

# The defining issues test of moral judgment development

Stephen J. Thoma and Yangxue Dong

*The University of Alabama*

ABSTRACT

A measure of moral judgment development, the Defining Issues Test (DIT) is described and the supporting evidence for the measure is summarized. We address these questions: what does the DIT measure; how does the measure work, and how has the measure been validated? The psychometric properties of the DIT are also presented. We suggest that the current evidence supports the DIT as a reliable and valid measure of the characteristic ways adolescents and adults comprehend moral issues.

**KEYWORDS:** defining issues test, moral judgment, development

The Defining Issues Test (hereafter the DIT) was first developed in the early 1970s (Cooper, Coder, Masanz and Anderson, 1974). Originally the measure was described as a paper and pencil alternative to Lawrence Kohlberg's (1969) semi-structured interview measure of moral judgment development (Rest, 1979). As such, the primary focus of the measure was an assessment of the understanding and interpretation of moral issues. Consistent with the Kohlbergian model, Rest viewed moral judgment development as a social and cognitive construct that progressed from a self-focused view of moral issues, through a group-based moral perspective, to a reliance on post-conventional moral principles. Also consistent with Kohlberg, Rest viewed moral judgments as primarily cognitive and a primary factor in the understanding of moral actions and emotions. In short and during the 70s the DIT was viewed as a measure designed to test Kohlberg's developmental sequence and contribute to the development of moral judgment theory in adolescent and adult populations.

Although different in structure from Kohlberg's interview assessment, Rest borrowed the basic components of the Kohlberg approach. Similar to Kohlberg's moral judgment interview the DIT used stories to focus the participant on a moral dilemma. Many of these stories were originally used by Kohlberg (e.g., the story of Heinz and the drug). Furthermore, many of the items used on the DIT were based on Kohlberg interview data. However, unlike the Kohlberg interview where an individual must produce a response, the DIT is a recognition measure. On the DIT, participants are required to rate and then rank 12 short issue statements. These statements represent the defining features of the moral dilemma as viewed from each of Kohlberg's six-stages (Rest, 1979). Specifically, participants taking the DIT read the story and then decide what the protagonist ought to do (e.g., on the Heinz dilemma the choices are "steal the drug," "not steal" or "can't decide"). Following this action

choice, 12 items are presented and rated in terms of importance on a 5-point scale (from great importance to no importance). Once completed, the participant is asked to consider the 12 items as a set and then rank the four items that best describe their understanding of how the protagonist ought to solve the dilemma. This process is repeated for the remaining stories.

The primary index of moral judgment development is derived from the four items ranked as most important. Rest and his colleagues demonstrated that the DIT scores produced results that were consistent with theoretical expectations based on Kohlberg's model (e.g., Kohlberg, 1969). As described below, research using the DIT supported Kohlberg's claim that moral judgment is developmental and increases rapidly across high school and college years. Additionally, the scores produced by the DIT were able to distinguish groups of individuals who could reasonably be expected to differ on moral judgment development, were able to demonstrate that the measure was sensitive to educational interventions, and could related to moral actions and choices. Thus, Rest claimed, one could measure moral judgment development without having to interview individuals, interpret and score their verbal protocols.

It should be noted that in addition to the similarities between the DIT and Kohlberg's method, the DIT also shares some commonalities with measures derived from the Model of Hierarchical Complexity (MHC). Both approaches acknowledge their theoretical and methodological ties to Kohlberg and Piaget while modifying the assessment process and the definition of underlying constructs (e.g., Commons & Pekker, 2007). Additionally, both approaches yield measures that are claimed to be developmental and reflect a hierarchical integration of increasingly complex information. These similarities notwithstanding, there are some significant differences in both the measurement process and intended outcomes. As mentioned above, the DIT uses a rating and ranking task to identify the characteristic way the individual interprets moral situations. These estimates of

moral judgments are described using a developmental model located solely within the moral domain. By contrast, the MHC uses multiple assessments (problem solving tasks, reasoning about vignettes and interviews) to generate an estimate of the individual's developmental stage, which encompasses moral judgments but is not limited to them. As a more general measure of development, the MHC emphasizes the participant's generalized ability to integrate information using estimates of performance on increasingly difficult tasks (Commons & Pekker, 2007). Thus, the MHC represents a measurement system that is not tied to any particular domain but can be used to inform our understanding of the moral domain.

### What does the DIT measure?

The original interpretation of the DIT and what it measured reflected its association with Kohlberg's model. In this view, the DIT was a user-friendly methodological alternative to the interview method. However, this characterization no longer holds (Thoma, 2002; 2006). As the theoretical foundation of the measure evolved from Kohlberg's model to Rest's Four Component Model (Rest, 1983) the interpretation of what the DIT measures also changed. These changes are outlined below.

