

## A Multilevel Examination of the Influence of Institutional Type on the Moral Reasoning Development of First-Year Students

Understanding the relationship between educational contexts and the students nested within these contexts has been the subject of scholarly inquiry for many years. Theoretical considerations explaining this relationship have spanned many schools of thought, ranging from Lewin's (1951) algorithmic representation of behavior as a function of the interaction between personality and environment, to ecologists' more organic theoretical descriptions of person-environmental dynamics (see Bronfenbrenner, 1979), to organizational models designed to understand the impact of institutional covariates on student-level outcomes (see Berger & Milem, 2000). Based on the seemingly ubiquitous, interdisciplinary importance of interrogating contexts for their potential in explaining human development, statistical methods have been developed to account for individuals nested in their natural, albeit often non-random environments. Despite the emergence of more sophisticated statistical techniques, they remain underutilized by student development scholars, who, ironically, are often charged with understanding how educational contexts shape student learning and development (McEwen, 2003; Stage, 1991; Strange, 1994).

The purpose of this paper was to examine institutional type and its influence on explaining the moral reasoning development of first-year students. Hierarchical linear modeling was used to account for students

The author gratefully acknowledges the Wabash Center of Inquiry in the Liberal Arts for supporting this study.

*Matthew J. Mayhew is an Associate Professor of Higher Education at New York University.*

nested within 19 institutions. It is anticipated that this study will provide researchers with compelling evidence for adopting multilevel research designs for answering questions concerning the roles colleges play in shaping student development.

### *Theoretical Overview*

The relationship between educational contexts and moral reasoning development was an important consideration of Lawrence Kohlberg (1976), the theorist frequently associated with developing the first, testable theory of moral development for adults. Kohlberg reflected his understanding of this importance in his constructivist approach to theory-building, in his turn to social stimulation as the theoretical explanation for progression through developmentally-sequenced stages, and in his orientation toward understanding the relationship between theory and educational practice.

Like Piaget (1965), Kohlberg adopted a constructivist approach to nature of development where “the individual is always inventing or constructing new responses to each situation encountered” (Colby & Kohlberg, 1987, p. 4). This approach underscored Kohlberg’s rationale for adopting a sequential stage theory to describe the development of moral reasoning where “each new stage of development represents a qualitative reorganization of the individual’s pattern of thought, with each new re-organization integrating within a broader perspective the insights achieved at the prior stages” (p. 5). For Kohlberg, exposure to external, often unfamiliar stimuli served as the catalyst for evolving notions of justice and fairness from a system that serves the self to one that serves known others to one that serves anonymous others.

The mechanisms examined for inducing moral reasoning growth included stimuli that forced individuals to re-examine existing ideas of justice in light of the new information presented by a social, and often unfamiliar encounter. In Kohlberg’s words, “moral development depends upon stimulation defined in cognitive-structural terms, but this stimulation must also be social, the kind that comes from moral decision-making, moral dialogue, and moral interaction” (Kohlberg, 1976, p. 49). Individuals who engage the challenging work that accompanies encountering unfamiliar social stimuli progress to and through more advanced stages of moral reasoning; alternatively, individuals who disengage from such challenge may retreat to less sophisticated forms of moral reasoning (Mayhew & Engberg, 2010).

Vested in this idea of social stimulation as the mechanism for moral growth, Kohlberg (1976) developed his theory with an eye towards edu-

cational practice, indicting his peers for not doing the same. Kohlberg (1981) attributed part of the problem with existing psychological research to the “psychologist’s fallacy” (p. 37), the faulty assumption that the variables psychologists deemed as important for research were also those valued by educators in the design of their learning environments. According to Kohlberg (1981), psychological theory and measurement were at risk of losing their potential impact when constructed and tested in the absence of application. Although psychologists may err in trying to subsume the complexity of educational contexts and processes under well-established psychological theories and measures, educational researchers may err in relying too heavily on individual characteristics, to the exclusion of accounting for the role of context, for explaining moral reasoning development.

It is clear that Kohlberg valued the influence of educational contexts in shaping moral reasoning development. He credited contexts and their embedded social processes for informing his constructive approach to theory building, for helping individuals progress through moral reasoning stages, and for providing insight into the variables important to consider when interrogating moral development. One such variable, institutional type, has often been used as one way to examine context and its influence on student outcomes (see, for example, Berger & Milem, 2000). As a gross indicator of context, it has often been assumed that students will report vastly different experiences based on exposure to the educational practices (and social stimuli) associated with different institutional types (see Pascarella, Wolniak, Seifert, Cruce, & Blaich, 2005). This study makes this assumption as well, presuming that students will demonstrate developmental differences in moral reasoning based on the type of institution attended.

#### *Literature Review*

A series of studies have examined institutional type for its potential in explaining moral reasoning development, however, most of these studies have developed an ad hoc approach to exploring this relationship or have not included samples from enough institutions to make substantive claims about the relationship between institutional type and moral reasoning development. This review will include a brief description of these studies and will conclude with the study’s conceptual framework and governing research questions.

A handful of studies explored moral reasoning development as a function of institutional type (Burwell et al., 1992; Good & Cartwright, 1998; McNeel, 1991, 1994; Pascarella & Terenzini, 1991; Shaver, 1985, 1987). The most comprehensive studies were performed by Pascarella

and Terenzini (1991) and then by McNeel (1994) who each conducted ad hoc analyses of data gathered from a number of smaller studies and then created effect sizes from these data to draw conclusions about moral reasoning differences by institutional type. For example, Pascarella and Terenzini (1991) found that moral reasoning varied significantly by institutional type, with students from church-affiliated liberal arts colleges scoring the highest, followed by students at public research universities, two-year colleges, private liberal arts colleges, private universities, and public comprehensive universities, respectively. Similar results were reported by McNeel (1994) who conducted a meta-analysis using 22 samples comprised of students from seven liberal arts colleges, three universities, and two Bible colleges. Large effect sizes were reported for students enrolled in liberal arts colleges; large or moderate average effect sizes were reported by students enrolled in universities; and no to moderate effect sizes were reported for students enrolled at the Bible colleges. Marking a departure from these ad hoc analyses, this study was intended to capture differences among types as part of its design, thereby attending to many of the concerns associated with ad hoc research designs, including meta-analyses (see LeLorier, Gregoire, Benhaddad, Lapierre, & Derderian, 1997).

