine if there is a different taxonomy for males and females and also for the combined group. Further, the intention is to determine, if we can standardize our examination and reporting of *P* score results, and also to determine if separate results should be reported for female and male clusters. This may result in a more meaningful reporting of the *P* score.

We commence with a review of the DIT literature, highlighting its value and contribution to the business ethics literature. We also explore confusions resulting from the reporting of *P* scores according to gender, age, work experience, and ethics training. The methodology and analytical procedures are then explained and this is followed by a discussion of results and recommendations for future research using the DIT in business settings.

Moral reasoning ability: a brief overview

The DIT measures moral reasoning ability, which is typically defined as ''the set of cognitive skills a person employs to reason about a moral problem'' (Elm and Nicholls, 1993, p. 818). The term was first

introduced by moral psychologist, Lawrence Kohlberg (1969) referring to a specific variable he created in his theory of Cognitive Moral Development (Rest and Narvaez, 1994). Cognitive Moral Development (hereafter CMD) theory is based on the seminal work of Piaget (Goolsby and Hunt, 1992; Piaget, 1965; Rest, 1983), who hypothesized that the skills involved in ethical decision-making develop over time (Castleberry et al., 1993) and that moral reasoning ability develops in sequential and distinctive cognitive stages (Fraedrich et al., 1994; Kohlberg, 1984; McDonald and Pak, 1996; Trevino, 1992).

According to Kohlberg (1976), there are a total of six different stages of CMD, classified under three higher-sequential levels of moral development: pre-conventional, conventional, and post-conventional (Colby and Kohlberg, 1987). Details of each of the stages are provided in Table I.

Moral reasoning ability became the measurable variable representing an individual's stage of CMD development. CMD theory proposes that the level of an individual's moral reasoning ability is closely linked to the eventual action taken, and the chosen action is likely to be more ethical as the level of moral reasoning ability increases (Colby and Kohlberg, 1987; Kohlberg, 1976). Rest (1979) developed a four-component model as a framework for positioning moral reasoning within the context of ethical decision-making.

The reasons influencing actual decisions is the important aspect of moral reasoning. Individuals will generally be found to justify their actions based on their level of moral development. For example, someone reasoning at the pre-conventional level (stages 1 and 2), will justify their decision on the basis of self-interest, whereas someone able to reason at the post-conventional level (stages 5 and 6) are more likely to consider their response through principled morality (Kohlberg, 1976). Further, it is proposed that individuals who operate at lower levels of moral reasoning ability are unaware of alternative courses of action they could take, due to their inability to think at a higher-level. Importantly, previous research suggests that the measurement of moral reasoning ability is a predictor of the likelihood to engage in unethical actions (Kohlberg, 1976; Rest, 1979, 1986a) and research has generally supported the positive relationship between CMD and ethical behavior (O'Fallon and Butterfield, 2005).

The Defining Issues Test

Although Kohlberg (1969) developed his own instrument to measure moral reasoning ability (called the Moral Judgment Interview), the DIT is the most widely accepted (and superior) device for measuring moral reasoning ability (Gibbs and Widaman, 1982; Goolsby and Hunt, 1992; Narvaez and Bock, 2002; Rest, 1986a; Rest et al., 1999). Even Kohlberg has said the DIT provides a broader spectrum and provides greater scoring reliability than the Moral

TABLE I

Six stages of Cognitive Moral Development

Level 1 – Pre-Conventional Level: Focus is on self
Stage 1: Obedience. You do what you are told to avoid punishment
Stage 2: Instrumental egoism and simple exchange: Let's make a deal or only consider the costs or benefits to self
Level 2 – Conventional level: focus is relationships
Stage 3: Interpersonal concordance: Be considerate, nice and kind and you'll make friends
Stage 4: Law and duty to social order: Everyone in society is obligated to and protected by the law
Level 3 – Post-Conventional Level: Focus is personally held principles
Stage 5: societal consensus: You are obligated by whatever arrangements are agreed to and by due process and procedure. Focus is on determining law or rule on grounds of equity and equality
Stage 6: Non-arbitrary social cooperation: Rational or impartial people would view cooperation as moral. Fairness of law or rules is derived from general principles of just and right as determined by rational people.

Source. Adapted from Rest and Narvaez (1994)

Judgment Interview (Rest, 1979). The DIT also overcomes issues related to the ability to articulate one's reasoning. For example, Narvaez and Bock (2002, p. 298) explain why the self-complete written test is preferred over the interview in stating ''The DIT does not measure the more competent end of the 'zone of proximal development' in which verbal articulation of one's perspective is required. To obtain a high score on a measure requiring verbal production such as the MJI, one must be able to explain one's reasoning logically and coherently, an ability that is facilitated by training in moral philosophy, but not necessarily by everyday life. The DIT tests the other, less competent, end of the 'zone,' that is apparent when assistance (such as words on a page) is available.'' Hence, the DIT may be considered most appropriate for measuring moral reasoning ability.

Finally, the DIT is a measure of tacit knowledge, which is knowledge that drives most human behavior (Narvaez and Bock, 2002). As such, in agreement with Wimalasiri (2004) we may assume that both personal and social variables would have a moderating influence on measured moral reasoning ability. Given the plethora of DIT research, researchers might assume findings in relation to basic demographic and business variables to be consistent.

Personal characteristics and the DIT

Despite the extensive usage of the DIT, there remains considerable confusion about the exact nature of the influences of individual demographic characteristics of age and gender on the level of moral reasoning ability.

Gender

Gender has been highlighted as the most widely studied demographic variable in relation to the DIT (Ford and Richardson, 1994). The theory of CMD does not propose the existence of a difference in moral reasoning ability based on gender and research exists to support this claim. For example, Archer and Waterman (1988) concluded that women and men do not differ in terms of their CMD. Walker (1984) conducted a meta-analysis of studies of moral judgment and found no gender differences. Ponemon (1992) found no significant relationship between gender and the level of moral reasoning ability when he examined US accountants. Paradice and Dejoi (1991) came to a similar conclusion in their study of moral reasoning ability of business students. Thoma (1986) performed a meta-anlalysis of 56 DIT studies and found little support for the contention that moral systems based upon a justice orientation disadvantaged women. Ryan (2001), and Skoe and von der Lippe (2002) found no support for sex differences, as did Abdolmohammadi et al. (2003) in a recent study in which the *P*-scores of male and female auditors were not found to be significantly different.

However, there has also been research demonstrating that differences do exist between males and females in relation to moral reasoning ability. Rest (1986a) acknowledged this but concluded from his synthesis of many DIT studies that any overall gender differences are 'trivial.' Certainly, Bernardi et al. (2004) reported some findings within their research where female students were found to use a higher-level of moral reasoning ability than male students.