*Changes related to the underlying developmental model.* Early in the development of the DIT, Rest questioned Kohlberg's acceptance of a strong stage model of development in which individuals move from stage to stage one stage at a time. Instead, the DIT supported a developmental model that defines growth as a gradual shift from lower to more complex conceptions of social/moral cooperation. Furthermore, DIT researchers assume that at any given time there are multiple conceptions available to the individual. Thus, appropriate measurement strategies must assess not only which conceptions are available, but the most preferred system.

Additionally, in the 1990s, DIT researchers adopted a schema view of moral judgment development. A transition that signaled an abandonment of cognitive operations as the defining features of moral stages that was so central to Kohlberg's stage definitions (Kohlberg, 1984). This schema-based model represented moral development as a developmentally ordered set of schemas which define the network of knowledge that is organized around particular life events and exist to help individuals understand new information based on prior experiences (Rest, Narvaez, Bebeau, & Thoma, 1999). In this view moral schemas are contextual, automatic, and less reflective than Kohlberg's stages. Consistent with this view is a companion position suggesting that schemas may not be explicitly understood by the individual and may operate at the tacit level. Thus DIT researchers argue that the DIT is best viewed as a device for activating moral schema (Narvaez and Bock, 2002).

The schemas activated by the DIT are further claimed to be the most general and context-free system for interpreting moral situations. These schemas are labeled as "bedrock schemas" to distinguish the level of assessment provided by the DIT from more context depended interpretive systems. More specifically, the schemas measured by the DIT are viewed as a default system that is evoked when other, more automatic and context-specific, interpretive systems fail or provide incomplete or inconsistent information.

*Micro vs. macro morality.* It has been helpful to maintain a distinction between micro morality, or the morality of everyday exchanges, and macro-morality, or reasoning which focuses on society-wide considerations (e.g., Rest, Narvaez, Bebeau, & Thoma, 1999; Thoma, 2002). DIT researchers argue that the DIT assesses macro morality. In this view, what is assessed are default or bedrock schema that capture an individual's understanding of social cooperation in terms of justice and fairness within the context of law, the mechanisms of government and other social institutions. Although one can conceptually distinguish micro and macro morality, in practice one must assume that they overlap. However, DIT researchers claim that everyday morality is much more contextually dependent than macro morality and influenced by multiple interpretive systems that include but are not limited to the default system measured by the DIT.

Although DIT researchers make a distinction between macro and micro morality this does not imply that the impact of macro-morality on moral function is limited. Indeed, the significance of macro moral processes in adolescence and adulthood is often noted (e.g., Adelson, 1971; Torney-Purta, 1990). In fact, Rest and colleagues argue that the DIT measurement system assumes that the major developmental shifts during adolescence and beyond are the growing understanding of macro-moral conceptions of social cooperation in conventional and post-conventional terms (Rest, Narvaez, Bebeau, & Thoma, 1999).

### How does the DIT work?

As mentioned above the DIT presents participants with a moral dilemma and then asks them to rate and rank 12 items for each dilemma. Each of the items raise particular issues that define the central features of the dilemma based on different moral schema considerations. These items do not present a complete rationale and interpretation of the dilemma but provide the gist of an explanation using a sentence fragment approach. The sentence fragment approach was adopted because early on in the development of the DIT it was noted that items which contained more detailed interpretations of the dilemmas yielded poor developmental indices in part because these items were prone to reinterpretation and idiosyncratic responding (Rest, 1979). By contrast, the use of sentence fragments are particularly well suited to trigger a schema because the fragment provides just enough information to suggest an interpretation, and the individual must fill in the necessary information to fully make sense of the item. Thus, DIT items which match the participant's preferred schema are rated as important and are candidates for being ranked as most important. However, if the item does not make sense or is viewed as too simplistic, then the item is rated as less important and will not be ranked. In short, DIT researchers assume that the rating and ranking of items across stories provide an index of the participant's preferred schema and more generally, represent how the participant generally approaches moral decisions beyond the DIT.

### How does the DIT measure moral judgment development?

In addition to altering the developmental model underlying the measure, DIT researchers also have focused on how best to define the developmental dimension measured by the DIT. In its original

conception, the DIT assessed a developmental dimension defined in terms of Kohlberg's stages as they were described in the early 70s. More recently, however, the fit of Kohlberg's model to DIT data has been assessed. Based on empirical studies using large and diverse samples including some with as many as 44,000 participants, the description of what the DIT measures have changed.