The other studies linking institutional type to moral reasoning development corroborate results from the meta-analyses, suggesting that students at liberal arts colleges were more likely than their counterparts to achieve higher moral reasoning scores (see Burwell et al., 1992; Good & Cartwright, 1998; Shaver, 1985, 1987). What is interesting about these studies is, in their approach to understanding institutional type, they include special focus institutions, like Bible colleges, as part of their small college samples. This study adopts a different approach to institutional sampling by including a more representative sample of institutions spanning Carnegie classifications, including research comprehensive, doctoral-granting institutions; regional, master's colleges and universities; liberal arts, baccalaureate colleges; and community colleges offering associate's degrees.

This study adds to and departs from other studies of institutional type and moral reasoning in a number of ways. First, this study is distinctive in its approach to institutional sampling, accounting for a range of institution types, including community colleges. This study is also unique in its use of hierarchical linear modeling to explain the effects of institutional type on moral reasoning development. This analytic strategy enables researchers to answer questions concerning the amount of variance explained in moral reasoning development by institutional differences, if these institutional differences can be explained by including

indicators of institutional type, and if, after adding student-level covariates, institutional type continues to explain any remaining variance in moral reasoning development.

### *Conceptual Framework*

The conceptual framework adopted for this study drew from a variety of empirical sources, including King and Mayhew's (2002, 2005) comprehensive literature review of over 250 articles on moral reasoning development, and Berger and Milem's (2000) framework for examining the role of institutional covariates in explaining student-level outcomes.

According to comprehensive literature reviews performed by King and Mayhew (2002, 2005), the following student-level variables were included to isolate the impact of educational contexts (e.g., institutional type) on moral reasoning development: gender, race, political orientation, cognitive motivation, precollege tested academic preparation, year in school, and college major. As a longitudinal study of first-year students, the precollege measure of moral reasoning development was also included as a student-level covariate. Consistent with suggestions forwarded by Berger and Milem (2000), institutional type was positioned as a structural-demographic institutional covariate predicting moral reasoning development. Based on this framework, the following research questions guided this study:

1. Does average moral reasoning development at the end of the first year in college vary significantly between institutions?
2. How does institutional type explain these institutional level differences?
3. After controlling for student-level covariates, does institutional type still explain significant proportions of the variance in moral reasoning scores at the end of the first year in college?

### *Methods*

#### *Sample*

The institutional sample consisted of full-time, first-year students at 19 four-year and two-year colleges and universities from across the United States. Institutions were selected based on their interest in participating in the Wabash National Study of Liberal Arts Education (WNSLAE), a project that uses a longitudinal research design to examine the effects of liberal arts colleges and liberal arts experiences on the outcomes conceptually associated with exposure to and participation in a liberal arts education (King, Kendell Brown, Lindsay, & Van-

Hecke, 2007). Of particular importance to this study was the decision to intentionally select institutions representing differences by way of institutional type. According to the 2007 Carnegie Classification of Institutions, three of the participating institutions were considered research universities, three were regional universities that did not grant the doctorate, two were two-year community colleges, and 11 were liberal arts colleges. These institutions ranged in selectivity from non-selective to highly selective. It is important to note that liberal arts colleges were purposefully overrepresented due to WNSLAE's focus on liberal arts colleges.

The students in the sample were full-time, first-year undergraduates. Data were collected in two waves. For the first round of data collection occurring in the fall of 2006, 4,501 students participated. These students were then approached for the second round of data collection which occurred in the spring of 2007. Of the initial pool of students, 3,081 participated in the follow-up, for a 68.5% longitudinal response rate. This sample represented 16.2% of the total population of full-time, first-year students at the 19 institutions.

A matrix sampling approach was used in order to curb the potential effects of survey fatigue (Groves, Fowler, Couper, Lepkowski, Singer, & Tourangeau, 2004), with half of the sample completing the Defining Issues Test, Version 2 (DIT2; Rest, Narvaez, Thoma, & Bebeau, 1999). Approximately 50% of students who returned for the second data collection completed the DIT2. After purging meaningless DIT2 scores (see Rest et al., 1999, for a discussion concerning the use of the M-score as an internal validity check on respondents who endorse certain items based on their pretentiousness rather than their meaning), data for 1,469 students remained and were used for subsequent analyses.

Of these 1,469 students, the majority (53.3%) were enrolled in liberal arts colleges. Just over a quarter of students (26.2%) were enrolled at research universities, 16.7% were enrolled at regional institutions, and the remaining 3.7% were enrolled in community colleges.

In addition, over half (57%) were female and over four fifths (80%) self-identified as White. Of the 20% identifying as students of color, 8% self-identified as Asian/Asian American, 5% self-identified as Black/African-American, 5% self-identified as Latino/a, under 1% self-identified as Native American, and the remaining 2% self-identified as another race, including non-U.S.-resident or unknown. Defined as ACT or ACT-equivalent, students' precollege tested academic preparation was 25.38 ( $SD = 4.58$ ). In addition, students tended to politically self-identify as liberal more than conservative and to report higher need for cognition scores than average.