Still others have reported differences that are statistically significant. For example, using samples of students and practitioners from a variety of business disciplines, Clarke et al. (1996), Mason and Mudrack (1997), Eynon et al. (1997) and Goolsby and Hunt (1992) have all found a significant difference between the measured level of moral reasoning ability for males and females. In each case, females were found to exhibit higher moral reasoning ability than males.

Given these conflicting results, a definitive conclusion as to the existence of a moderating effect of gender on ethical decision-making is far from resolved. This makes it important for researchers to be aware of the possibility that the gender balance of the sample may influence the findings from any research sample. As reported by O'Fallon and Butterfield (2005), the only consistent finding is that, on average, males do not exhibit a higher level of moral reasoning ability than females.

Age

Despite one of the basic premises of CMD being that moral judgment increases with age (Rest, 1986a),

conflicting results have also been found with regard to age and the level of moral reasoning ability (Ford and Richardson, 1994). Longitudinal studies have supported this premise (Colby and Kohlberg, 1987; Rest, 1986a). In support of the theoretical position, Dawson (1997), Deshpande (1997), Ponemon (1992) and Ruegger and King (1992) have all found older respondents to exhibit a higher level of moral reasoning ability and choose a more ethical course of action. On the other hand, Ho et al. (1997), Ponemon (1993), Paradice and Dejoi (1991) and Mason and Mudrack (1997) could not find any significant relationship between age and the level of moral reasoning ability and Goolsby and Hunt (1992) could only find a significant relationship between age and females with regards to moral reasoning ability. Contrary to the theorized relationship, other studies (e.g., Bigel, 2000; Clarke et al., 1996; Elm and Nichols, 1993; Eynon et al., 1997) have actually reported moral reasoning ability declining with age. These latter findings are particularly pertinent, as the sampled groups were accountants and business managers, who typically gain increased responsibility for ethical decisions as they grow older.

In their synthesis of moral reasoning research, O'Fallon and Butterfield (2005) confirm these conflicting results with regards to age, finding 14 studies reported no significant differences for different age categories, 10 studies that reported a positive relationship, with another six reporting a negative relationship.

In summary, within the extensive DIT research reported, results for personal demographic factors such as gender and age remain confusing and inconclusive. Hence, does this confusion also exist for social (external) factors? We move to examination of DIT findings in relation to two basic social factors, (ethics intervention training and actual work experience) and how they impact on moral reasoning ability.

Social factors and the DIT

Ethics training

The success of ethics training programs has also led to some surprising results in moral reasoning ability studies. Even the positive findings have been qualified. For example, Eynon et al. (1997) found support for ethics training having a significant positive effect on *P* score in a student sample. They did not investigate the exact nature of the ethics training and also speculated on the likelihood of those students choosing an ethics course as an elective being more positively disposed to, and more receptive to, ethical training which could account for the increase in the *P* scores. Bebeau and Thoma (1994) found fourth-year medical students' moral reasoning skills improved following ethics training. In a meta-analysis of 55 studies of education interventions utilizing the DIT as the measure of change in moral reasoning ability, Schlaefli et al. (1985) found that only programs involving dilemma discussion and psychological development programs (such as learning about CMD theory) had a significant, albeit small, positive effect on *P* scores, with the largest effect being for adults aged 24+. They also found that short courses (<3 weeks) and academic courses had no significant effect on moral reasoning ability.

There are also other findings reported where ethics training has not worked, both at the higher education level and in the work environment. Izzo (2000) found that compulsory ethics training did not improve the moral reasoning ability of sales people, while Self et al. (1998) found that university ethics instruction was not associated with continual improvement in an individual's moral reasoning ability. In addition, Rau and Weber (2003) found that an ethics intervention only resulted in a change in moral reasoning ability related to one of two ethical dilemmas employed in their research.

Work experience

The impact of work experience on moral reasoning ability has received less attention within DIT studies. However there have been some studies noting the effect of organizational aspects on moral reasoning ability. For example, in their review of moral reasoning literature, O'Fallon and Butterfield (2005) found numerous instances where work experience influenced ethical decision-making, even though the effect was marginal. Both Ponemon (1990) and Shaub (1994) found that higher ranked Certified Public Accountants displayed lower levels of moral reasoning than their less experienced counterparts.

However, a later study by Thorne et al. (2003) found no significant relationship between moral reasoning and work tenure.

Hence, researchers cannot yet conclude that there is consistent knowledge in relation to the demographic variables of gender and age, nor with the social variables of ethics training and work experience. This supports Mudrack's (2003) contention as to "what is wrong" with the *P* score?

One issue of course could relate to the lack of research generally into Rest's (1986a) ethical decision-making framework upon which the measure of moral reasoning ability is based. O'Fallon and Butterfield (2005) highlight the fact that the model has never been tested in its entirety. However, before this can be rectified, it is important to further explore the immediate influences on moral reasoning ability. One way to do this is to look at DIT results from different analytical angles. One such angle would be cluster analysis.

Cluster analysis

Cluster analysis is a multivariate statistical technique that can be used to group individuals or objects into clusters based on particular characteristics that they possess (Hair et al., 1995). When clustering individuals, the ultimate goal is to arrive at clusters of people with homogeneous characteristics who thereby exhibit small within-cluster (internal) variation, but at the same time exhibit large between cluster (external) variation (Aldenderfer and Blashfield, 1984; Hair et al., 1995).

The main advantage of cluster analysis is that it enables the researcher to define a cluster variate (i.e., the characteristic variables included in the comparison) which then determines the commonalities and differences among and between groups and leads to natural groupings (Hair et al., 1995). Furthermore, this approach provides an opportunity to explore structures existing in data prior to attempting to explain why they exist. Finally, a taxonomy can be developed to help describe a population.

As identified above, in studies utilizing the DIT, there is confusion as to the impact of the most basic demographic variables of age and gender, as it relates to the *P* score and what is proposed in CMD theory. There has also been surprising findings about the impact of social variables such as ethics training and work experience on moral reasoning ability. This could be due to the existence of natural clusters within the sample populations that have not been exposed. As an exploratory technique, cluster analysis is certainly suited to exploration of the reasons for this phenomena occurring so frequently within samples. Further, if the different moral reasoning ability groupings can be defined, appropriate education and training can be provided to those groups in most need, and we can perhaps uncover what is required to move individuals from a lower moral reasoning ability group to a higher one.

Rather than just report the average *P* score as well as test for differences between gender and age groups, cluster analysis offers the opportunity to develop a more concise and understandable description of respondent scores. Further, the interaction effects between the variables might also provide a better picture of how the *P* score results might be reported and examined.

