

The distinction between moral judgment development and verbal ability: some relevant data using socio-political outcome variables

Stephen J. Thoma^{a*}, Pitt Derryberry^b and Darcia Narvaez^b

^a*Department of Educational Studies in Psychology, Research Methodology, and Counseling, 310 Carmichael Hall, University of Alabama, Tuscaloosa, AL 35487, USA;* ^b*Department of Psychology, Tate Page Hall 224, Western Kentucky University, 1906 College Heights Blvd. #21030, Bowling Green, KY 42101-1030, USA*

Critics of moral judgment measures question whether the obtained pattern of relationships between moral judgment scores and outcome variables might be better explained by verbal/general ability. To address this concern, we assess the degree to which moral judgment development reduces to verbal ability using the Defining Issues Test (DIT). Our findings suggest that DIT scores describe a latent variable that is distinct from verbal ability. Second, we show that verbal ability cannot account for the relationship between moral judgments and social political attitudes and actions. These results are consistent with the developmental view that moral judgment development scores are related to, but do not reduce to, cognitive and ability measures.

Keywords: moral judgment; verbal ability

Critics of the moral judgment construct in the Kohlbergian tradition have long speculated that verbal ability may actually account for the pattern of results found between moral judgment measures and external criteria (e.g., Kurtines & Grief, 1974). More recently, Lubinski (2004) suggests that moral judgment development "... make(s) intuitive sense and ostensible capture(s) unique psychological subtleties, but, when compared with preexisting measures, (they) frequently fails to add value" (p. 100). This claim is based on earlier work that directly questioned whether moral judgment development is distinct from general cognitive ability (Sanders, Lubinski, & Benbow, 1995). That is, could the psychological phenomena "explained" by moral judgment measures be just as easily explained by measures of cognitive ability (e.g., verbal ability, cognitive aptitude, and school achievement)? And if measures of moral judgment development do not add any significant information beyond these much used measures (some measure of "g"), then – practically speaking – why go through the time and trouble of collecting moral judgment development information – especially when general ability measures are routinely collected and available in school files of participants?

Whether or not moral judgment is distinct from verbal ability is an important concern when interpreting the often-noted finding that gifted students score higher on measures of moral judgment than their peers (e.g., Derryberry & Barger, 2008). If it is the case that measures of moral judgment are primarily assessing verbal ability

*Corresponding author. Email: Sthoma@bamaed.ua.edu

applied to moral content, then finding higher moral judgment scores for gifted groups is best understood using models that focus on ability. However, if the two constructs are distinct, then findings of group differences on moral judgment development measures require models that take into account how ability and social experiences interact to promote moral judgment development. That is, the question of whether the two constructs differ is central to the question of why gifted students obtain higher moral judgment scores.

Lubinski (2004) and Sanders et al. (1995) concerns about the viability of the moral judgment development construct reflect a general skepticism within individual differences psychology over the distinctiveness of many psychological constructs. Particularly suspect are constructs in personality and social psychology. This point of view is stated clearly by Sanders et al. (1995): "In the social sciences, measures do not always assess what they purport to measure, and the causal determinants of our most favorite constructs and outcomes do not always fit prior expectations (p. 498)." Following Humphreys (1991) and Gordon (1987), Lubinski (2004) and Sanders et al. (1995) suggest that a host of psychological phenomena in the social sciences involving many constructs could be more simply explained in terms of general intelligence. Similarly, Lykken (1991) suggests that Kohlbergian moral judgments probably can be reduced to verbal ability. In short, Sanders et al. (1995) suggest that much of the theoretical underbrush of social science – involving constructs like creativity, ego development, moral reasoning, self-efficacy and gender expectations – could be cleaned up if researchers had first controlled for measures of general cognitive ability.

It should be noted that Sanders et al. (1995) are not the first to suggest that verbal or general ability drives moral judgment scores. Indeed, concerns over the role of verbal ability in the assessment of moral judgment development have a long history (Kurtines & Grief, 1974; Lind, 2003, 2004; Skoe & von der Lippe, 2002). With respect to Kohlberg's measure of moral judgment development (Colby & Kohlberg, 1987), critics have wondered whether the scoring rules for moral stage assignments was unduly influenced by the verbal sophistication of the participant (Kurtines & Grief, 1974). Ironically, concerns over the role of verbal ability in the scoring of Kohlberg's interview measure was one of the major issues motivating the development of the Defining Issues Test (DIT) as an objective measure of moral judgment (Rest, 1979; Thoma, 2002, 2006). Given this history and sensitivity to concerns over the role of verbal ability in moral judgment research, and contrary to the Lubinski (2004) and Sanders et al. (1995) statements mentioned above, it is not surprising that many early validity studies using measures of moral judgment did include measures of ability (Kohlberg, 1969, Rest, 1979). For instance, in 1979, there was sufficient data to conclude that DIT scores were consistently related to measures of cognitive ability with correlations typically in the .25–.50 range (Rest, 1979). The existence and magnitude of these relationships was interpreted as supporting the DIT. As a measure informed by Kohlberg's cognitive developmental theory of moral development, an important expectation driving these validity studies using the DIT was that DIT scores should be more strongly related to cognitive measures than they were related to personality measures. Even accounting for differential reliability estimates for the personality measures, relationships between DIT scores and cognitive measures were consistently stronger than with personality measures (correlations typically ranging from .15 to .30) (Rest, 1979, 1986).