Specifically, empirical estimates of the ways in which DIT items cluster suggest that the six stages described by Kohlberg do not fit the data. Instead, the obtained number of item clusters suggests three distinct groupings: Stage 2 and 3, Stage 4, and Stage 5 and 6. The finding of three distinct clusters is especially clear when the assessment is based on a heterogeneous sample including participants ranging from high school through the adult years (e.g., Thoma and Rest, 1999). That is, empirically, the best fitting scheme based on DIT data is no longer the six Kohlberg stages. Instead a three level model loosely informed by Kohlberg's model seems more appropriate.

It seems plausible that the obtained clusters are due in part to the adolescent and adult populations typically studied by DIT researchers and perhaps the properties of the DIT itself. However, empirically, it seems clear that participants taking the DIT tend to view items representing Stages 2 and 3 as less important reasoning than items in other clusters. Taken together, the stage 2 and 3 items are not often ranked; although attraction to these items is growing (e.g., Thoma, Bebeau & Dong, in preparation). That is, items that highlight self-preservation, self-interest, and personal relationships are viewed together as personal concerns that are not as central as other more-system wide issues represented by the stage 4 items and those that form the post conventional cluster. Unlike the stage 2 and 3 cluster, the stage 4 and postconventional items are often ranked and viewed as highly important. These findings support the view that the DIT items are assessing moral judgment development at the macro-moral level since the power of the DIT derives from the Stage 4 conventional items and the post-conventional items.

**Interpreting the three clusters of items.** The three clusters of items suggest that the DIT measures three distinct moral schemas that are developmentally ordered. These schema are labeled: the Personal Interests schema (combining elements of Kohlberg's descriptions of Stages 2 and 3); the Maintaining Norms schema (derived from Kohlberg's definition of Stage 4); and the Post-conventional schema (drawing from Kohlberg's Stages 5 and 6—and equivalent to the items forming the original summary index called the P score). A description of each schema is presented below.

**Personal interest schema.** Rest, Narvaez, Bebeau, and Thoma (1999) describe the main focus of the personal interest schema as highlighting a perspective that attends the gains and losses each individual may personally experience within a moral dilemma. Similarly, no attention is given to the larger social systems within this schema. Overall, as viewed through a personal interest lens, the social world is a loosely tied network of micro-moral considerations linking close relationships and individual interests. The Personal Interest Schema is fully developed by the time participants are able to reliably complete the DIT (typically defined as a 9<sup>th</sup> grade reading level). Unfortunately, the DIT can say little about the development of the schema within childhood, except to say that empirically, adolescent and older participants recognized it as, at best, a secondary consideration.

**The maintaining norms schema.** The Maintaining Norms schema is representative of a society-wide moral perspective. Within the maintaining norms perspective the moral basis of society is understood in terms of how cooperation can be organized on a society-wide basis. However, drawing heavily from the description of Kohlberg's stage 4, the organization of society this schema prioritizes is based on an understanding of rules, roles and the importance of authorities. In addition to Kohlberg's description of stage 4, the Maintain Norms Schema is also informed by Adelson's (1971) conception of the adolescents' developing understanding of political thought and in particular, Adelson's views on adolescent authoritarianism.

More specifically the Maintaining Norms schema has been defined as having the following characteristics: (a) a perceived need for generally accepted social norms to govern a collective; (b) the necessity that the norms apply society-wide, to all people in a society; (c) the need for the norms to be clear, uniform, and categorical (i.e., that there is "the rule of law."); (d) the norms are seen as establishing a reciprocity (each citizen obeys the law, expecting that others will also obey); and (e) the establishment of hierarchical role structures, of chains of command, of authority and duty (e.g., teacher-pupil, parent-child, general-soldier, doctor-patient, etc.—see Rest, Narvaez, Bebeau, & Thoma, 1999, p. 37).

In short, the Maintaining Norms schema prioritizes the established social order and promotes its maintenance as a moral obligation. Consistent with Kohlberg's stage 4, the Maintaining Norms schema support the view that without law there would be no order, people would act on their own special interests with the result a chaotic and lawless society. This schema, does not provide any additional rationale for defining morality beyond simply asserting that an act is prescribed by the law, is the established way of doing things, or is the established Will of God.

**Post-conventional schema.** Compared to Kohlberg's view of the postconventional stages, DIT researchers assume a different definition of what constitutes a post-conventional system. Avoiding ties to any given philosophical theory or tradition, DIT researchers describe the essential features of Post-conventional thinking in more general terms. In this view, postconventional thinking suggests all moral obligations are to be based on criteria that emphasize shared ideals, are fully reciprocal, and are open to scrutiny (i.e., subject to tests of logical consistency, experience of the community, and coherence with accepted practice—(See Rest, Narvaez, Bebeau and Thoma 1999, p. 38 for a more detailed description).

