

Arrested Development? Comparing Educational Leadership Students With National Norms on Moral Reasoning

Educational Administration Quarterly
2015, Vol. 51(4) 511–542
© The Author(s) 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0013161X14539807
eaq.sagepub.com



Jennifer L. Greer¹, Linda J. Searby²,
and Stephen J. Thoma³

Abstract

Purpose: The public expects school leaders to be moral exemplars, yet prior research indicates that teachers and, more recently, school principals may score lower than other career groups on a widely used measure of moral reasoning, the Defining Issues Test. Moreover, little empirical research has been conducted on educators during leadership preparation in graduate school. The purpose of this research was to create a baseline profile for moral reasoning in educational leadership/administration graduate students in one Southern state and to compare their scores with a composite national average for graduate students across disciplines to see if educational interventions are needed. **Research Design:** In the summer of 2012, the updated Defining Issues Test–2 was offered via e-mail as an online questionnaire to 539 master’s, educational specialist,

¹University of Alabama, Birmingham, AL, USA

²Auburn University, Auburn, AL, USA

³University of Alabama, Tuscaloosa, AL, USA

Corresponding Author:

Jennifer L. Greer, University of Alabama at Birmingham, UAB Lister Hill Library, G03F, 1720
Second Avenue South, Birmingham, AL 35294-0013, USA.

Email: jlgreer1@uab.edu

and doctoral students in the five advanced-track schools in the study state. The questionnaire also gathered data on the students' demographics and the virtual test-taking environment. **Findings:** Complete data for 113 respondents (a 21% response rate) show that the educational leadership/administration graduate students prefer a maintaining norms schema when solving moral problems. In fact, participants scored significantly lower ($M = 30$) on postconventional (advanced) moral thinking than the national norm for graduate students across disciplines ($M = 41$) and a historical average for graduate students ($M = 53$) on the Defining Issues Test. **Conclusions:** Researchers urge evidenced-based educational interventions for this group—specifically, Rest's four-component model, which addresses all four moral psychological processes: sensitivity, judgment, motivation, and character. They also recommend beginning ethics instruction with professional identity development. **Classification:** This is an empirical study.

Keywords

moral reasoning, educational leadership, ethics education, Four Component Model, professional identity development

Introduction

Moral dilemmas arise daily for educational leaders, whether they are engaged in teacher evaluations or student discipline, management of school funds, or negotiations of community controversies (Foster, 2004; Frick, Faircloth, & Little, 2012; Shapiro & Stefkovich, 2011; Starratt, 2004; Strike, Haller, & Soltis, 2005). "As soon as one ethical crisis passes, there's likely to be another on the horizon," observes C. E. Johnson (2009, p. 224). In today's dynamic and high-stakes environment, if experience alone built moral muscle, educators would be exemplars. Nevertheless, ethical failings by educational leaders make the news frequently. Headlines range from "Educators Implicated in Atlanta Cheating Scandal" (Koebler, 2011) and "Nine Schools Cited for Exam and Credit Irregularities" (Phillips, 2012) in the K-12 setting to "Gaming College Rankings" (Perez-Pena & Slotnik, 2012) and "Malone U. President Steps Down Amid Plagiarism Accusations" (Laster, 2010) in higher education.

Anecdotal data aside, of cause for concern is prior research indicating that teachers and, more recently, K-12 school principals may score lower than other career groups on a standard test of moral reasoning used in higher education and the professions. For several decades, researchers (Bloom, 1976; Chang, 1994; Cummings, Dyas, Maddux, & Kochman, 2001; McNeel, 1994; Yeazell & Johnson, 1988) have reported that teacher education students score

lower than peers on the Defining Issues Test (DIT), one of the most widely used and well-validated assessments of life span development (Bebeau, 2002; Rest, 1979; Rest, Narvaez, Bebeau, & Thoma, 1999; Thoma, 2006). More recently, two studies (Slavinsky, 2006; Vitton & Wasonga, 2009) found lower-than-expected scores in public school principals. In the latter study, 60 elementary school principals in the Midwest were found to be just below moral reasoning levels for the general adult population but well below others who have attained graduate degrees. Some scholars maintain that evidence of disciplinary variations in moral reasoning is inconclusive (Derryberry, Snyder, Wilson, & Barger, 2006; King & Mayhew, 2002; Livingston, Derryberry, King & Vendetti, 2006), yet others remain concerned about teachers' and principals' moral development and the effectiveness of preparation programs in higher education (Cummings et al., 2001; Cummings, Harlow, & Maddux, 2007; Cummings, Maddux, & Cladianos, 2010; Cummings, Wiest, Lamintina, & Maddux, 2003; Vitton & Wasonga, 2009).

Since formal education is the strongest predictor for advanced moral reasoning (Rest et al., 1999), it is surprising to find low DIT scores in a highly educated group of people such as school principals. This is especially true since research shows that graduate and professional schools offer ideal opportunities for moral growth as measured by the DIT (Bebeau & Monson, 2008). In fact, nearly three decades of intervention studies (King & Mayhew, 2002; Rest & Narvaez, 1994; Rest et al., 1999; Thoma, 2006) document that more mature people (e.g., college students and adults) experience greater increases in moral reasoning in a formal education environment than do younger people (high school students).

Purpose of this Study and Research Questions

Given that moral reasoning is a requisite capacity for educational leaders, the primary purpose of this quantitative study was to use the Defining Issues Test-2 (DIT-2) to create a baseline profile of moral problem solving in educational leadership/administration (EDL/EDA) graduate students in one Southern state. This target population included master's students, educational specialist students, and doctoral students in the five advanced degree-granting institutions in the state. In addition, researchers sought to compare the EDL/EDA scores on moral reasoning with national norms for graduate students across disciplines. Two questions guided this study: What are the characteristic moral problem-solving schemas of EDL/EDA graduate students in one state in the South, based on their scores on the DIT-2? How do the characteristic moral problem-solving schemas of EDL/EDA graduate students in one state in the South compare with national norms, based on a historical composite of scores on the DIT?

Review of the Literature

Evidence of low scores for educators on moral reasoning date back to the late 1970s (Bloom, 1976), but recently, Cummings and colleagues (Cummings et al., 2001; Cummings et al., 2003; Cummings et al., 2007; Cummings et al., 2010) spent a decade studying moral reasoning in teacher education students in the western United States. After first documenting lower-than-expected scores in 2001, the researchers became concerned with a lack of improvement in moral reasoning as evidenced by DIT P scores (for advanced or post-conventional reasoning) for these students from their freshman to senior year. These are years when most college students experience significant growth in this area of moral development (Rest et al., 1999). In 2003, Cummings and colleagues published an extensive review of 526 teacher preparation courses in 30 college programs, finding 90% of them dedicated to hands-on, task-oriented methods (not critical thinking). Teachers, the authors wrote, “risk becoming technicians instead of morally engaged people who think critically about and reflect upon their ethical and moral responsibilities to students” (p. 167). In 2007, Cummings and colleagues published a comprehensive review of the moral reasoning literature on teachers, confirming a pattern of low scores on the DIT and a scarcity of effective intervention studies. Subsequently, Cummings and colleagues (2010) piloted a successful moral reasoning intervention in an online course for teachers in training. Employing the updated DIT-2, the researchers documented posttest gains in students in moral reasoning after 5 weeks of instruction. Meanwhile, two recent DIT studies (Slavinsky, 2006; Vitton & Wasonga, 2009) documented lower-than-predicted P scores for moral judgment among school principals in Connecticut and the Midwest, respectively. Attempting to explain the low scores for elementary school principals in their study, Vitton and Wasonga (2009) suggest three contributing factors: fixed, change-resistant mental maps (or schemas) on the part of almost half the educators; changing regulatory and school environments that create more complex ethical dilemmas; and inadequate preparation in moral leadership.

Moral Development: A Leadership Imperative and Professional Requisite

Although prior DIT studies of moral reasoning have focused on educational practitioners, little empirical research has been conducted on educational administrators during one of the most critical phases of leadership preparation—graduate school. This dearth of data is puzzling for several reasons.

First, many schools of education have kept pace with trends in political science, business, and organizational studies that position ethical action as a leadership imperative. Across the disciplines, in fact, many scholars consider leadership a transformational practice with ethics at its heart (Bass & Riggio, 2006; Bertram Gallant, 2011; Bolman & Gallos, 2011; Burns, 1978, 2003; Ciulla, 2004, 2005; Fullan, 2003). Leading education scholar Howard Gardner (2008) heralds “the ethical mind” as one of the “5 minds for the future” (p. 127) while conceding that it is the last—and most difficult—one to develop. Education reformers, such as the late William F. Foster (1986), situate moral leadership on a high, almost spiritual, plane: “Each administrative decision carries with it a restructuring of a human life; this is why administration at its heart is the resolution of moral dilemmas” (p. 33). Educational leadership coaches, such as Fullan (2003), highlight the importance of the administrator’s transformational influence on others: “The principal with a moral imperative can help realize it only by developing leadership in others” (p. xv).

