

Research Methods in Taxation Ethics: Developing the Defining Issues Test (DIT) for a Tax-Specific Scenario

Elaine Doyle
Jane Frecknall-Hughes
Barbara Summers

ABSTRACT. This paper reports on the development of a research instrument designed to explore ethical reasoning in a tax context. This research instrument is a version of the Defining Issues Test (DIT) originally developed by Rest [1979a, *Development in Judging Moral Issues* (University of Minnesota Press, Minneapolis, MN); 1979b, *Defining Issues Test* (University of Minnesota Press, Minneapolis, MN)], but adapted to focus specifically on the environment encountered by tax practitioners. The paper explores reasons for developing a context- (and profession-) specific test, and details the manner in which this was undertaken. The study on which it is based aims to compare the reasoning of tax practitioners in the tax-specific context and in the general social context covered by the original DIT, and to compare this with the reasoning of non-specialists in these two contexts. The paper therefore also considers the issues that arise when using such tests to compare reasoning in different domains or to compare groups. The focus on instrument development to measure ethical reasoning in a specific domain will contribute to the literature on research methods in the area of the DIT and will facilitate cross-study comparisons.

KEY WORDS: Defining Issues Test, context-specific scenario, ethical reasoning, research methods, tax practitioners

Introduction

The Defining Issues Test (DIT) was developed by Rest (1979a, 1986) following earlier work in cognitive development theory by Kohlberg (1973) and has been extensively employed as a measure of the ethical reasoning process.¹ However, while the DIT is “a broad, general measure of moral reasoning” (Fisher, 1997, p. 143), acceptable in dealing with personal issues in a social context (Fraedrich et al.,

1994), concern has been expressed that it does not, and cannot, fairly represent the reasoning used when facing ethical dilemmas in a business context (see for example, Dellaportas et al., 2006; Elm and Nichols, 1993; Fraedrich et al., 1994; Trevino, 1986, 1992; Weber, 1990; Welton et al., 1994).

These concerns have led to the development, in a number of areas, of instruments based on the DIT, but containing context-specific scenarios.² However, many of the papers reporting use of a context-specific DIT have only been able to devote a small amount of coverage to the actual process of test development, as such papers have been focussed primarily on delivering results of empirical research in domain-specific ethics rather than explicitly on methodology or methods. As interest in ethics in financial domains continues to grow, it is useful to detail the process underlying the development of a specific ethical reasoning test. The aim of this paper is therefore to identify the issues of concern in developing a context-specific DIT, and in the use of such instruments in domain and group comparisons, using a study focussing on the context of tax practitioners as an example.³ As it stands, the development of context-specific instruments does show some variation, and because of the primary focus on ethics, this variation has not been explored. We therefore feel that an explication of the development of a context-specific instrument will make a significant contribution to the DIT literature. It is also important in opening a discussion of methodology that will maximise the value of studies comparing ethical reasoning between social and professional contexts, and between professional contexts. Such work presents some methodological challenges if comparisons across studies are to remain valid.

The remainder of the paper is structured as follows. Section “[Overview of the DIT](#)” provides an overview of the DIT and its development; Section “[The need for context-specific scenarios](#)” discusses the need for context-specific scenarios; Section “[Issues in the development of context specific scenarios](#)” examines issues relevant to their development, such as the number and nature of scenarios, the ‘items for consideration’ by participants (which test the ethical reasoning levels invoked) and the order in which they should appear; Section “[Measures from the DIT](#)” considers the measures of ethical reasoning derived from the DIT, such as the P score, U score, etc.; Section “[The approach to the development of tax-specific scenarios for this study](#)” details the approach to the development of specific scenarios for this study; Section “[The development of the ‘items for consideration’ for each scenario](#)” looks at the development of ‘items for consideration’ in the context of this study; Section “[Procedural and design issues in the context of group comparison](#)” outlines issues to be considered for group comparisons; and Section “[Conclusion](#)” offers conclusions.

Overview of the DIT

Rest developed the DIT in 1979 (Rest, 1979b) using Kohlberg’s cognitive development theory as a basis. Kohlberg had developed a model of ethical cognition based on interview studies in which participants

responded in an open-ended way to ethical dilemmas. His stage-sequence theory sets out three developmental levels. Individuals move upward through these three levels, beginning at the ‘pre-conventional morality’ level, onwards to the ‘conventional morality’ level, sometimes reaching the final and highest level known as ‘post-conventional morality’. Within each of the three levels there are two developmental steps, resulting in a total of six stages. The second stage in each level is a more advanced and organised form of the first.

The DIT is a self-administered, multiple-choice questionnaire, making use of the same ethical dilemmas used by Kohlberg in his original analysis (for example ‘Heinz and the Drug’, see Appendix A). Rest (1979a) developed the items for the questionnaire based on an interpretation of the stages in Kohlberg’s stage-sequence theory (see Table I). Although Kohlberg contended that, at any point in time, an individual will be at one of the six levels of ethical development, Rest posits that, while one stage might dominate an individual’s reasoning, he or she is never simply at one stage of cognition. Rest therefore views ethical development as a shifting distribution of responses from lower levels on the stage-sequence to higher levels. As an example, an individual may reason predominantly at stage four but also utilise some stage three and five reasoning. As ethical development takes place, he or she will reduce his/her reliance on stage three reasoning and increase reliance on stage five. The DIT therefore

TABLE I
Six stages of ethical reasoning

Pre-conventional: focuses on the individual	
Stage one	The morality of obedience: do what you are told
Stage two	The morality of instrumental egoism and simple exchange: let’s make a deal
Conventional: focuses on the group and relationships	
Stage three	The morality of interpersonal concordance: be considerate, nice and kind: you’ll make friends
Stage four	The morality of law and duty to the social order: everyone in society is obligated to and protected by the law
Post-conventional: move towards universalism and thus strive towards general norms	
Stage five	The morality of consensus-building procedures: you are obligated by the arrangements that are agreed to by due process procedures
Stage six	The morality of non-arbitrary social cooperation: morality is defined by how rational and impartial people would ideally organise cooperation

Adapted from Rest (1994).

assumes that a person can operate at many stages at once, and rather than attempting to assess the stage to which a person 'belongs', it instead measures the comprehension and preference for the principled level of reasoning (i.e., reasoning at stages five and six) (Rest et al., 1999).

Participants taking the DIT are presented with either six (long-form DIT) or three (short-form DIT) ethical dilemmas stated in third-person form.⁴ The dilemmas are presented as narratives describing the circumstances of the third party who is having to make a decision on how to act in the scenario. After reviewing the dilemmas, participants choose what the actor should do in the circumstances from three options offered: take the action, do not take the action, or "cannot decide". They are then asked to rate the importance of 12 considerations relating to the particular dilemma, indicating how important each is (in their opinion) in making the decision described in the scenario, using a five-level scale (great importance, much importance, some, little or no importance). The 12 statements were constructed by Rest to include considerations that would be prevalent at particular stages of ethical judgement development in each situation. Once the 12 items have been rated, the participant is asked to select the four items that he/she considers to be of most importance to the decision and to rank these in order of importance. The first of the DIT scenarios, 'Heinz and the Drug', is set out in Appendix A as an example.

In scoring the DIT, points are allocated to the considerations chosen as the four most important in each scenario. Four points are given for the most important, through to one point for the least important. These points are assigned to the stage of ethical reasoning which that consideration represents, thus producing a distribution of responses by level. This provides an overview of the range of stages involved in that participant's ethical reasoning. The points corresponding to the highest modes of ethical reasoning (stages five and six) are used to construct a single measure known as the 'P' score (standing for 'principled moral thinking') for each participant (Rest, 1994). The P score measures the percentage of stages five and six responses made by an individual for the entire three or six cases. Results of the P scores are expressed as a continuum from 0 to 95 (it cannot reach 100 owing to

the fact that, on three stories, there is no fourth possible principled item to choose from). Since the Rest (1979a) model is developmental and sequential, a higher P score implies a lower percentage of reasoning at lower levels. Thus the P score measures the percentage of a participant's thinking at a principled level.