Hence, in answering Mudrack's (2003) question as to "what is wrong" with the *P* score as it has been utilized to measure ethical disposition in business settings, we explore the use of cluster analysis as an assistant to explain the clusters of moral reasoners. We believe we can move the DIT research forward to try to really understand how unethical decisions are made by exposing the complexity and nature of moral reasoning ability. As such we explored the following proposition:

Cluster analysis may be utilized to develop meaningful categories of moral reasoning, which provides a better understanding of moral reasoning ability according to age and gender, work experience, and ethics training.

Method

Data was collected via a self-complete survey and administered to a convenience sample of undergraduate and postgraduate business students from a single Australian university. The sample was deemed appropriate for two reasons. Firstly, these students are current (postgraduates) and future (undergraduate) business professionals and can be considered

representative of the future business population (Grunbaum, 1997). Secondly, as many students are also working at the same time as studying, they do have current business work experience. In fact, many of the students studying at the university do so on a part-time basis and work full-time. Hence, comparisons can be made based on work experience. In addition, the DIT instrument has been extensively used previously to measure the moral reasoning ability of university students. For example, King and Mayhew (2002) report finding over 500 DIT studies utilizing student samples.

Questionnaires were distributed across a range of core business courses at both the undergraduate and postgraduate levels to ensure capture of a broad spectrum of students pursing a variety of different business majors. The questionnaire consisted of the DIT test as well as demographic questions capturing data about age, gender, work experience, and previous ethics training.

The DIT consists of a series of short standardized vignettes relating to general social dilemmas. While the full DIT contains six vignettes, a three-vignette version of the DIT was utilized for this research. The three vignette version is popular among researchers, particularly where there is a concern regarding likely response rates (e.g., Bay and Greenburg, 2001; Earley and Kelly, 2004; Eynon et al., 1997; Goolsby and Hunt, 1992; Ho et al., 1997) and has a reported very high degree of correlation with the longer version of the DIT (usually between 0.91 and 0.94) (Rest, 1986a). The choice of stories was based on their applicability to the Australian environment, necessitating minimal cultural adaptation of moral dilemma topics. Despite the fact that the DIT is widely adopted in moral reasoning assessment, we conducted pilot testing so as to measure the 'Australianisation' of chosen scenarios. This included spelling changes, removal of gender specification of scenario characters (e.g., 'he' changed to 'the doctor'), and 'Australianising' the location and character name in one scenario.

There are three processes involved in completing the DIT. Respondents read each vignette separately and are asked to choose one of three courses of action that should be taken in relation to the dilemma, that is, what should and should not be done with an option for answering 'can't decide.' Respondents then rate the importance (on a 5-point likert scale from 'great importance' to 'no importance') of each of 12 issues presented in determining their preferred course of action. Each issue is a prototypical statement representing one of the six stages of CMD, as defined by Kohlberg (1969). Respondents are expected to endorse the statements according to their developed level of reasoning. After rating the issues, respondents are then asked to select and rank the four statements they believe are most important in making their determinations about their chosen course of action.

From this data an index called the *P* score (standing for ''principled morality'') is calculated. Rest (1986a, p. 2) describes this score as ''the relative importance a subject gives to principled moral considerations in making a decision about moral dilemmas.'' When a respondent includes a statement reflecting principled morality in the four most important statements, a weighted score (on the basis of importance rank) is assigned. The *P* score represents the percentage of total possible scores (0 to 95) assigned to stage 5 and 6 statements (according to Kohlberg's CMD theory), with higher-scores indicating a higher level of moral reasoning ability equating with the ability to reason at the higher stages of CMD. To ensure accuracy, Rest's (1993) guidelines were followed for survey administration and scoring calculations during the data collection and analysis stages.

Reliability checks are also included as part of the test, and this research strictly adhered to Rests' rules for consistency (Rest, 1986a, 1993). The first test involves the calculation of an *M* score, which stands for 'Meaningless.' The *M* index is an internal reliability check for the researcher to detect non-thoughtful respondents. *M* items were written to sound lofty and pretentious but have no assigned meaning (Rest, 1986a). The *M* items are not representative of any stage of thinking, and individuals who score too highly on these items are considered to be unreliable respondents and as such are discarded from the data set.

A second inbuilt check on subject reliability is called the 'Consistency Check.' Each respondent's ratings are compared with their rankings. It is expected that the rankings should correspond to the ratings. Logically, an item ranked as most important in Part C of the test, should not have any other items rated above it in Part B. Rest (1986a) considers those

respondents who are inconsistent in following this logic cannot be considered reliable as the inconsistency is most likely due to such factors as careless responding, random checking, or misunderstanding instructions. Rest (1986a) lists a series of rules and cut-off points for the inconsistency checks. Although these rules and cut-off points have been empirically derived (Rest, 1979), researchers may vary these according to individual needs. Rest also claims a typical response loss of 5–15% due to the adoption of respondent reliability checks (Rest, 1986a).

As well as the *P* score, a score for each of the stages of CMD can also be calculated for each individual. These scores are generally ignored in the literature as most studies just report an overall measure of the level of moral reasoning ability. However, we considered this would provide a useful validity assessment for the *P* score. Consistent with Rest's (1986b) approach a variable was created scoring each respondent between 1 and 6 in accordance with their main stage of moral thinking, which was called 'CMD score.' This variable was then correlated with actual *P* score to provide the degree to which one variable explained the other.

A series of t-tests were then performed to examine for significant differences between *P* scores for males and females, those who had and had not received previous ethics training, and those who did and did not have work experience. ANOVA was utilized to examine for any significant differences in *P* score according to the different age categories. Regression analysis was used to explore the explanatory effect of any statistically significant impacting variables.

Cluster analysis was then used to synthesize the cases into a smaller number of groups to explore and interpret different groupings of respondents according to *P* score, age, gender, work experience, and ethics training. Three different cluster analyses (CA) were conducted. To begin, the whole dataset was cluster-analyzed (CA1). Then in order to explore for differences between males and females, the dataset was split in two according to gender, and the male (CA 2) and female groups (CA 3) were analyzed separately. The nature of each of the identified clusters was explored and is reported below.

Results

The initial response of 369 returned questionnaires was reduced to 232 usable responses following eliminations due to non-targeted disciplines (37), incompleteness (31), and inconsistency checks (69). Non-targeted disciplines included students enrolled in 'sport' and 'health' programs who had enrolled in the courses as electives. The reasonably high level of loss due to incompleteness can be attributed to the DIT requirement for every question of the test to be completed. There is also no provision for missing values treatment with the DIT. While the further loss of 18.7% of responses due to the DIT inconsistency checks is slightly outside the expected bounds suggested by Rest (1993), it is less than many previously reported losses such as 26% (Paradice and Dejoi, 1991) and 31% (Eynon et al., 1997).