Second, national consortia—such as the University Council for Educational Administration (2013), the Interstate School Leaders License Consortium, and the Educational Leadership Constituent Council—consider ethical action a professional requisite, embedding ethics language into their codes and standards for administrators. For example, language from the new University Council for Educational Administration’s “Code of Ethics for the Preparation of Educational Leaders” directly addresses the goal of developing advanced ethical judgment by “[fostering] the capacity to critique and challenge the status quo within the field of educational leadership.” Hence, in theory, leadership preparation programs in higher education are accountable for students’ moral growth.

Nevertheless, as Pritchard (1999) points out, such codes “are not necessarily static documents” (p. 404), nor are they simply “[algorithms] for decision making” (p. 405). Interpretation of codes, their application in new and unforeseen contexts, the need to reform them—all these actions call for critical and independent professional judgment. Similarly, from their empirical study of moral reasoning in principals, Vitton and Wasonga (2009) conclude that leadership preparation calls for more than “a set of codes and standards” or traditional instruction; it requires engaged interventions that “may lead aspiring school leaders to new, different, and more comprehensive ways of thinking” (p. 112).

Theoretical Framework: Four-Component Model of Moral Functioning

In addition to demonstrated need among educators, there is established educational psychology theory and a body of evidence around the DIT to suggest

that graduate and professional school may present one of the best opportunities to shape the moral development of future leaders. Built on the foundational moral psychology work (observational) of Jean Piaget (1932/1965) and the moral judgment studies (interview based) of Lawrence Kohlberg (1969), the DIT was developed as a quantitative measure of moral reasoning in the 1970s by cognitive psychologist James R. Rest (1979) at the University of Minnesota. In one of several important breaks with Kohlberg, Rest (1983) identified moral reasoning to be just one of four processes involved in moral functioning, which he described as the four-component model (FCM). The new model spawned an entire research program in moral development—the Minnesota approach (Thoma, 2002)—which has, in turn, inspired research programs on one or more of the four components. Given that the contributions of Piaget and Kohlberg have been oft documented and discussed, they are described briefly here; it is the newer work of the neo-Kohlbergians (Rest et al., 1999) over the last half century that is primarily summarized in this review.

In his seminal work *The Moral Judgment of the Child* (1932/1965), developmental psychologist Jean Piaget first articulated a modern empirical view of morality as a cognitive developmental process. To gather his data, Piaget conducted a series of observations of and conversations with children along the developmental range, interacting early in life with their parents and later with their peers. On the basis of everyday interactions (e.g., games of marbles), he theorized that children passed through two primary phases of moral development. The first is a morality based on “relations of constraint” (p. 395), in which preschoolers learn to unquestioningly follow the rules of adults or risk getting caught and punished. The second phase is a morality based on “relations of cooperation” (p. 395), in which 7- to 10-year-olds learn to negotiate and finesse the rules out of a sense of fairness, autonomy, and respect for one another as equals. Piaget believed that this peer-inspired “functional equilibrium” (p. 399) laid the groundwork for democratic cooperation in a larger society. Although Piaget has his critics, he is credited with introducing the idea that moral judgment is a cognitive developmental process, which has wide currency today (Lapsley, 1996). Indeed, writes Lapsley, Piaget inspired “what is arguably one of the most important theories in the history of psychology” (p. 40)—the moral stage theory of Lawrence Kohlberg.

In the 1960s, Kohlberg, a graduate psychology student at the University of Chicago, extended Piaget’s theory with work on moral judgment in adults. He created a method of interviewing people about their reactions to complex moral dilemmas, calling it the moral judgment interview method (Kohlberg, 1969). Using this method, discourse analysis, and a complicated scoring

protocol, Kohlberg documented six stages of moral reasoning, from novice to expert: Stage 1, heteronomous morality (egocentric); Stage 2, individualistic morality (transactional); Stage 3, interpersonal morality (maintaining personal norms); Stage 4, social system morality (maintaining local society norms); Stage 5, human rights and social welfare morality (advocate for universal values beyond borders); and Stage 6, principled morality (a hypothetical justice-for-all goal). For 20 years, Kohlberg and his colleagues conducted hundreds of studies and held fast to stage theory, on the basis of longitudinal studies, interventions with pre- and posttests, and studies across cultures and nations. Yet they drew heavy criticism on a variety of issues (Lapsley, 1996; Rest et al., 1999; Rest & Narvaez, 1994; Thoma, 2006), including gender bias (Kohlberg's early studies focused mostly on men), methodology (his scoring method was constantly criticized and revised), and overemphasis on structuralism (his obsession with a progression through hard stages). It was left to one of Kohlberg's graduate students, James R. Rest, to create a new paradigm that took our thinking about moral development to the next level.

Rest reformulated and extended his mentor's work by designing a quantitative instrument, the DIT, to more accurately and reliably assess moral reasoning in a variety of people (Rest et al., 1999; Thoma, 2006). As the DIT evolved, Rest and his team broke with Kohlberg's stage theory of explaining growth in moral reasoning and, instead, embraced F. C. Bartlett's (1932/1995) schema theory. The neo-Kohlbergians theorized that people construct morality as they mature and develop and that they prefer one of three schemas to make quick decisions about complex moral dilemmas: a personal interest schema, a maintaining norms schema, or a postconventional (advanced) schema. Unlike stage theory, schema theory more accurately integrates social cognition with moral cognition and allows researchers to measure more subtle indices of development (Bebeau & Thoma, 2003; Narvaez & Bock, 2002). Importantly, subsequent analysis of a megasample ($N = 44,000$) confirmed that the items on the DIT did indeed cluster around these three schemas instead of six stages (Rest, Thoma, & Edwards, 1997).

In developing the DIT, Rest and the neo-Kohlbergians addressed multiple measurement problems with Kohlberg's studies, creating a much more accurate, sophisticated, and gender-neutral tool that produced consistently reliable and valid results (Rest, 1979, 1986; Rest et al., 1999; Thoma, 2006). For example, Carol Gilligan (1982) famously raised the issue of gender bias in research by Kohlberg, who relied heavily on male study participants and used a qualitative interview and scoring method that were also subject to criticism (Rest et al., 1999; Rest & Narvaez, 1994). However, based on hundreds of studies involving thousands of women and men, there is no empirical evidence to indicate that the quantitative schema-based DIT is biased against

women (Rest et al., 1999; Thoma, 1986, 2006; Walker, 1984). This finding holds true for the DIT-2, which was developed by Rest with his wife, respected psychologist Darcia Narvaez. When gender differences do occur on the DIT, women produce higher mean scores (Bebeau, 2002); yet, typically, these differences are not statistically significant. This finding of gender neutrality has been born out repeatedly in the literature. In 1986, for example, Thoma conducted a meta-analysis of 56 DIT studies with more than 6,000 participants and found that gender differences accounted for a scant .002 of the variance in scores, in contrast with the effect of formal education, which was 250 times more powerful. In sum, the evidence shows that the DIT overcomes any traditional gender biases of other methodologies.

Today, the DIT is considered the standard instrument for assessing moral reasoning (King & Mayhew, 2002; Rogers, 2002; Thoma, 2006). Indeed, Rogers, in his 2002 retrospective evaluation of its use in college studies, praised the DIT for being grounded in “landmark empirical and conceptual integrations of the literature” (p. 325) and for being “top of the class” (p. 326) on a critical criterion for educators—namely, sensitivity to interventions. As previously noted, however, Kohlberg and Rest both faced criticism that there was more to moral action than simply moral reasoning as measured by the DIT. Unlike Kohlberg, Rest (1983) agreed and conducted an extensive, bottom-up review of the empirical literature. From that review, he conceptualized a total of four psychological processes involved in moral functioning known as Rest’s FCM: sensitivity (empathy and recognition of an ethical issue), moral reasoning (cognitive schemas for problem solving), motivation (prioritizing moral concerns over pragmatic ones), and character (the competence and perseverance to follow through on moral action in the face of difficulty). Today, neo-Kohlbergians offer compelling evidence (Bebeau & Monson, 2008; Bebeau, Rest, & Narvaez, 1999; Bebeau & Thoma, 2013; Rest & Narvaez, 1994; Thoma, 2006; Thoma & Bebeau, 2013) that the DIT can be used in a synergistic approach with FCM-based curricula to stimulate moral and professional development (see Figure 1).