Similarly, the magnitude of the correlations with cognitive measures was welcomed because at these levels, it did not appear that cognitive measures could

account for DIT scores. The latter interpretation was buttressed by additional studies that statistically controlled verbal ability, and showed that DIT scores were uniquely related to other criterion variables (e.g., Rest, 1979, 1986).

Despite this support, concerns about the role of verbal ability in the measurement of moral judgments continue to be raised (Lind, 2003, 2004; Lubinski, 2004; Moran & Joniak, 1979; Sanders et al., 1995; Skoe & von der Lippe, 2002). To moral judgment theorists generally, and DIT researchers in particular, the main response to these criticisms is to highlight the range of data supporting measure of moral judgments and then to point out the logical difficulty in reinterpreting these data using some other variable such as verbal ability or political reasoning (Rest, Narvaez, Bebeau, & Thoma, 1999; Thoma, 2006). For instance, DIT researchers suggest that the DIT measures moral judgment development and not some other variable because DIT scores are supported by six basic criterion clusters: (1) relationships with moral comprehension measures, (2) correlations with social political attitudes, (3) links to moral behaviors, (4) the ability to differentiate moral expert groups from others, (5) are sensitive to educational interventions, and (6) demonstrate longitudinal gains (Rest, 1983, 1986; Rest, Thoma, & Edwards, 1997; Thoma, 2006; Thoma, Narvaez, Rest, & Derryberry, 1999). These researchers go on to suggest that if a challenge to the interpretation of DIT scores is offered – such as the verbal ability claim – then these alternative variables should account for the relationships between moral judgment and representative variables from the six criterion clusters. Following this view, Thoma et al. (1999) summarized the validity data associated with the DIT, and assessed whether verbal/general ability (or appropriate proxies for verbal ability) could plausibly account for one or more of the validity criterion clusters. Thoma et al. (1999) demonstrated that the relationship between moral judgment and the six criterion variables remained after controlling for ability.

This support notwithstanding, it is important to note that the Thoma et al. (1999) review was a post hoc secondary analysis. As such, the analyses were constrained by whether or not the necessary data were available and therefore, the resulting interpretation could be subject to a selection bias. The later issue is significant in that studies reporting verbal ability may be different in some way from studies that did not collect or report this information. Thus, it might be claimed that the conclusions reached in the Thoma et al. (1999) review are premature. The present study, therefore, is an attempt to complement the review-based information on the distinctiveness of DIT scores by directly assessing the joint relationships between DIT scores, verbal ability and an outcome variable directly implicated as a source of DIT score validity: social and political reasoning including political liberalism and conservatism, humanitarian liberalism, religious fundamentalism and attitudes toward human rights (e.g., Rest et al. 1999).

The choice of social political outcome variables is particularly important given the strong theoretical connections between these variables and moral judgment development assumed by proponents of the DIT (Narvaez, Getz, Rest, & Thoma, 1999). Unlike other descendants of Kohlberg's theory, which have tended to focus on interpersonal morality or "micro" morality (e.g., Walker, Pitts, Hennig, & Matsuba, 1995) DIT researchers have highlighted the "macro" moral aspects of Kohlberg's theory. That is, the aspect of moral judgment development that focuses on the individuals' growing sense of community, its laws and structures. They note that early on, Kohlberg (e.g., Kohlberg, 1969) suggested the utility of moral judgment development in explaining political attitudes, reactions to Martin Luther King,

and discussions of the Vietnam War among other issues of the day. However, this focus on macro morality was downgraded when Kohlberg moved to a college of education and his interest shifted to moral education and individual development. DIT researchers have recently argued that Kohlberg's early speculations have special merit and the moral judgment construct can contribute to our understanding of the macro controversies of society (e.g., Narvaez et al., 1999). In making this claim, DIT researchers have elevated the social political validity criterion and thus, its importance in supporting the DIT as a measure of moral judgment development.

Method

Participants

Participants for this study were 154 college students from a southeastern university in the United States. Of these participants, 146 supplied valid data across measures. Although both sexes were included in the sample, more women than men (64%) participated in the study. Age ranges were typical for college students with a mean age range of 19.88 ($SD=1.48$).