Based on these descriptions, one can observe that the main source of variance in the DIT is provided by the differences between maintaining norms (conventionality) and Postconventionality. These differences are what Kohlberg regarded as the distinction between Stage 4 and Stage 5; and later Adelson's described as the development of political thought). Although the focus of the DIT measurement system is more directly on the shift from maintaining norms to postconventional thinking than prior models (e.g., Kohlberg's system), the significance of this shift is noteworthy. For instance, the distinction between conventionality and post-conventionality is what tends to drive so many public policy disputes such as the reactions to the wars in Iraq and Afghanistan, how best

to stimulate an economy, minority rights, religion in the schools, medical policy, and so on. Further and perhaps most importantly given the events following 9/11, conventional and post-conventional reasoning addresses the divide between religious fundamentalism and secular modernism (see Marty & Appleby, 1993).

**Indices derived from the DIT.** For many years, the summary index derived from DIT data was the P score. This score is based on the participant's ranking of post-conventional items. The P score has been criticized for at least two reasons: treating qualitative data as continuous, and for failing to incorporate subject responses to non postconventional items. There is an extensive literature about the first criticism (e.g., Rest, 1979, Rest, 1986, Rest, Thoma, Narvaez, & Bebeau, 1997). In general, the argument advanced by DIT researchers acknowledges the qualitative distinctions between different conceptions of moral thinking as represented by the moral schema. However, the use of a continuous score like the P score, signals the view that the assessment process also is quantitative and should be concerned with the rates of participant responses across the types of moral thinking. P scores, therefore represent the participant's relative location on the developmental continuum (defined by qualitatively different markers). In short, as P scores increase we assume that the participant's developmental location is shifting toward higher levels of moral judgment development.

The second criticism of the P score focuses on the fact that the DIT scoring process does not use all of the participant information available to it. As mentioned previously, P scores only focus on the postconventional schema items and do not attend to other schema scores in the scoring process. The fact that the DIT's main index of development fails to use information on the full complement of schema information has been a concern from many since it violates all of the basic tenants of classical measurement theory (e.g., Loevinger, 1976). Although the P score has been used for many years with general success, there have been a number of attempts to improve on P by supplementing the P score with information from other items. Over the last 10 years, a new index, N2, has been developed and become the primary index of the DIT. The N2 score is best viewed as a modified P score. It uses the P score as its starting point and then adjusts the P score based on the participants' ability to discriminate between P items and lower stage items. The N2 score increases in a positive direction if the individual discriminates high and low items. That is, rates the postconventional items as more important than the personal interest items. Similarly, N2 scores decrease when the participant does not discriminate between postconventional and personal interest items or prefers the personal interest items over the postconventional items. Given that the P and N2 score have a similar starting point, it is not surprising that the correlations between them are high and range from the mid-80s – lower 90s (see Rest, Thoma, Narvaez, & Bebeau, 1997). Empirical comparisons of the two scores indicate that the N2 is an improvement over the P score in older and presumably more developed individuals. Thus, the N2 score should be most helpful in discriminating at the high end of the developmental scale. Current best practice recommendations encourage researchers to use the N2 scores as their summary index when focusing on graduate and professional school populations index as it should be an improvement over the

P scores. By contrast, P scores and N2 scores tend to behave very similarly in high school and college samples suggest that (Office for the Study of Ethical Development, personal communication).

**Additional measures derived from the DIT and DIT-2.** Beginning in the 1990s there was an interest in developing measures that could broaden our picture of moral judgment development in ways not captured by the schema and summary scores. These measures include an index of developmental phases which describe individuals as either consolidated or transitional in their developmental profile. The second cluster of measures includes indices that can be derived from responses to the DIT and address related but non-moral development constructs. These variables include assessments of social and political attitudes and choices.

**Developmental phase indicators.** This index was created to explore the role of consolidation and transition on moral judgment development. Thoma and Rest (1999) created a method for assessing developmental phase indicators based upon Snyder and Feldman's (1984) description of developmental phases in development and drawing from Walker and Taylor's (1991) application of the developmental phase notion within the moral judgment domain. Thoma and Rest (1999) measured the degree to which participants were transitional based on a schema profile that indicated little preference for the various stage-based items and, thus, presents a flat response profile. By contrast, a consolidated pattern was indicated when the participant presented a clear preference for a particular schema-based items and, thus, a peaked response profile.