Specifically, research documents that moral reasoning is enhanced through the right kinds of educational experiences, including a popular intervention method—“dilemma discussion” of abstract and field-specific cases (Bebeau, 2002; Cummings et al., 2010; King & Mayhew, 2002; Rest & Narvaez, 1994; Rest et al., 1999). Bebeau has used the FCM to study moral development in the professions (medicine, dentistry, and law) for more than 30 years and believes that graduate students are ideal candidates for growth:

Students in professional education are intellectually mature and though they may come to professional education with low P scores on measures such as the

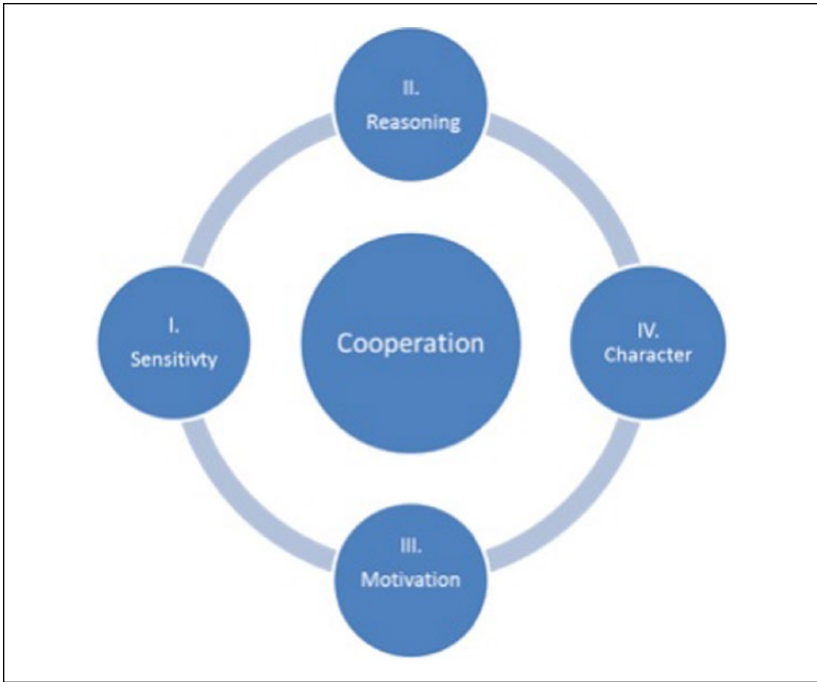


Figure 1. Rest’s four-component model of moral functioning.

Note. In his extensive, bottom-up review of the literature, James Rest identified four nonlinear, nonoverlapping psychological processes—sensitivity, reasoning, motivation, and character or competence—that work together in support of cooperative moral action. Effective moral development education and experiences address all four components; recent four-component model research on students in professional schools shows that Component 3, motivation, activates professional identity and may be the best entry point for graduate students. Based on a description from Thoma and Bebeau’s (2013) “Moral Motivation and the Four Component Model.”

DIT, they often learn quickly to construct well-reasoned arguments and to apply criteria for judging the adequacy of an argument. (Bebeau & Monson, 2008, p. 568)

However, Bebeau is not convinced that dilemma discussion is the right place to begin (Bebeau & Monson, 2008), indicating that it could be counterproductive if students have not first established professional identities.

In 2002, almost 30 years after being introduced, the DIT was the subject of an extensive review involving 172 studies of moral reasoning among college students (King & Mayhew, 2002). This review documented that

“dramatic gains in moral judgment” are possible in college (p. 247), but it also cautioned that this growth “occurs in context.” For example, more growth was not found in students taking traditional accounting courses or spending time in Greek fraternities; it was, however, documented in students who have participated in well-designed ethics interventions, who have been exposed to rich and different experiences followed by critical reflection, or who have networks of independent friends (not closed cliques). To design effective interventions, advised King and Mayhew (2002), one must understand more than the gross context, such as the discipline, field, or the institutional setting; one must know “the specific content and curricular approaches that make up any given academic discipline” (p. 255).

This current study, then, sought to introduce the DIT and its conceptual model, the FCM, as a theoretical framework for research on moral reasoning and moral development in educational leadership preparation. Our hope is to lay the groundwork for future studies on the EDL/EDA graduate school context and suggest possible interventions, where needed.

Method

This study was initially conceived as a “census study” (Creswell, 2008, p. 394) with the goal of creating a moral reasoning profile of EDL/EDA graduate students and informing higher education faculty and administrators in one state in the South. Hence, the choice of instrument, the delivery method, and the timing of the study were all carefully considered to facilitate ease of participation by graduate students enrolled in EDL/EDA programs in the five doctoral-granting institutions.

Research Design and Participants

In the summer of 2012, the primary investigator began working with institutional gatekeepers in the five target institutions to offer the DIT-2 (as an electronic questionnaire) via e-mail to 539 master’s, educational specialist, and doctoral students in the state. Through the gatekeepers, researchers secured institutional permissions and student e-mails. Participation in the study was entirely voluntary (not linked to any course or grade). Both the institutions and the participants were ensured confidentiality, based on protocols approved by the researchers’ Institutional Review Board. After receiving a short introductory e-mail from an EDL/EDA program coordinator at their institution, students received the first in a series of invitation

e-mails from the primary investigator, with an explanation of the study, the board's protocol information, the host research institution, and the link to the questionnaire on SurveyMonkey.

Participants included EDL/EDA graduate students from the five doctoral-granting universities in the state: an urban medical center and public research university, a public research and teaching university, a public land grant and research university, a private faith-affiliated university, and a historically Black university. In all, of the 539 students contacted, 205 responded to the e-mail invitation, with 10 electing not to participate. Of the 195 respondents who initially indicated that they would answer the questionnaire, 113 actually completed it, for a final response rate of approximately 21%. Nonresponders did not supply sufficient data for analysis, so it is not known why they stopped taking the questionnaire or how they would have responded.

Instrument: Defining Issues Test

The DIT (Rest, 1979) consists of five hypothetical dilemmas, each followed by 12 action items. Participants must first decide what the protagonist in the dilemma should do; then, they must rate and rank the items in terms of importance in their interpretation of the dilemma. Historically, the summary score of the DIT has been the P score, calculated from ranking data and attending to items keyed to Kohlberg's (1969) Stages 5 and 6, the highest on his scale. A newer score, N2, developed in the late 1990s, adjusts the P score up or down on the basis of an individual's ability to discriminate between higher and lower levels when rating items as important or not important. The N2 score has been described as a useful tool for assessing older adults, presumably at the higher end of the developmental scale (Thoma, 2006); hence, it is also reported in this study of graduate students.

As described previously, the constructs measured by the DIT have been reinterpreted as reflecting decisions based on schemas and not stages (Rest et al., 1999). According to large-sample analyses, the DIT measures three developmentally ordered schemas: personal interest (incorporating aspects of Kohlberg's Stages 2 and 3), maintaining norms (closely aligned with Kohlberg's Stage 4), and postconventional (the traditional P score, aligned with Kohlberg's Stages 5 and 6). The validity and reliability of the DIT is fully discussed in work by Rest et al. (1999). Additional questions regarding demographics and possible distractions in the test-taking environment were included as part of the DIT-2 for norming purposes. Information and materials related to the DIT and DIT-2, including a manual, can be obtained from the Office for the Study of Ethical Development (2013) at the University of Alabama.

Table 1. Comparison of Educational Leadership/Administration Graduate Students With National Norms: Defining Issues Test–2 Scores for Respondents and 2005-2009 Composite Sample.

Score	M	SD
Educational leadership/administration graduate students ($N = 113$)		
Personal interest (Stage 2/3)	25.50	11.27
Maintain norms (Stage 4)	39.16	12.33
Postconventional (P score)	29.98	13.70
N2 score (N2 score)	29.27	13.69
Composite sample from 2005-2009 national norms ($N = 15,496$) ^a		
Personal interest (Stage 2/3)	20.61	11.46
Maintain norms (Stage 4)	34.07	14.36
Postconventional (P score)	41.06	15.22
N2 score (N2 score)	41.33	14.47

^aOffice for Ethical Study of Development.

Results

In this study, researchers set out to create a baseline profile of moral reasoning scores and preferred schemas for educational leadership graduate students in one Southern state and to compare this profile with a national norm for graduate students across disciplines. The research was prompted by concerns that educators, as a profession, may perform below average on a standard test of moral reasoning, the DIT, as well as by research showing that graduate school offers a critical window of opportunity for professional and moral growth.

EDL/EDA Students' Baseline Profile and Comparison With National Norms

For the first question, regarding a baseline profile of moral reasoning capacities, the sample's DIT-2 scores and summary statistics were provided by the Office for the Study of Ethical Development as part of the scoring service for use with the instrument and the office's ongoing norming process. As can be seen in Table 1, mean scores for the entire sample ($N = 113$) clustered around the maintaining norms schema ($M = 39.16$, $SD = 12.33$), indicating that it is the group's preferred schema for default decision making. The remainder of the scores are almost split between the more advanced postconventional schema ($M = 29.98$, $SD = 13.70$) and the less advanced personal interest

schema ($M = 25.50$, $SD = 11.27$). As expected, the N2 score ($M = 29.27$, $SD = 13.69$) was highly correlated with the P score ($M = 29.98$, $SD = 13.70$), given the observed value ($r = .89$) in this sample.