Instruments and variables

Operationalized variables are indicated by a brief label given in all capital letters.

DIT (Developmental Indexes)

P score. The original DIT first constructed in the early 1970s consists of six stories or dilemmas; each followed by 12 items (as "issues" of the moral dilemma). Participants are first asked what should the protagonist in the story do, and are then asked to rate and rank the items in terms of their importance in solving the moral dilemma. For over 25 years the summary score of the DIT most frequently used has been the "P" score, calculated from ranking data ("which of these 12 items is the most important"), attending to items designed as Stages 5 and 6. The P score is interpreted as the relative importance given to Postconventional (i.e., Stages 5 and 6) moral considerations. The validity and reliability of the DIT is fully discussed in Rest et al. (1999; see also Thoma, 2006).

P2 score. The P2 score is derived from DIT2 data. The revised or DIT2 version of the DIT differs from the original in two main ways. First, the content of the dilemmas is different. In the place of the six original stories, the revised test consists of five dilemmas: (a) a father contemplates stealing food for his starving family from the warehouse of a rich man hoarding food; (b) a newspaper reporter must decide whether to report a damaging story about a political candidate; (c) a school board chair must decide whether to hold a contentious and dangerous open meeting; (d) a doctor must decide whether to give an overdose of pain-killer to a frail patient; and (e) college students demonstrate against US foreign policy. Participants take the DIT2 in the same way as before: first a decision is made concerning the action choice, then 12 items are rated in terms of importance, followed by the ranking task. Summary scores are computed as before and have similar meaning (Thoma, 2006).

Social political ideology measures

FUNDA (Religious fundamentalism). Brown & Lowe's Inventory of Religious Belief (1951) is a 15-item measure that uses a five-point, Likert-type scale. Its items differentiate between those who believe and those who reject the literalness of Christian tenets. It includes items like: "I believe the Bible is the inspired Word of God" (a positively keyed item); "The Bible is full of errors, misconceptions and contradictions" (a negatively keyed item); "I believe Jesus was born of a Virgin"; and "I believe in the personal, visible return of Christ to earth." Scores on the Brown-Lowe range from 15 to 75. High scores indicate strong literal Christian belief. Criterion group validity is good between more and less fundamentalist church groups (Brown & Lowe, 1951; Getz, 1985). Test-retest reliability has been reported in the upper .70s. Spearman-Brown reliability has been found in the upper .80s (Brown & Lowe, 1951). In Narvaez et al. (1999), Cronbach alpha was .95 for the entire group of participants.

POLCON (Self-identification with political conservatism). Participants were asked to identify their political identity on a five-point scale (political conservatism scale), 1=liberal and 5=conservative

HUMLIB (Humanitarian liberalism). The score represents agreement on the DIT dilemmas with the action choices of liberal philosophy students. Scores can range from 0 (no agreement) to 6 (agreement on all six dilemmas with liberal philosophy students).

ATHRI (Opinions about public policy issues). The Attitudes toward Human Rights Inventory was constructed by Getz (1985). Participants are asked to agree or disagree (on a five-point scale) with statements on public policy issues such as abortion, euthanasia, homosexual rights, due process rights of the accused, free speech, women's roles, and the role of religion in public schools. The ATHRI poses issues drawn from the American Constitution's Bill of Rights, similar to the large-scale studies of American attitudes of civil liberties by McClosky and Brill (1983). The ATHRI contains 40 items, 10 of which are platitudinous, "apple pie" statements of a general nature with which everyone tends to agree. Here are two examples of the platitudinous, non-controversial items: "Freedom of speech should be a basic human right" and "Our nation should work toward liberty and justice for all." In contrast, 30 items are controversial, specific applications of human rights, for example: "Books should be banned if they are written by people who have been involved in un-American activities." And, "Laws should be passed to regulate the activities of religious cults that have come here from Asia." During initial validation, a pro-rights group (from an organization that had a reputation for backing civil liberties) and a selective-about-rights group (from a group with a reputation for backing civil liberties selectively) were enrolled for a pilot study ($N=101$) with 112 controversial items (Getz, 1985). Thirty of the items that showed the strongest divergence between groups were selected for the final version of the questionnaire, along with 10 items that expressed platitudes with which there was not disagreement (see Getz, 1985, for further details on pilot study). Therefore, with the ATHRI we have a total of 40 human rights issues that are related to civil libertarian issues. Scores range from 40 to 200. High scores represent advocacy of civil liberties. Cronbach alpha was .93 for the entire group of participants in Narvaez et al. (1999), and hence for these analyses, the single, general score from the ATHRI is used.

In summary, the social/political attitude cluster included a general indicator of political orientation, religious orientation, attitudes toward human rights, and Humanitarian Liberalism. These various indicators are representative of the types of outcome measures used in studies typically identified as indicators of the political attitude criterion cluster (Thoma et al. 1999).