### *Measures*

The DIT2 is an objective test of moral reasoning based on Rest's revision of Kohlberg's cognitive-developmental theory of moral development (Rest, 1979, 1986; Rest, Narvaez, Thoma, & Bebeau, 1999). On the DIT2, participants are presented with five moral dilemmas and are then asked to read a list of 12 items that include prototypic reasoning for each of the stages of moral development. Participants are then asked to rate how important they think each item is in making the decision regarding this dilemma, to indicate what their decision would be, and then to rank the four most important items. This process yields a DIT2 score, called "N2," and is the index used for this study. The index accounts for ranked items that reflect postconventional moral reasoning and for rated items that reflect respondents' preferences for higher versus lower stage reasoning. Based on early testing administrations (Rest, Thoma, Narvaez, & Bebeau, 1997), the DIT2 shares similar reliability correlations as the original DIT (long form), with test-retest reliability estimates ranging from 0.71 to 0.82.

The DIT2 also includes items that serve as reliability checks that help to ensure that the data students provide accurately reflect their intended responses. Examples of these checks include procedures for handling student responses to meaningless items, student responses to items with all 3s or 4s, students who fail to rate 3 or more items, or students who fail to rank more than 6 items (see Rest et al., 1997).

The Need for Cognition Scale (NCS; Cacioppo, Petty, & Kao, 1984) was also administered to students and purports to measure an individual's tendency "to engage in and enjoy effortful cognitive activity" (Cacioppo, Petty, Feinstein, & Jarvis, 1996, p. 197). Such activities are integral to what Cacioppo et al. (1996) call "cognitive motivation" (p.197).<sup>1</sup> The NCS is an 18-item scale with high internal consistency; alpha levels are typically greater than 0.85. High-scoring individuals enjoy engaging in effortful thought activities while low-scoring individuals tend to dislike engaging in such activities.

### *Variables*

The variables used to construct the hierarchical linear models are presented in Tables 1 and 2. They include structural-demographic organizational characteristics (e.g., institutional type), student demographic and precollege covariates (i.e., race, gender, political orientation, tested precollege academic preparation [ACT or equivalent, including COMPASS score], cognitive motivation, Time 1 moral reasoning score),<sup>2</sup> and the outcome of interest (Time 2 moral reasoning score).



A number of analytical procedures common to Level-1 and -2 variables were performed before including them in the hierarchical linear models. First, indicator variables for categorical variables with more than two levels (e.g., institutional type) were created; in each case, the variable serving as the reference group represented the majority of respondents or institutions. Second, discrete dichotomous variables (e.g., gender) were recoded into "0" and "1" categories, with "0" representing the category with the majority of respondents or institutions. All continuous variables were standardized before including them in the analytic models.

### *Analyses*

A number of descriptive and exploratory analyses were performed to ensure that all continuous variables used in the model were normally distributed and assumptions of linearity and homogeneity were not violated. In addition, these series of analyses explained the relationships among independent variables and between these variables and the criterion, the Time 2 moral reasoning score. See Table 3 for a description of pre- and posttest moral reasoning scores by institutional type.

A 2 (Time) x 4 (Institutional type) mixed-model repeated-measures ANOVA was also used to test for within- and between-subjects variability among Time 1 moral reasoning, Time 2 moral reasoning, and a combined categorical variable indicating institutional type. Time 1 and Time 2 moral reasoning scores were used to create the 2-level factor, "Time." Results from this analysis were used to inform construction of the hierarchical linear models.

A number of preliminary steps were performed to build the multilevel models. First, fully unconditional models were explored, with resulting parameter estimates used to compute the intraclass correlation coefficient (ICC) for the model. Results from this procedure demonstrated whether any proportion of the variance in the criterion significantly varied across institutions.

For the next part of the analysis, a Level-2 model was constructed to answer one of the auxiliary research questions: How does institutional type explain these institutional level differences? For this model, institutional-level variables were included in an effort to explain the overall grand mean on Time 2 moral reasoning. Indicator variables for institutional type were uncentered.

Next, a Level-1 model was constructed, including only student-level variables. The parameter estimates produced in the variance components analysis associated with this model were used, in conjunction with those parameter estimates in the combined model, to derive final variance estimates (Raudenbush & Bryk, 2002).



TABLE 1  
Level-2 Operational Definitions and Descriptive Statistics ( $N = 19$ )

Institutional type	Operational Definition	Mean	<i>SD</i>	Min	Max
Liberal arts colleges <sup>a</sup>	1 = Liberal arts colleges; 0 = All others	0.58	0.51	0.00	1.00
Research universities	1 = Research universities; 0 = All others	0.16	0.37	0.00	1.00
Regional universities	1 = Regional universities ; 0 = All others	0.16	0.37	0.00	1.00
Community colleges	1 = Community colleges; 0 = All others	0.11	0.32	0.00	1.00

<sup>a</sup> Indicates reference group.

TABLE 2  
Level-1 Operational Definitions and Descriptive Statistics ( $n = 1,469$ )

	Operational Definition	Mean	<i>SD</i>	Min	Max
Outcome Time 2 moral reasoning	Continuous (Z-score): Higher score indicates higher posttest moral reasoning.	0.00	1.00	-2.68	2.67
Gender Male	1 = Male; 0 = Female	0.66	0.047	0.00	1.00
Race White	1 = White/Caucasian, 0 = All others	0.80	0.40	0.00	1.00
African American/ Black	1 = African American/Black, 0 = All others	0.05	0.21	0.00	1.00
Latino(a)	1 = Latino(a), 0 = All others	0.05	0.21	0.00	1.00
Asian/Pacific Islander	1 = Asian/Pacific Islander, 0 = All others	0.08	0.27	0.00	1.00
Native American	1 = Native American, 0 = All others	0.00	0.05	0.00	1.00
No race given	1 = No Race Given, 0 =All others	0.02	0.13	0.00	1.00
Pretest Time 1 moral reasoning	Continuous (Z-score): Higher score indicates higher pretest moral reasoning.	0.00	1.00	-2.54	2.97
Cognitive motivation Need for Cognition	Continuous (Z-score): Higher score indicates higher need for cognition.	0.00	1.00	-3.79	2.36
Precollege academic preparation	Continuous (Z-score): ACT or equivalent with higher score indicating higher precollege academic preparation.	0.00	1.00	-2.96	2.12
Political identification	Continuous (Z-score): Higher score indicates more conservative political orientation.	0.00	1.00	-1.99	2.35