The EDL/EDA students' mean scores were then compared with national norms through a composite sample of graduate (master's/doctoral) student DIT scores from 2005 to 2009, supplied by the Office for the Study of Ethical Development. The composite sample included graduate students ($N = 15,496$), all native English speakers, from a variety of disciplines and all regions of the country. Independent samples t tests were conducted to compare the P scores and N2 scores of the EDL/EDA state sample with the P scores and N2 scores of the national composite sample. There were significant group differences between the state and national samples. For the P scores, there was a significant difference between the EDL/EDA sample ($M = 29.98$, $SD = 13.70$) and the national composite sample ($M = 41.06$, $SD = 15.22$), $t(112) = -8.60$, $p < .00$. For the N2 scores, there was a significant difference between the EDL/EDA sample ($M = 29.27$, $SD = 13.69$) and the national composite sample ($M = 41.33$, $SD = 14.47$), $t(112) = -9.34$, $p < .00$. Effect sizes for P scores and N2 scores were .77 and .86, respectively, indicating practical significance. (Effect sizes of greater than 0.33 standard deviations are typically considered to be practically meaningful.) Taken together, these two tests indicate that the EDL/EDA students scored significantly lower on postconventional moral thinking than the national composite sample. Moreover, the modal value for the EDL/EDA students is associated with the maintaining norms schema, whereas the modal value for the national composite is associated with the postconventional schema. This contrast in preferred schemas highlights the ethical decision-making differences between the EDL/EDA sample and the larger population of graduate students.

Demographics and Online Test-Taking Environment

Although no research questions addressed demographics, these optional questions were also analyzed, according to descriptive statistics, in an attempt to better understand the DIT P scores. Of the 82 respondents to demographics, more than two thirds (71%) self-identified as White, with 28% identifying as Black and one as Asian or Pacific Islander. The average age of the group was 41 years, with two thirds (66%) female and one third (34%) male. The majority (58%) sought a master's degree, with most of the remaining (41%) pursuing a doctorate (PhD or EdD) and with five students working on an educational specialist degree (EdS). All respondents considered English their primary language, which was important for comparative analysis with

national norms, and all but one was a U.S. citizen. Finally, students were asked to indicate a direction of their political views, from “very conservative” to “very liberal.” Nearly half the sample (48%) identified as somewhat or very conservative, with 20% identifying as neither or politically neutral and with 32% as somewhat or very liberal.

Of the demographics, only one—self-identified political view—was a statistically significant factor in influencing moral reasoning. A one-way between-subjects analysis of variance was conducted to compare the effects of political attitudes (independent variable) on DIT P scores (dependent variable) for the liberal ($n = 31$), neither ($n = 19$), and conservative ($n = 47$) subgroups. There was a significant effect of political attitudes on P scores at the $p < .05$ level, $F(2, 94) = 4.72, p = .01$. Given the statistical significance with this variable, researchers computed a Tukey post hoc test to compare each of three conditions to every other condition. Post hoc comparisons based on the Tukey honest significant difference test indicate that the mean score for the liberal subgroup ($M = 36.32, SD = 15.43$) is significantly different from that of the neutral subgroup ($M = 25.16, SD = 14.05$) and the conservative subgroup ($M = 28.72, SD = 12.02$). However, the neutral subgroup does not significantly differ from the conservative subgroup. Taken together, these results suggest that high levels of political liberalism have a positive effect on DIT P scores and high levels of political conservatism have a negative effect on DIT P scores.

Finally, as indicated in the Method section, the DIT-2 was recently adapted to the online environment, and the instrument includes a standard series of questions for respondents to help researchers assess the problem of distraction in the virtual test-taking environment. This is an especially important consideration with a cognitively complex test such as the DIT-2 (Bebeau & Thoma, 2003). The majority of participants (91%) said that they took the test in one sitting, with 88% indicating that they took it in the same way or almost the same way as they would in the classroom. The most notable possible distraction was having the television on (37%), with 20% of respondents reporting lesser interruptions, including receiving phone calls or text messages, replying to text messages, and engaging in conversation. Although this information was surprising—given that the EDL/EDA students routinely administer academic tests and would not allow similar conditions for their students—further statistical analysis of these data failed to show any statistically significant distractors or effect on P scores.

Discussion

In this empirical study, researchers used the DIT to establish a baseline profile of moral reasoning for a sample ($N = 113$) of EDL/EDA graduate students

in one Southern state, and they compared that profile with a national norm of graduate students across disciplines ($N = 15,496$). For the Southern state sample of educators, the researchers document an average DIT P score of approximately 30, very low by historical averages (53) in graduate students across disciplines. Researchers also document an EDL/EDA Southern state preference for conventional moral thinking (or desire to maintain norms). Just as troubling is the finding that the EDL/EDA sample's mean P score is significantly lower than that of the national sample (41), which prefers the more advanced decision-making schema—postconventional thinking. Given that formal education, such as an undergraduate college degree, is one of the strongest predictors for advanced moral reasoning, the question becomes, what has happened in this Southern state to arrest the development of its education graduate students?

These two key findings and the suggestion of arrested development are explored in this discussion, as well as implications for research and practice. If ethics are truly “the heart of leadership” (Ciulla, 2004), then interventions are warranted for this group but not the traditional ones. Although the final sample size was modest, researchers evaluated one fifth of the state's future educational leaders, who will eventually lead a much larger number of teachers, students, parents, and other community stakeholders. With such low scores, the central questions for higher education administrators in this state are not tactical ones about what to teach in a single ethics course and how to teach it; rather, they are more urgent and strategic: What works in ethics education for graduate students and professionals? What can we learn from proven models in other fields, such as medicine, dentistry, and law? How do we integrate proven models for fostering ethical growth into the field of educational leadership?

Low EDL/EDA P Scores: Arrested Development or Potential for Growth?

Historically, P scores increase by educational level, with senior high students averaging in the 30s, college seniors in the 40s, students graduating from professional programs in the 50s, and moral philosophy/political science doctoral students in the 60s (Bebeau & Thoma, 2003). At first glance, then, the EDL/EDA sample's mean DIT score, 30, appears much lower than what would historically be expected for the average American adult (40) and is on par with that of a senior high school student (30). In fact, only 10 respondents in the EDL/EDA sample have individual P scores of 50 or above, which would approximate a historic average (53) for graduate and professional students (Rest, 1979; Rest et al., 1999). Notably, however, the average P score

for the national composite group of graduate students (41) is also lower than historic averages. This lower score for the composite group confirms an overall decline in scores on the DIT over the last two decades (Thoma, 2012, personal conversation). Neo-Kohlbergians attribute this steady decline in moral reasoning to larger societal influences (discussed below).

Regardless of this larger trend, the EDL/EDA sample's scores are still significantly lower than the current national norm for graduate students. For example, only 26 of the EDL/EDA students (23% of this sample) had individual P scores around or above the average (41) for graduate students in the 2005-2009 national composite. In addition, three recent comparable studies on educators of similar educational and occupational levels provide evidence for interpreting the EDL/EDA sample's low P score. Ducut (2005) found a mean P score of 40.5 for EdD students ($N = 60$) at Pepperdine University, a religiously affiliated school in California. Slavinsky (2006) found a mean P score of 42 for Connecticut school principals ($N = 64$). Vitton and Wasonga (2009) found a mean P score of 39 for elementary school principals ($N = 60$) in the Midwest. Mean scores for these three studies, then, are still 9 to 12 points above this EDL/EDA sample's approximate average, 30. In sum, the majority of the participants in the EDL/EDA sample have not kept pace with the moral development of their peers at the national level, indicating substantial room for growth.

Practically speaking, the low P scores for the sample are worrisome because moral reasoning is not just an ethical capacity. Moral reasoning has been linked to prosocial behaviors in education, from critical reflection on practice and facilitative classroom management to student-centered teaching and respect for diversity (Chang, 1994; Cummings et al., 2007; Cummings et al., 2010; McNeel, 1994; Reiman, 2002; Vitton & Wasonga, 2009). Researchers who have recently documented low P scores in educators (Cummings et al., 2007; Cummings et al., 2010; Vitton & Wasonga, 2009) have argued that postconventional moral reasoning equips leaders to take a more sophisticated, multiethical approach to solving moral dilemmas in today's diverse schools and communities. Additionally, myriad studies (across the disciplines) reviewed by Rest and colleagues (1999) have linked high DIT scores with a superior understanding of moral concepts, higher scores on other developmental instruments, professional behaviors assessed for job performance, and better recall and reconstruction of moral arguments. Just as important, low levels of moral reasoning have been linked to negative outcomes (e.g. professional disciplinary problems) in some fields. For example, Baldwin and Self (2006) found a link between low DIT scores and medical malpractice claims. Bebeau found a similar connection in a study of 41 dentists recommended for ethics assessment by their state board. In eight

cases where disciplinary action was taken for providing substandard specialty care, “seven of the eight had moral reasoning scores below the mean for dental graduates, and five of the eight had very low scores (DIT P scores in the low 30s)” (Bebeau & Monson, 2008, p. 574).