The need for context-specific scenarios

The use of scenarios⁵ (sometimes known as vignettes) in business ethics research is very common. Indeed, Weber (1992) observes that scenarios have played an integral role in empirical business ethics research. Cavanagh and Fritzsche (1985) and Weber (1992) have carried out reviews of business ethics research using scenarios as a data-gathering technique, and both reviews have strongly supported the continued use of scenarios. Cavanagh and Fritzsche (1985) suggest that one of the advantages of using scenarios is that they give the researcher the ability to structure a research question in a manner that incorporates the complex issues which arise when making decisions in real-world environments.

However, Fraedrich et al. (1994) consider it highly questionable to assume that ethical behaviour in a business context is identical to ethical behaviour in a personal or social scenario. Because the situational factors present differ, they called for more work to be done to investigate ethical reasoning in a business context. Their view is shared by many researchers in the field of business ethics (see for example, Dellaportas et al., 2006; Elm and Nichols, 1993; McGeorge, 1975; Trevino, 1986; Weber, 1990; Welton et al., 1994).

Elm and Weber (1994) express concern about the validity of using hypothetical social dilemmas in the DIT to assess the ethical reasoning of managers, because they lack any business content. They suggest that presenting participants with broad social scenarios may result in the participants abandoning their managerial role when they are taking the test. Instead, they take on the mantle of a person in society at large and thereby compromise results. Elm and Weber (1994) expect that using dilemmas common in a business context would give rise to different ethical reasoning levels from dilemmas drawn from the broader socio-ethical context.

Jackall (1988) suggests that what is ethically acceptable at work may not be acceptable at home or outside the corporation. He argues that organisations develop their own structure of internal rules that override the ethical rules of society at large and that managers follow these organisational rules when they are at work but not outside work. The finding that managers use lower-level ethical reasoning to resolve business problems is consistent with an understanding of human behaviour based on cultural anthropology. Cultural anthropology finds that individuals play different roles that allow them to accept different values, norms and behaviours in different life domains (e.g. home and work).

Elm and Weber (1994) also raise the possibility of the existence of a 'story pull' effect in the DIT. This effect suggests that individuals use different levels of reasoning depending on their familiarity and experience with a particular scenario (Freeman and Giebink, 1979; Magowen and Lee, 1970; Weber, 1990; all cited by Weber, 1991). Familiarity is deemed to 'pull' or elicit a particular level of response (Rest, 1986). The influence of 'story pull' may account for significant differences in the participants' responses to the ethical dilemmas presented, depending on their ability to empathise with the character in the scenario or understand fully the problem contained in the dilemma, owing to similar personal experience (Elm and Weber, 1994). Straughan (1985) argues that hypothetical dilemmas necessarily lack first-hand urgency, which is an essential ingredient of genuine ethical experiences.

Ponemon (1990) contends that the DIT is based on conflicts in human life, which are on a higher plane than client needs. Kaplan et al. (1997) found that the ethical reasoning of taxpayers in a tax-paying context was lower than their reasoning on broader DIT issues, supporting Ponemon's contention.

Freeman and Giebink (1979) showed evidence of higher ethical reasoning for a remote dilemma than for more familiar dilemmas, concluding that the level of ethical judgement rendered was affected by the nature of the dilemma. Weber (1990, 1991) also found higher ethical level responses for the less familiar scenario than for the more familiar scenario placed in a business context, confirming the findings of Freeman and Giebink (1979). In his 1990 study, Weber investigated the ethical reasoning abilities of

managers using three dilemmas contained in the DIT. Two of the dilemmas were altered to reflect dilemmas that were typically faced in a corporate context. Managers reasoned at lower levels on these two dilemmas compared with the original and broader unaltered social dilemma. This finding was also supported by Welton et al. (1994).

Tolleson et al. (1996) conclude that remoteness from the subject matter has a significant impact on ethical reasoning and behaviour. Similarly, Kenny and Eining (1996) use attribution theory to posit that individuals behave differently when they are involved as actors, as opposed to being merely observers in the ethical dilemma.

Thus, many researchers contend that the familiarity embodied in realistic dilemmas set in a business context would serve to elicit a more accurate representation of the ethical reasoning of managers (Dellaportas et al., 2006; Ponemon, 1993; Weber, 1991). "Ethical behaviours tend to be situation specific and measures of ethical beliefs must also be situation specific" (Reckers et al., 1994, p. 826).

Others have risen to the challenge and have developed and used context-specific test instruments containing dilemmas particular to the specific business area they were investigating. These include accounting-specific tests (Thorne, 2000, 2001; Welton et al., 1994), broad business dilemmas and auditing scenarios (Massey, 2002; Weber, 1991), management vignettes (Lovisky et al., 2007) and a test of tax compliance (Fisher, 1997). Welton et al. (1994) suggest that accounting-specific dilemmas provide a better measure of ethical reasoning from a professional perspective, because they are more likely to elicit responses that reflect the level of reasoning used in resolving accounting dilemmas.

Similarly, researchers carrying out studies of professional groups have concerns that scenarios really need to be profession specific. This has led to the development of a number of profession-specific test instruments, including ethical sensitivity tests for dentists (Bebeau et al., 1985; Newell et al., 1985), counsellors (Volker, 1984) and nurses (Waithe et al., 1989), as well as Dennis-Escoffier and Fortin's (1991) tax ethics test, all cited by Welton et al. (1994).

Context-specific instruments to assess ethical reasoning are based on the DIT. The difference between these and the original test is in the nature of

the dilemmas presented to participants and in the related ‘items for consideration’ following each dilemma. This use of the DIT as a basis for context-specific instruments offers an opportunity to compare results across studies, and this in turn brings methodology into focus: instruments need ideally to be developed and administered in such a way that comparability with the original DIT and with each other is valid.

Issues in the development of context-specific scenarios

Rest (1979a, Chap. 4) gives very specific guidance on how the DIT is constructed and the rationale behind the structure of the test. These properties need to be preserved in context-specific tests, otherwise the integrity of the test will be damaged. Issues relating to construction of the DIT are explored below.

The number of scenarios

Rest warns that the use of a single test dilemma is not sufficient, recommending that enough data be collected “so that a general picture emerges that is not overly influenced by an idiosyncratic subject-situation interaction or by fluctuations in the subject’s thinking” (Rest, 1979a, p. 69). He also suggests building in some diversity of the test stimuli by using a variety of kinds of dilemmas, so that the P score is based on a representative sampling of possible situations. It is important that the administration of the test and the subsequent scoring of the instrument be as standardised as possible, to ensure that any differences in P score are due to actual developmental differences.

The most appropriate number of scenarios in a particular study is dependent on the purpose of the research and the length and context of the scenarios themselves (Weber, 1992). Too few scenarios may cause the responses to be biased by the limited scope of the issues. However, too many may lead to information overload and fatigue for the respondent (Weber, 1992). For studies making comparisons between contexts, one aim may be to produce an equality between the richness of the

two contexts, so that different biases and influences on the results are not present in the contexts to be compared.

Previous studies have taken a variety of approaches. Fisher (1997) combined the three-scenario version of the DIT and a three-scenario tax-paying test in his study, while both Thorne (2000) and Lovisky et al. (2007) combined the long-form DIT (six scenarios) with a six-scenario context-specific test. However, Thorne later reduced her context-specific test to four scenarios in the interests of time availability and subject fatigue. Welton et al. (1994) used four accounting scenarios in their test (no DIT scenarios were used in the study). Massey (2002) used the short-form DIT together with four auditing scenarios in her study. The studies utilising both the long-form DIT and a six-scenario context-specific test involved using students as participants and class time for the completion of the instruments (taking 2 h on average to complete) (Lovisky et al., 2007). Studies aimed at investigating the ethical reasoning of professionals have to be cognisant of the time it will take to complete the instrument. In this context, and bearing in mind our concerns in relation to comparability, we have opted to use the short-form version of the DIT together with a three-scenario tax-context test.