The typical respondent was found to be under the age of 25 (83.6%), female (51.7%), and studying full-time (77.1%). Just over half (51.7%) of the respondents had some form of work experience with 27% having current or previous full-time work experience. 18.7% of respondents had previously undertaken some form of ethics training.

Using SPSS and following the DIT Guide book (Rest, 1993) instructions carefully, *P* scores were calculated for each individual response. The overall mean *P* score for the sample was 38.99 with a standard deviation of 15.47 and a range of 6.67–80.0. The distribution of *P* scores was found to be normal, based on a Q-Q plot of normality. In examining the reported average *P* scores of previous research, it was found that this result falls within the average range reported of between 34.2 and 43.1. Hence, the results falling within this range was deemed to provide validity for the research.

The 'CMD score' variable was then created which scored each respondent according to their exhibited main level of CMD. This was then correlated with the *P* score. Finding a large correlation of 0.67 demonstrated that the *P* score was highly related to the level of thinking (Cohen, 1988), thereby demonstrating validity for the *P* score as a suitable representative of the measured stage of CMD of respondents.

The t-test utilized to determine if a significant difference could be found between genders for *P*

score indicated that females (mean = 41.56, SD = 16.8) had scored significantly higher on average than males (mean = 36.61, SD = 13.4) as a group ($t = -2.42_{(226)}$, $p < 0.05$). This finding is consistent with other reported research which has found that females will generally report higher average *P* scores than males (Eynon et al., 1997; Forte, 2004; O'Fallon and Butterfield, 2005). However, these results are contrary to CMD theory and Rest's (1986a) findings.

ANOVA was then used to assess the effects of age on the level of moral reasoning ability in the first instance, passing the Levene's test for homogeneity between the variances. Even though there were some differences reported for different age categories, no significant differences were found among any of the age categories ($F_{(8,23)} = 1.77$, $p > 0.05$). However, we did note that moral reasoning ability did seem to rise with age, and then suddenly decline for the two highest age categories. Although this finding does support previous research (e.g., Elm and Nicholls, 1993; Eynon et al., 1997), it provides a conflict with the premise that moral reasoning ability increases with age, although it must be considered in context of the two highest age categories containing very small sample sizes (7 and 3 respectively). Despite the finding of no significant differences, we also explored gender distribution across the different age categories. The means are plotted in Figure 1, where it can be clearly seen that the distribution of *P* scores for males and females in this sample has some differences. While there is a steady increase in *P* scores for females (discounting the single respondent in the 50+ age), males do appear to exhibit steadily declining moral reasoning ability from their 30's or at best remaining static.

T-tests were also used to explore for significant differences between *P* score and ethics training and *P* score and work experience. Both tests also indicated no significant differences between *P* score and ethics training ($t_{(228)} = -.79$, $p > 0.05$) or *P* score and work experience ($t_{(201)} = -1.34$, $p > 0.05$). Hence, for the entire sample population we cannot say that either previous ethics training or work experience impacts *P* score. This supports previous findings (e.g., Izzo, 2000; O'Fallon and Butterfield, 2005).

We then performed a multiple regression analysis to determine the explanatory value of the variables on the level of moral reasoning ability. Although the results of the multiple regression indicated a significant impact on *P* score ($F = 4.52$, $p < 0.05$), less than ten percent (8.5%) of the variance in *P* score is explained by this group of variables. An examination of the *t*-values also indicates that only gender ($t = 3.54$, $p < 0.05$) and age ($t = 2.26$, $p < 0.05$) contribute to the *P* score, with gender explaining 6% of the variance and age only 2.6%.

For the whole sample, the cluster analysis was performed using a variate set of *P* score, age, gender, work experience, and ethics training. A summary of the results of the cluster analysis is provided in Table II, panel a.

Overall, six clusters were exposed, with a good spread of sample respondents within each cluster. Clusters were not all distinguished by their *P* score. Some were differentiated more by gender, age, work experience, and/or ethics training. Gender was a distinguishing factor with 5 out of the 6 groups single gender groups, with three male only clusters and two female only clusters. While females were easily distinguished as either high or low *P* score, the male segments were distinguished by a high *P* score group, but then two lower *P* score groups distinguished by work experience. There was also a mixed group (although mostly male), exhibiting low *P* scores, a mix of ages but all with work experience. These were called 'middle of the road', and while this cluster would disappear in the split-gender

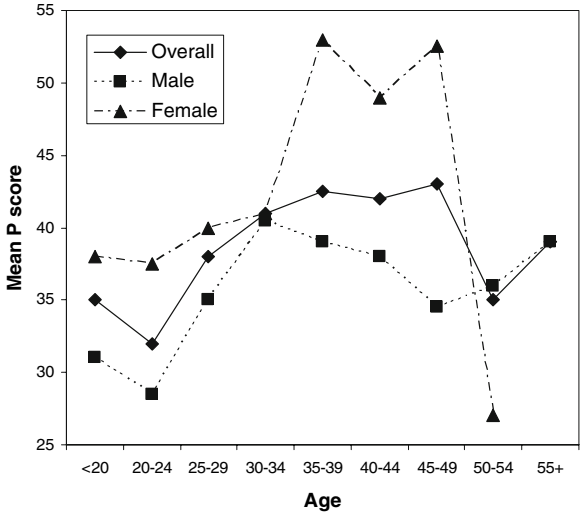

Fig. 1 Mean plots for *P* scores by age

TABLE II

Cluster analysis results

Cluster Group	*N*	*P* Score (SD)	Gender	Age	Work experience	Ethics training	Cluster name
(a) Distinguishing characteristics for each identified cluster – overall sample							
1	45	33.9 (13.2)	Male	<20–49	N★	Y/N	Males without work experience
2	20	36.5 (15.3)	Male	25–55+	Y	Y★	Males with ethics training and work experience
3	30	40.1 (13.1)	Male	30–39★	Y	N	Mid-career males
4	31	49.5 (17.9)★	Female★	20–49	Y	Y/N	High *P* score females
5	42	39.3 (16.1)	Female★	<20–44	N	N	Low *P* score females
6	31	34.4 (13.5)	Mixed	25–29, 40–55+	Y	N	Middle of the road
(b) Female sample							
1	49	45.9 (17.1)		25+	Y	Y/N	High MRA, work experience
2	31	37.4 (16.3)		<20–29	N	N	Low MRA, no work experience
(c) Male sample							
1	39	35.5 (12.8)		<20–49	N	N	Inexperienced low *P* score males
2	29	33.5 (15.3)		<20–55+	Y/N	Y	Low *P* score ethics training
3	54	37.8 (14.0)		25–55+	Y	N	Mid-career males

★ Indicates most significant important distinguishing characteristic

clustering, it was interesting to note at this stage that work experience is related to lower *P* scores.