Educators Prefer Status Quo: National Sample Prefers Postconventional Thinking

Unlike the national interdisciplinary sample of graduate students who prefer the postconventional schema (discussed below), the EDL/EDA students employ the maintaining norms schema—conventional, hierarchal, by-the-book decision making—as a default mode. Their preference for the status quo allows for moral certainty, uniform application of policy, and a sense of doing one’s duty (Rest et al., 1999). At first glance, these goals seem favorable. Yet, when exaggerated and unchallenged (as is common in a group of like-minded people), conventional thinking may place social order over civil liberties and human rights, resulting in a strongly authoritarian approach resistant to change and biased against those perceived to be different (Narvaez & Bock, 2002). For example, an educational leader with a maintaining norms preference may understand the importance of professional “fairness” but may treat others fairly only if they belong to his or her social group. Alternatively, such a rules-based conventional schema may assist decision making to meet federal and state legal mandates but provide little support for real-life problem solving in “best interests of students” dilemmas (Frick et al., 2012). As a result, conventional moral thinking offers today’s educational leader little flexibility for resolving complex questions that are not in the policy manual, which may itself be inadequate, underresourced, and in need of reform. Hence, if one returns to the University Council for Educational Administration’s (2013) “Code of Ethics for the Preparation of Educational Leaders,” one wonders how many of these EDL/EDA students are prepared to meet even the first goal: “the capacity to critique and challenge the status quo within the field of educational leadership.”

In the national sample, however, students prefer the postconventional schema as their default. This schema prioritizes moral ideals and relies on theoretical frameworks for resolving complex moral issues (Bebeau & Thoma, 2003; Rest et al., 1999; Thoma, 2006). Postconventional thinkers respect social norms but place a primacy on moral criteria, in contrast with more pragmatic claims; they draw on shared ideals that are fully reciprocal, not hierarchal, and hence not bound by status or class. They are given to self-reflection, making decisions that are open to scrutiny, based on logical criticism or the collective experience of the community (Thoma, 2006, p. 79).

Personal Interest	Maintaining Norms	Postconventional
<ul style="list-style-type: none"> • What do I gain or lose? • High School • DIT-P scores 20-30s 	<ul style="list-style-type: none"> • How do we keep the status quo? • College (4 years) • DIT-P scores in the 40s 	<ul style="list-style-type: none"> • How do we all live better together? • Graduate School • DIT-P scores 50s & 60s

Figure 2. Moral schema theory.

Note. Based on a description from Rest, Narvaez, Bebeau, and Thoma's (1999) *Postconventional Moral Thinking: A Neo-Kohlbergian Approach*.

Although they obey the law, postconventional moral thinkers recognize the possibility of an unjust rule or law and will work to overturn it (American examples include 19th-century child labor laws or 20th-century segregation laws). Such a thinker draws a widening circle of cooperation around himself or herself to encompass all members of the school community. "Individuals with the full use of postconventional tools are able to function at the highest levels of solving moral dilemmas within the community" (Narvaez & Bock, 2002, p. 306; see Figures 2 and 3).

Given valuable lessons from prior research on moral exemplars (Colby & Damon, 1992; Rule & Bebeau, 2005; Walker & Frimer, 2007), case summaries were conducted on the top 10 scorers in the EDL/EAD sample—those with P scores ranging from 48 to 66. Although providing only a snapshot of the students' moral capacities, these summary data are interesting. The top scores were split by gender, with four master's students, four PhD/EdD students, and two EdS students. Politically, seven were liberal, two were conservative, and one was neutral. The top two scorers, with P scores of 66, were both women on doctoral tracks; one self-identified as very conservative, the other as somewhat liberal. It is not possible to state, even to speculate, why the top scorers in this study scored so high; these questions are beyond the scope of this study. Still, this snapshot of DIT scores suggests that additional qualitative research is needed in EDL/EDA to identify moral exemplars who stand out among their peers in all four neo-Kohlbergian components of moral development—sensitivity, reasoning, motivation, and action. Scholars such as the aforementioned Colby and Damon (1992), Rule and Bebeau (2005), and Walker and Frimer (2007) provide excellent models for such studies in other disciplines.

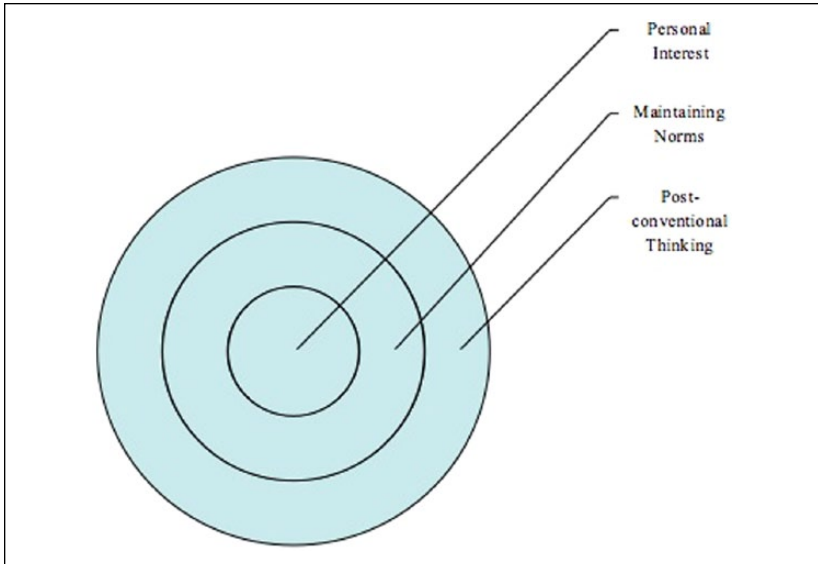


Figure 3. Postconventional thinking: A widening circle of cooperation. Postconventional thinking is an advanced developmental moral schema built around one core concept: cooperation (Rest, Narvaez, Bebeau, & Thoma, 1999). The postconventional decision maker thinks beyond the range of self or a homogeneous circle of society to seek diverse viewpoints that enrich and challenge his or her perspective. He or she prioritizes moral criteria over the pragmatic, appeals to shareable ideals, and practices full reciprocity in application of norms. He or she welcomes scrutiny or new evidence for decision making, questions the status quo, and engages in critical reflection. In groups, he or she is more facilitative than authoritative, working to build logical and coherent consensus. Based on descriptions from Rest and colleagues' (1999) *Postconventional Moral Thinking: A Neo-Kohlbergian Approach*.

Factors Possibly Influencing the Low Scores for Moral Reasoning

A variety of demographic factors—especially education, gender, and political attitudes—have been studied extensively for possible influences on moral reasoning, as measured by DIT scores (Rest et al., 1999; Thoma, 2006). Although demographics were not the focus of the original research questions in this study, subsequent analysis showed that political attitudes do contribute to differences in the EDL/EDA participants' schema preferences, confirming prior research showing that fundamentalist or ideological views correlate with lower P scores (Narvaez, Getz, Rest, & Thoma, 1999; Rest et al., 1999). Although this ideological factor has not been studied extensively in educational leaders, Vitton and Wasonga (2009) found a conservative political

view to be the only significant demographic factor depressing P scores in their study of elementary school principals in the Midwest. Additionally, in the present study, given that nearly 40% of respondents answered the questionnaire with the television on, data regarding the online test-taking environment were further examined for potential distractions. None of these, however, proved to be statistically significant. Finally, some prior studies point to lower-than-average DIT scores for educators as a subgroup, and the bulk of the literature shows that context does indeed influence moral reasoning. Hence, the ethical culture of the K-12 school environment and the curriculum in preparation programs may stimulate or arrest growth in moral reasoning. These factors are discussed in turn.

Contexts that Stimulate or Arrest Moral Development

For the past two decades, the United States' documented preoccupation with individualism and narcissism—qualities linked with the personal interest schema—have concerned DIT researchers as they watched average P scores decline nationally across all contexts (Thoma, 2012, personal communication). Neo-Kohlbergians who have studied moral reasoning and cultural psychology believe that low DIT scores could be a result of “reciprocal dynamics” (Rest et al., 1999, p. 180) between culture and moral development, where the former fosters or inhibits the growth of the latter. College campuses—once ideal environments for moral growth—may no longer provide such learning grounds. Recent studies show that “empathy and moral reasoning among college students are decreasing whereas egocentrism is up” (Narvaez, 2010). Such a decline is not limited to students, writes Bertram Gallant (2011), who argues for a systems approach to reforming the ethical culture of higher education. At the K-12 level, societal forces may also stymie moral growth among school leaders. “The culture of public schools is more likely to reflect [society’s] vices than to counter them” (Strike, 2008, p. 131). Indeed, today’s schools, writes Lindle (2004), seldom offer “a safe space” for critical thinking: “For many school administrators, the political and logistical realities of their schools suppress deeply reflective thinking by leaders, teachers, or pupils” (p. 170). Given today’s high-stakes testing pressures, one wonders—in the K-12 environment, where many of the EDL/EDA sample participants currently work, is conventional thinking more expedient and efficient at advancing many children through the system than postconventional thinking that aims to ensure “the best interests of the student” for all children?

