A cross-sectional study of undergraduates' ethical reasoning skills at a liberal arts institution

Joel R. TerMaat (1), Kristopher J. Williams (2), and Christopher D. Wentworth (1)

(1) Department of Engineering and Physics, Doane University, (2) Director of Institutional Effectiveness, Doane University

Abstract

Liberal arts institutions, given their inherent purpose, are arguably well-positioned to promote the ethical development of science and engineering students. When assessed using the Defining Issues Test (DIT2) corresponding to the Neo-Kohlbergian model of moral development, other researchers have typically observed dramatic gains during college, particularly pronounced at liberal arts institutional types. A cross-sectional DIT2 study sampling across the residential student population was conducted to investigate the localized and modern impact at the home liberal arts institution. Other considerations due to gender, political identification, retention mortality, and major/division are discussed.

Keywords

Engineering Ethics Education, Liberal Arts Institutions, Defining Issues Test

Introduction

Understanding the complex ethical dimensions can be considered vital to preparing the next generation of engineers for professional practice. Most professional societies, such as ASME or ASCE, for example, have adopted codes of ethics to govern the decision-making and behavior of their members [1], [2]. However, the increased specialization of technology and often complex and unpredictable societal impacts present new challenges that may go beyond the reductive application of codes. Developing critical thinking skills, recognizing diverse perspectives, and applying macroethical considerations will become increasingly important to engineering practice. These skills are now part of the standard definition for liberal arts education in American higher education [3], and this education is central to the mission of the nation's liberal arts colleges [4]. But do these institutions truly affect the ethical development of their graduates, particularly engineering majors? This question may be especially appropriate given the technical credit hour requirements (e.g., ABET) of professional programs transplanted into a traditionally liberal arts context that may reduce the breadth of a liberal arts education.

Due to their validity and ease of use, the Defining Issues Test (DIT) and the updated DIT2 are widely used assessments of moral reasoning within the neo-Kohlbergian framework [5]. These assessments involve participants reading scenarios of moral dilemmas and then rating and ranking items related to the dilemmas in terms of their moral importance. The multiple-choice nature of the assessments allows them to be used with minimal expense by educators unfamiliar with interview-based assessments. The DIT was validated according to six criteria by over 400 published studies [5]. The DIT2 updates the scenarios, improves on validity characteristics, and

requires eliminating fewer participants than the DIT [6]. Two indices have been defined to give a quantitative measure of moral reasoning development: the older P score, which measures the extent to which a person prefers Postconventional moral thinking, and the newer N2 score, which measures both the preference for Postconventional thinking and the degree to which Personal Interest items (the lowest developmental stage) receive lower ratings [7]. DIT scores (P and N2 developmental indices) are reliable and significantly related to moral comprehension, prosocial behaviors, and desired professional decision-making [8]. Both scores measure the correspondence of items to post-conventional moral reasoning schema.

The use of the DIT assessments to study moral reasoning development has shown that college education does produce gains. In fact, "the gains in moral judgment are among the largest, most impressive gains of all tested variables influenced by the college experience" [9], and liberal arts colleges stand out among institutional types showing gains [10]. The reasons for better gains for students attending liberal arts institutions remain an open question. Possible reasons include student self-selection, smaller class size, and specific curriculum elements, but we do not yet know (see p. 396 in [10]). There is evidence that educational interventions that "encourage experiential learning, reflection, group work, active learning, and decision-making are generally more effective in promoting moral judgment than those that use more traditional pedagogies" (see p. 404-405 in[10]). Co-curricular interventions can also be effective in developing moral judgment [10].

 The research reported in this paper aimed to verify the underlying assumption of strong gains at the home liberal arts institution (Doane University). The gross indicator of institutional type may need to be more accurate, as these colleges range from highly-selective elite institutions to smaller, non-selective colleges that serve a broad spectrum of students. Additionally, generational shifts could influence results from previous studies. Secondly, this initial study may provide direction for further exploration into course/program/college contexts that underlie the broad institutional indicator.