Applications of the developmental phase index were consistent with theoretical expectations. For instance and consistent with the findings reported by Walker & Taylor, (1991), change in moral judgments varied as a function of consolidation and transition. Specifically, participants associated with a greater rate of change on DIT summary scores were disproportionately in the group who were moving from a transitional to consolidated phase. Furthermore, Thoma and Rest (1999) found that moral information is more central in the decision-making process during the consolidation phase regardless of developmental level. More recently, developmental phase has been shown to relate to the time it takes to arrive at decisions about moral issues (Thoma, Narvaez, Endicott & Derryberry, 2001). This work found that subjects identified as consolidated took longer to judge the moral issues suggesting a deeper processing of these issues. Further, Derryberry and Thoma (2005) found that developmental phase indicators moderated the link between moral judgment and action. In general, the common finding across these studies is that developmental phase information moderates the relationship between DIT scores and other variables theoretically linked to moral judgment development. These findings indicate that if an effect is observed using the DIT, the same effect will be stronger if computed on participants in the consolidated groupings (Thoma, 2006).

**Non-moral judgment measures derived from the DIT.** A second set of variables was developed to provide additional information about non-moral constructs by using participant responses to DIT items. For the most part these variables are proxies of non-moral constructs and are useful because they are an efficient way to gather additional information without relying on other measures and the added time demands on participants (Thoma, 2002). At

present these variables capture the following information: a) the degree of decisiveness on the DIT story action choices; b) agreement with action choice decisions made by a group of graduate students in philosophy and political science who achieved the highest scores on the DIT; and c) a proxy measure of religious orthodoxy orientation.

**Number of can't decides.** The Can't Decide variable is an index of the decisiveness with which an individual selects action choices on the DIT. The procedure used to compute this variable is straightforward and represents a simple count of the can't decide choices. That is, for each of the 6 (or 5 on the DIT2) stories, the DIT asks the participant to choose an action choice for the story protagonist. For example following the Heinz dilemma on the DIT-1, the participant is asked whether Heinz should steal the drug to save his wife or should not steal the drug. A can't decide option is also available. By simply counting the can't decide choices the resulting index ranges from 0 – 6 on the DIT and 0-5 on the DIT-2. The interest in the can't decide index is based on the view that indecision is in part a result of the ease with which participants can process moral information. Following from the Thoma and Rest (1999) study, there is the additional expectation that developmental phase and indecision should be related such that transitional phases should be associated with increased indecision. This expectation is based on the view that transitional phases are associated with multiple and potentially conflicting interpretations of moral situations and issues resulting in more indecision. These expectations have been noted in recent norming studies (e.g., Thoma, Bebeau, Dong, Wiu & Jiang, 2011).

**Humanitarian/liberal perspective.** The humanitarian/liberal index represents a proxy variable for a humanitarian and liberal perspective on moral issues. This index was created based on the observation that professionals in political science and philosophy obtained the highest P scores for any group similarly assessed. These scores were so high in fact, that for many years this group was used as an “expert” group and used to describe the upper end of the DIT measurement system (Rest, 1979). More recently, and upon closer inspection of this group it was found that not only were these participants obtaining high scores on the DIT, but they were also very consistent in their action choices. As a group these participants supported the position that Heinz should steal the drug for his dying wife. They also endorsed the view that the neighbor should not turn in the escaped prisoner now leading an exemplary life; that the principal should keep the student newspaper open even though they published controversial topics; that the doctor should provide an overdose of a pain killer to a coherent terminally ill patient; that a repair shop owner should hire the minority applicant even if some customers complain and stop patronizing the shop; and that students were justified in occupying the administration building as part of a protest. The clear endorsement patterns suggested a variable in which participants responses to the action choice portion of the DIT assessment is compared to the choices of this “expert” group (Rest, 1979, Thoma 2002). For the DIT -1 the score can range from 0 (no matches) to six (all matches). Across a number of studies the basic finding is that the relationship between moral judgment development and the humanitarian/liberalism scores are curvilinear. The form of

this relationship indicates that high scores are associated with personal interest and post-conventional schemas and lower scores are related to the maintaining norms schema.

**Religious orthodoxy.** The Religious Orthodoxy score is based on a particular rating and ranking pattern of an item on the doctor's dilemma (or the cancer dilemma on the DIT-2). The particular story containing this item is similar on both versions of the DIT and addresses the question of whether or not the physician ought to provide a drug to a dying woman that will hasten her death. The target item is one that highlights the idea that only God should determine whether one should live or die. By focusing on the ratings and ranking of this item it was noted that a resulting summary index is strongly related to the total scores on religious orthodoxy measures such as the Brown and Lowe Inventory of Religious Beliefs (1951) (Thoma, Bebeau, Dong, Liu, and Jiang, 2011, Narvaez, Getz, Rest, & Thoma, 1999).