The short-form DIT includes the Heinz, Prisoner, and Newspaper scenarios. These three scenarios have the highest correlation of any three-story set within the full six-story version of the DIT (Rest, 1986). The P score from the short version is reported to correlate 0.93 with the P score based on six scenarios using a sample of 160 subjects, while on a larger sample of 1,080 subjects, the correlation is 0.91 (Rest, 1986, Sect. 5, p. 7). Rest (1986) also confirms that the shorter version has substantially the same properties as the six-scenario form.

The nature of the scenarios

In choosing scenarios, the aims of the study must be considered. The use of the DIT in different contexts suggests two obvious approaches, addressing different research questions, and supporting different types of comparison.

First, the types of dilemma present in different contexts are of theoretical interest in themselves.

Different contexts may give rise to different mixes of ethical issues, invoking different distributions of ethical reasoning stages. This issue may be best addressed by studying ethical dilemmas typically faced by individuals in the context (e.g. tax practitioners) and identifying the ethical reasoning stages in those dilemmas. DIT-style instruments addressing this sort of research question could be used to check which sorts of ethical reasoning are most often invoked, providing a complementary analysis of the ethical landscape for that context.

Secondly, instruments such as the DIT are also well suited to another research issue, namely, whether the context *of itself* alters the level of ethical reasoning applied by participants. From this viewpoint, it may be useful to consider dilemmas which have some echo in those used in the original DIT, thus identifying context-specific differences in ethical reasoning on similar issues.

Whichever approach is appropriate, it is important that the scenarios be as realistic as possible, so that they invoke the domain being studied clearly, as indicated in the discussion in Section “[The need for context-specific scenarios](#)”.

‘Items for consideration’

The ‘items for consideration’ used in the DIT are designed to reflect factors that would be considered by individuals at particular stages of ethical development. Rest suggests that the considerations following the ethical dilemma should be short and cryptic (“[s]hort statements discourage spurious interpretations”; Rest, 1979a, p. 90), using fragments of a line of ethical argument that would be recognisable by someone who understood the line of thinking, rather than providing a complete logical argument. If a participant is already thinking about the scenario along the lines suggested in the item for consideration, they will be attracted to that item (Rest et al., 1999). The items should also be comparable in length, sentence complexity and vocabulary, and should avoid building an argument for one course of action over another. Rest’s justification for this latter point is that, if a participant does not understand a consideration, nothing in the wording itself should advocate the choice of the consideration over another. In practice, this means

that development of the considerations has to be undertaken with extreme care, so that issues are addressed as neutrally as possible.

In addition to items for consideration derived from the dilemma, Rest also includes meaningless but complex-sounding items (Rest called these ‘M’ items – standing for ‘meaningless’). The M items are scattered throughout the list of considerations as a trap for participants who opt for particular considerations because of their apparent complexity and sophistication, rather than because of their meaning. Participants are alerted in the instructions to the fact that there will be pretentious-sounding but meaningless considerations in the instrument, and are instructed to rank these items very low when they come across them. If a participant chooses items that sound complex regardless, the likelihood is that they will select many M items. An M score is calculated for each participant completing the test instrument, and should participants attain too high an M score (greater than eight for the long-form DIT, Rest, 1990a, p. 13),⁶ their test is regarded as invalid and is discarded from the sample. This is a useful test of genuine response and it would seem sensible to include it in any instrument.

Rest also specifically includes some items (known as ‘A’ items) that reflect an ‘anti-establishment’ orientation. These considerations assume an understanding of stage four reasoning, but they blame the existing authority for being hypocritical and inconsistent with its own rationale. The A viewpoint condemns tradition and the existing social order for its arbitrariness or its corruption by the rich for the exploitation of the poor. While the A orientation is critical of the status quo, it offers nothing constructive in its stead (Bebeau and Thoma, 2003). The A point of view was popular in the 1960s, when it was suggested that the anti-establishment orientation may have represented a transition phase between conventional morality and principled morality (in other words a stage 4½). More recently, however, this has not been common (Bebeau and Thoma, 2003; Rest, 1990a), and some of the context-specific DIT-based tests that have been developed have omitted A items. However, there is a need to include A items in the tax-specific DIT, as views on taxation and ethical issues may well reflect an anti-establishment orientation, for example, if participants are opposed to the imposition of (certain) taxes or have adverse experience of dealing with tax authorities.

Ordering of 'items for consideration'

Rest (1979a) places a great degree of importance on the ordering of the considerations that follow an ethical dilemma. Although it is common in other sorts of instrument, Rest does not see random ordering as optimal. Instead, he recommends that low stage considerations appear early on the list, followed by a mix of considerations. Considerations at stages five and six should only be introduced after the fourth consideration on the list, and even then, be mixed with considerations at stages three and four. The rationale for this non-random ordering is to introduce lower stage ideas early in the list, so that less ethically developed participants can find their own ideas quickly and might therefore be less likely to project their thinking to a higher stage item. If the less ethically advanced participants read some considerations that they understand early on the list, they will then be more comfortable with rejecting the higher stage statements along with the M items (discussed above). Rest assumes that more ethically advanced participants will also understand the early low stage considerations and see their relevance and logic. However, they will view them as inadequate ways of constructing the dilemma. Rest suggests that they will therefore wade through the remainder of the list until they arrive at the higher stage items, realising that what is necessary is a principle for establishing a hierarchy of claims, such as the higher stage items suggest without actually spelling out. Rest posits that the more ethically developed participants will recognise in the higher stage considerations the mode of reasoning that they would employ when facing the dilemma. The less ethically developed participants, however, will not see anything more in the high stage considerations than in the M considerations.

Fisher (1997) points out that, while the lack of randomisation in the ordering of considerations has been challenged (Martin et al., 1977), strong support for Rest's assumptions is provided by Lawrence (1987) and Carroll and Rest (1981). Fisher (1997) follows the advice given by Rest in his ordering of the items for consideration in the Taxpayer Compliance Issues Test (TCIT). When undertaking comparative studies, comparability of treatment in each context is important, as any variations can otherwise be put forward as explanations for any

differences found in the results. Therefore, if the original DIT is used as one comparison context, following Rest's approach would be indicated for the instruments developed to cover any other contexts.

The number of 'items for consideration' at each stage

The DIT does not contain equal numbers of considerations at each stage of Kohlberg's stage sequence. The long form of the DIT uses six scenarios, each with 12 considerations. No stage one considerations are included, on the basis that the reading level and general personal development required to take the test instrument is such as virtually to eliminate the possibility of finding stage one participants (Rest, 1979a). Similarly, there very few stage two considerations, on the assumption that the participants have also advanced beyond this stage. It is expected that participants' reasoning will cluster around stages three, four and the principled stages, and these should have roughly equal number of considerations (Rest, 1979a).

Rest (1979a, p. 94) goes on to argue that the considerations and the number at each level should be designed on the basis of the real issues within each dilemma, rather than trying to achieve a balance of considerations. He suggests that, if considerations are contrived solely for the purpose of balancing the numbers of considerations for each stage, the instrument will be less functional, because participants will not be attracted to artificial considerations. He compares overly artificial considerations with distracter items in a multiple-choice test: so silly that no one ever selects them.

While the issues raised by contrived items are clearly important, our development of a tax-specific instrument indicates that there is often a wide range of potential items for consideration from which to choose. It may well be possible to produce an instrument that is both realistic and that mirrors the distribution of stages in the DIT.