Additionally, this study sought to expand upon previous research that looked narrowly at an ethics class intervention for undergraduate engineering majors [11]. In that work, the pre/post-DIT2 scores indicated a positive benefit to the engineering ethics course. For example, the DIT2 N2-score mean increased from 26.05 (pre) to 32.39 (post) in the Spring 2022 class (N = 14, p-value of 0.074). The effectiveness of this class intervention was additionally supported by Spring 2023 results (unpublished repeated study), which similarly showed an N2-score pre/post mean increase from 24.8 to 31.2 (N = 16, p-value of 0.077). While those results are encouraging at a focused engineering course level, this new cross-sectional study enables a broader institutional view of undergraduates' ethical reasoning while also providing a framing context for the previous engineering-specific results.

Methods

A cross-sectional study of the institution's residential student population was performed in Fall 2022. Random samples of names from freshmen, sophomore, junior, and senior populations were emailed a link to a DIT2 Qualtrics survey. Students who voluntarily opted into and completed the survey were each provided a small remuneration gift card for their good faith effort. Survey results were processed at the Center for Ethical Development and returned to the

authors for further analysis. Of the 73 returned survey results, N = 15 each for the freshmen, sophomore, and junior respondents and N = 28 for the senior population. This study was approved under the institutional IRB F22 005 DC IRB HS.

The norms provided for reference and discussion are taken from the 2011-2020 norms compiled by the Center for Ethical Development [12].

Results and Discussion

The results of the Fall 2022 study, along with the 2011-2020 DIT2 norms, are presented in Table 1 and graphically in Figure 1 and Figure 2. The populations' standard deviations were similar across the institution and the norms as would be anticipated. At the same time, the standard errors are markedly different due to the large sample size of the norms.

Table 1. Summary data of the Fall 2022 cross-sectional study with comparison to 2011-2020 DIT2 norms. N is the number of students completing the DIT2 instrument. SD is the standard deviation.

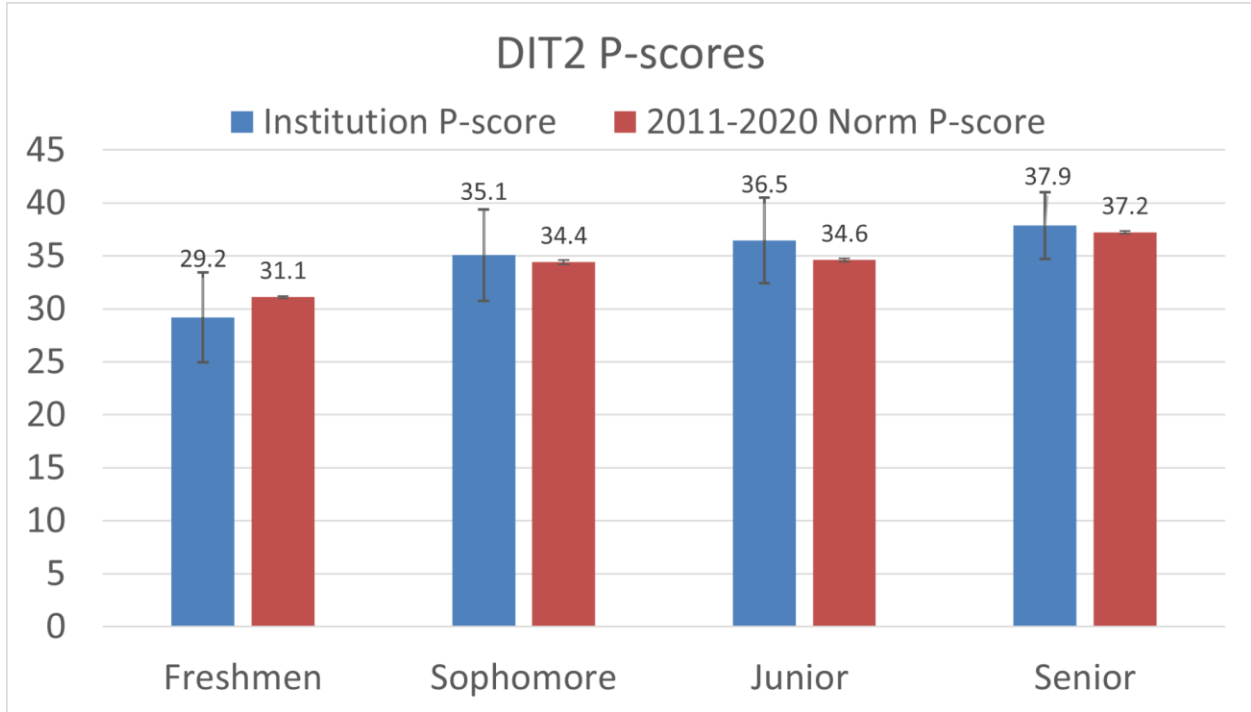| Results	N	P-score	SD		N	N2-score	SD
Freshmen	15	29.2	16.4		15	28.2	16.7
Sophomore	15	35.1	16.8		15	35.9	17.4
Junior	15	36.5	15.7		15	35.3	14.9
Senior	28	37.9	16.7		28	35.6	17.5
DIT-2 Norms	N	P-score	SD		N	N2-score	SD
Freshmen	18985	31.1	15.3		18976	30.26	15.19
Sophomore	7116	34.4	15.7		7111	33.67	15.40
Junior	8751	34.6	15.8		8750	34.03	15.57
Senior	14894	37.2	16.0		14887	36.96	15.68