TABLE 3  
Means and Standard Deviations for Pre- and Posttest Moral Reasoning by Institution Type

	Pretest Moral Reasoning Mean ( <i>SD</i> )	Posttest Moral Reasoning Mean ( <i>SD</i> )
Institution type		
Liberal arts colleges	37.271 (14.395)	42.151 (14.751)
Research universities	40.802 (13.756)	45.137 (13.991)
Regional universities	32.654 (13.822)	37.396 (14.285)
Community colleges	26.054 (13.049)	28.981 (14.219)
Total	37.003 (14.474)	41.644 (14.865)

Proceeding with the analysis, the combined model was built. To isolate the amount of variance explained by institutional type, variances of student-level variables were fixed and centered at the grand-mean, with one exception. The student-level variable measuring tested precollege academic preparation (ACT or equivalent), a proxy for institutional selectivity, was group-mean centered and left free to vary at Level 2. Reintroducing the Level-1 means for precollege academic preparation (ACT or equivalent) as Level-2 controls isolate the effects of this variable on explaining the relationship between Level-2 variables (i.e., institutional type) and the criterion (i.e., Time 2 moral reasoning scores) while maintaining statistical power at Level 2 (see Enders & Tofighi, 2007). In short, this step helps ensure that the variance explained in the criterion by institutional type is not an artifact of precollege academic preparation, the proxy measure for institutional selectivity.

#### *Limitations*

A series of limitations are worth noting for this study. First, with only 19 institutions, parameters were estimated without using robust standard errors as their use is only appropriate for datasets comprised of a moderate to high number of Level-2 units (Raudenbush & Bryk, 2002). Also, the restriction in institutional sample size reduced the number of

Level-2 variables that could be included in the model; certainly other institutional-level variables would be worth exploring for their potential to predict moral reasoning development. Finally, development was measured at only two time points, making it difficult to assess whether the change in moral reasoning would be sustained over time.

## Results

### *Model Construction: Time 2 Moral Reasoning*

Results from the 2 (Time)  $\times$  4 (Institutional type) mixed-model repeated-measures ANOVA indicated that moral reasoning development was related to, but not dependent upon, institutional type. Specifically, results showed that the main effect for institutional type was significant  $F(3, 1465) = 34.252, p < 0.001, \eta^2 = 0.066$ . Thus, there were overall differences in the moral reasoning gain scores of students enrolled in research universities ( $M = 42.970$ ), compared to those at liberal arts colleges ( $M = 39.711$ ), compared to those enrolled at regional universities ( $M = 39.711$ ), and compared to those enrolled at community colleges ( $M = 27.517$ ).

A significant main effect for Time was also observed,  $F(1, 1465) = 80.386, p < 0.001$ , though the effect was weak ( $\eta^2 = 0.055$ ). Moral reasoning scores after the first year in college ( $M = 41.644$ ) were significantly higher than before the first-year in college ( $M = 37.003$ ).

However, the interaction term, Time  $\times$  Institutional Type, failed to reach statistical significance,  $F(3, 1465) = 0.601, p > 0.05$ . This suggests that moral reasoning change scores were not dependent upon institutional type and that developmental trajectories in moral reasoning were similar, in magnitude and direction, for students enrolled in institutions representing each type.

Taken together, these results suggest that differences in moral reasoning were related to institutional type, but that the gains shown by students were similar, in magnitude and direction, regardless of type. Collectively, these results demonstrated a need for further examination of the relationship between institutional type and moral reasoning development. Findings from the HLM analysis are now presented.

### *Fully Unconditional Model*

Results from the fully unconditional model indicated that a statistically significant proportion of the variance in Time 2 moral reasoning was explained by the differences between institutions ( $\chi^2 = 278.337, p < 0.001, N_{\text{student}} = 1,469, N_{\text{institution}} = 15$ ). The reliability estimate reached 0.903, indicating a high degree of stability across parameter estimates

for each institution. The intraclass correlation coefficient for this model reached 0.1856, suggesting that 18.56% of the variance explained in the outcome measure was due to institution-level differences.

#### *Level-2 Model Only*

Results from the Level-2 model showed that reliability estimates fell within acceptable conventions, with lambda reaching 0.903. When compared to students enrolled at liberal arts colleges, those students at community colleges ( $b = -0.889, p < 0.01$ ) and regional institutions ( $b = -0.394, p < 0.05$ ) were significantly less likely to achieve higher Time 2 moral reasoning scores. No significant differences were observed for students enrolled at liberal arts colleges when compared to those enrolled at research universities. Not accounting for student-level covariates, institutional type explained 12.62% of variance in the criterion.

#### *Level-1 Model Only*

The Level-1 model consisted of the following variables: the intercept representing institutional differences, the intercept representing student-level differences, gender, race, political orientation, cognitive motivation, precollege academic preparation (ACT or equivalent), and Time 1 moral reasoning. After controlling for all other model covariates, women were significantly more likely to show gains in moral reasoning than men ( $b = -0.134, p < 0.001$ ). Time 1 moral reasoning scores significantly predicted Time 2 scores ( $b = 0.554, p < 0.001$ ) as did need for cognition scores ( $b = 0.116, p < 0.001$ ), and precollege academic preparation ( $b = 0.178, p < 0.001$ ). Reliability estimates for the Level-1 model intercept reached 0.736 and for precollege academic preparation reached 0.119.