As was the case with the number of scenarios used in context-specific instruments, analysis of the items for consideration in previous studies showed a variety of approaches. The researchers closely examined the context-specific tests that were developed in the areas of accounting/auditing (Massey, 2002; Thorne,

2000; Welton et al., 1994), management (Loviscky et al., 2007) and tax-paying (Fisher, 1997), as well as the original DIT and DIT-2 instruments.⁷ The published papers describing context-specific tests do not usually identify the ethical reasoning stage of each of the items for consideration included in the context-specific scenarios. However, having spent time developing a deeper understanding of the finer points of each of the stages, and using the items for consideration in the DIT as a guide [the stage for each consideration is identified in the DIT manual (Rest, 1990a)], a list of the considerations developed at each stage for the accounting/auditing and tax-paying specific tests was compiled, with the aim of getting a better understanding of the test construction process.

This analysis demonstrated that there was no standard way of developing instruments. In developing the TCIT, Fisher (1997, p. 150) included more stage two items (13.8% versus 6.9%) and fewer stage four items (19.4% versus 26.4%) than those contained in the DIT, arguing that the difference was slight and the number of issue statements assessing each stage compares favourably with the DIT. Massey (2002) lists 16 items for consideration after each of her four audit scenarios and uses a weighting system in order to compare auditing P scores for four scenarios with DIT P scores for three scenarios. No specific detail as to how either the accounting specific test scenarios or the items for consideration were developed is given by Welton et al. (1994) or by Thorne (2000). Neither of these papers gives any detail on the number of items at each stage of reasoning that were used in the context-specific studies compared with the DIT, or whether the number of items at each level varied. In Loviscky et al.'s (2007) management test, the number of issues representing the various stages of ethical judgement was standardised for each of the six scenarios. The list of items for consideration contained one meaningless item, one stage two item, three stage three items, three stage two items, three stage five items and one stage six item.

It appears that, while the Loviscky et al. (2007) and Welton et al. (1994) tests all omitted A items, the Massey (2002), Thorne (2000) and Fisher (1997) tests included some. All five of the context-specific tests discussed above included M items.

The variety of approaches indicates that analysis, and consideration of the implications of differences

between instruments will be a necessary part of comparative studies. If an objective of the instrument is to make comparisons across contexts to see if the context itself affects ethical reasoning on a particular issue, issues can arise if the number of items at each stage or the percentage of items at stages five and six differ from those in the original DIT. This is because the P index is the percentage of the overall score arising from these stages. If a set of scenarios has a lower number of stage five and six items than another set, this would limit the percentage of the overall score that could arise from these stages. Although weighted calculations can address this to some extent, differing contexts for judgment are still present. Parducci (1968) shows that the way a particular stimulus is judged can depend on the other stimuli in its context. The ethics of a particular action⁸ (e.g., 'poisoning a neighbour's dog whose barking bothers you') is judged less harshly when it was in a set including more serious acts (such as 'murdering your mother without justification or provocation') than when it was in a set including less serious acts (such as 'keeping a dime you find in a telephone booth'). Matching as far as possible the ethical reasoning levels of items for consideration to those of the DIT reduces the chances of any differences found in the study arising from this effect.

An example scenario from our study (that mirroring 'Heinz and the Drug', Appendix A) is included in Appendix B. The stages of moral reasoning included and the order in which stages are introduced in the list of 'items for consideration' are as in the original dilemma.

Measures from the DIT

The P score, noted above, is the basic score emerging from a DIT-based study. This score measures the percentage of points assigned to stages five and six by the participant's responses. There are, however, other measures which can provide important insights into reasoning in a particular context.

The U score

In an attempt to measure the extent to which individuals utilise ethical reasoning in their

decision-making and behaviour, Thoma et al. (1991) developed the U score. They focussed on the correlation between the items that the participant selects for consideration after each DIT scenario and the action choice that the participant makes at the end of each dilemma and before reading the considerations (i.e. whether the actor in the dilemma should take action, refrain from action or whether the participant is uncertain as to what the actor should do). The action choice that a participant selects is unrelated to the items for consideration and does not form part of the DIT P score for the participant. However, each of the items for consideration following the scenario implies an action choice one way or another. Once the participant has selected his/her preferred items for consideration, an implied action choice can be identified. This implied action choice is based solely on justice reasoning. If the actual decision made by the participant is consistent with the decision implied by the items selected, then it is reasonable to assume that justice reasoning is driving both decisions. If, on the other hand, there is a mismatch between the actual and the implied decision, then it seems that the participant used a non-justice-based system to arrive at the action choice. Hence, the P score of the individual may not be an accurate predictor of his/her behaviour. This might occur, for instance, in an individual who uses a religious doctrine to decide what is ethically right even if this contradicts what he/she thinks is fair or just. Thoma et al. (1991) propose that participants who employ justice reasoning can be identified by comparing an actual and implied decision based on DIT responses. This is called the U score (U standing for 'utiliser'). The U score represents an estimate of the agreement between the actual and the implied decisions of the participant. This estimated agreement in turn increases the predictability of ethical judgement to behaviour (Thoma, 1994). This again is a useful calculation in any test, and this score provides another potential point of comparison between contexts. It is possible that actual and implied decisions relate in different ways in different contexts. For example, one context may see higher agreement between the two decisions, or implied decisions may seem similar in two contexts, but actual decisions are not. Consideration of the U score is, therefore, useful in providing insight into the differences between thought processes and motivations across contexts.

Consistency check

Apart from the internal reliability check using the M score as outlined above, there is also a second check on subject reliability called the 'consistency check'. There is always a danger that a participant will not read the considerations properly, but instead merely tick random considerations. Rest (1979a) built a consistency check into the DIT to detect random ticking. After rating the importance of the 12 considerations that have, in their opinion, the most significant influence on the resolution of each of the scenarios, participants are then asked to rank their top four considerations in order of importance (most important, second most important, third most important and fourth most important). The consistency check then compares a participant's rankings of his/her top four considerations with his/her ratings of all 12 considerations. If a participant has ranked a consideration in first place of importance, then the rating of that consideration should be as high or higher than the ratings of all other considerations. Any discrepancies that occur are counted (e.g. the number of items that are rated higher than the item ranked as most important), and if the number of inconsistencies exceeds a certain limit, the test is removed from the sample. Rest indicates that the invalidation of tests on the basis of both the M score check and the consistency check is in the 2–15% range (Rest, 1990b, p. 15).

The approach to the development of tax-specific scenarios for this study

Fredrickson (1986) sets out a detailed methodology for developing realistic scenarios. He argues that realism in all aspects of the scenario is crucial for stimulating the interest and, therefore, involvement of the respondent. This involvement enables the researcher to represent more closely real-world conditions and elicit more realistic responses. The role that interviews play in developing realistic scenarios is considered critical in providing the researcher with a detailed knowledge of the industry needed to prepare scenarios that will stimulate the respondents' participation (Fredrickson, 1986).

There is a lack of information as to how the scenarios contained in context-specific tests based on

the DIT were developed for other studies in the business domain. Welton et al. (1994) refer to having the wording of their scenarios validated by a DIT scholar. Thorne (2000) adapted her scenarios from prior accounting ethics literature to the format and style of the traditional DIT, while a survey of audit seniors' perceptions about various ethical situations was used to assist in the selection of four topics which formed the basis of an audit-context test (Massey, 2002). Consultation with experts (professional, academic or both) is also cited as having assisted in the development of realistic scenarios (Loviscky et al., 2007; Massey, 2002; Thorne, 2000).

A starting point in developing any instrument is the relevant literature. While we have concentrated thus far on the DIT literature and development within it of context-specific issues, there may also be a literature based on the ethics of the particular context being studied. In the case of our study, there is further literature, not DIT related or connected with ethical reasoning, but indicative of areas in taxation likely to give rise to ethical issues for tax practitioners. Marshall et al. (1998) identified broad areas that are considered in the Australian literature as particularly problematic, among which are included: confidentiality and whistle-blowing; failure to maintain technical competence, do adequate research or make reasonable enquiries into a client's business; ability to communicate 'bad news' to a client; conflicts of interest between the personal gain of the practitioner and the client or between different clients; failure to correct an error made by a tax authority; and many issues concerning different aspects of tax-avoidance schemes and aggressive interpretations of tax law.