Figure 1. P-scores from the Fall 2022 study by grade level with comparison to DIT2 norms.
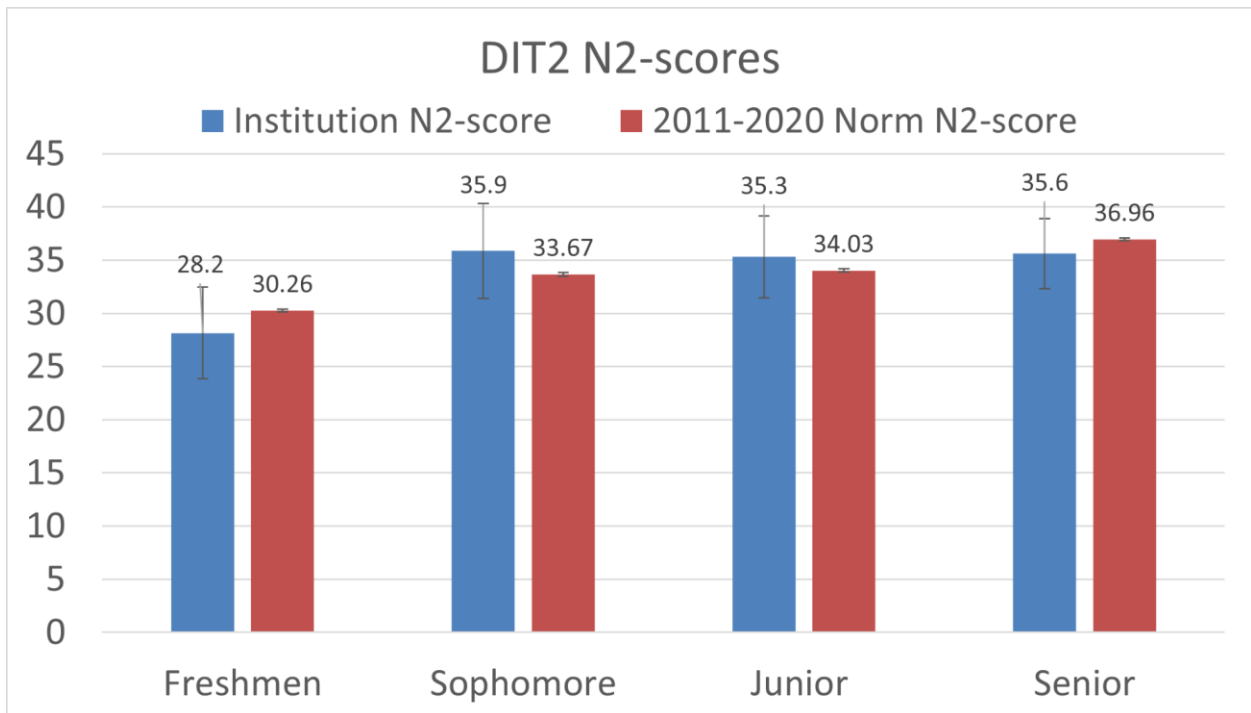(Mean +- S.E.)