When the sample was then separated by gender, and cluster analysis performed for the males and females separately, results became even more distinguishable. Females were easily clustered by high and low *P* score and work experience (Table II, panel b). Males were clustered by *P* score to a lesser degree, with ethics training and work experience significantly more important (see Table II, panel c). Male cluster 1 were distinguished by no work experience and no ethics training but exhibiting a mid-range *P* score, while cluster 2 exhibited the lowest *P* scores but consisted of all males who had received ethics training. Cluster 3 consisted of the highest-male *P* scores while only including males with work experience and no ethics training. It would appear that work experience seems to be influencing male *P* scores and ethics training has not had a positive impact on the *P* scores of males.

Discussion

Although there have been few uses of the DIT with Australian samples, the reported average *P* score in this study falls within the reported range of DIT studies generally, thus supporting the use of the *P* score as an overall measure of moral reasoning ability and validity for the use of the DIT. Validity is further demonstrated for the DIT with the *P* score highly correlating with the individual's most commonly utilized level of CMD thinking. The calculation of the *P* score is quite complicated and is not merely a recorded number according to the main level of thinking. It is in fact a measure of the degree to

which the respondent is thinking at stages 5 and 6 (i.e., principled thinking). Hence, it was important to demonstrate that the *P* score is a good surrogate measure for the actual level of CMD thinking.

In terms of gender, our results are similar to those reported elsewhere. We did find a significant difference between males and females on the *P* scores even though they are not in line with the original theoretical position. Not only were females higher overall in their level of moral reasoning ability, but also the range was smaller and the SD considerably smaller as well. This would suggest that unknown factors may be influencing moral reasoning ability for males but is not affecting females in the same way. Moral reasoning ability of males generally appears to have a more complex explanation. This suggests that further exploration of the moral reasoning ability of males and females as separate samples is warranted, as there may be different antecedents to moral reasoning ability according to gender. In other words, models consisting of variables influenced by different moderators and/or mediators might better represent each gender.

Age was a much more confusing variable in terms of results and these results certainly provide support for Pasupathi and Staudinger's (2001) claim that the relationship between age and moral reasoning ability is complex. Although we found no significant difference between *P* scores for different age categories, we did find that past the age of 50, the *P* score declined. Once this age category was removed, the relationship between age and *P* score was demonstrated with the overall sample. However, while we believe that our results might have been an artifact of the small sample size for this age category, there remains enough evidence in the literature to support these findings (e.g., Clarke et al., 1996; Elm and Nichols, 1993; Eynon et al., 1997). In particular, Eynon et al. (1997), Elm and Nichols (1993) and Clarke et al. (1996) arrived at similar conclusions about the decline in moral reasoning in older age categories: some factor within the work environment was promoting lower levels of moral reasoning ability. Further exploration of this phenomena has been called for (O'Fallon and Butterfield, 2005), although no empirical evidence is yet available to provide an explanation. For example, both Bigel (2000) and Forte (2004) have suggested that an increase in cynicism among older disappointed workers could affect moral reasoning ability.

Our findings also suggest that while females continue to grow in moral reasoning ability until their 50's, males seem to grow only until their early 30's (on average), and then, at best, tend to remain static in the moral thinking, perhaps even steadily declining. This finding supports Bay's (2002) suggestion that females have the ''potential to increase ethical understanding at a different rate and in a different way'' and requires considerable further exploration within the business ethics literature, as it would suggest that older females have particular importance in understanding the nature of ethical business practices. In support of this, Akaah (1989) posited an improvement in ethical decision-making in business as more women move into executive positions.

The cluster analysis was then able to provide a picture of likely segments exhibiting certain similar characteristics and this is where our most significant findings can be reported. We believe that the cluster analysis provides a basis for explaining anomalies that have existed in previous research findings. The gender differences are confirmed and further evidence is provided emphasizing the need for closer examination of males and females as separate subsamples, if we are to further understand moral reasoning ability and the antecedents to unethical behavior. A number of researchers have found that females and males reason differently. Most notably, Gilligan (1982) posited that males take a justice orientation toward resolving conflicts whereas females take a more caring orientation, which could account for differences in *P* scores.

Further, when the male and female sub-samples were explored separately, different clusters and cluster differentiators for males and females were uncovered. Whereas females can be easily clustered by the level of moral reasoning ability, this was not the same for males, where it was found that work experience and ethics training were more important in distinguishing the clusters. In fact, ethics training seemed to have had a deleterious effect on moral reasoning ability with average *P* score lowest for males who had received previous ethics training. It would appear that at least with the sample used in this research, the ethical training needs of males are not being met. The results also support Arlow's (1991) position that

ethics instruction should consider the differing needs of males and females.

The results of the multiple regression indicated that overall, these four characteristics examined in this research accounted for a very small proportion of the variance in *P* score for the overall sample. This also suggests that there are other factors playing a greater role in determining the level of moral reasoning. These need to be explored to gain a better picture of moral reasoning ability as it pertains to ethical decision-making in business.

Further, it would appear that social factors might actually impact on moral reasoning to a greater degree than we think. A number of researchers have already suggested and investigated the impact of a variety of social factors, most pertinently related to the work environment. For example, Bigel (2000) found a negative relationship between job tenure and *P* score, speculating on a moderating influence of cynicism as job tenure increased. While our results do not take into account job tenure, this may provide some explanation for finding a decline in *P* score generally in the highest age categories. In an experiment into the impact of competition on CMD, Reall et al. (1998) found that the introduction of a competitive environment lowered moral reasoning, providing evidence that organizational climate and culture will influence moral reasoning ability. Forte (2004), Hunt and Jennings (1997) and Robertson and Rymon (2001) also report the influence of unethical climate on moral reasoning ability.

The results also suggest that greater consideration may need to be given to customized ethics training rather than utilizing a standardized approach which might have a detrimental impact, particularly on males. This approach has some prior support in the ethics literature. For example, Solberg et al. (1995) found that ethics training should be based upon individual needs and Eaton and Giacomino (2001) determined that ethics training is more likely to impact upon females' values rather than males' values. Implications for managers include the necessity to explore avenues for trialing different types of ethics training and even offer some different choices for staff. At the same time, it is imperative to develop a particular understanding of what ethics training is most acceptable to those exhibiting lower levels of moral reasoning ability. As has previously been found (Eynon et al., 1997), those who have an interest in moral reasoning ability, and improving their ethical decision-making generally, will respond better to ethics training. Hence, approaches need to be found to create ethics training that is enticing to everyone. This is a challenge for business firms and may involve the use of rewards and incentives schemes to 'sell' the virtue of ethics. However, this discussion is beyond the scope of this research and in itself provides a fruitful avenue for further investigation.