Ability measures

ACT scores. With the participant's permission, university records were accessed and standardized ability test scores were recorded (American College Testing (ACT) scores). ACT provides a composite and scale scores for English, math, social studies, and natural sciences. Given that the primary interest of this study, the English subscale was selected as the primary index of verbal ability. In order to have multiple indicators of verbal ability, we also selected the social studies sub test as it was more closely related to verbal ability in our preliminary analysis.

GPA. In addition, we obtained participants' grade point average, which was also used to assess the ability variable. Thus, three observed variables were available to estimate verbal ability.

Procedures

Participants were offered course credit as an incentive for completing the measures, which were complete in regularly scheduled class periods. The presentation of DIT1 and DIT2 was counterbalanced. Furthermore, and due to the potential fatigue involved in taking both measures in one setting, two separate class periods were utilized. Monitors were present during the two testings in order to insure an appropriate test-taking environment and to answer any questions. With participant's approval, university records were accessed in order to record the cognitive measures (i.e., GPA, ACT-verbal, and ACT-social studies).

Results

Table 1 presents descriptive data on the observed variables representing the three clusters: moral judgment, social/political attitudes and ability. The observed means and standard deviations are typical for samples drawn from the Southeast. For instance, the P score values are lower than published norms; however, these values are often lower in the southeastern states (e.g., Rest, 1986). Similarly, scores on the social/political measures are more conservative and orthodox than in other regions (Rest et al., 1999). Ability measures, on the other hand, are fairly standard for students in large state-supported universities (US averages are approximately ACT scores of 21). Interestingly, the variances on the various measures do suggest some range restriction. For example, a more typical P score standard deviation is in the 13–15-point range (Thoma & Bebeau, 2004).

Table 2 presents correlations between the observed measures. These values corroborate the view that there is some restriction of range in the estimates. For example, the correlation between DIT1 and DIT2 P scores range in the high .70s for more heterogeneous samples (Rest et al., 1999). Similarly, relationships between the

Table 1. Descriptive statistics for the observable variables.

Variables	<i>M</i>	<i>SD</i>
P	29.63	12.42
P2	30.64	13.19
ATHRI	141.67	13.39
FUNDE	64.08	8.66
POLCON	3.22	1.01
HUMLIB	2.31	1.21
ACTENG	23.00	4.03
ACTSS	23.43	4.66
GPA	2.84	0.63

Note: Data is based on 146 participants. Key: P, DIT1 P score; P2, DIT2 P score, ATHRI, Attitudes Toward Human Rights Inventory; FUNDE, Inventory of Religious Belief; POLCON, Political identity item; HUMLIB, Humanitarian Liberalism scale; ACTENG, American College Testing English subscale; ACTSS, American College Testing Social studies subscale; GPA: grade point average.

Table 2. Zero-order correlations.

	P	P2	ATHRI	FUNDE	POLCON	HUMLIB	ACTEN	ACTSS
P	1.0							
P2	.53	1.0						
ATHRI	.29	.30	1.0					
FUNDE	.16	.12	.30	1.0				
POLCON	.11	.05	.30	.25	1.0			
HUMLIB	.16	.12	.25	.10	.19	1.0		
ACTEN	.29	.32	.18	.09	.04	.08	1.0	
ACTSS	.16	.19	.18	.10	.10	.13	.66	1.0

Note: Correlations greater than $|.16|$ are statistically significant ($p < .05$). Correlations in bold represent the within cluster correlations. $N=191$. Key: P, DIT1 P score; P2, DIT2 P score, ATHRI, Attitudes Toward Human Rights Inventory; FUNDE, Inventory of Religious Belief; POLCON, Political identity item; HUMLIB, Humanitarian Liberalism scale; ACTENG, American College Testing English subscale; ACTSS, American College Testing Social Studies subscale.

social political variables tend to be higher in more representative samples (Narvaez et al., 1999). This difficulty notwithstanding, the pattern of correlations between verbal ability, P scores, and ATHRI are consistent with other studies (Rest et al. 1999).

Structural equation modeling (SEM) was used to assess the joint relationships between moral judgment, social political, and verbal ability variables. This procedure has the advantage over more traditional means of testing the presence or absence of relationships – e.g., regression – because SEM takes into account measurement error. By contrast, regression and other general linear model approaches assume error-free estimates of the latent constructs. Thus, SEM is increasingly the method of choice in estimating hypothesized relationships between variables (e.g., Schumacker & Lomax, 2003).

It is typical to approach the SEM modeling process in two steps: the measurement and the structural modeling phases. In the measurement phase, the focus of the analysis is on whether the latent or underlying variables are adequately assessed by the

observed variables. This step is particularly important in the current study, since there is some dispute about whether verbal ability measures and DIT scores represent a single latent variable (e.g., verbal ability) or separate variables (e.g., moral judgment and verbal ability).