Figure 2. N2-scores from the Fall 2022 study by grade level with comparison to DIT2 norms.
(Mean +- S.E.)

A one-way ANOVA was performed to compare the effect of student grade level on the P-score. The results were not statistically significant at the 0.05 level, which is partially explainable as we see the scores increase from freshman to sophomore year but do not change substantially after that. We then ran a t-test comparing Freshman P-scores (mean=29.20, SD=16.40) to all upper-class student P-scores (mean=36.77, SD=16.29) and found a difference with t(71)=1.6058, p=0.11.

We then also performed a one-way ANOVA to compare the effect of student grade level on the N2-score. The results were not statistically significant at the 0.05 level. We then ran a t-test comparing Freshman N2-scores (mean=28.15, SD=16.71) to all upperclass student N2-scores (mean=35.61, SD=16.53) and found a difference with t(71)=1.5543, p=0.12.

Our data offer some preliminary evidence of improving (Neo-Kohlbergian) moral reasoning skills from the freshman to upper-class level, though not at a high level of statistical significance. An increase in scores would be consistent with the work of other researchers, who observed dramatic gains at liberal arts colleges [10]. But before such an interpretation, additional analysis is warranted. Is the low freshman score potentially a result of the small opt-in sample? The small sample sizes, when sub-grouped even further, preclude detailed analysis and conclusions. However, some preliminary insights arose during further analysis and are provided hereinafter.

Acknowledging limitations of sample sizes and variability, a statistically non-significant increase in scores is observed from freshmen to sophomore populations, with primarily level values observed thereafter. This appears to suggest gains in moral development in the first year of college at the home institution. We noted the relatively low index scores of the freshmen population compared to the freshmen norm (for example, P-scores of 29.20 vs 31.10). Despite those initially low scores, the limited data seems to indicate first-year growth of those students with respect to ethical reasoning skills, which is an explicit mission emphasis of the home liberal arts institution.

These analyses indicate that there is something worth studying in students that occurs between the first and second years of college. It would be interesting to have the same set of students retake the assessment to see if changes occurred within that group of students during the past year.

With regards to sex identification, freshmen and senior sample populations were essentially balanced and directly comparable, whereas sophomore and junior samples were slightly more female.   Sex differences did not appear to influence the data. Overall (across all grade levels) analysis by sex followed typical norms, with males (N = 33) scoring below females (N = 40) in both P-scores (32.97 and 37.07, respectively) and N2-scores (32.07 and 35.73, respectively).

Political identification (1- very liberal, 3 – neither liberal nor conservative, 5 – very conservative) also did not appear to be a relevant factor, even though upper-level students trended more towards slightly liberal identification. Both freshmen and sophomore results were slightly conservative at an average of 3.1, yet had a wide difference in P and N2 scores. As such, political leanings did not appear to factor into the grade-level findings. One open research question that has been mentioned by others as a possible explanation for liberal arts institutional effects is whether more open-minded and/or liberal students naturally select liberal arts

institutions [9]. While the argument may have some validity at selective institutions or the liberal arts college archetype of decades past, we do not believe that statement characterizes the home institution composition.

Mortality (as in college retention/drop-out) of participants warranted special consideration in the cross-sectional (not longitudinal) study. Is the improvement in sophomores caused by lower-scoring students preferentially dropping out? While one may expect this to be visible in the norms as well, this effect may be more pronounced at specific institutions depending on their admission and retention rate characteristics. To investigate the effect of first-semester dropouts, the participant results were cross-referenced to census data for the subsequent Spring 2023 semester. Of the 15 freshmen respondents, only one participant (who scored above those freshmen averages) did not enroll the following semester. Preferential mortality was, therefore, not considered a relevant explanatory factor in the lower freshmen means.