Although some researchers (Derryberry et al., 2006; King & Mayhew, 2002; Livingston et al., 2006; Rest, 1986) have asserted that evidence of disciplinary differences is inconsistent, others have offered evidence for negative

“reciprocal dynamics” in the field of education. As early as 1976, Bloom reported low P scores—an average of 30—for education graduate students ($N = 82$). In 1988, Yeazell and Johnson repeated these concerns when they found a mean P score of 43 ($n = 33$) for education graduate students, with no significant difference in P scores of 38 for education undergraduate students ($n = 38$). After several empirical studies on teachers and a review of the DIT literature on educators, Cummings and associates (2007) posited two possible reasons for persistent low scores. First, they wondered if education, as a practitioner-oriented field, attracts and self-selects for students who are less inclined to be critical ethical thinkers when compared to those in other fields, such as the humanities. Second, the researchers believed that standard teacher training curriculum is too technical and task oriented to allow students to develop their critical ethical thinking skills. The authors based their conclusion on an exhaustive review of curriculum for elementary school teachers. Similarly, in a study of moral reasoning in principals, Vitton and Wasonga (2009) mused that inadequate preparation in moral leadership could be one of the reasons why the group did not activate postconventional schemas 40% of the time. The authors suggested that two additional factors contributed to the principals’ overall low moral judgment scores: fixed mental maps and changing regulatory and school environments that create more complex ethical dilemmas. Vitton and Wasonga’s conclusions seem to support the neo-Kohlbergians’ ideas about “reciprocal dynamics,” whereby culture influences moral judgment, which in turn solidifies culture, making it resistant to change.

Importantly, numerous studies of the DIT show that educational context matters a great deal in the success or failure of moral development education. For example, high P scores are correlated with rich, intellectual environments where moral principles/priorities are challenged, tested, and debated from multiple points of view (Rest et al., 1999). Low P scores are correlated with traditional academic settings where authority is not challenged, conformity is rewarded, and technical knowledge is prioritized over liberal arts. For example, in a recent meta-analysis of DIT data and educational context (Maeda, Bebeau, & Thoma, 2009), researchers found that medical students could expect to score an average of 7.1 points higher on the DIT’s P scale than other graduate students, when all other conditions are controlled. The authors observed that the data for their study were gathered during a time of rapid growth of professional ethics education in the field of medicine. Earlier, Bebeau and Monson (2008) reviewed 33 moral reasoning studies of educational interventions in medicine, dentistry, law, and veterinary science and found that professional schools had cause for concern: “Professional school educational programs do *not* promote moral judgment development unless the program includes a *well-validated* ethics curriculum” (p. 570).

Additionally, while EDL/EDA graduate students are often older than traditional graduate students (the mean age was 41 in our sample), note that age is much less of a predictor of DIT scores than the context of formal education. For example, college students in their 20s have higher moral reasoning scores than older adults in their 50s with no formal education (Rest et al., 1999). Research also shows that DIT scores increase while a person is involved in formal education, then plateau when that person leaves the educational environment. Importantly, however, there is no empirical evidence to show that individuals regress, or move backward, in their moral reasoning capacities as measured by the DIT. Hence, while the question of “nontraditional students” in the EDL/EDA field is a valid one, it probably turns on their preparation or pathway to the profession, rather than their average age. As we indicate in our recommendations, these methods of preparation and pathways offer areas for future study.

Limitations of the Study

These findings must be viewed in light of several limitations. First, the response rate to the questionnaire was 21% ($N = 113$), and respondents came from the five advance-degree granting institutions for educational leadership and administration in one state in the South. Hence, given the sample's size and demographics, these scores are not necessarily generalizable to other EDL/EDA students, including those outside these schools or this state. Second, the study involved an online questionnaire with a complex cognitive test, the DIT-2, and additional questions, all of which took from 45 to 60 minutes to complete. Of the 195 students who initially responded yes, a smaller number, 113, finished the assessment. It is not known how the scores of the noncompleters would have affected this sample's overall mean P scores. Third, although the majority of completers (91%) said that they finished the questionnaire in one setting, at least 20% took it with potential distractions, including nearly 40% with the television on. Although these distractions did not prove to be statistically significant, they could have influenced student scores in ways that are not known.

Recommendations for Practice: Move Toward What Works

Moral development in the professions does not rest on a one-shot required ethics course (Bebeau & Monson, 2008) but an “ethics across the curriculum” approach as well as an “ethics beyond school professional development” approach. Moreover, instruction must address more than growth in moral reasoning. Evidence from many intervention studies in other

professions supports a holistic, integrated moral development strategy, such as Rest's FCM, which encompasses all four psychological processes identified in moral behavior—sensitivity, reasoning, motivation, and character. Because the model involves multiple processes, not prescribed content, the FCM can be used as a curricular framework to promote moral development in support of discipline-specific learning objectives. Such a curricular revision should not be “patchwork,” however; rather, it should be foundational, according to a recent well-described and well-documented FCM-based ethics intervention for teachers at Winthrop University (L. E. Johnson, Vare, & Evers, 2013, p. 110).

In addition to offering a curricular framework, the FCM literature suggests signature pedagogies that involve many authentic instruction techniques that faculty already use, as described in Bebeau and Monson (2008). For example, to enhance moral sensitivity (Rest Component 1), the authors use stimulus case studies, analysis of cause-consequent chains, and empathy activities that “present clues to a moral problem without ever signaling what issue is at stake” (p. 569); these Component 1 activities are assessed with student written analysis and evaluation rubrics, along with more specific quantitative assessments, such as the Dental Ethical Sensitivity Test or the Racial Ethical Sensitivity Test. To develop moral reasoning (Rest Component 2), they use dilemma discussion, problem-solving, and role-taking activities; these Component 2 activities are assessed with the DIT-2 and a newer quantitative instrument, the Intermediate Concept Measure, developed for assessing moral reasoning with profession-specific dilemmas. To activate moral motivation (Rest Component 3), or the prioritization of moral action, they use discipline-specific role-plays and simulations, as well as self-reflection on professional principles and commitments that foster professional identity formation. To assess these Component 3 activities, they use the Role Concept Essay, the Professional Role Orientation Inventory, and a series of qualitative short-answer questions with a rubric that describes how a student's professional identity is evolving along a continuum of three types: independent operator, team-oriented idealist, and self-defining professional (Bebeau & Thoma, 2013, p. 488). To promote development in moral character and competence (Rest Component IV), they employ action and implementation plans (emerging out of problem solving), along with performance assessments and case simulations. They recommend spreading the case simulation type of assessment for Component IV throughout the curriculum so different members of the faculty have an opportunity to assess the character and competence of students. Such an integrative approach “sends a powerful signal of the valued placed on ethics education by the school.

One important key to success in fostering moral development in the professions may be the sequencing of instruction (Bebeau & Monson, 2008; Bebeau & Thoma, 2013). Although dilemma discussion courses to enhance moral reasoning are among the most common starting places, recent research shows that such discussions with graduate and professional students can do them a “disservice” if they have not already begun to form a new professional identity (Bebeau & Monson, 2008, p. 575). If, for example, a graduate student is asked to have an opinion on an ethical dilemma before he or she understands his or her role in relation to professional and societal expectations, “it may encourage a defensive stance on personal moral values, rather than open reflection upon what it means to become professional, and, in effect, exploring whether the profession’s value system and one’s own are congruent” (p. 575). For instance, students of educational leadership in their first course in the preparation program have not yet had extensive field experiences or shadowing opportunities with practicing school leaders, so they are not yet viewing themselves as administrators. Hence, it is difficult for them to predict, even with role-plays and case studies, how they will process an ethical decision-making dilemma. Instead, Bebeau and Monson recommend that professional students need to begin with a strong grounding in professional identity development, which activates moral motivation (Rest Component 3). Research shows that most students “do not come to professional school with a clear vision of society and professional expectations, and are not likely to intuit them from the general educational process” (p. 575). Thus, professional identity development and moral development go hand in hand.

The suggestion of a need for evidenced-based moral development instruction naturally gives rise to questions about a school’s existing programs and whether they foster or inhibit ethical growth. As previously noted, Cummings et al. (2003) and a panel of experts reviewed more than 500 courses—across 30 institutions—and found 90% of the curricula to be methods or skills driven and not conducive to developing critical ethical judgment. Although no such study has been conducted in the state where the EDL/EDA sample was assessed, faculty at the five studied schools may want to consider a self-reflective program review for two reasons. First, many educational leadership certification programs in the state were redesigned according to recommendations by the Southern Regional Education Board (2006)—namely, that graduate course work in this state be more experiential and less theoretical, helping educational leaders focus on maximizing student achievement in an era of high-stakes testing in K-12. Questions remain about whether the current curricula best serve the educational goals of all graduate students in the programs—master’s, EdS, EdD, and PhD. Second, new research at the

national level suggests that the old debate of practitioner vs. scholar may be giving way to a new paradigm for doctoral studies that captures the best of both educational priorities: the professional. Extensive studies of doctoral education and the education doctorate are underway by the Carnegie Foundation for the Advancement of Teaching and Learning (Shulman, 2010). Scholars such as Shulman (2010) are increasingly believing that a perceived divide between professional schools and arts-and-sciences doctoral programs is both “distracting” and “dysfunctional” (p. 3). “Properly understood, doctoral programs in the arts and sciences are truly (or should be) programs of professional education” (p. 2). This potential new direction for doctoral studies makes the work of Bebeau and colleagues, who have studied moral development in the professions for more than three decades, even more timely for schools of education to consider embracing.