#### *Combined Model*

The combined model included all Level-1 covariates, institutional type, and the intercept representing the grand mean. For Level-1 predictors, most variables were centered using the grand-mean with variances fixed, with the exception of precollege academic preparation which was group-mean centered and left free to vary at Level 2. The indicator variables for institutional type remained uncentered. Reliability estimates reached 0.610 for the model intercept and 0.150 for the slope representing precollege academic preparation; this indicates stability in mean estimates for each institution included in the model (see Raudenbush et al., 2000).

Table 4 presents results for the combined model. Indicators for community college ( $b = -0.409, p < 0.01$ ) remained statistically significant,

TABLE 4  
Standardized Regression Coefficients for the Combined Model Predicting Moral Reasoning Development

	$\lambda$	b	SE
Random effects			
Intercept	0.610	0.018	0.040
Institutional type			
Research (Liberal arts)		0.067	0.079
Regional (Liberal arts)		-0.156+	0.085
Community college (Liberal Arts)		-0.409**	0.128
Fixed effects			
Gender			
Male		-0.133***	0.040
Race			
African American/Black (White)		-0.175+	0.090
Latino(a) (White)		-0.119	0.092
Asian/Pacific Islander (White)		-0.051	0.071
Native American (White)		-0.636	0.404
No race given (White)		0.028	0.142
Need for Cognition			
Cognitive motivation		0.114***	0.020
Precollege academic preparation	0.150	0.178***	0.29
Political orientation		-0.033+	0.020
Pretest			
Moral reasoning Time 1		0.549***	0.021

+ $p < 0.10$  \* $p < 0.05$ , \*\* $p < 0.01$  \*\*\* $p < 0.001$ .  
Parentheses indicate reference group for dummy coded variables.

after accounting for student-level covariates. Students enrolled at community colleges were significantly less likely to demonstrate gains in moral reasoning than were students enrolled at liberal arts colleges.

Indicators for regional institutions approached statistical significance ( $b = -0.156$ ,  $p < 0.09$ ), suggesting a statistical trend that students enrolled at regional universities were less likely to demonstrate gains in moral reasoning than were students enrolled at liberal arts colleges. Differences in effect magnitude and significance levels between the Level-2 only model and the combined model were likely due to controlling for institutional selectivity by allowing the student-level variable measuring precollege academic preparation to vary at Level-2. Consistent with the results reported in the Level-2 model, no significant differ-

ences were observed for students enrolled at research universities when compared with those enrolled at liberal arts colleges.

Turning to student-level variables, gender, precollege academic preparation, cognitive motivation, and Time 1 moral reasoning reached statistical significance. Women were statistically more likely than men to achieve higher Time 2 moral reasoning scores ( $b = -0.133, p < 0.001$ ). Students more likely to enjoy engaging in effortful thinking were significantly more likely to achieve higher moral reasoning scores than students less likely to enjoy engaging in effortful thinking ( $b = 0.114, p < 0.001$ ). Students with more precollege academic preparation were also significantly more likely to demonstrate developmental gains in moral reasoning ( $b = 0.178, p < 0.001$ ). Time 1 moral reasoning scores were also significantly predictive of Time 2 scores ( $b = 0.549, p < 0.001$ ). It should also be noted that a statistical trend was observed for certain indicators for race: when compared to White students, students self-identifying as African American/Black ( $b = -0.175, p < 0.053$ ) were less likely to show moral reasoning gains at the end of their first-year in college.

The total amount of variance explained by this model reached 49.22%. The total amount of variance explained by student-level covariates reached 48.10%. Institutional type explained an additional 1.10% of the total variance in the criterion. Interestingly, the slope for precollege academic preparation, the conceptual proxy for institutional selectivity, only explained an additional 0.2% of the variance in Time 2 moral reasoning. Variance component estimations for this model, ( $\chi^2 = 38.477, p < 0.05, N_{\text{student}} = 1,469, N_{\text{institution}} = 15$ ), suggests that additional institutional covariates, other than institutional type, should be explored for their potential to explain more variance in Time 2 moral reasoning.

### *Discussion*

This study takes an important step toward analyzing educational contexts and their influence on moral reasoning development. By accounting for students nested in their natural, albeit, nonrandom learning environments, this study answers theoretical calls for understanding the increasingly complicated role that colleges play in spurring moral growth.

One interesting and fairly surprising finding emerged from examining the amount of variance explained in moral reasoning development as a function of institutional-level covariates. That the interclass correlation coefficient for Time 2 moral reasoning reached 18.26% is evidence that the educational experiences that spur moral reasoning vary from institution to institution. Although unified in the call to graduate

“responsible citizens,” administrators may need to design experiences distinctive to their campus environments for optimizing moral growth. Due to institutional sample size restrictions, this study only investigated institutional type as a possible determinant of moral reasoning development. Researchers performing other studies, with more robust samples, may want to interrogate other institutional variables of interest, such as those suggested by Berger and Milem (2000). Such variables include, but are not limited to, peer socialization and organizational behaviors (i.e., bureaucratic, collegial, political, symbolic, and systemic).

Turning to institutional type, several findings are particularly noteworthy. Results from this study support the contention that moral reasoning gains are related to, but not dependent upon, institutional type. As a *block* of covariates comprised of indicator variables, institutional type did little to explain moral reasoning development, beyond student-level covariates included in the model. That this block of indicators was not significantly predictive on Time 2 moral reasoning in the combined model, but was in the Level-2 model, is probably due to the combined model’s inclusion of Time 1 moral reasoning as a student-level covariate. Including Time 1 moral reasoning may have obscured the potential effect of institutional type, as a block of variables, on explaining significant portions of the variance in Time 2 moral reasoning. Future research is needed to untangle these relationships; specifically, scholars interested in the transition from high school to college may want to investigate the effects of moral reasoning on college choice by examining end of high school moral reasoning scores as determinants of college selection.