#### How do we know the DIT measures moral judgment development?

One historical advantage of the DIT research program is the focus on different approaches to validate a measure of moral judgment development (e.g., Thoma, 2002; Thoma, 2006). Given this focus it is not surprising that the empirical support for the DIT as a measure of moral judgment development are many and varied (see, Rest, 1979, 1986, Rest and Narvaez, 1994, Rest, Narvaez, Bebeau & Thoma, 1999; Thoma, 2006; Thoma, Bebeau, Dong, Liu, & Jaing, 2011). These authors note that a well-articulated set of validity criterion was essential in the development of the DIT. Additionally, these studies contributed to the theoretical shifts mentioned in previous sections. Furthermore, these criterion studies served as the proving ground for new indexes like the N2 score. That is, to support any modification to the scoring or the addition of new indices, the proposed changes were required to yield significantly better trends across criteria and studies than the trends produced by current variables.

These same validity criteria were helpful in addressing criticisms of the DIT. For example, when Sanders, Lubinski and Benbow (1995) concluded that the DIT actually measured verbal ability, DIT researchers were able to find studies that represented the different types of validity criteria and also contained a measure of verbal ability or some reasonable proxy of it (e.g., Thoma, Derryberry, & Narvaez, 2009). The evaluation of these different criticisms was tested by a strategy whereby studies were reanalyze while controlling for verbal ability. Using this approach, the question asked is whether DIT scores can still produce age trends, differentiate known groups, relate to political attitudes and choices and so on when verbal ability is controlled. In response to Sanders, Lubinski and Bebbow's challenge, Thoma, Narvaez, Rest, & Derryberry, (1999) found that when verbal ability was statistically controlled for the dominant trends remained. That is, verbal ability could not account for findings using DIT scores.

The specific criteria used to validate the DIT include: (a) differentiation of various age/education groups; (b) longitudinal gains; (c) correlation with cognitive capacity measures; (d) sensitivity to moral education interventions; (e) correlation with behavior and professional decision making; and (f) predicting to political choice and attitude.

*Differentiating age/educational groups.* The main approach used in these studies is to assess whether or not the DIT is able to distinguish groups which ought to differ on a measure of moral judgment development. For instance, graduate students in political science and philosophy should score higher than other graduate students who are not so well versed in political and ethical theory. Similarly, college students should score higher than high school students and so on. More recently, large composite samples (thousands of subjects) show that 30% to 50% of the variance of DIT scores is attributable to level of education in samples ranging from junior-high education to Ph.D.s (Thoma, 1986).

The longitudinal gains criteria suggest that a measure of moral judgment development ought to produce evidence of upward movement across time. This criterion follows from the claim that a developmental measure ought to describe change in an upward manner. For instance, a 10-year longitudinal study on the DIT indicates upward change in summary scores for both men and women, for college students and people not attending college, and for people from diverse walks of life (Rest, 1986). A review of a dozen studies comparing freshman to senior college students ( $n = 755$ ) shows effect sizes (expressed as Cohen's  $d$  statistic) of .80 ("large" gains). In short, of all of the variables studied in college student samples, the DIT produces some of the most dramatic longitudinal gains (Maeda, Thoma, Bebeau & You, 2009; Rest & Narvaez, 1994).

Criterion 3 proposes that DIT scores ought to be related to measures of moral comprehension and other cognitive measures. However, relationships with cognitive measures should not be excessive and as such, raise the possibility that DIT scores are actually measuring general cognitive skills. Nor should cognitive measures subsume the relationship between DIT scores and other criterion variables (as claimed by the Sanders, Lubinski, and Benbrow, 1995 study mentioned above). Overall, the existing literature indicates that DIT scores are significantly related to measures of cognitive capacity and moral comprehension, to recall and reconstruction of post-conventional moral argument, to Kohlberg's measure, and to other cognitive developmental measures (Rest, 1979; 1986; Thoma 2006).

The fourth criterion focuses on whether the DIT is sensitive to specific experiences that ought to stimulate development. Intervention studies are the prototype for this criterion (e.g., presence or absence of a dilemma discussion condition). For example, Rest, (1986) describes a review of over 50 intervention studies reports an effect size for dilemma discussion interventions to be .41 ("moderate" gains), whereas the effect size for comparison groups was only .09 ("small" gains).

The fifth criterion suggests that DIT scores ought to be linked to moral actions and desired professional decision making outcomes. For instance, one review reports that 32 out of 47 measures of moral action were statistically significant (Rest, 1986). Furthermore, Rest & Narvaez (1994) linked DIT scores to many aspects of professional decision-making.