Analysis by discipline is of particular research interest, as one might hypothesize that liberal arts majors may score higher than other (STEM, education, business) majors due to the nature of their programs. While analysis by major was not possible in the small samples, data were analyzed by rough college divisions. Student majors were categorized as Business, Education, Humanities/Social Sciences, or STEM. A one-way ANOVA was performed to compare the student's major division on the P-score and N2-score. No statistically significant difference between means was found at the 0.05 level.

As part of the general education requirements, Doane University requires students to complete courses in several areas that emphasize learning outcomes related to moral reasoning skills: Liberal Arts Studies (LAR 101, LAR 202, LAR 303 courses), Community & Identity (FAKCI), In Search of Meaning & Values (FAKMV), Global & Cultural Context (FAKGC), Human Creativity (FAKHC), and Rhetorical Communication (FAKRC).

Multiple regression analyses were conducted to examine the relationships between DIT2 scores and the number of courses students had taken in the general education areas specified. Independent variables were coded for liberal arts courses by experience with the courses LAR 101, LAR 202, and LAR 303 (1="taken course", 0="not taken") and for other courses by the number of courses taken (varied from 0 to 5). The results from these analyses are in Table 2 and Table 3.

Table 2. Results of multiple linear regression for P-score on course data.

P-model	Coefficients	Standard Error	t value	Sig.	
(Intercept)	22.363	7.700	2.904	0.00505	**
LAR101	8.437	7.013	1.203	0.23337	
LAR202	2.621	5.423	0.483	0.63054	
LAR303	9.467	4.266	2.219	0.03002	*
FAKCI	-1.166	2.155	-0.541	0.59019	
FAKRC	0.119	2.049	0.058	0.95388	
FAKGC	1.196	2.207	0.542	0.58969	
FAKMV	-6.497	3.553	-1.829	0.07213	
FAKHC	3.890	1.751	2.221	0.02988	*

For the P regression model, the fitting statistics were

Residual standard error: 15.66 on 64 degrees of freedom
Multiple R-squared: 0.1944, Adjusted R-squared: 0.09371
F-statistic: 1.931 on 8 and 64 DF, p-value: 0.07044

Table 3. Results of multiple linear regression for N2-score on course data. * indicates the result is statistically significant at the 0.05 level or better.

N2-model	Coefficients	Standard Error	t value	Sig.	
(Intercept)	20.655	7.899	2.615	0.0111	*
LAR101	9.616	7.194	1.337	0.1861	
LAR202	1.474	5.564	0.265	0.7919	
LAR303	9.754	4.377	2.229	0.0294	*
FAKCI	-0.251	2.211	-0.114	0.9099	
FAKRC	-1.680	2.103	-0.799	0.4273	
FAKGC	1.030	2.264	0.455	0.6507	
FAKMV	-5.551	3.645	-1.523	0.1328	
FAKHC	3.748	1.796	2.086	0.0409	*

For the N2 regression model, the fitting statistics were

Residual standard Error: 16.06 on 64 degrees of freedom,
Multiple R-squared: 0.1798
Adjusted R-squared: 0.07727
F-statistic: 1.754 on 8 and 64 degrees of freedom, p-value: 0.1031

From Table 2 and Table 3, we can see that taking LAR 303 and the number of Human Creativity courses are significantly associated with the N2-score and P-score. As students typically take the LAR 303 course as juniors and seniors, the association may be attributable more to the age or time spent in school than the course in particular. Surprisingly, the variable for the number of In Search of Meaning & Value courses (FAKMV) has a negative correlation with N2-score and P-score. We believe that the descriptions of the learning outcomes for this set of courses more closely align with the type of reasoning that would show an increase in the N2-score and P-score. However, this may support other findings that have shown that learning about ethical reasoning is not necessarily associated with developing ethical reasoning skills.

As each model has relatively low R^2 and adjusted R^2 values, we must consider other experiences of the students that might better explain the variations in the scores rather than the general education courses taken.