During the second step in the analysis, the relationships between latent variables are measured and the fit of the model assessed. In the present study, this second step in the analysis assesses whether the relationship between moral judgment and verbal ability can account for the link between moral judgment and the latent variable representing social political attitudes. A finding of this nature would be troubling for proponents of the DIT because it suggests that verbal ability can account for the relationship between moral judgment and other variables.

Table 3 presents the measurement phase of the modeling process assuming one, two, and three latent variables. In column 1, a one latent variable model is fitted to the nine variables. This model in effect suggests that all of the moral, verbal, and social political variables define a single latent variable. There is no theoretical support for this model but it is used here for comparative purposes.

One of the most researched aspects of SEM is how to estimate the fit between the observed covariance structure and the hypothesized model (Schumacker & Lomax, 2003). At present, there is no single statistic that identifies the best fitting model in SEM. It is suggested, therefore, that a variety of fit statistics be used as each has its strengths and weaknesses. Selected below are four common fit statistics: chi-square, goodness-of-fit index (GFI), the adjusted goodness-of-fit statistic (AGFI), and the standardized root mean squared residual (SRMR). The chi-square is the original fit statistic that has remained a mainstay of SRM analyses. Low chi-square values associated with high probability values are indicative of a well-fitting model.

Chi-squares have been shown to have some weaknesses, especially with regard to larger sample sizes. To offset this problem, the GFI was developed and has been shown to be less sensitive to sample size concerns. GFI values greater than .90 are considered evidence for an acceptable fit. The AGFI adjusts the GFI for degrees of freedom. AGFI values of .80 and above indicate good fit. Finally, the SRMR is a standardized measure of the residuals obtained by computing the difference between the sample covariance matrix and the model defined covariance matrix. If the difference in the actual and estimated covariance matrices is small, good fit is assumed. Values of .10 and below are considered acceptable. Applying these criteria to the one latent variable model clearly indicates a lack of fit as no fit criterion falls within acceptable ranges.

The two latent variable model (column 2) is defined by separating the social political variables from the ability and moral judgment variables. This model is consistent

Table 3. Fit statistics for the three explanatory models.

Fit statistics	Model 1	Model 2	Model 3
Degrees of freedom	27	26	25
Chi-square	107.96	71.98	20.74
(<i>p</i> -value)	.00	.00	.65
GFI	.84	.90	.97
AGFI	.73	.83	.95
SRMR	.13	.10	.05

Key: GFI, goodness-of-fit index; AGFI, the adjusted goodness-of-fit statistic; SRMR, the standardized root mean squared residual.

with the view advanced by Sanders et al. (1995) that DIT scores are really an indirect measure of verbal ability since the measurement model tests whether verbal and moral judgment variables belong on the same latent variable. Although the fit statistics indicate that the two latent variable model is an improvement over the one latent variable model, the overall pattern of fit statistics suggests a mixed interpretation.

Column 3 presents the three latent variable model. In this model, ability, social political and moral judgment measures are all claimed to define separate latent variables. This model represents the model adopted by moral judgment development researchers who argue that DIT scores are related but distinct from verbal ability. Inspection of the fit statistics clearly indicates that three latent variables are implicated in the data. Figure 1 presents the three latent variable model. The standardized path coefficients (interpreted in the same way as beta coefficients in multiple regression) indicate that all of the observed variables significantly define their respective latent variables. In addition, the correlations between latent variables (represented by the curved doubled headed arrows) are consistent with previous estimates of these relationships (i.e., moral and verbal: $r=.35$, moral and political: $r=.44$, verbal and political:

