Comparison of the DIT2 and EERI instruments for assessing the development of ethical reasoning of engineering students

Joel R. TerMaat

Dr. Joel TerMaat is an Assistant Professor of Engineering and chair of the Engineering & Physics department at Doane University.

Kristopher J. Williams

Christopher D. Wentworth

WORK-IN-PROGRESS: Comparison of the DIT2 and EERI instruments for assessing the development of ethical reasoning of engineering students

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:

Measuring the ethical reasoning of engineering students, despite its importance to the profession and accreditation requirements, remains a challenge for many undergraduate programs. Incorporation of standardized instruments into program assessment is considered advantageous due to their reliability, validity, and ease of implementation. The well-established Defining Issues Test (DIT2) is one such instrument with generalized moral dilemmas applicable to all, whereas the similarly structured Engineering Ethical Reasoning Instrument (EERI) provides discipline-relevant scenarios to specifically probe the ethical understanding of engineers. Using a pre/post-intervention methodology, both tests were administered in a semester-long, integrated design/ethics engineering course to facilitate direct comparisons by using the same set of respondents. The pre and post testing results for each test were analyzed using a pairwise, two-tailed t-test for statistical significance. More pronounced differences were observed in the N2-scores and with the EERI for the engineering-specific course. Comparison between the two instruments indicated a positive, and hence somewhat transferable, relationship between the generic and specific instruments. The utility of both instruments for short-term interventional and longitudinal programmatic assessment is discussed.

1.0 INTRODUCTION

1.1 Significance of Engineering Ethics

The technological decisions made by engineers in their work can have profound impact on the public and society. This reaches beyond the extreme "disaster-ethics" where failures in engineering competence or duties result in significant loss of life or property. Engineers, in their day-to-day capacities, often make design and project decisions that incrementally, but assuredly, influence how technology and society interact. As our world becomes more reliant and integrated with technology, it is critical that the next generation of engineering professionals be able to identify their ethical responsibilities and act accordingly.

Engineering has undergone considerable advancements since the industrialization of our society, moving from a primary responsibility to the employer to a profession holding the safety and health of the public as paramount. Recently, an emphasis on sustainability in considering the effects on future generations of the public, and inherent value of the environment has arisen in engineering ethics. Engineering students should be able to identify not only the immediate impacts of their work, but also the indirect, unintended, and future consequences. Clearly, ethical decision-making is a challenging proposition for engineers as they must manage diverse constituencies along with technologies yet to exist. Engineering students, who could safely be described as leaning towards technology optimism in the aggregate, must have ethics emphasized in their undergraduate education alongside technical competence to prepare them for 21st-century practice.

The importance of professional ethics is also clearly recognized by professional engineering societies and accreditation bodies, illustrated by The National Society of Professional Engineers (NSPE) Model Rules [1] and ABET's Student Outcome #4 [2].

Discipline-specific societies (ASCE, ASME, IEEE, etc.) all have similar codes of ethics that govern their professional members.

1.2 Moral Development Theory & Assessment

With its significance established, how exactly is morality defined and promoted in students? Various models have been proposed, with a prominent model being psychologist Lawrence Kohlberg's cognitive-development approach [3]. Following the groundwork of John Dewey and Jean Piaget, Kohlberg outlines distinct stages associated with preconventional, conventional, and postconventional levels. The preconventional level is associated with physical and self-oriented consequences and rarely survives past adolescence. The conventional level stages, which are pervasive, recognize interpersonal concordance and social law and order. The post-conventional (or "principled") level includes critically examined, mutable social contracts and universal ethical principles, yet is only achieved in a minor percentage of adults.

Kohlberg's cognitive and stage-based theories were not without their criticism. Rather than discard the ideas, a "Neo-Kohlbergian" model was proposed by Rest and colleagues [4]. Cognition is still central to the new theory, with the primary modification being from distinct, stair-case stages to transitory moral schemas organized into 1) Personal Interest, 2) Maintaining

Norms, and 3) Postconventional schemas. Schemas are understood to be general knowledge structures that exist in a person's long-term memory. The newly proposed schemas are not developmentally rigid, do not directly assess cognitive operations, and are not necessarily universal. The Maintaining Norms schema stresses clear, uniform, and categorical norms with established reciprocity and roles that govern society. In contrast, the postconventional schema emphasizes the primacy of moral criteria and full reciprocity of shared ideals that are open to scrutiny.

The Defining Issues Test (DIT and the updated DIT2) are established and widely accepted assessment measures of moral reasoning skills, at least as defined in relation to Neo-Kohlbergian moral schema theory [5]. For example, the validity of the DIT has been extensively assessed [6]. Both tests employ a set of moral dilemmas and response questionnaires designed to activate and measure moral schemas of participants with direct quantitative scoring. Both tests provide a numerical index that is supposed to describe a test participant's moral reasoning development. The first proposed index was called the P index and was based on the relative importance participants give to principled moral considerations in making a moral decision. Criticisms of the P index led to defining a new index called the N2 index that made expanded use of both rating and ranking test questions. DIT scores (P and N2 developmental indices) have been shown to be reliable and significantly related to moral comprehension, prosocial behaviors, and desired professional decision-making [7].