Implications for Research

First, more DIT studies are needed on this population—in other institutions, states, and regions of the country—to provide a baseline profile of the moral reasoning schemas of future educational leaders and to help confirm or dismiss a possible gap for educators as a field. Second, researchers should consider using the DIT as a pretest/posttest measure to assess the value of existing EDL/EDA ethics courses for growth in moral reasoning. Also, given that the entire graduate school experience should result in significant moral reasoning growth (Rest, 1986; Rest & Narvaez, 1994; Rest et al., 1999), administrators could test EDL/EDA students as they come into a program and again when they graduate. Data from the DIT, which is a test of life span development, could eventually be used in longitudinal studies to track students’ moral reasoning growth through the program and through their profession. Rogers (2002) argues effectively for tracking moral reasoning, along with the development of moral self, beyond college, and Bebeau has conducted extensive research with working professionals (Bebeau & Monson, 2008). At the graduate and professional level, students themselves would benefit from knowing how they compare with others in their cohort and where they need to focus their energies for self-directed learning (Bebeau & Monson, 2008). One note about method of delivery for the DIT: Although it is available online, the DIT is cognitively demanding and more time-consuming (40+ minutes) than most web-based surveys (15 minutes). Researchers may want to consider face-to-face group administration as a first mode of delivery for the DIT to help ensure higher response rates and minimize the influence of distractions.

Second, more qualitative research addressing profession-specific dilemmas is needed to develop valid and reliable assessments to use with the DIT

and to evaluate the effectiveness of moral development and FCM instruction in the field of EDL/EDA. However, this qualitative research should be carefully designed to address new questions. For example, the advantage of the DIT is that it is a cognitive psychological test of problem-solving macromoral dilemmas, not a self-report survey that invites biased assessment of how ethical a person thinks he or she is. Furthermore, based on the compelling evidence for moral schema theory, the DIT issue items are believed to accurately map participants' macromoral reasoning because they emerged from analysis of qualitative interview data on thousands of people, women and men, in all walks of life. For this reason, additional qualitative interviews with study participants regarding the DIT dilemmas might provide little new information. However, as suggested, focused qualitative research on the EDL/EDA graduate student population would help educational researchers develop multiple modes of assessment as well as profession-specific dilemmas for instruments to holistically evaluate student needs and learning (see FCM strategies and tools in the previous section). Given recent counsel from neo-Kohlbergians to begin instruction with moral motivation (Rest Component 3; Bebeau & Monson, 2008), researchers are encouraged to consider the use of qualitative methods in conjunction with the professional identity methodology well described by Bebeau and Thoma (2013).

Conclusion

This study documents significantly low scores for moral reasoning in a sample ($N = 113$) of educational leadership graduate students in five doctoral-degree granting institutions in one Southern state. The scores are based on a well-validated and easily administered test of life span development: the DIT-2. To our knowledge, no other researchers have attempted a census-type study to assess the ethical thinking capacities of the future educational leaders in one state. Because the design of higher education programs is influenced by state and regional leaders, such a perspective is important for leaders and policy makers at all levels.

These findings confirm other prior research, albeit inconclusive, that describes three decades of concern regarding arrested moral development in education students. In search of new solutions, the authors grounded their questions, methods, and findings in an evidenced-based process approach—namely, Rest's FCM, which has been employed to promote moral development in other professions, such as medicine, dentistry, and law. Given the relatively small sample size, some might dismiss these findings as small scale or isolated and hence not particularly worrisome. Considering the low scores, others might see these findings as good news and use them to politically

condemn an educational system already under fire. The authors, however, see these findings as an opportunity for growth for the educational leadership profession. To that end, the researchers propose a blueprint for future action research/improvement studies that would serve to develop and assess professional identities, moral sensitivity, moral reasoning/judgment, and ethical character/competencies in educational leadership. As the late James Rest (1986) wrote,

people who develop in moral judgment are those who love to learn, who seek new challenges, who enjoy intellectually stimulating environments, who are reflective, who make plans and goals, who take risks, who see themselves in the larger social contexts of history and institutions, and broader social trends, who take responsibility for themselves and their environs. (p. 57)

Few would question the need to nurture more of these educational leaders in any state.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Note

1. Information about the national composite sample and materials related to four-component model, the Defining Issues Test, and the Defining Issues Test–2, including a manual, can be obtained from the Office for the Study of Ethical Development (2013) at the University of Alabama, <http://www.ethicaldevelopment.ua.edu/>.

References

- Baldwin, D. C., Jr., & Self, D. J. (2006). The assessment of moral reasoning and professionalism in medical education and practice. In D. T. Stern (Ed.), *Measuring medical professionalism* (pp. 75-94). New York, NY: Oxford University Press.
- Bartlett, F. C. (1995). *Remembering: A study in experimental and social psychology*. New York, NY: Cambridge University Press. (Original work published 1932)
- Bass, B., & Riggio, R. (2006). *Transformational leadership*. Mahwah, NJ: Erlbaum.
- Bebeau, M. J. (2002). The Defining Issues Test and the four-component model: Contributions to professional education. *Journal of Moral Education, 31*(3), 271-295.

- Bebeau, M. J., & Monson, V. E. (2008). Guided by theory, grounded in evidence: A way forward for professional ethics education. In L. P. Nucci & D. Narvaez (Eds.), *Handbook of moral and character education* (pp. 557-582). New York, NY: Routledge.
- Bebeau, M. J., Rest, J. R., & Narvaez, D. (1999). Beyond the promise: A perspective on research in moral education. *Educational Researcher*, 28(4), 18-26.
- Bebeau, M. J., & Thoma, S. J. (2003). *Guide for DIT-2*. Minneapolis, MN: University of Minnesota, Center for the Study of Ethical Development. Retrieved from <http://www.ethicaldevelopment.ua.edu/>
- Bebeau, M. J., & Thoma, S. J. (2013). Moral motivation in different professions. In K. Heinrichs, F. Oser, & T. Lovat (Eds.), *Handbook of moral motivation: Theories, models, and applications* (pp. 475-498). Rotterdam, Netherlands: Sense.
- Bertram Gallant, T. (2011). *Creating the ethical academy: A systems approach to understanding misconduct and empowering change in higher education*. New York, NY: Routledge.
- Bloom, R. B. (1976). Morally speaking, who are today's teachers? *Phi Delta Kappan*, 57, 624-625.
- Bolman, L. G., & Gallos, J. V. (2011). *Reframing academic leadership*. San Francisco, CA: Wiley.
- Burns, J. M. (1978). *Leadership*. New York, NY: Harper & Row.
- Burns, J. M. (2003). *Transforming leadership*. New York, NY: Grove Press.
- Chang, F. Y. (1994). School teachers' moral reasoning. In J. R. Rest & D. Narvaez (Eds.), *Moral development in the professions: Psychology and applied ethics* (pp. 71-83). Hillsdale, NJ: Erlbaum.
- Ciulla, J. B. (2004). *Ethics, the heart of leadership*. Westport, CT: Praeger.
- Ciulla, J. B. (2005). The state of leadership ethics and the work that lies before us. *Business Ethics: A European Review*, 4(4) 323-335.
- Colby, A., & Damon, W. (1992). *Some do care: Lives of moral commitment*. New York, NY: Free Press.
- Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (3rd ed.). Upper Saddle River, NJ: Pearson Education.
- Cummings, R., Dyas, L., Maddux, C. D., & Kochman, A. (2001). Principled moral reasoning and behavior of preservice teacher education students. *American Educational Research Journal*, 38(1), 143-158.
- Cummings, R., Harlow, S., & Maddux, C. D. (2007). Moral reasoning of in-service and pre-service teachers: A review of the research. *Journal of Moral Education*, 36(1), 67-78.
- Cummings, R., Maddux, C. D., & Cladianos, A. (2010). Moral reasoning of education students: The effects of direct instruction in moral development theory and participation in moral dilemma discussion. *Teachers College Record*, 112(3), 621-644.
- Cummings, R., Wiest, L., Lamintina, D., & Maddux, C. (2003). Teacher education curricula and moral reasoning. *Academic Exchange Quarterly*, 7(1), 163-168.