A second stage is to explore issues with practitioners to develop an understanding of the types of ethical dilemmas that are actually encountered in practice in order to inform the development of realistic scenarios. This was done in our study by means of semi-structured interviews. Semi-structured interviews allow the interviewer to prompt the interviewee to consider different types of dilemma, but then allow the interviewee to develop unprompted consideration of the ethical issues that could emerge from different dilemmas considered.

After undertaking a thorough review of the scenarios already developed for previous studies and the dilemmas discussed with and shared by interviewees,

the researchers selected four dilemma types with a view to moulding them into short vignettes. In our case, four scenarios were developed, although only three would be used in the new test. It was considered prudent to develop an extra scenario to allow for the scenario that was considered least suitable by the pilot test participants to be eliminated. Given that the targeted respondents to the tax-context DIT work in a variety of organisational types and sizes, following Sims (1999), the scenarios were written in a general manner to allow the respondents to imagine themselves in that situation. As another aim of this study was to allow comparison between tax practitioners/specialists and non-tax specialists on specific issues, particular attention was paid to language, and the use of jargon was eliminated. Scenarios were written in a form that would be as clear as possible to non-tax specialists. The wording was tested in a pilot study to check this.

An important issue in development was the refinement of scenarios. Once the four scenarios were drafted, four tax academics who had formerly worked in tax practice were asked to provide feedback as to how they could be strengthened. They were also asked to suggest other scenarios that might be suitable for use in a tax-specific DIT. On the basis of this feedback, three of the original scenarios were amended slightly, one of the scenarios was rejected and two new scenarios were drafted.

One of the difficulties in designing the scenarios was to balance realism with the need to represent a genuine ethical problem. The scenarios also had to be easy to read and understand. Freeman and Giebink (1979) found significant differences in participants' responses to a non-business scenario, depending upon the participants' familiarity with the issue. Weber (1992) warns that scenario-based research is vulnerable to the criticism of presenting the participant with a task framed by unrealistic circumstances, making a generalisation difficult. However, in order to represent a genuine ethical problem, the scenarios could not be confined to issues that occur on a day-to-day basis in tax practice and for which there is an industry-wide accepted solution. A balance had to be achieved between realism and ethical difficulty. The DIT scenarios themselves describe scenarios that, while conceivable, are not common.

In order to ensure that the scenarios were realistic in a tax practice context, and following the example

of Reckers et al. (1994), who had their scenarios reviewed by a number of professionals, a survey instrument was developed including the five scenarios emerging after consultation with tax academics. The survey was distributed to 30 Irish tax practitioners working in firms of different sizes. Using a five-point Likert scale, the survey presented tax practitioners with each of the five scenarios in turn and asked for their views on the following.

- How easy the scenario was to read and understand (1 = difficult to read and understand, 5 = easy to read and understand).
- The degree to which the scenario presented an ethical dilemma for the tax practitioner concerned (1 = no dilemma at all, 5 = a serious ethical dilemma).
- The realism of the scenario (1 = very unrealistic, 5 = very realistic).

The survey then asked the tax practitioner to list the types of issues that he/she might consider if he/she were faced with the scenario and had to decide upon a course of action. Finally, the practitioner was asked for additional comments.

Although all scenarios scored highly in terms of readability and understandability (average score of 4.67 out of 5), there was some variety in terms of how they were rated on presenting a dilemma (from 3.3 to 4.5 out of 5) and on realism (from 2.9 to 3.57 out of 5). The two scenarios with the combined lowest scores for degree of dilemma presented and realism were eliminated. The scenarios eliminated were the two that had been developed after consulting with tax academics, leaving the three originally drafted scenarios. These three scenarios were amended further to reflect some of the comments made by practitioners in the survey. Finally, one entirely new scenario was developed after in-depth discussions with two of the tax practitioners who completed the survey and were anxious to contribute to the study. This new scenario was based on a real situation that occurred recently in Ireland. Therefore, while it was not tested on the 30 pilot tax practitioner participants, its realism cannot be questioned.

Four robust scenarios therefore emerged from all of the consultation and discussion outlined above, three of which had been through a number of

rounds of alteration based on the consultations and the survey mentioned above. A scenario which had been developed involving the discovery of a serious mistake made by the tax practitioner was eventually eliminated because it was quite similar to one of the other three.

The surviving three scenarios were finally further refined so that, to the maximum degree possible, they mirrored the three original DIT scenarios used in the context of the psychological responses triggered by each story. The first scenario was refined to mirror that of 'Heinz and the Drug'. In both these scenarios an irrevocable loss will occur, unless the actor does something to prevent it (see Appendix A. In 'Heinz and the Drug' the actor faces the potential death of his wife. In the tax-specific scenario, the 'death' envisaged is the demise of a company and the loss of the employment it provides). The second of the tax-context scenarios evokes a response similar to that triggered in the 'Escaped Prisoner' dilemma. In both cases the actor is faced with having to decide whether or not to inform on someone who has done something illegal, but has also contributed to society in a very positive manner. The third scenario and 'The Newspaper' scenario involve an issue that has been agreed in advance, although public outcry has added another dimension to the situation.

The development of the 'items for consideration' for each scenario

Once the scenarios themselves were roughly drafted, the next stage was to develop the list of 'items for consideration' for each scenario.

In order to develop the items for consideration for her test, Thorne (2000) engaged four chartered accountants and former auditors to read each case and identify the key factors that would influence their resolution of each case. The factors were then formulated into a format consistent with the items for consideration in the DIT. M and A items were also generated to accommodate the use of the traditional DIT as a template for the accounting-specific instrument. Fisher (1997) drafted the 12 issue statements following each of his taxpayer dilemmas from issues raised spontaneously during informal interviews with graduate students, tax practitioners,

tax faculty and taxpayers. The issue statements were refined consistent with factors deemed important in previous tax compliance research and to items contained in the DIT. Fisher also included both A and M issue statements in his test.

It is critical that the 12 considerations drafted to follow each of the tax-specific scenarios echo the different stages of ethical reasoning. The items assess reasoning from stage two to stage six. It is acknowledged that it is inappropriate to include stage one items, as the level of reading necessary to complete the test will eliminate any person at this level of reasoning (Fisher, 1997). The items themselves must be relevant to the tax-specific scenarios developed and flow naturally. However, they also need to be drafted consistently with the wording and format of the DIT items for consideration at each stage of ethical development.

Following the example of Fisher (1997), the items for consideration for the tax practice scenarios were informed by comments gathered during interviews with tax practitioners, as to their perceptions of the dilemmas they raised. Four tax academics (who had formerly worked as practitioners and were also referred to above) also contributed their views on the issues a tax practitioner might consider, when faced with the specific dilemmas raised. Finally, the survey completed by 30 tax practitioners elicited their views on the issues a tax practitioner might consider when faced with the five dilemmas that had been drafted for the study. The final scenario, however, was not included in the survey, as it was only developed after the survey had been done.

More than 12 considerations were initially drafted for each scenario, so that the 12 most appropriate could be selected. Efforts were also made to write statements of comparable length, sentence complexity and vocabulary at each stage.

Once the finalised tax scenarios and the items for consideration for each were drafted, a panel of three tax academics and one currently practising tax practitioner were asked independently to code the items. The coders were asked to indicate the stage that they considered each of the items represented, using Table I as a guide. While there was a very high level of consistency between the stages identified by the independent coders and the researchers (80%), the items for consideration were altered slightly to address any ambiguity in the stages identified by the panel.