Overall, the results of the analysis indicate that we do not yet fully understand what influences moral reasoning ability. Simply reporting *P* scores does not provide the depth of information required to make strategic business decisions about how to train staff, what type of ethics training is required for whom, or whether, training programs should be individualized. We do not know the impact of organizational factors such as climate and internal competition on the decisions made by employees. And we do not know how we might be inadvertently causing a lowering of ethical standards through ethics training in the work environment. While the *P* score provides an overall impression of what might be happening with regards to ethical thinking of individuals within a firm, this should be mitigated by the gender balance, and the type of training and previous and current work experiences of employees. Hence, we need to revisit and extend Rest's original model through investigating the role of personal and social influences on moral reasoning ability.

Conclusion

The aim of this research was to enhance knowledge of moral reasoning ability as it impacts on ethical decision-making by providing a clearer picture of what the DIT *P* score is telling us. As such we explored the gender and age variables and utilized cluster analysis to try to find segments of *P* score exhibiters who displayed similar characteristics. The constitution of a sample population will have the effect of generalizing *P* scores. Finding differences in results for gender can be explained by different clusters exhibiting different characteristics. Further, males and females also cluster differently. Hence, when you have a heterogeneous population, you will get generalized results. If we are to examine the

impact of variables such as age, gender, work experience, and ethics training we need to explore deeper into the nature of moderating and mediating influences given the small amount of variance explained by different variables. Not only have we uncovered the considerable differences between males and females, but we have also provided an explanation for confusing and often conflicting results.

However, the results reported in this research have some associated limitations. Being based on an Australian sample provides an initial issue for generalizability of the results. While utilizing a sample of business students who do naturally have the characteristics of business people, they are not all currently working in the business world and many of those who did indicate that they were working will not be in senior ranks of organizations. Hence, collection of further data from business and global samples will assist in providing further validity to these results through replication. Such datasets should also include a larger spread across all age categories by gender to ensure adequate data points in all cells analyzed. In addition, specific data regarding the nature of ethics training was not captured in this study. Although, in line with previous research (e.g., Eynon et al., 1997), we asked respondents if they had completed an ethics course, we did not request detail on the specific length of any such courses, the type of courses or how these courses were presented. Accordingly, future research should consider deeper investigation into the effects of different aspects of ethics training (e.g., course duration and course type) on moral reasoning ability.

It needs to be understood that while evidence has been found linking moral reasoning ability to ethical action, this research cannot say that the clusters found would automatically make certain types of decisions. Further research is required to examine the possibility of such links. In this regard, we concur with O'Fallon and Butterfield (2005) who highlighted a lack of empirical examination of Rest's full ethical decision-making model. Also, consistent with O'Fallon and Butterfield's (2005) recommendations, we believe that our research further demonstrates the need for investigating the existence of antecedents, additional consequences and most particularly moderating variables, so as to extend Rest's model.

The results of this research also provide other directions by which DIT research may be propelled forward to provide a deeper understanding of ethical decision-making in business. The DIT is a very sensitive instrument responding to demographic variables such as age and gender. We propose that rather than merely making generalized comparisons, this instrument would be better engaged for exploring the moderating influences on moral reasoning ability among genders and different clusters of people. Further research is particularly recommended in terms of exploring the identified moral reasoning clusters and their actual actions. Explorations of actual behaviors taken, or the creation of experimental conditions to determine causality should be explored. In addition, we now call for the exploration of antecedents and specific consequences (such as actual behaviors) for each of the determined clusters, especially if males and females are examined separately.

Moderating influences should include both personal and social factors. Personal factors might include orientation, for example 'caring' versus 'justice' as suggested by Gilligan (1982) for females and males respectively. Social influences could include job tenure and satisfaction, organizational ethical climate and culture, peer pressure and competitive environment as discussed above.

Such investigations will provide a much deeper understanding of ethical decision-making in business. This might lead to more specifically targeted training and development which will make a bigger difference to the behaviors exhibited, rather than the general approach that is taken currently due to lack of knowledge of who is making what decision and why. We believe that this is a major contribution of this article.

The DIT is also highly applicable to longitudinal research, which can take two forms. Firstly, the knowledge about moral reasoning in business could be gained through following the progressive development of moral reasoning in a group of individuals through their college education and into work life. Many of the questions about what influences moral reasoning ability could be uncovered and explored. Secondly, the DIT can be used as a pre- and post-measure of the impact of ethics training (as shown in previous studies such as Dellaportas, 2006). As an example, McDonald (2004) provides detail on how to implement ethics training in the business curriculum, and calls for testing of the proposed model.

This would be a suitable use for the DIT as a pre- and post-measure of the success of business ethics training programs.

Finally, we believe that what we have demonstrated is the innate ability of the *P* score to be very sensitive to a number of known and still unknown factors which impact on moral reasoning ability. We have provided a further avenue for the DIT to explore how business decisions are made and how ethical decision-making should be encouraged and taught. Should we be able to distinguish differences between males and females, we will be able to advance our understanding on how to best deal with ethical issues in business contexts.

References

Abdolmohammadi, M., W. J. Read and D. P. Scarbrough: 2003, 'Does Selection-Socialization Help to Explain Accountants' Weak Ethical Reasoning?', *Journal of Business Ethics* **42**(1), 71–81.

Abdolmohammadi, M. and J. Sultan: 2002, 'Ethical Reasoning and the Use of Insider Information in Stock Trading', *Journal of Business Ethics* **37**, 165–173.

Abdolmohammadi, M., D. Gabhart and M. Reeves: 1997, 'Ethical Cognition of Business Students Individually and in Groups', *Journal of Business Ethics* **16**, 1717–1725.

Akaah, I.: 1989, 'Differences in Research Ethics Judgments Between Male and Female Marketing Professionals', *Journal of Business Ethics* **8**, 375–381.

Aldenderfer, M. and R. Blashfield: 1984, *Cluster Analysis* (Sage Publications, Newberry Park).

Allen, W., P. Bacdayan, K. Kowalski and R. Mathew: 2005, 'Examining the Impact of Ethics Training on Business Student Values', *Education and Training* **47**(3), 170–182.

Archer, S. L. and A. S. Waterman: 1988, 'Psychological Individualism: Gender Differences or Gender Neutrality?', *Human Development* **31**, 65–81.

Arlow, P.: 1991, 'Personal Characteristics in College Students' Evaluations of Business Ethics and Corporate Social Responsibility', *Journal of Business Ethics* **10**, 63–69.

Armstrong, M.: 1987, 'Moral Development and Accounting Education', *Journal of Accounting Education* **5**, 27–43.