Despite the limited collective effects the block of variables comprising institution type exerted on Time 2 moral reasoning, it appears as though students enrolled in liberal arts colleges were more likely to demonstrate developmental gains than those in community colleges. Such a finding may be an artifact of the study’s sampling design, over-representing liberal arts colleges with stakeholders intentionally choosing to participate in the study. Alternatively, this study may provide evidence similar to previous work linking type to moral reasoning (Burwell et al., 1992; Good & Cartwright, 1998; McNeel, 1991, 1994; Pascarella & Terenzini, 1991; Shaver, 1985, 1987), indicating that the liberal arts college may offer more educational opportunities for spurring moral growth than those offered by community colleges.

It should also be noted that although significant differences in moral reasoning scores did not exist between students attending liberal arts colleges and those attending research universities, students from both of these types of institutions demonstrated higher moral reasoning scores



than those enrolled in either regional or community colleges. High school students with higher moral reasoning scores may be more likely to choose a college based on its commitment to educating the whole student—a commitment often expressed in the broad-based curricular requirements associated with liberal arts colleges or in the wealth of involvement opportunities offered at research universities. These findings suggest the need for future explorations into each institutional type for its distinctive approach to providing educational experiences that help students expand their notions of justice and fairness.

The findings concerning gender and moral reasoning echo those from previous studies. It appears as though exposure to college influences women's moral reasoning scores to a greater and more positive degree than it does for men. This finding resonates with the growing national concern over men and their college experiences, with men entering college at lower rates (U.S. Department of Education, 2001; Wilson, 2007) and becoming less involved (Kinzie, Gonyea, Kuh, Umbach, Blaich, & Korkmaz, 2007) than women. Stakeholders interested in addressing these trends are encouraged to use structural frames for examining questions of gender differences: what types of institutional supports are needed to facilitate men's moral reasoning development?

Turning to race, results from this study highlight a statistical trend that moral reasoning development may differ based on students' self-identified race. Specifically, White students were more likely to show moral reasoning gains than students identifying as Black or African American. Until now, existing studies of the effects of racial identification on moral reasoning development have either examined samples insufficient in size to make claims concerning subgroup differences (see Gongre, 1981; Hurtado, Mayhew, & Engberg, 2003; Johnson, Insley, Motwani, & Zbib, 1993; Katz, 2001; Locke & Tucker, 1988; Loviscky, 2000; Lupfer, Cohen, Bernard, & Brown, 1987; Mayhew & Engberg, 2010; Mayhew & King, 2008; Murk & Addleman, 1992) or have adopted conceptual frames that include a robust number of demographic covariates, in addition to race, that may obscure its effects (see Mayhew, Siefert, & Pascarella, 2010). Champions of Kohlberg's work may be able to use the results of this study as an impetus for reexamining existing claims about moral reasoning development, specifically in regard to its universal application and subsequent measurement. Using empirically based evidence for pursuing this line of inquiry would be a worthwhile endeavor, offering the type of critical theoretical refinement Kohlberg had always embraced (see Kohlberg, 1991).

Cognitive motivation, the enjoyment of engaging in effortful thinking, and precollege academic preparation were positively related to

moral reasoning development. Theoretically, these findings corroborate others linking cognitive to moral development (see Mentkowski & Associates, 2000; Stepp, 2002; Taylor, Waters, Surbeck, & Kelley, 1985; Wanshaffe, 2001) as well as assumptions forwarded by Kohlberg (1976) and Rest et al. (1999) suggesting that cognitive development may be a necessary but insufficient condition for making moral reasoning gains. However, the underlying mechanisms explaining how complex forms of cognition engender moral reasoning growth still remains somewhat of a mystery; as King and Mayhew (2004) note, “Few [of these] studies attempted to identify the underlying influences in the cognitive domain that might affect moral reasoning (or vice versa); doing so would help explain why the relationships between cognitive complexity and moral reasoning are so strong” (p. 409). Scholars interested in exploring change mechanisms may need to adopt mixed methods designs to unpack some of the reasons students offer for making developmental gains in moral reasoning by way of gains in variables measuring forms of cognition.

#### *Implications*

This study has implications for scholars and practitioners interested in moral reasoning development. First, this study adopts a multilevel approach for examining educational contexts and their influence on moral reasoning development. That 18.56% of the variance in Time 2 moral reasoning can be explained by institutional level differences provides evidence for more research designed to account for students clustered in their natural learning environments. Accounting for student clustering is an important step toward providing practitioners with the empirical information they need to design the most effective learning environments that help students make moral gains.

Another implication from this study comes from its use of cross-comparative frames for examining the role of institutional type, and to a lesser degree, self-identified race, in explaining moral reasoning development. What is the utility of adopting such an approach? With regard to institutional type, results from this study suggest that differences among type exist, but that these differences had little to do with the nature and direction of moral reasoning gains, especially after accounting for student-level covariates. All students in this study, regardless of type, made moral reasoning gains; perhaps more important than comparing students by institution type would be to ask research questions designed to uncover the distinctive ways certain institutions approach their responsibilities to provide students with a moral education. How does moral reasoning development occur in community colleges?

In regional institutions? With a more focused scope, results from studies guided by more specific research questions may be more useful for administrators interested in shaping particular learning environments in ways that optimize moral reasoning development.

A similar case could be made for race. Given the amount of scholarship suggesting that student needs, perceptions, and experiences are often related to their lived, raced experiences (see, as examples, Gurin, Dey, Hurtado, & Gurin, 2002; Hurtado, Milem, Clayton-Pedersen, & Allen, 1998), future questions guided by findings from this study and that are concerned about the relationship between race and moral education in college may need to focus on the distinctive ways educators approach developing environments poised for encouraging moral growth for African American students, for example. The philosophies underscoring this approach are similar to the ideas adopted by critical race theorists like Teranishi (2007), who argues that because “racial groups are neither equally nor consistently comparable ... the approaches to studying differing student populations are not universally applicable” (p. 39).