Finally, criterion six focuses on the link between DIT scores and social/ political variables. In this cluster, the assumption is that DIT scores should be significantly linked to political attitudes and political choices. This view follows from the position that the

DIT is a measure of macro-morality. As mentioned previously, an understanding of macro-morality addresses an understanding of society-wide institutions and their role in promoting social cooperation through laws and the political process. In a review of several dozen correlates between political attitude and DIT scores it was found that they typically correlate in the moderate range (Thoma, Narvaez, Rest, & Derryberry, 1999, Crowson, DeBacker, & Thoma, 2005). When DIT scores were combined in multiple regression with measures of cultural ideology, the overall prediction increased to up to two-thirds of the variance in opinions about controversial public policy issues. These issues include abortion, religion in the public school, women's roles, rights of the accused, rights of homosexuals, civil liberties, the rights of minorities, and free speech issues. Given that these issues are among the most hotly debated of our time, the DIT has the potential to contribute to our understanding of individual differences in political preferences and attitudes.

In addition to these validity criteria, DIT researchers also focused on traditional standards for tests and measures such as acceptable psychometric evidence as well as response stability across different test-taking sets. In addition, DIT scores show discriminate validity from a host of competing variables such as verbal ability/general intelligence and from conservative/liberal political attitudes (Thoma, Derryberry, & Narvaez, 2009; Thoma, Narvaez, Rest & Derryberry, 1999). Moreover, the DIT is equally valid for males and females since gender accounts for less than one half of a percent of the variance of the DIT, whereas education is 250 times more powerful in predicting DIT variance (Thoma, 1986).

## » SUMMARY

The DIT has evolved significantly over its 35-year history and from its roots in the Kohlbergian model. Consistent with other contemporary models of development the shift to a neo-Kohlbergian position resulted in some modifications and rejections of traditional assumptions. Unlike other models that expanded the measurement system to broaden their focus beyond moral functioning (e.g., the MHC approaches), DIT researchers maintained their interest in moral functioning. Specifically, the DIT is claimed to measure default schema by which individuals interpret moral issues. Focusing on the macro-moral level, these default schemas inform the individual's understanding of social structures and their mechanisms. Further, it is claimed that the development of these schemas is ordered such that, starting during the second decade of life, a focus on understanding and maintaining norms gives way to a post-conventional understanding.

These assumptions have been supported by the six validity criteria clusters that contain multiple indicators and cohorts. The results of these analyses clearly support the view that the DIT measures a developmental construct within the moral domain. Further, existing evidence suggests that the measure is particularly good at assessing the shift from a conventional/maintaining norms perspective to a post conventional view of social cooperation. These findings suggest that the DIT will continue to offer the field a theoretical model and research strategy that serves to further moral judgment research. ■