Following some of the prior research involving the development of a context-specific DIT in the accounting and auditing area (Massey, 2002; Thorne, 2000; Welton et al., 1994), we have also had the tax-context test instrument validated by two recognised DIT scholars. This step involved the scholars reviewing the tax scenarios and considerations, and independently classifying each of the 'items for consideration' according to the stages of ethical development. This provided a check on the reliability with which the items were rated, and was useful in addressing any remaining ambiguity in the classification of items. Some further modifications were made to the scenarios and the items for consideration on the basis of the feedback from our two DIT scholars.

Table II summarises the process of developing both scenarios and items for consideration in a tax practice context.

Procedural and design issues in the context of group comparison

As indicated above, the intention is to administer the tax-specific scenarios developed here to tax practitioners working in Ireland and the UK and, as a second stage in the research, to administer the instrument to 'ordinary people', who are not employed in the tax or accounting professions. In both stages the short-form DIT and the tax-specific scenarios will be administered. In order to allow comparisons to be made between the ethical reasoning that takes place in the two contexts and in the two different groups of participants, steps have to be taken to ensure that extraneous variables that might otherwise provide explanations for the results of the comparisons are eliminated as far as is possible.⁹ These issues are not always clearly discussed in reports of studies of ethical reasoning, so we provide the details of our approach here to broaden the discussion on design considerations.

Some steps to achieve the elimination of extraneous variables have already taken place in the development of the scenarios and items for consideration. Dilemmas were chosen which mirrored issues in the short-form DIT. The number of considerations at each stage of ethical reasoning was also matched, thus addressing the potential for 'context effects' in the sense put forward by Parducci, and as discussed above. As the

TABLE II
Summary of scenario development process

-
- 1 Literature review of extant studies in accounting and tax utilising scenarios or indicative otherwise of ethical issues facing taxation practitioners
 - 2 Identifying some of the common themes
 - 3 Interviews with practitioners on the types of ethical scenarios encountered by practitioners in a tax context
 - 4 Both the literature review and the interviews led to four draft scenarios being developed
 - 5 These were discussed with four tax academics and amended as a result of their feedback. One of the original scenarios was dropped and two new scenarios developed
 - 6 The five scenarios that made it through the consultation process above were sent to 30 tax practitioners to elicit their view on: (1) the clarity and readability of the scenario, (2) the degree to which it would present an ethical dilemma in tax practice and (3) the realism of the scenario
 - 7 The practitioners surveyed above were also asked to list the types of issues they would consider if faced with each of the five dilemmas (this was asked in order to gather data that would inform the development of the ‘items for consideration’)
 - 8 On the basis of the results of the survey of tax practitioners, three of the five scenarios were further strengthened, two were dropped and a new scenario involving an aggressive tax interpretation was developed
 - 9 Three of the four surviving scenarios were selected on the basis that they best matched the type of psychological responses triggered in the three original DIT scenarios being used
 - 10 Items for consideration were drafted to follow each scenario based on issues raised by tax practitioners during the interviews mentioned above in step 3, tax academics during the discussions mentioned in step 5 and tax practitioners during the survey mentioned in step 7
 - 11 A panel of three tax academics and one practising tax practitioner independently coded the items for consideration following each tax scenario, linking each back to the ethical reasoning stage detailed in Table I
 - 12 Two recognised DIT scholars reviewed the scenarios and items for consideration
 - 13 The completed instrument consisting of instructions, three DIT scenarios, three tax context scenarios and a general information questionnaire was pilot-tested
-

ordering of items for consideration in the DIT instrument is specifically non-random, a similar approach was taken in the tax-specific DIT and the orders were matched to the short-form DIT used.

If comparisons are to be made between contexts, another issue that needs to be addressed is the potential for order effects. If all participants were to answer the short-form DIT scenarios first and then the tax scenarios, differences might arise as a result of the first set priming participants to think in a particular way. A counter-balanced design is therefore being used, with one version of the questionnaire having the DIT scenarios first and the other, the tax scenarios. This will allow us to check for order effects and, if they are present, adopt an averaged view in comparative analysis. As the literature does not indicate that the scenarios for the short-form DIT are presented in different orders to sets of participants, further order randomisation within each context was not used, for comparability purposes.

Conclusion

This paper has discussed in detail the development of a context-specific research instrument, designed to test the ethical reasoning processes of tax practitioners and, in due course, ‘ordinary people’ in the tax context. The research instrument adapts the general socio-ethical reasoning test, the DIT, as developed by Rest (1979a, b) into a context-specific instrument using tax-based scenarios, in line with the recognition in the literature of a need for domain-specific scenarios.

In order to allow for comparisons of reasoning in the tax versus the general social context, the tax DIT developed uses the three original scenarios developed by Rest for the short-form DIT in combination with three specially developed tax scenarios.

The paper explores the methodological issues in developing context-specific versions of the DIT in detail, including a consideration of the issues that arise in comparative studies. The considerations that shaped the

development of the original DIT by Rest have been carefully explored and utilised to produce tax-specific scenarios which mirror as exactly as possible the three original scenarios in the short-form DIT, in terms of psychological triggers, and the number and nature of scenarios. The process of validating different elements in the research instrument, by reference to other context-specific DITs has been explored, and we have added to this literature by giving explicit detail on how the tax-specific tests have been validated.

Our analysis of previous studies shows a variety of approaches in the number of scenarios, items for consideration and items for consideration at each stage of reasoning. This variation suggests that a discussion of methodological issues for context-specific and comparative work would be both timely and useful in the development of future instruments. In our explication of the process, and in highlighting implications for group comparisons and cross-study comparisons, we provide a starting point for this discussion in the comparatively focussed context-specific DIT literature. We thus provide information that can facilitate future comparisons of ethical reasoning between social and professional contexts, and between professional contexts.

Appendix A

DIT scenario one: Heinz and the Drug (Rest, 1986)¹⁰

(Note that the stages placed after the items for consideration are not present in the instrument used with participants).

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

Should Heinz steal the drug?

Should steal it Can’t decide Should not steal it

Rate the following 12 items in terms of importance	Great	Much	Some	Little	No
1. Whether a community’s laws are going to be upheld. (Stage 4)					
2. Isn’t it only natural for a loving husband to care so much for his wife that he’d steal? (Stage 3)					
3. Is Heinz willing to risk getting shot as a burglar or going to jail for the chance that stealing the drug might help? (Stage 2)					
4. Whether Heinz is a professional wrestler, or has considerable influence with professional wrestlers. (M item)					
5. Whether Heinz is stealing for himself or doing this solely to help someone else. (Stage 3)					
6. Whether the pharmacist’s rights to his invention have to be respected. (Stage 4)					
7. Whether the essence of living is more encompassing than the termination of dying, socially and individually. (M item)					
8. What values are going to be the basis for governing how people act towards each other. (Stage 6)					
9. Whether the pharmacist is going to be allowed to hide behind a worthless law which only protects the rich anyway. (A item)					
10. Whether the law in this case is getting in the way of the most basic claim of any member of society. (Stage 5)					
11. Whether the pharmacist deserves to be robbed for being so greedy and cruel. (Stage 3)					
12. Would stealing in such a case bring about more total good for the whole society or not? (Stage 5)					

From the list of questions above, select the four most important:

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

Appendix B

Tax DIT scenario one: capital allowances

Anne is a tax practitioner with an accounting firm. She is working on a capital allowances claim to benefit one of her firm’s corporate clients that is in financial distress. Despite profitable trading, the client has suffered severe cashflow problems as a result of adverse economic conditions. The capital allowances claim relates to a new factory building and will

significantly reduce taxable corporate profits (and thus the tax the client has to pay). To be eligible for capital allowances the factory has to be in use at the end of the client’s financial year. Without the reduction in tax from the capital allowances, it is unlikely that the company will survive, which will result in 5,000 employees losing their jobs.

It is now a month since the client’s financial year end and Anne has asked the financial controller for documentary evidence that the factory was in use at the end of the financial year. The financial controller sends her a copy of the minutes of the latest directors’ board meeting. The last item on the board minutes notes that the factory premises became fully operational on the last day of the financial year. However, Anne is convinced that this was not the case as she drives past the factory every evening and it is clearly unoccupied. However, she also knows that the company will not survive if the capital allowances cannot be claimed.