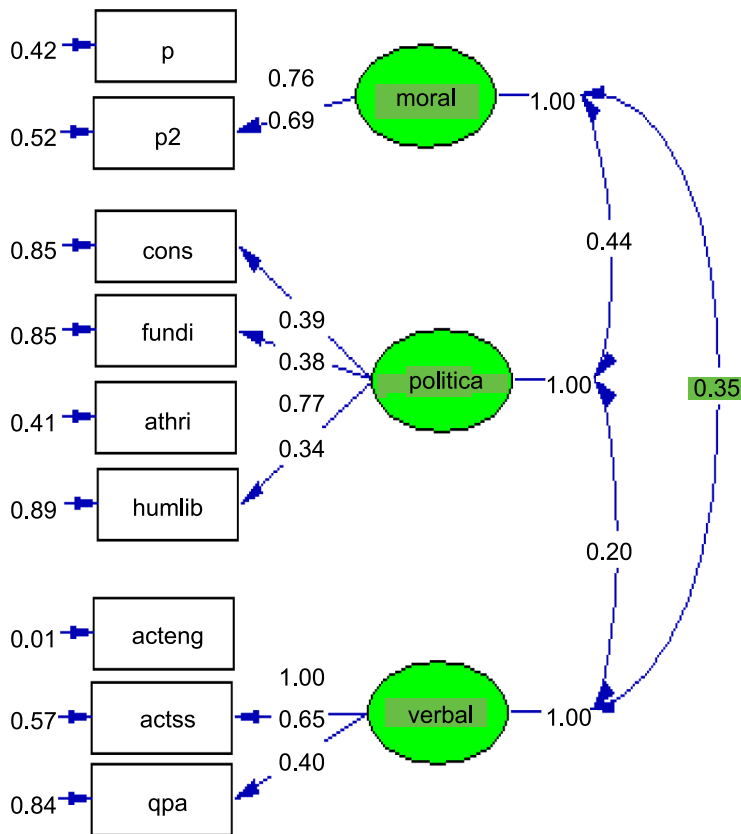


Figure 1. Path coefficients and reliability estimates generated by the three latent variable model. Key: P, DIT1 P score; P2, DIT2 P score, ATHRI, Attitudes Toward Human Rights Inventory; FUNDE, Inventory of Religious Belief; POLCON, Political identity item; HUM-LIB, Humanitarian Liberalism scale; ACTENG, American College Testing English subscale; ACTSS, American College Testing Social studies subscale; GPA, grade point average.

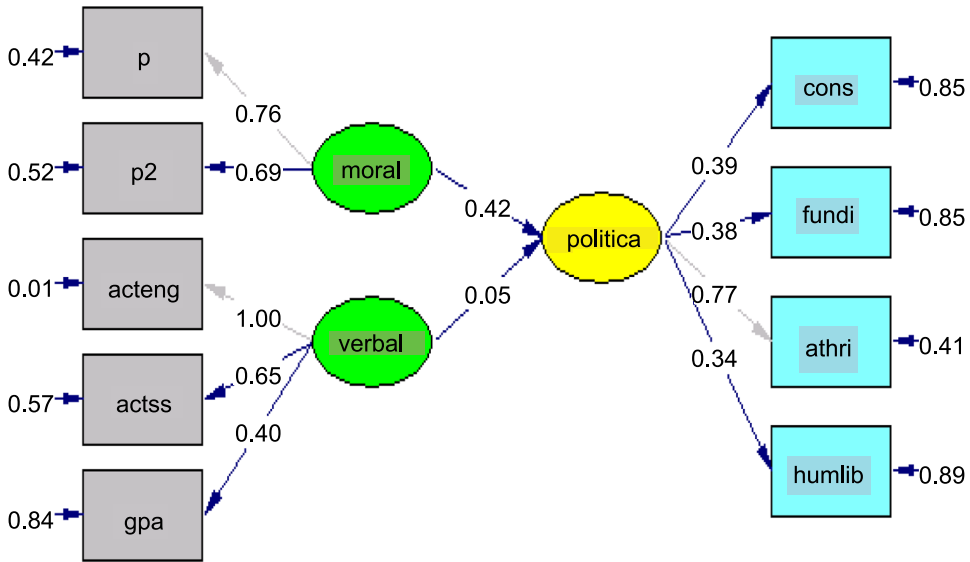


Figure 2. Structural model linking moral judgment development and verbal ability with the social political criterion construct. Key: P, DIT1 P score; P2, DIT2 P score, ATHRI, Attitudes Toward Human Rights Inventory; FUNDE, Inventory of Religious Belief; POLCON, Political identity item; HUMLIB, Humanitarian Liberalism scale; ACTENG, American College Testing English subscale; ACTSS, American College Testing Social studies subscale; GPA, grade point average.

$r=.20$). Thus, from the measurement phase of the analysis it does not appear that DIT scores can be subsumed by measures of verbal ability.

Finding a well fitting measurement model, it is then appropriate to move to the structural modeling phase in which relationships between the latent variables are assessed. The proposed structural model describes the moral and verbal latent variable both predicting social political attitudes. As the paths from these latent variables to the political variable are assessed simultaneously, we can estimate whether the verbal latent variable can account for the relationship between moral and political variables. Figure 2 presents this model and the path estimates as well as the fit statistics (GFI=.97; AGFI=.95; RMSR=.05). Of special interest are the two paths from the verbal and moral latent variables to the social political variable. The standardized path coefficients suggest a larger and statistically significant path from the moral to political variable whereas the verbal to political path is small and not statistically significant. In addition, the fit statistics indicate acceptable fit of the model to the data. Therefore, results of the structural phase of the analysis clearly suggest that verbal ability did not account for the relationship between moral judgment and social political variables.