### *Conclusion*

Theoretical calls by student development scholars and moral psychologists adopting Kohlbergian frames for understanding moral reasoning have charged researchers with designing studies that directly interrogate educational contexts and their effects on moral reasoning development. With its use of hierarchical linear modeling to analyze the relationship between institutional type and moral reasoning development, this study takes a small step toward answering these calls. Moral reasoning development was associated, but not conditioned upon institutional type; such a finding urges researchers to explore other institutional covariates for explaining moral reasoning gains.

### *Notes*

<sup>1</sup> The Need for Cognition scale was used to assess and ultimately control for cognitive motivation for this study. Unlike measures of academic motivation (see Pascarella et al., 2005), cognitive motivation, as measured by the Need for Cognition scale, has been empirically linked to moral reasoning (see Crowson, 2004) and thus serves as a better control for isolating the effects on institutional type on moral reasoning development.

<sup>2</sup> As a study of first-year students, year in school and college major had limited variability, due to many students reporting similar ages and not knowing their major at the time of survey administration. Thus, these two variables were excluded from the process of model construction.

## References

- Bensimon, E. M. (2007). The underestimated significance of practitioner knowledge in the scholarship on student success. *The Review of Higher Education, 30*(4), 441–469.
- Berger, J. B., & Milem, J. F. (2000). Organizational behavior in higher education and student outcomes. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 15, pp. 268–338). New York, NY: Agathon Press.
- Blakeney, C. D., & Blakeney, R. F. (2004). Epilogue. In V. Siddie Walker & J. R. Snarey (Eds.), *Race-ing moral formation: African American perspectives on care and justice* (pp. 147–150). New York, NY: Teachers College Press.
- Brabeck, M. (1983). Moral judgment: Theory and research on differences between males and females. *Developmental Review, 3*, 274–291.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Burwell, R. M., Butman, R. E., & Van Wicklin, J. (1992). Values assessment at three consortium colleges: A longitudinal follow-up study. Houston, NY: Houghton College. Retrieved from ERIC Document Reproduction Service. (ED345635)
- Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*(2), 197–253.
- Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of the need for cognition scale. *Journal of Personality Assessment, 48*, 306–307.
- Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgment* (Vol. 1). New York, NY: Cambridge University Press.
- Crowson, H. M. (2004). *Do cognitive motives predict moral judgment development?* (Unpublished manuscript).
- Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multi-level models: A new look at an old issue. *Psychological Methods, 12*, 121–138.
- Gilligan, C. (1977). In a different voice: Women's conceptions of the self and morality. *Harvard Educational Review, 47*, 481–517.
- Gongre, W. M. (1981). *A study of moral judgment, using the Defining Issues Test, for three ethnic groups at Bacone College* (Unpublished doctoral dissertation). Oklahoma State University, Stillwater, OK.
- Good, J. L., & Cartwright, C. C. (1998). Development of moral judgment among undergraduate university students. *College Student Journal, 32*, 270–276.
- Groves, R. M., Fowler, F. J. Jr., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey methodology*. Hoboken, NJ: Wiley-Interscience.
- Gurin, P., Dey, E. L., Hurtado, S., & Gurin, G. (2002). Diversity and higher education: Theory and impact on educational outcomes. *Harvard Educational Review, 72*(3), 330–336.
- Hurtado, S., Mayhew, M. J., & Engberg, M. E. (2003, November). *Diversity in the classroom and students' moral reasoning*. Paper presented at the annual meeting of the Association for the Study of Higher Education, Portland, OR.

- Hurtado, S., Milem, J. F., Clayton-Pederson, A. R., & Allen, W. R. (1998). Enhancing campus climates for racial/ethnic diversity: Educational policy and practice. *The Review of Higher Education, 21*(3), 279–302.
- Johnson, J. L., Insley, R., Motwani, J., & Zbib, I. (1993). Writing performance and moral reasoning in business education? *Journal of Business Ethics, 12*(5), 397–406.
- Katz, A. D. (2001). *The influence of intergroup dialogues on undergraduates' identity formation and moral development schemas* (Unpublished master's thesis). California State University, Long Beach, CA.
- King, P. M., Kendall Brown, M., Lindsay, N. K., & Van Hecke, J. R. (2007). Liberal arts student learning outcomes: An integrated perspective. *About Campus: Enriching the Student Learning Experience, 12*, 2–9.
- 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.
- King, P. M., & Mayhew, M. J. (2005). Theory and research on the development of moral reasoning among college students. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 19, pp. 375–440). Edison, NJ: Agathon Press.
- Kinzie, J., Gonyea, R., Kuh, G., Umbach, P., Blaich, C., & Korkmaz, A. (2007). *The relationship between gender and student engagement in college*. Paper presented at the annual meeting of the Association for the Study of Higher Education, Louisville, KY.
- Kohlberg, L. (1976). Moral stages and moralization: The cognitive-developmental approach. In T. Lickona (Ed.), *Moral development and behavior: Theory, research, and social issues* (pp. 31–53). New York, NY: Holt, Rinehart and Winston.
- Kohlberg, L. (1981). *The meaning and measurement of moral development*. Worcester, MA: Clark University Press.
- Kohlberg, L. (1991). My personal search for universal morality. In L. Kuhmerker (Ed.), *The Kohlberg legacy for the helping professions* (pp. 11–17). Birmingham, AL: R.E.P. Books.
- LeLorier J., Gregoire, G., Benhaddad, A., Lapierre, J., & Derderian, F. (1997). Discrepancies between meta-analyses and subsequent large randomized controlled trials. *New England Journal of Medicine, 337*, 536–542.
- Lewin, K. (1951). *Field theory in social science*. New York, NY: Harper and Row.
- Locke, D. C., & Tucker, D. O. (1988). Race and moral judgment development scores. *Counseling and Values, 32*, 232–235.
- Lovisky, G. E. (2000). *Construct validity of a managerial defining issues test* (Unpublished doctoral dissertation). The Pennsylvania State University, State College, PA.
- Lupfer, M. B., Cohen, R., Bernard, J. L., & Brown, C. M. (1987). The influence of level of moral reasoning on the decisions of jurors. *Journal of Social Psychology, 127*(6), 653–667.
- Mayhew, M. J., & Engberg, M. E. (2010). Diversity and moral reasoning: How negative diverse peer interactions affect the development of moral reasoning in undergraduate students. *The Journal of Higher Education, 81*, 459–488.
- Mayhew, M. J., & King, P. M. (2008). How curricular content and pedagogical strategies affect moral reasoning development in college students. *Journal of Moral Education, 37*(1), 17–40.