Bay, D.: 2002, 'A Critical Evaluation of the Use of the DIT in Accounting Ethics Research', *Critical Perspectives in Accounting* **12**, 159–177.

Bay, D. and R. Greenburg: 2001, 'The Relation of the DIT and Behavior: A Replication', *Issues in Accounting Education* **16**, 367–380.

Bebeau, M. J. and S. J. Thoma: 1994, 'The Impact of a Dental Ethics Curriculum on Moral Reasoning', *Journal of Dental Education* **58**(9), 684–692.

Bernardi, R. A, R. L. Metzger, R. G. S. Bruno, M. A. W. Hoogkamp, L. E. Reyes and G. H. Barnaby: 2004, 'Examining the Decision Process of Students' Cheating Behavior: An Empirical Study', *Journal of Business Ethics* **50**(4), 397–414.

Bigel, K.: 2000, 'The Ethical Orientation of Financial Planners Who are Engaged in Investment Activities: A Comparison of United States Practitioners Based on Professionalization and Compensation Sources', *Journal of Business Ethics* **28**, 323–337.

Castleberry, S., W. French and B. Carlin: 1993, 'The Ethical Framework of Advertising and Marketing Research Practitioners: A Moral Development Perspective', *Journal of Advertising* **22**(2), 39–46.

Clarke, P., N. Hill and K. Stevens: 1996, 'Ethical Reasoning Abilities: Accountancy Practitioners in Ireland', *Ibar* **17**, 94–109.

Cohen, J.: 1988, *Statistical Power Analysis for the Behavioral Sciences* (Lawrence Erlbaum Associates, Hillsdale, NJ).

Colby, A. and L. Kohlberg: 1987, *The Measurement of Moral Judgment: Theoretical Foundations and Research Validations and Standard Issue Scoring Manual, Vols, 1 and 2* (Cambridge, MA, Cambridge University Press).

Dawson, L.: 1997, 'Ethical Differences Between Men and Women in the Sales Profession', *Journal of Business Ethics* **16**, 1143–1152.

Dellaportas, S.: 2006, 'Making a Difference with a Discrete Course on Accounting Ethics', *Journal of Business Ethics* **65**(4), 391–404.

Deshpande, S.: 1997, 'Managers' Perception of Proper Ethical Conduct: The Effect of Sex, Age, and Level of Education', *Journal of Business Ethics* **16**, 79–85.

Earley, C. and P. Kelly: 2004, 'A Note on Ethics Educational Interventions in an Undergraduate Auditing Course: Is There an ''Enron Effect''?', *Issues in Accounting Education* **19**(1), 53–71.

Eaton, T. and D. Giacomino: 2001, 'An Examination of Personal Values: Differences Between Accounting Students and Managers and Differences Between Genders', *Teaching Business Ethics* **5**(2), 213–229.

Elm, D. and M. Nichols: 1993, 'An Investigation of the Moral Reasoning of Managers', *Journal of Business Ethics* **12**, 817–833.

Eynon, G., N. Hill and K. Stevens: 1997, 'Factors that Influence the Moral Reasoning Abilities of Accountants:

Implications for Universities and the Profession', *Journal of Business Ethics* **16**, 1297–1309.

Ford, R. and W. Richardson: 1994, 'Ethical Decision Making: A Review of the Empirical Literature', *Journal of Business Ethics* **13**, 205–221.

Forte, A.: 2004, 'Business Ethics: A Study of the Moral Reasoning of Selected Business Managers and the Influence of Organizational Ethical Climate', *Journal of Business Ethics* **51**, 167–173.

Forte, A.: 2005, 'Locus of Control and the Moral Reasoning of Managers', *Journal of Business Ethics* **58**(1/3), 65–77.

Fraedrich, J., D. Thorne and O. Ferrell: 1994, 'Assessing the Application of Cognitive Moral Development Theory to Business Ethics', *Journal of Business Ethics* **13**, 829–838.

Gibbs, J. and K. Widaman: 1982, *Social Intelligence: Measuring the Development of Sociomoral Reflection* (Prentice-Hall, Englewood Cliffs).

Gilligan, C.: 1982, *In a Different Voice: Psychological Theory and Women*'s Development (Harvard University Press, Cambridge, MA).

Goolsby, J. and S. Hunt: 1992, 'Cognitive Moral Development and Marketing', *Journal of Marketing* **56**(1), 55–68.

Grunbaum, L.: 1997, 'Attitudes of Future Managers towards Business Ethics: A Comparison of Finnish and American Business Students', *Journal of Business Ethics* **16**, 451–463.

Hair, J., R. Anderson, R. Tatham and W. Black: 1995, *Multivariate Data Analysis* (Prentice Hall, Englewood Cliffs, NJ).

Ho, F., S. Vitell, J. Barnes and R. Desborde: 1997, 'Ethical Correlates of Role Conflict and Role Ambiguity in Marketing: The Mediating Role of Cognitive Moral Development', *Journal of the Academy of Marketing Science* **25**, 117–126.

Hunt, S. and S. Vitell: 1986, 'A General Theory of Marketing Ethics', *Journal of Macromarketing* **48**, 5–16.

Hunt, T. and D. Jennings: 1997, 'Ethics and Performance: A Simulation Analysis of Team Decision Making', *Journal of Business Ethics* **16**, 229–239.

Izzo, G.: 2000, 'Compulsory Ethics Education and the Cognitive Moral Development of Salespeople: A Quasi-Experimental Assessment', *Journal of Business Ethics* **28**, 223–241.

Kaynama, S., A. King and L. Smith: 1996, 'The Impact of a Shift in Organizational Role on Ethical Perceptions: A Comparative Study', *Journal of Business Ethics* **15**, 581–590.

King, P. and M. Mayhew: 2002, 'Moral Judgment Development in Higher Education: Insights From the Defining Issues Test', *Journal of Moral Education* **31**(3), 247–270.

Kohlberg, L.: 1969, *Stage and Sequence: The Cognitive-Development Approach to Socialization Theory and Practice* (Rand McNally, Chicago).

Kohlberg, L.: 1976, 'Moralization: The Cognitive-Development Approach', in T. Likona (ed.), *Morality: Theory, Research and Social Issues* (Holt, Rinehart & Winston, New York).

Kohlberg, L.: 1984, *Essays in Moral Development, (Volume 1): The Philosophy of Moral Development* (Harper and Row, New York).

Marnburg, E.: 2001, 'The Questionable Use of Moral Development Theory in Studies of Business Ethics: Discussion and Empirical Findings', *Journal of Business Ethics* **32**, 275–283.

Mason, E. and P. Mudrack: 1997, 'Do Complex Moral Reasoners Experience Greater Ethical Work Conflict?', *Journal of Business Ethics* **16**, 1131–1318.