Discussion

The results of this study suggest that moral and verbal ability measures are not redundant sources of information at least as they relate to social political variables. Specifically, DIT scores were found to define separate latent variables and verbal ability variables could not account for the relationship between the moral judgment variable and social political variable. Taken together, these data support the view

advanced by moral judgment development researchers that DIT scores are related to, but are separate from, verbal and other cognitive measures. Moreover, our findings are consistent with the previous literature reviews that reach the same conclusions (e.g., Rest, 1979, 1986; Thoma et al. 1999). Thus, there is convincing data from a variety of sources that contradicts the criticism of the DIT advanced by Sanders et al. (1995) and others (Emler, Renwich, & Malone; 1983, Lind, 2004; Skoe & von der Lippe, 2002).

Given that Sanders et al. provide the most targeted case for critiquing the DIT, it is interesting to focus on why it is that the results of our study are so different from those of Sanders et al. (1995). In their paper, Sanders et al. (1995) report that of 62 criterion variables that they studied, the DIT correlates significantly only with the cognitive variables and with none of the non-cognitive personality trait measures. As mentioned in the introduction of this paper, correlations between DIT scores and personality trait measures are typically small and unstable (Rest, 1979). For instance, Rest (1979) states (pp. 197–198),

Of approximately 150 correlations between the DIT and personality trait measures, most are nonsignificant and only rarely are any in the .50s and .60s range, whereas the converse was true with the correlations of the DIT with cognitive and developmental measures ... The DIT is more related to cognitive processes than to personality traits ... Personality traits do not exhibit developmental trends, whereas the DIT does.

Thus, Sanders et al. (1995) confirm that there is little stable relationship between personality trait measures and the DIT; partialling out shared variance with general cognitive ability produces even lower correlations with the DIT.

More importantly, and what we feel represents the root cause of the different outcomes between the two studies, is the lack of a theoretical justification for the criterion variables selected by Sanders et al. (1995). Moral judgment is not an all-purpose construct that should be correlated with everything else. Indeed, as Blasi (1980) cautioned, one must be very clear to select criterion variables that are theoretically justified and relevant for the population under study. Regarding the specific measures selected by Sanders et al. (1995), it is not at all obvious how broad-based family measures and leisure activities should be linked to moral judgment, especially given the more current targeted studies that attend to specific family factors related to growth (e.g., Walker & Taylor, 1991) and environmental correlates of change during high school and college (e.g., Deemer, 1986). In short, the main difference between the two conclusions may be the result of using markedly different criterion variables. In the present case, the social political variables used as the primary criterion variable represents a well-established criterion cluster that has been used to validate the DIT (e.g., Thoma et al. 1999). That is, there is both theoretical and empirical evidence to suspect that DIT scores ought to be related to these variables. By contrast, Sanders et al. (1995) selected variable that were not theoretically linked to moral judgment development and thus have limited utility in assessing the distinctiveness of the moral judgment construct.

Finally, as a cognitive developmental measure, it is not surprising that moral judgment development is related to ability. However, less clear is an explanation for this effect. Recent work by Derryberry and his colleagues (e.g., Derryberry, Wilson, Snyder, Norman, & Barger, 2005) suggests that verbal ability and moral judgment development are related because linguistic information is central to the process of inferring and attending to moral issues within complex social situations. Those individuals who are

particularly gifted in linguistic skills presumably are able to abstract more information from a social exchange than do others who are less able. More generally, cognitive skills associated with the processing of information and decision making are both theoretically and empirically linked to moral judgment development (e.g., Rest, 1979).

Furthermore, it is not surprising that advanced moral judgment development is often associated with gifted populations (e.g., Narvaez, 1993). Indeed, the developmental advantage associated with giftedness is striking. For instance, Derryberry and Barger (2008) found a sample of gifted adolescents (ages 12–16) over a standard deviation higher on DIT scores when compared to an average group of older college students. Taken together, the links between ability and moral judgment development suggest that the way one generally processes information is powerfully related to moral judgment development. However, the work presented here suggests that this relationship does not imply that moral judgment scores are simply the result of ability applied to moral content.