- Derryberry, W. P., Snyder, H., Wilson, T., & Barger, B. (2006). Moral judgment differences in education and liberal arts majors: Cause for concern? *Journal of College & Character*, 52(4), 1-10.
- Ducut, J. (2005). *Nurturing the ethical reasoning of leaders: The illumination of ethics education, demographics and teaching methods* (Unpublished doctoral dissertation). Pepperdine University, Malibu, CA.
- Foster, W. P. (1986). *Paradigms and promises: New approaches to educational administration*. Amherst, NY: Prometheus Books.
- Foster, W. P. (2004). The decline of the local: A challenge to educational leadership. *Educational Administration Quarterly*, 40(2), 176-191.
- Frick, W. C., Faircloth, S. C., & Little, K. S. (2012). Responding to the collective and individual "best interests of students": Revisiting the tension between administrative practice and ethical imperatives in special education leadership. *Educational Administration Quarterly*, 49(2), 207-242.
- Fullan, M. (2003). *The moral imperative of school leadership*. Thousand Oaks, CA: Corwin.
- Gardner, H. (2008). *Five minds for the future*. Boston, MA: Harvard Business School.
- Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.
- Johnson, C. E. (2009). *Ethical challenges of leadership: Casting light or shadow*. Thousand Oaks, CA: Sage.
- Johnson, L. E., Vare, J. W., & Evers, R. B. (2013). Let the theory be your guide: Assessing the moral work of teaching. In M. N. Sanger & R. D. Osguthorpe (Eds.), *The moral work of teaching and teacher education: Preparing and supporting practitioners* (pp. 92-112). New York, NY: Teachers College Press.
- King, P. M., & Mayhew, M. J. (2002). Moral judgment development in higher education: Insights from the Defining Issues Test. *Journal of Moral Education*, 31(3), 247-270.
- Koebler, J. (2011). Educators implicated in Atlanta cheating scandal. *US News and World Report Education*. Retrieved from <http://www.usnews.com/education/blogs/high-school-notes/2011/07/07/educators-implicated-in-atlanta-cheating-scandal>
- Kohlberg, L. (1969). Stage and sequence: The cognitive developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory* (pp. 347-480). Chicago, IL: Rand McNally.
- Lapsley, D. K. (1996). *Moral psychology*. Boulder, CO: Westview Press.
- Laster, J. (2010). Malone U. president steps down amid plagiarism accusations. *Chronicle of Higher Education*. Retrieved from <http://chronicle.com/article/Malone-U-President-Steps-Down/64328/>
- Lindle, J. C. (2004). William P. Foster's promises for educational leadership: Critical idealism in an applied field. *Educational Administration Quarterly*, 40(2), 167-175.
- Livingston, G., Derryberry, W., King, A., & Vendetti, M. (2006). Moral development consistency? Investigating differences among academic majors. *Ethics & Behavior*, 16(3), 265-287.

- Maeda, Y., Bebeau, S. J., & Thoma, M. J. (2009). Understanding the relationship between moral judgment development and individual characteristics: The role of educational contexts. *Journal of Educational Psychology, 101*(1), 233-247.
- McNeel, S. P. (1994). College teaching and student moral development. In J. R. Rest & D. Narvaez (Eds.), *Moral development in the professions: Psychology and applied ethics* (pp. 27-49). Hillsdale, NJ: Erlbaum.
- Narvaez, D. (2010). The science of morality? Not so fast. *Moral Landscapes*. Retrieved from <http://www.psychologytoday.com/blog/moral-landscapes/201009/science-morality-not-so-fast>
- Narvaez, D., & Bock, T. (2002). Moral schemas and tacit judgment or how the Defining Issues Test is supported by cognitive science. *Journal of Moral Education, 31*(3), 297-314.
- Narvaez, D., Getz, I., Rest, J. R., & Thoma, S. J. (1999). Individual moral judgment and cultural ideologies. *Developmental Psychology, 35*(2), 478-488.
- Office for the Study of Ethical Development. (2013). *The DIT and the DIT-2*. Tuscaloosa, AL: University of Alabama. Retrieved from <http://www.ethicaldevelopment.ua.edu/the-dit-and-dit-2>
- Perez-Pena, R., & Slotnik, D. (2012). Gaming college rankings. *New York Times*. Retrieved from <http://www.nytimes.com/2012/02/01/education/gaming-the-college-rankings.html?pagewanted=all>
- Phillips, A. M. (2012). Nine schools cited for exam and credit irregularities. *New York Times*. Retrieved from <http://www.nytimes.com/schoolbook/2012/02/23/nine-schools-cited-for-exam-and-credit-irregularities/>
- Piaget, J. (1965). *The moral judgment of the child* (M. Gabaian, Trans.). New York, NY: Free Press. (Original work published 1932)
- Pritchard, M. S. (1999). Kohlbergian contributions to educational programs for the moral development of professionals. *Educational Psychology Review, 11*(4), 395-409.
- Reiman, A. J. (2002). *A comparison of four-year longitudinal studies of postconventional moral judgment reasoning in teacher education and other selected undergraduate samples*. Paper presented at the American Educational Research Association, New Orleans, LA.
- Rest, J. R. (1979). *Development in judging moral issues*. Minneapolis: University of Minnesota.
- Rest, J. R. (1983). Morality. In P. H. Mussen (Series Ed.), J. Flavell, & E. Markham (Vol. Eds.), *Handbook of child psychology: Vol. 3. Cognitive development* (4th ed., pp. 556-629). New York, NY: Wiley.
- Rest, J. R. (1986). *Moral development: Advances in research and theory*. New York, NY: Praeger.
- Rest, J. R., & Narvaez, D. (1994). *Moral development in the professions: Psychology and applied ethics*. Hillsdale, NJ: Erlbaum.
- Rest, J. R., Narvaez, D., Bebeau, M. J., & Thoma, S. J. (1999). *Postconventional moral thinking: A neo-Kohlbergian approach*. Mahwah, NJ: Erlbaum.
- Rest, J. R., Thoma, S. J., & Edwards, L. (1997). Designing and validating a measure of moral judgment: Stage preference and stage consistency approaches. *Journal of Educational Psychology, 89*(1), 5-28.

- Rogers, G. (2002). Rethinking moral growth in college and beyond. *Journal of Moral Education*, 31(3), 325-338.
- Rule, J. T., & Bebeau, M. J. (2005). *Dentists who care: Inspiring stories of professional commitment*. Carol Stream, IL: Quintessence.
- Shapiro, J. P., & Stefkovich, J. A. (2011). *Ethical leadership and decision making in education: Applying theoretical perspectives to complex dilemmas*. New York, NY: Routledge.
- Shulman, L. (2010). Doctoral education shouldn't be a marathon. *Chronicle of Higher Education*. Retrieved from <http://chronicle.com/article/Doctoral-Education-Isnt-a/64883/>
- Slavinsky, R. A. (2006). *Cognitive moral development of Connecticut public school principals as measured by Rest's Defining Issues Test* (Unpublished doctoral dissertation). University of Connecticut, Storrs.
- Southern Regional Education Board. (2006). *Schools can't wait: Accelerating the redesign of university principal programs*. Retrieved from http://publications.sreb.org/2006/06V04_Schools_Cant_Wait.pdf
- Starratt, R. J. (2004). *Ethical leadership*. San Francisco, CA: Jossey-Bass.
- Strike, K. A. (2008). School, community and moral education. In L. P. Nucci & D. Narvaez (Eds.), *Handbook of moral and character education* (pp. 117-133). New York, NY: Routledge.
- Strike, K. A., Haller, E. J., & Soltis, J. F. (2005). *Ethics of school administration: Professional ethics in education series*. New York, NY: Teachers College Press.
- Thoma, S. (1986). Estimating gender differences in the comprehension and preference of moral issues. *Developmental Review*, 6, 165-180.
- Thoma, S. (2002). An overview of the Minnesota approach to research in moral development. *Journal of Moral Education*, 31(3), 225-245.
- Thoma, S. (2006). Research on the Defining Issues Test. In M. Killen & J. Smetena (Eds.), *Handbook of moral development* (pp. 67-91). Hillsdale, NJ: Erlbaum.
- Thoma, S. J., & Bebeau, M. J. (2013). Moral motivation and the four-component model. In K. Heinrichs, F. Oser, & T. Lovat (Eds.), *Handbook of moral motivation: Theories, models, and applications*. (pp. 49-67). Rotterdam, Netherlands: Sense.
- University Council for Educational Administration. (2013). *UCEA code of ethics for the preparation of educational leaders*. Retrieved from <http://ucea.org/ucea-code-of-ethics/>
- Vitton, C. J., & Wasonga, T. A. (2009). Between Kohlberg and Gilligan: Levels of moral judgment among elementary school principals. *Leadership and Policy in Schools*, 8(1), 92-116.
- Walker, L. J. (1984). Sex differences in the development of moral reasoning: A critical review. *Child Development*, 55, 677-691.
- Walker, L. J., & Frimer, J. A. (2007). Moral personality of brave and caring exemplars. *Journal of Personality and Social Psychology*, 93(5), 845-860.
- Yeazell, M. I., & Johnson, S. F. (1988). Levels of moral judgment of faculty and students in a teacher education program: A micro study of an institution. *Teacher Education Quarterly*, 15(1), 61-70.

Author Biographies

Jennifer L. Greer is a lead instructor in the professional development program in the Graduate School at the University of Alabama at Birmingham. A former journalist and Peace Corps volunteer in adult education, she taught English as a Second Language in K-12 before joining the university. She designs and teaches ethics and writing curricula for graduate students, research staff, and postdoctoral fellows.

Linda J. Searby is an associate professor of educational leadership in the School of Education at Auburn University. She has 26 years of experience in K-12 teaching and administration as well as 9 years in higher education. The recipient of a 2012 Dean's Award for Excellence in Mentorship at the University of Alabama at Birmingham, she researches and publishes on mentoring and leadership.

Stephen J. Thoma is a professor and program coordinator of educational psychology in the School of Education at the University of Alabama in Tuscaloosa. The author of numerous articles on neo-Kohlbergian research, he was named a Paul W. Bryant Professor for research at the university for 2010-2011. He also serves as the executive director of the university's Office for the Study of Ethical Development.