Should Anne file a tax return claiming capital allowances for the financial year?

She should Can’t decide She shouldn’t

Rate the following 12 items in terms of importance	Great	Much	Some	Little	No
1. Would it be fair to all the taxpayers who have met the legal requirements to claim capital allowances if one client is permitted to claim allowances without meeting the criteria?					
2. What impact will the company’s demise have on the accounting firm Anne works for?					
3. Under self-assessment, once Anne has the proper documentation on file her position is covered regardless of whether the building is actually ‘in use’.					
4. Whether Anne’s notification juxtaposes immediate Revenue authority cognisant of the client’s actions.					
5. Whether Anne and the financial controller are close friends.					
6. Isn’t a tax practitioner required to file an accurate tax return?					
7. Aren’t capital allowances the essence of alternative displacement?					
8. Which values best determine how tax practitioners should interact with their clients and engage with the tax legislation?					
9. Whether a tax system that includes random and meaningless definitions ought to be completely abandoned.					
10. Whether the saving of 5000 jobs will bring about the best result for society as a whole.					
11. Whether the firm’s reputation will be damaged if the claim is subsequently challenged.					
12. Whether it is socially acceptable for management inadequacies to deprive employees of their opportunity to earn a living.					

From the list of questions above, select the four most important:

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

Notes

¹ It has been used in more than 1,000 studies (Rest, 1987, cited by Fisher, 1997) and there are more than 400 published articles and books on the subject (Rest et al., 1999). Although it is possible to draw a distinction between ‘ethics’ and ‘morality’ (see, for example, Singer, 1994, p. 7), one should note that the terms ‘ethical’ and ‘ethics’ are frequently used interchangeably with ‘moral’ and ‘morality’ without distinction of meaning throughout the literature in this area. In this paper we use ‘ethical’ and ‘ethics’ as far as possible, except where directly quoting from sources, or where a term has become standardised, such as in describing the ‘conventional morality’ levels in relation to Kohlberg’s work (see Section “[Overview of the DIT](#)”).

² For example, in business-related areas there are: accounting-specific tests (Thorne, 2000, 2001; Welton et al., 1994), broad business dilemmas and auditing scenarios (Massey, 2002; Weber, 1991), management vignettes (Lovinsky et al., 2007) and a test of tax compliance (Fisher, 1997).

³ Tax practitioners (sometimes also referred to as ‘tax preparers’, ‘tax advisers’ or ‘tax agents’) are often criticised for advising clients how to decrease their tax liabilities, especially where such advice involves the use of schemes which, while they do not contravene the word of the law, do breach its spirit or intention. As schemes may then result in governments collecting less tax revenue to use for socio-economic purposes, there is an undoubted ethical aspect to practitioners’ work. Their work therefore provides a context of interest for research into ethical reasoning in a professional situation.

⁴ The complete version of the DIT comprises six dilemmas. These are: (i) Heinz and the Drug: examines whether Heinz should steal a drug that might save the life of his wife who is dying from cancer; (ii) Student Take-Over: examines university students’ freedom of speech and their right to protest; (iii) 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; (iv) The Doctor’s Dilemma: examines the issue of euthanasia; (v) Webster: examines discrimination against minorities; and (vi) Newspaper: examines freedom of speech as it relates to the press. The short version of the DIT contains only three of these dilemmas (Heinz, Prisoner and Newspaper stories).

⁵ Defined as “short descriptions of a person or a social situation which contain precise references to what are thought to be the most important factors on the decision-making or judgement-making processes of respondents” (Alexander and Becker, 1978, p. 94).

⁶ On the short form of the DIT, if a subject’s raw M score is greater than 4, then the questionnaire should be invalidated (Rest, 1990b).

⁷ DIT-2 is a further development of the DIT. However, it has not yet been as widely used and validated in empirical studies as the original DIT.

⁸ Measured on a 1–5 scale, where 1 is ‘not particularly bad’ and 5 is ‘extremely evil’.

⁹ This is a standard consideration in research study design, and issues to consider in the context of within-subjects and between-subjects studies are discussed in detail in methodology texts such as Heiman (1999).

¹⁰ The Heinz scenario has been slightly altered from the original Rest (1986) version in order to update the language slightly for an Irish and UK jurisdiction context. The original dollar figure mentioned in the scenario has been changed to pounds and the word ‘druggist’ has been replaced by ‘pharmacist’.

Acknowledgements

We gratefully acknowledge comments made by participants at the Tax Research Network Conference in Galway, Ireland, in September 2008, the European Business Ethics Conference in Antalya, Turkey, in October 2008 and at the American Accounting Association North East Region Conference in Newport Rhode Island, USA, in November 2008. Particular thanks are due to Professors Dawn Massey and Linda Thorne for their expert feedback on our tax context scenarios and considerations. We are grateful to the anonymous *Journal of Business Ethics* referees and the editor for very helpful comments made during the reviewing process. Thanks are also due to the Centre for the Study of Ethical Development in Alabama (formerly Minnesota) for granting us permission to use the DIT in our research.

References

- Alexander, C. and H. J. Becker: 1978, 'The Use of Vignettes in Survey Research', *The Public Opinion Quarterly* **42**(1), 93–104.
- Bebeau, M., J. Rest and C. Yamoor: 1985, 'Measuring the Ethical Sensitivity of Dental Students', *Journal of Dental Education* **49**(4), 225–235.
- Bebeau, M. and S. Thoma: 2003, *Guide for DIT-2*, 3rd Edition (Center for the Study of Ethical Development, Minnesota).
- Carroll, J. and J. Rest: 1981, 'Development in Moral Judgment as Initiated by Rejection of Lower-Stage Statements', *Journal of Research in Personality* **15**(4), 538–544.
- Cavanagh, S. and D. Fritzsche: 1985, 'Using Vignettes in Business Ethics Research', in L. Preston (ed.), *Research in Corporate Social Performance and Policy* (JAI, Greenwich, CT), pp. 279–293.
- Dellaportas, S., B. J. Cooper and P. Leung: 2006, 'Measuring Moral Judgment and the Implications of Cooperative Education and Rule-Based Learning', *Accounting and Finance* **46**(1), 53–70.
- Dennis-Escoffier, S. and K. A. Fortin: 1991, Testing Moral Development in the Tax Curriculum. Paper Presented at the 43rd Annual Meeting of the Southeast Region American Accounting Association, Birmingham.
- Elm, D. R. and M. L. Nichols: 1993, 'An Investigation of the Moral Reasoning of Managers', *Journal of Business Ethics* **12**(11), 817–833.
- Elm, D. R. and J. Weber: 1994, 'Measuring Moral Judgment: The Moral Judgment Interview or the Defining Issues Test?', *Journal of Business Ethics* **13**(5), 341–355.
- Fisher, D.: 1997, 'Assessing Taxpayer Moral Reasoning: The Development of an Objective Measure', *Research on Accounting Ethics* **3**, 141–171.
- Fraedrich, J., D. M. Thorne and O. C. Ferrell: 1994, 'Assessing the Application of Cognitive Moral Development Theory to Business Ethics', *Journal of Business Ethics* **13**(10), 829–838.
- Fredrickson, J. W.: 1986, 'An Exploratory Approach to Measuring Perceptions of Strategic Decision Process Constructs', *Strategic Management Journal* **7**(5), 473–483.
- Freeman, S. J. M. and J. W. Giebink: 1979, 'Moral Judgment as a Function of Age, Sex and Stimulus', *The Journal of Psychology* **102**(1), 43–47.
- Heiman, G. W.: 1999, *Research Methods in Psychology*, 2nd Edition (Houghton Mifflin, Boston, MA).
- Jackall, R.: 1988, *Moral Mazes* (Oxford University Press, New York).
- Kaplan, S. E., K. J. Newberry and P. M. J. Reckers: 1997, 'The Effect of Moral Reasoning and Educational Communications on Tax Evasion Intentions', *Journal of the American Taxation Association* **19**(2), 38–54.
- Kenny, S. Y. and M. M. Eining: 1996, 'Integrating Ethics into Intermediate Accounting: An Experimental Investigation Incorporating Attribution Theory', *Research on Accounting Ethics* **2**, 159–186.
- Kohlberg, L.: 1973, *Collected Papers on Moral Development and Moral Education* (Laboratory of Human Development, Harvard University, Cambridge, MA).
- Lawrence, J.: 1987, 'Verbal Processing of the Defining Issues Test by Principled and Non-Principled Moral Reasoners', *Journal of Moral Education* **16**(2), 115–130.
- Loviscky, G. E., L. K. Trevino and R. R. Jacobs: 2007, 'Assessing Managers' Ethical Decision-Making: An Objective Measure of Managerial Moral Judgment', *Journal of Business Ethics* **73**(3), 263–285.
- Magowen, S. A. and T. Lee: 1970, 'Some Sources of Error in the Use of the Projective Method for the Assessment of Moral Judgment', *British Journal of Psychology* **61**(4), 535–543.
- Marshall, R. L., R. W. Armstrong and M. Smith: 1998, 'The Ethical Environment of Tax Practitioners: Western Australian Evidence', *Journal of Business Ethics* **17**(12), 1265–1279.
- Martin, M., M. Shafto and W. Van Deirse: 1977, 'Reliability, Validity and Design of the Defining Issues Test', *Development Psychology* **13**(5), 460–468.
- Massey, D. W.: 2002, 'The Importance of Context in Investigating Auditors' Moral Abilities', in B. M. Schwartz (ed.), *Research on Accounting Ethics*, Vol. 8 (JAI, Greenwich, CT), pp. 195–247.
- McGeorge, C.: 1975, 'The Susceptibility to Faking of the Defining Issues Test of Moral Development', *Developmental Psychology* **11**(1), 108.
- Newell, K., L. Young and C. Yamoor: 1985, 'Moral Reasoning in Dental Hygiene Students', *Journal of Dental Education* **49**(4), 79–84.
- Parducci, A.: 1968, 'The Relativism of Absolute Judgments', *Scientific American* **219**, 84–90.
- Ponemon, L. A.: 1990, 'Ethical Judgment in Accounting: A Cognitive-Developmental Perspective', *Critical Perspectives on Accounting* **1**(2), 191–215.
- Ponemon, L. A.: 1993, 'The Influence of Ethical Reasoning on Auditors' Perceptions of Management's Competence and Integrity', *Advances in Accounting* **11**, 1–29.