References

- Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin*, *88*, 593–637.
- Brown, D.G., & Lowe, W.L. (1951). Religious beliefs and personality characteristics of college students. *Journal of Social Psychology*, *33*, 103–129.
- Colby A., & Kohlberg, L. (1987). *The measurement of moral judgment*. Cambridge: Cambridge University Press.
- Deemer, D. (1986). *Moral judgment and life experience*. Unpublished doctoral dissertation, University of Minnesota.
- Derryberry, W.P., & Barger, B. (2008). Do contributors to intellect explain the moral judgment abilities of gifted youth? *Gifted Child Quarterly*, *52*, 340.
- Derryberry, W.P., Wilson, T., Snyder, H., Norman, A., & Barger, B. (2005). Moral judgment differences between gifted youth and college students. *Journal of Secondary Gifted Education*, *17*, 6–19.
- Emler, N., Renwich, S., & Malone, B. (1983). The relationship between moral reasoning and political orientation. *Journal of Personality and Social Psychology*, *45*, 1073–1080.
- Getz, I. (1985). *The relation of moral and religious ideology to human rights*. Unpublished doctoral dissertation, University of Minnesota.
- Gordon, R.A., (1987). SES vs. IQ in the race–IQ delinquency model. *International Journal of Sociology and Social Policy*, *7*, 30–96.
- Humphreys, L.G. (1991). Limited vision in the social sciences. *American Journal of Psychology*, *104*, 333–353.
- Kohlberg, L. (1969). Stage and sequence. The cognitive–developmental approach to socialization. In D.A. Goslin (Ed.), *Handbook of socialization theory and research*. Chicago, IL: Rand McNally.
- Kurtines, W., & Grief, E. (1974). The development of moral thought: Review and evaluation of Kohlberg’s approach. *Psychological Bulletin*, *81*, 453–479.
- Lind, G. (2003). Does religion foster or hamper morality and democracy? Invited plenary lecture at the meeting of the Association for Moral Education, Krakow, Poland, July 20, 2003.
- Lind, G. (2004). The meaning and measurement of moral judgment competence revisited. A dual aspect model. In D. Fasko & W. Willis (Eds.), *Contemporary philosophical and psychological perspectives on moral development and education*. Cresskill, NJ: Hampton Press.
- Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman’s (1904) “‘General intelligence,’ objectively determined and measured”. *Journal of Personality and Social Psychology*, *86*, 96–111.
- Lykken, D.T. (1991). What’s wrong with psychology anyway? In D. Chicchetti & W. Grove (Eds.), *Thinking clearly about psychology* (pp. 3–39). Minneapolis, MN: University of Minnesota Press.

- McClosky, H., & Brill A. (1983). Dimensions of tolerance: What Americans believe about civil liberties. New York: Russell Sage Foundation.
- Moran, J.J., & Joniak A.J. (1979). Effects of language on preferences for responses to a moral dilemma. *Developmental Psychology*, *15*, 337–348.
- Narvaez, D. (1993). High achieving students and moral judgment. *Journal for the Study of the Gifted*, *16*, 268–279.
- Narvaez, D., Getz, I., Rest, J.R., & Thoma, S. (1999). Individual moral judgment and cultural ideologies. *Developmental Psychology*, *35*, 478–488.
- Rest, J.R. (1979). *Development in judging moral issues*. Minneapolis, MN: University of Minnesota Press.
- Rest, J.R. (1983). Morality. In P. Mussen (Series Ed.), *Handbook of child psychology, Vol. 3, Cognitive development* (4th Ed., pp. 556–629). New York: Wiley.
- Rest, J.R. (1986). *Moral development: Advances in research and theory*. New York: Praeger.
- Rest, J.R., Narvaez, D., & Bebeau, M., & Thoma, S.J. (1999). *Post conventional moral thinking: A neo-Kohlbergian approach*. Mahwah, NJ: LEA.
- Rest, J.R., Thoma, S.J., & Edwards, L. (1997). Devising and validating a measure of moral judgment: Stage preference and stage consistency approaches. *Journal of Educational Psychology*, *89*, 5–28.
- 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*, 498–504.
- Schumacker, R.E., & Lomax, R.G. (2003). *A beginner's guide to structural equation modeling*. Mahwah, NJ: LEA.
- Skoe, E.A., & von der Lippe, A.L. (2002). Ego development and the ethics of care and justice: The relations among them revisited. *Journal of Personality*, *70*, 485–508.
- Thoma, S.J. (2002). An overview of the history of the Minnesota approach to morality research. *Journal of Moral Education*, *31*, 4–19.
- Thoma, S.J. (2006). Research on the defining issues test. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (pp. 67–91). Mahwah, NJ: Lawrence Erlbaum, Associates.
- Thoma, S.J., & Bebeau, M. (2004). *Guide to the DIT-1*. Available from the Center for the Study of Ethical Development, The University of Minnesota, Minneapolis, MN.
- Thoma, S.J., Narvaez, D. Rest, J.R., & Derryberry, P. (1999). Does moral judgment reduce to political attitudes or verbal ability? Evidence using the defining issues test. *Educational Psychology Review*, *11*, 325–343.
- Walker, L.J., Pitts, R.C., Hennig, K.H., & Matsuba, M.K. (1995). Reasoning about morality and real-life moral problems. In M. Killen & D. Hart (Eds.), *Morality in everyday life: Developmental perspectives* (pp.371–407). Cambridge: Cambridge University Press.
- Walker, L.J., & Taylor, J.H. (1991). Stage transitions in moral reasoning: A longitudinal study of developmental processes. *Developmental Psychology*, *27*, 330–337.

Copyright of High Ability Studies is the property of Routledge 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.