Conclusion

The limited study conducted at the home institution appears to provide current evidence of the positive benefit of a liberal arts education on the ethical development of undergraduates. Even though these study results are encouraging, there are still considerable research avenues to

explore to better understand what specifically about liberal arts colleges promotes moral reasoning skills. Specific interventions/experiences can be tied more so than others to moral advancement; however, the investigation into any one single factor in isolation is unlikely to provide an adequate explanation. At a minimum, a complex set of experiences and co/extra/curricular factors preferentially associated with liberal arts institutions is likely responsible for consistently strong gains in ethical reasoning skills. While we do not claim that it is necessarily the traditional liberal arts experience in toto, it may be difficult to replicate such a complex network of factors at other institutional types.

References

[1] ASME, "Ethics in Engineering," 2023. https://www.asme.org/about-asme/governance/Ethics-in-Engineering (accessed Aug. 22, 2023).
[2] ASCE, "Code of Ethics," 2020. https://www.asce.org/career-growth/ethics/code-of-ethics (accessed Aug. 22, 2023).
[3] C. G. Schneider, "Making Excellence Inclusive: Liberal Education and America's Promise," *Lib. Educ.*, vol. 100, no. 4, 2014.
[4] E. M. Lang, "Distinctively American: The Liberal Arts College," *Daedalus*, vol. 128, no. 1, pp. 133–150, 1999.
[5] J. R. Rest, S. J. Thoma, and M. J. Bebeau, *Postconventional moral thinking: A neo-Kohlbergian approach.* Psychology Press., 1999.
[6] J. R. Rest, D. Narvaez, S. J. Thoma, and M. J. Bebeau, "DIT2: Devising and testing a revised instrument of moral judgment," *J. Educ. Psychol.*, vol. 91, pp. 644–659, 1999, doi: 10.1037/0022-0663.91.4.644.
[7] J. Rest, S. J. Thoma, D. Narvaez, and M. J. Bebeau, "Alchemy and beyond: Indexing the Defining Issues Test," *J. Educ. Psychol.*, vol. 89, pp. 498–507, 1997, doi: 10.1037/0022-0663.89.3.498.
[8] S. J. Thoma, "Research on the defining issues test," in *Handbook of moral development*, Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, 2006, pp. 67–91. [Online]. Available: https://www.google.com/books/edition/Handbook_of_Moral_Development/4CV5AgAAQBAJ?hl=en&gbpv=1&dq=%22Research%20on%20the%20defining%20issues%20test%22%20S.J.%20Thoma&pg=PA67&printsec=frontcover
[9] J. R. Rest, "Chapter 1 Background: Theory and Research," in *Moral Development in the Professions: Psychology and Applied Ethics*, Hillsdale, N.J., USA: Lawrence Erlbaum Associates Publishers, 1994, pp. 1–26.
[10] P. M. King and M. J. Mayhew, "Theory and Research on the Development of Moral Reasoning Among College Students," in *Higher Education: Handbook of Theory and Research*, J. C. Smart, Ed., in Higher Education: Handbook of Theory and Research. Dordrecht: Springer Netherlands, 2004, pp. 375–440. doi: 10.1007/1-4020-2456-8_9.
[11] Joel R. TerMaat, Kristopher J. Williams, and Christopher D. Wentworth, "Comparison of the DIT2 and EERI instruments for assessing the development of ethical reasoning of engineering students," presented at the 2023 ASEE Rocky Mountains Section Conference, GOLDEN, CO, May 2023.
[12] N. Gungordu, G. Nabizadehchianeh, E. O'Connor, W. Ma, and D. I. Walker, "Moral reasoning development: norms for Defining Issue Test-2 (DIT2)," *Ethics Behav.*, pp. 1–18, May 2023, doi: 10.1080/10508422.2023.2206573.

Joel R. TerMaat, Doane University

Joel R. TerMaat is an Associate Professor of Engineering and Chair of the Department of Engineering and Physics at Doane University.

Kristopher J. Williams, Doane University

Kristopher J. Williams is Associate Professor of Mathematics and Director of Institutional Effectiveness at Doane University.

Christopher D. Wentworth, Doane University

Christopher D. Wentworth is Professor Emeritus of Physics at Doane University.