- Reckers, P., D. Sanders and S. Roark: 1994, 'The Influence of Ethical Attitudes on Taxpayer Compliance', *National Tax Journal* **47**(4), 825–836.
- Rest, J.: 1979a, *Development in Judging Moral Issues* (University of Minnesota Press, Minneapolis, MN).
- Rest, J.: 1979b, *Defining Issues Test* (University of Minnesota Press, Minneapolis, MN).
- Rest, J.: 1986, *DIT Manual*, 3rd Edition (University of Minnesota Centre for the Study of Ethical Development, Minneapolis, MN).
- Rest, J.: 1987, *Guide for the Defining Issues Test: How to Use the Optical Scan Forms and the Center's Scoring Service* (Center for the Study of Ethical Development, Minneapolis, MN).
- Rest, J.: 1990a, *DIT Manual*, 3rd Edition (University of Minneapolis Press, Minneapolis, MN).
- Rest, J.: 1990b, *Guide for the Defining Issues Test* (Centre for the Study of Ethical Development, Minneapolis, MN).
- Rest, J.: 1994, 'Background: Theory and Research', in J. R. Rest and D. Narvaez (eds.), *Moral Development in the Professions* (Erlbaum, Hillsdale, NJ), pp. 1–26.
- Rest, J., M. Bebeau and S. Thoma: 1999, *Postconventional Moral Thinking: A Neo-Kohlbergian Approach* (Lawrence Erlbaum, Mahwah, NJ).
- Sims, R.: 1999, 'The Development of Six Ethical Business Dilemmas', *Leadership & Organization Development Journal* **20**(4), 189–197.
- Singer, P.: 1994, 'Introduction', in P. Singer (ed.), *Ethics* (Oxford University Press, Oxford), pp. 4–13.
- Straughan, R.: 1985, 'Why Act on Kohlberg's Moral Judgments?', in S. Modgil and C. Modgil (eds.), *Lawrence Kohlberg: Consensus and Controversy* (Falmer, Philadelphia), pp. 149–161.
- Thoma, S.: 1994, 'Moral Judgments and Moral Action', in J. R. Rest and D. Narvaez (eds.), *Moral Development in the Professions: Psychology and Applied Ethics* (Lawrence Erlbaum, Hillsdale, NJ), pp. 199–211.
- Thoma, S., J. Rest and M. L. Davison: 1991, 'Describing and Testing a Moderator of the Moral Judgment and Action Relationship', *Journal of Personality and Social Psychology* **61**(4), 659–669.
- Thorne, L.: 2000, 'The Development of Two Measures to Assess Accountants' Prescriptive and Deliberative Moral Reasoning', *Behavioral Research in Accounting* **12**(1), 139–169.
- Thorne, L.: 2001, 'Refocusing Ethics Education in Accounting: An Examination of Accounting Students' Tendency to Use Their Cognitive Moral Capability', *Journal of Accounting Education* **19**(2), 103–117.
- Tolleson, T., B. D. Merino and A. G. Mayer: 1996, 'Applying Behavioral Models as Prescriptions for Ethics in Accounting Practice and Education', *Research on Accounting Ethics* **2**, 21–49.
- Trevino, L. K.: 1986, 'Ethical Decision Making in Organizations: A Person–Situation Interactionist Model', *Academy of Management Review* **11**(3), 601–617.
- Trevino, L. K.: 1992, 'Moral Reasoning and Business Ethics: Implications for Research, Education, and Management', *Journal of Business Ethics* **11**(5–6), 445–459.
- Volker, J.: 1984, Counseling Experience, Moral Judgment, Awareness of Consequences and Moral Sensitivity in Counseling Practice. Unpublished Doctoral Dissertation, University of Minnesota, Minneapolis.
- Waithe, M., L. Duckett, K. Schmitz, P. Crisham and M. Ryden: 1989, 'Developing Case Situations for Ethics Education in Nursing', *Journal of Nursing Education* **28**(4), 175–180.
- Weber, J.: 1990, 'Managers' Moral Reasoning: Assessing Their Responses to Three Moral Dilemmas', *Human Relations* **43**(7), 687–702.
- Weber, J.: 1991, 'Adapting Kohlberg to Enhance the Assessment of Managers' Moral Reasoning', *Business Ethics Quarterly* **1**(3), 293–318.
- Weber, J.: 1992, 'Scenarios in Business Ethics Research: Review, Critical Assessment, and Recommendations', *Business Ethics Quarterly* **2**(2), 137–160.
- Welton, R. E., R. M. Lagrone and J. R. Davis: 1994, 'Promoting the Moral Development of Accounting Graduate Students: An Instructional Design and Assessment', *Accounting Education* **3**(1), 35–50.

Elaine Doyle
 Kemmy Business School,
 University of Limerick,
 Limerick, Ireland
 E-mail: elaine.doyle@ul.ie

Jane Frecknall-Hughes
 The Open University Business School,
 Walton Hall, Milton Keynes, MK7 6AA, U.K.
 E-mail: j.frecknall-hughes@open.ac.uk

Barbara Summers
 Leeds University Business School,
 University of Leeds,
 Maurice Keyworth Building, Leeds, LS2 9JT, U.K.
 E-mail: bs@lubs.leeds.ac.uk

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

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