McDonald, G.: 2004, 'A Case Example: Integrating Ethics into the Academic Business Curriculum', *Journal of Business Ethics* **54**, 371–384.

McDonald, G. and P. Pak: 1996, 'It is all Fair in Love, War and Business: Cognitive Philosophies in Ethical Decision-Making', *Journal of Business Ethics* **15**, 973–996.

Mudrack, P.: 2003, 'The Untapped Relevance of Moral Development Theory in the Study of Business Ethics', *Journal of Business Ethics* **42**, 225–236.

O'Fallon, M. and K. Butterfield: 2005, 'A Review of the Empirical Ethical Decision-Making Literature: 1996–2003', *Journal of Business Ethics* **59**, 375–413.

Narvaez, D. and T. Bock: 2002, 'Moral Schemas and Tacit Judgment or How the Defining Issues Test is Supported by Cognitive Science', *Journal of Moral Education* **31**(3), 297–314.

Paradice, D. and R. Dejoi: 1991, 'The Ethical Decision-Making Processes of Information Systems Workers', *Journal of Business Ethics* **10**, 1–21.

Pasupathi, M. and U. M. Staudinger: 2001, 'Do Advance Moral Reasoners also Show Wisdom? Linking Moral Reasoning and Wisdom-Related Knowledge and Judgement', *International Journal of Behavioral Development* **25**(5), 401–415.

Piaget, J.: 1965, *The Moral Judgment of a Child* (The Free Press, New York).

Ponemon, L.A.: 1990, 'Ethical Judgement in Accounting: A Cognitive Development Perspective', *Critical Perspectives in Accounting* **1**, 191–215.

Ponemon, L.: 1992, 'Ethical Reasoning and Selection-Socialization in Accounting', *Accounting Organization and Society* **17**(3–4), 239–258.

Ponemon, L.: 1993, 'Can Ethics be Taught in Accounting?', *Journal of Accounting Education* **11**, 185–209.

Rau, S. E. and J. Weber: 2003, 'Can the Repeated Prisoners' Dilemma Game be Used as a Tool to Enhance Moral Reasoning?', *Teaching Business Ethics* **7**(4), 395–416.

Reall, M., J. Bailey and S. Stoll: 1998, 'Moral Reasoning ''On Hold'' During a Competitive Game', *Journal of Business Ethics* **17**, 1205–1210.

Rest, J.: 1979, *Development in Judging Moral Issues* (University of Minnesota Press, Minneapolis).

Rest, J.: 1983, 'Morality', in P. Musson (ed.), *Manual of Child Psychology (Vol. III), Cognitive Development* (Wiley, New York), pp. 556–629.

Rest, J.: 1986a, *Moral Development: Advances in Research and Theory* (Praeger, New York).

Rest, J.: 1986b, *DIT Manual for Defining Issues Test,* 3rd Edition (University of Minnesota Center for the Study of Ethical Development, Minneapolis).

Rest, J.: 1993, *Guide to Using the DIT, Version 1.3*. (Center for the study of Ethical Development, University of Minnesota).

Rest, J. and D. Narvaéz: 1994, *Moral Development in the Professions: Psychology and Applied Ethics* (Lawrence Erlbaum Associates, Hillsdale, N.J).

Rest, J., D. Narveaz, M. Bebeau and S. Thoma: 1999, *Postconventional Moral Thinking; A neo-Kohlbergian Approach* (Erlbaum, Mahwash, NJ).

Robertson, D. and T. Rymon: 2001, 'Purchasing Agents' Deceptive Behavior: A Randomized Response Technique Study', *Business Ethics Quarterly* **11**(3), 455–479.

Ruegger, D. and E. King: 1992, 'A Study of the Effect of Age and Gender Upon Students Business Ethics', *Journal of Business Ethics* **11**, 179–186.

Ryan, J. J.: 2001, 'Moral Reasoning as a Determinant of Organizational Citizenship Behaviours: A Study in the Public Accounting Profession', *Journal of Business Ethics* **33**(3), 233–244.

Shaub, M. K.: 1994, 'An Analysis of the Association of the Tradition Demographic Variables with the Moral Reasoning of Auditing Students and Auditors', *Journal of Accounting Education* **12**, 1–26.

Schlaefli, A., J. Rest and S. Thoma: 1985, 'Does Moral Education Improve Moral Judgment? A Meta-Analysis of Intervention Studies Using the Defining Issues Test', *Review of Educational Research* **55**(3), 319–352.

Self, D. J. and D. C. Baldwin, Jr.: 1998, 'Does Medical Education Inhibit the Development of Moral Reasoning in Medical Students? A Cross-Sectional Study', *Academic Medicine* **73**, s91–s93.

Self, D. J., M. Olivarez and D. C. Baldwin, Jr.: 1998, 'Clarifying the Relationship of Medical Education and Moral Development', *Academic Medicine* **73**, 517–520.

Skoe, E. E. and A. L. von der Lippe : 2002, 'Ego Development and the Ethic of Care and Justice: The Relations Among them Revisited', *Journal of Personality* **70**(4), 485–508.

Solberg, J., K. C. Strong and C. McGuire, Jr.: 1995, 'Living (not learning) Ethics', *Journal of Business Ethics* **14**(1), 71–81.

Thoma, S. J.: 1986, 'Estimating Gender Differences in the Comprehension and Preference of Moral Issues', *Developmental Review* **6**, 165–180.

Thorne, L., D. W. Massey and M. Magnan: 2003, 'Institutional Context and Auditor's Moral Reasoning: A Canada-US Comparison', *Journal of Business Ethics* **43**(4), 305–321.

Trevino, L.: 1986, 'Ethical Decision Making in Organizations: A Person-Situation Interactionist Model', *The Academy of Management Review* **11**(3), 601–617.

Trevino, L.: 1992, 'Moral Reasoning and Business Ethics: Implications for Research, Education and Management', *Journal of Business Ethics* **11**, 445–459.

Walker, L. J.: 1984, 'Sex Differences in the Development of Moral Reasoning: A Critical Review', *Child Development* **55**(3), 677–691.

Wimalasiri, J.: 2004, 'Contrasts in Moral Reasoning Capacity: The Fijians and the Singaporeans', *Journal of Business Ethics* **49**, 253–272.

White, R.: 1999, 'Are Women More Ethical? Recent Findings on the Effects of Gender Upon Moral Development', *Journal of Public Administration Research and Theory* **9**(3), 459–471.

Department of Marketing, Griffith Business School, Griffith University,
PMB 50 Gold Coast Mail Centre, Gold Coast, QLD, 9726, Australia
E-mail: c.herington@griffith.edu.au