## REFERENCES

- Adelson, J. (1971) The political imagination of the young adolescent. *Daedalus*, 100(4), 1013-1050.
- Brown, D. G., & Lowe, W. L. (1951). Religious beliefs and personality characteristics of college students. *Journal of Social Psychology*, 33, 103-129. doi:10.1080/00224545.1951.9921803
- Commons, M.L. & Pekker. (2007). *Hierarchical Complexity: A formal Theory*. Unpublished manuscript submitted for publication.
- Crowson, H. M., DeBacker, T., & Thoma, S.J. (2007). Are DIT scores empirically distinct from measures of political identification and intellectual ability? A test using post-9/11 data. *British Journal of Developmental Psychology*, 25(2), 197-211.
- Derryberry, P. & Thoma, S.J. (2005). Moral Judgment, Self-Understanding, and Moral Actions: The Role of Multiple Constructs. *Merrill-Palmer Quarterly*, 51(1), 67-92.
- Kohlberg, L. (1969). Stage and sequence. The cognitive- developmental approach to socialization. In Goslin (ed.) *Handbook of socialization theory and research*. Chicago: Rand McNally.
- Kohlberg, L. (1984). *Essays on Moral Development: The nature and validity of moral stages, Vol. 2*. San Francisco: Harper & Row.
- Loevinger, J. (1976). *Ego Development*. San Francisco: Jossey-Bass Publishers.
- Maeda, Y., Thoma, S. J., & Bebeau, M. (2009). Understanding the relationship between moral judgment development and individual characteristics: The role of educational contexts. *Journal of Educational Psychology*, 101(1), 233-247. doi:10.1037/a0013381
- Marty, M. E., & Appleby, R. S. (Eds.), (1993). *Fundamentalism and the State*. Chicago: University of Chicago.
- Narvaez, D. & Bock (2002). Moral Schemas and Tacit Judgments or How the Defining Issues Test is Supported by Cognitive Science. *Journal of Moral Education*. 31(3), 297-314. doi:10.1080/0305724022000008124
- Narvaez, D., Getz, I., Rest, J., & Thoma, S. (1999). Individual Moral Judgment and Cultural Ideologies. *Developmental Psychology*, 35(2), 478-488.
- Rest, J., Cooper, D., Coder, R., Masanz, J., & Anderson, D. (1974). Judging the important issues in moral dilemmas — an objective test of development. *Developmental Psychology*, 10(4), 491-501. doi:10.1037/h0036598
- Rest, J. (1979). *Development in judging moral issues*. Minneapolis, MN: University of Minnesota Press.
- Rest, J. (1983). *Morality*. In P. H. Mussen (Series Ed.) & J. Flavell & E. Markman (Vol. Ed.), *Handbook of child psychology: Vol. 3, Cognitive Development* (4th ed., pp. 556-629). New York: Wiley.
- Rest, J. (1986). *Moral development: Advances in research and theory*. New York: Praeger.
- Rest, J. & Narvaez, D. (1994). *Moral development in the professions*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rest, J., Narvaez, D., Bebeau, M., & Thoma, S. (1999). *Post-conventional moral thinking: A Neo-Kohlbergian approach*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Rest, J., Thoma, S. J., Narvaez, D., & Bebeau, M. J. (1997). Alchemy and beyond: Indexing the Defining Issues Test. *Journal of Educational Psychology*, 89 (3), 498-507.
- Rummelhart, D.E. (1980) Schemata: the building blocks of cognition, In R. Spiro, B. Bruce, & W. Brewer (Eds.), *Theoretical Issues in Reading Comprehension*, (pp. 33-58), Hillsdale, NJ: Erlbaum.
- Sanders, C.E., Lubinski, D., & Benbow, C.P. (1995). Does the Defining Issues Test measure psychological phenomena distinct from verbal ability? *Journal of Personality and Social Psychology*, 69(3), 498-504.
- Snyder, S. S., & Feldman, D. H. (1984). Phases of transition in cognitive development: Evidence from the domain of spatial representation. *Child Development*, 55(3), 981-989. doi:10.1111/1467-8624.ep12427036
- Thoma, S. J. (1986). Estimating gender differences in the comprehension and preference of moral issues. *Developmental Review*, 6(2), 165-180. doi:10.1016/0273-2297(86)90010-9
- Thoma, S. J. (2006). Research using the Defining Issues Test. In Killen and Smetana (Eds.), *Handbook of Moral Psychology*. L. Erlbaum: Mahwah, NJ.
- Thoma, S. J. (2002). An overview of the Minnesota approach in moral development. *Journal of Moral Education*, 31(3), 225-246. doi:10.1080/0305724022000008098
- Thoma, S. J., Derryberry, P. & Narvaez, D. (2009). The distinction between moral judgment development and verbal ability: Some relevant data using socio-political outcome variables. *High Ability Studies*, 20(2), 173-186. doi:10.1080/13598130903358535
- Thoma, S. J., Narvaez, D., Rest, J. & Derryberry, P. (1999). The distinctiveness of moral judgment. *Educational Psychology Review*, 11, 361-376.
- Thoma, S. J., Bebeau, M. J., & Dong, Y. Moral Judgment competency is declining over time: Evidence from 20 years of Defining Issues Test data. In preparation.
- Thoma, S. J., Narvaez, D., Endicott, L. & Derryberry, P. (2001). *Developmental phase indicators and moral information processing*. Paper presented to the American Educational Research Association, New Orleans, LA.
- Thoma, S. J., Bebeau, M. J., Dong, Y., Liu, W., & Jiang, H. (2011). Guide for DIT-2. Available from the Office for the Study of Ethical Development.
- Thoma, S. J., & Rest, J. (1999). The relationship between moral decision-making and patterns of consolidation and transition in moral judgment development. *Developmental Psychology*, 35(2), 323-333. doi:10.1037/0012-1649.35.2.323
- Torney-Purta, J. (1990). *Youth in relation to social institutions*. In S. Feldman & G. R. Elliott (Eds.), *At the threshold: The developing adolescent* (pp. 457-478). Cambridge, MA: Harvard University Press.
- Walker L. J. & Taylor, J. H. (1991). Stage Transitions in moral reasoning: A longitudinal study of developmental processes. *Developmental Psychology*, 27(2), 330-337. doi:10.1037/0012-1649.27.2.330