- Mayhew, M. J., Seifert T. A., & Pascarella, E. T. (2010). A multi-institutional assessment of moral reasoning development among first-year students. *Review of Higher Education, 33*(3), 357–390.
- Mayhew, M. J., Seifert T. A., & Pascarella, E. (in press). How the first year of college influences moral reasoning development for students in moral consolidation and moral transition. *Journal of College Student Development*.
- McEwen, M. K. (2003). The nature and uses of theory. In S. R. Komives, D. B. Woodard, Jr., & Associates, *Student services: A handbook for the profession* (pp. 153–178). San Francisco, CA: Jossey-Bass.
- McNeel, S. (1991). Christian liberal arts education and growth in moral judgment. *Journal of Psychology and Christianity, 10*(4), 311–322.
- McNeel, S. (1994). College teaching and student moral development. In J. Rest & D. Narvaez (Eds.), *Moral development in the profession: Psychology and applied ethics* (pp. 26–47). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Mentkowski, M., & Associates. (2000). *Learning that lasts: Integrating learning, development, and performance in college and beyond*. San Francisco, CA: Jossey-Bass.
- Murk, D. A., & Addleman, J. A. (1992). Relations among moral reasoning, locus of control, and demographic variables among college students. *Psychological Reports, 70*(2), 467–476.
- Pascarella, E., Wolniak, G. C., Seifert, T. A., Cruce, T. M., & Blaich, C. F. (2005). *Liberal arts colleges and liberal arts education: New evidence on impacts*. San Francisco, CA: Jossey-Bass/ASHE.
- Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students: Findings and insights from twenty years of research*. San Francisco, CA: Jossey-Bass.
- Piaget, J. (1965). *The moral judgment of the child*. (M. Gabain, Trans.). New York, NY: Free Press.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage Publications.
- Raudenbush, S. W., Bryk, A. S., Cheong, Y., & Congdon, R. T. (2000). *HLM 5: Hierarchical linear and nonlinear modeling*. Chicago, IL: Scientific Software International.
- Rest, J. R. (1979). *The impact of higher education on moral judgment development*. *Moral Research Projects* (Technical report no. 5). Minneapolis, MN.
- Rest, J. R. (1986). *Moral development: Advances in research and theory*. New York, NY: Praeger Press.
- Rest, J. R., Narvaez, D., Thoma, S. J., & Bebeau, M. (1999). DIT2: Devising and testing a revised instrument of moral judgment. *Journal of Educational Psychology, 91*(4), 644–659.
- Rest, J. R., Thoma, S. J., Narvaez, D., & Bebeau, M. J. (1997). Alchemy and beyond: Indexing the Defining Issues Test. *Journal of Educational Psychology, 89*(3), 498–507.
- Shaver, D. G. (1985). A longitudinal study of moral development at a conservative, religious, liberal arts college. *Journal of College Student Development, 26*(5), 400–404.
- Shaver, D. G. (1987). Moral development of students attending a Christian, liberal arts college and a Bible college. *Journal of College Student Development, 28*(3), 211–218.

- Siddle-Walker, V., & Snarey, J. (2004). *Race-ing moral formation: African American perspectives on care and justice*. New York, NY: Teachers College Press.
- Stage, F. K. (1991). Common elements of theory: A framework for college student development. *Journal of College and Student Development, 32*(1), 56–61.
- Stepp, J. O. (2002). *The relationship of moral development to critical thinking and family structure of matriculating college students* (Unpublished doctoral dissertation). The Southern Baptist Theological Seminary, Louisville, KY.
- Strange, C. (1994). Student development: The evolution and status of an essential idea. *Journal of College and Student Development, 35*(6), 399–412.
- Taylor, J. B., Waters, B., & Surbeck, E. (1985). Cognitive, psychosocial, and moral development as predictors of pre-service teachers' ability to analyze child behavior. *College Student Journal, 19*, 65–72.
- Teranishi, R. T. (2007). Race, ethnicity and higher education policy: The use of critical quantitative research. In F. Stage (Ed.), *Using quantitative data to answer critical questions* (pp. 37–49). San Francisco, CA: Jossey-Bass.
- U.S. Department of Education, National Center for Education Statistics. (2001). *Digest of education statistics 2000* (NCES 2001-034). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- Walker, L. J. (1984). Sex differences in the development of moral reasoning: A critical review. *Child Development, 55*, 677–691.
- Walker, L. J. (2006). Gender and morality. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (pp. 93–118). Mahwah, NJ: Lawrence Erlbaum.
- Wanshaffe, K. R. (2001). *Are formal operations necessary for moral reasoning or ego identity achievement?* (Unpublished master's thesis). San Jose State University, San Jose, CA.
- Wilson, R. (2007, January 26). The new gender divide. *The Chronicle of Higher Education, 53*(21), A36.



Copyright of Journal of Higher Education is the property of Ohio State University Press 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.