A number of discipline-specific ethical reasoning instruments have also been modeled after the DIT. For example, the Engineering and Science Issues Test (ESIT) [8] and the Engineering Ethics Reasoning Instrument (EERI) [9] use technical (rather than general) dilemmas contextually specific to engineering and science. These instruments and their underlying case studies, therefore, have particular utility for incorporation into engineering ethics instruction to promote and assess student ethical reasoning development.

1.3 Research Motivations

For accreditation, engineering programs are required to conduct assessment of ethics learning objectives and continuous improvement of its ethics instruction. Standardized assessments such as the DIT2 and EERI offer significant advantages for engineering instructors compared with use of rubrics or NCEES Fundamentals of Engineering Exam reporting.

Given the utility and similarity of the DIT2 and EERI, Cimino and Streiner [10] administered both tests to assess the effectiveness of ethical interventions in a first-year engineering course. In that study, EERI pretest, EERI post-test, and DIT2 post-test scores were compared with independent recruitment for each test. Surprisingly, a non-significant downshift was observed in the post-test EERI P-score and N2-indices. One possible explanation for the shift offered by Cimino and Streiner was the focus on engineering codes of ethics and its relation to maintaining norms [10]. While we would agree with that rationale, the slightly lower scores may also be attributable to the independency of respondents for each test coupled with the inherently wide variance of the surveys. It is noted that Streiner et al. have continued research on competitive game-based ethics instruction across several public universities, which includes detailed DIT2 and EERI comparative assessments [11].

In this work, we used the same set of participants for all four tests to facilitate direct pairwise comparisons of the pre- and post-intervention scores for the DIT2 and EERI. This approach may also offer additional insight into the relationship between the two tests. For example, is there a positive, transferable relationship between gains in one's ethical reasoning in the general vs. discipline-specific contexts? Additionally, side-by-side comparison of the two instruments facilitates insight into their relative strengths and appropriateness depending on one's assessment objectives. Lastly, this work may provide specific and contrasting data relative to a liberal arts institutional type.

2.0 METHODS

A semester-long study was conducted in the Spring of 2022 using a pre/post methodology for an integrated design/ethics course primarily taken by sophomore engineering majors. The DIT2 survey (paper format completed in-class) and EERI survey (online Qualtrics survey, as an assignment) were both administered in the first and final regular week of the semester. In all survey instances, students were asked to give a good faith effort in completion with no supplemental instructions provided beyond the survey content itself.

Pre/post surveys were processed at the Center for Ethical Development (DIT2 survey) or at Purdue University by the original EERI developer [9]. Of the class size of 19 students, only those students completing all four tests (2 pre and 2 post) were included in the analysis (N = 14). This study was originally performed as part of the program's internal assessment processes, and has been reviewed by the home Institutional Review Board under project S23 EX03 DC IRB HS.

The norms provided for reference and discussed are taken from the 2005-2009 norms compiled by the Center for Ethical Development [12].

3.0 RESULTS & DISCUSSION

The P-score and N2-score results for both the DIT2 and EERI are shown in Table 1. All developmental variables showed an increase in the post-test means. Recall that the newer N2-score, which incorporates the P-score but also includes the rejection of lower-level arguments, is generally considered the more powerful developmental index. The results of this study seem to additionally support that notion as the N2-score improvements were more pronounced than that of P-scores.

Table 1: DIT2 and EERI summary results from the semester-long course study. (N = 14)

	DIT2 P-score		DIT2 N2-score		EERI P-score		EERI N2-score	
	PRE	POST	PRE	POST	PRE	POST	PRE	POST
Mean	28.57	29.86	26.05	32.39	45.14	51.00	39.86	49.93
SD	11.24	14.24	13.39	16.49	12.84	15.20	17.65	19.46
p-value		0.698		0.074		0.175		0.040

The pre and post testing results were analyzed using a pairwise, two-tailed t-test for statistical significance, with P-values shown in the table. The slight increases in the P-scores, again representing “post-conventional” or “principled” thinking, were not statistically significant for either instrument. The N2-score for the DIT2 went from 26.05 pretest to 32.59 posttest, with a P-value of 0.074. The N2-score for the EERI went from 39.86 pretest to 49.93 posttest, with a P-value of 0.040.

Before continuing, it is important to acknowledge the limitations of the study. The sample size (N = 14) was a consequence of the class size typical for the institution. Coupled with relatively large standard deviations associated with these types of instruments, any conclusions should be made prudently. While definitive statements cannot be made, there are a few further points worth discussing.

The non-significance of the P-scores is not unsurprising given the short duration (single semester course) in the maturation of post-conventional thinking. One aspect of the course is the focus on adherence to an explicit professional codes of ethics. This could possibly be considered more associated with the “maintaining norms” schema (Neo-Kohlbergian) or “law-and-order” orientation (Kohlberg’s stage 4), rather than higher post-conventional schema. While engineering codes of ethics should continue being emphasized, the primacy of the underlying moral criteria and deeper discussions of case studies in which ethics codes are equivocal may benefit student development.

A comparison of the DIT 2 scores to national norms [10] is shown in Figure 1. We were encouraged by the observation that the pretest N2-score was much lower at 26.05, while the post-test score was 32.39, which approached parity with the norms. This suggests that students may have benefited mostly from critical reasoning to reject lower-level rationale rather than sole post-conventional P-score achievement.

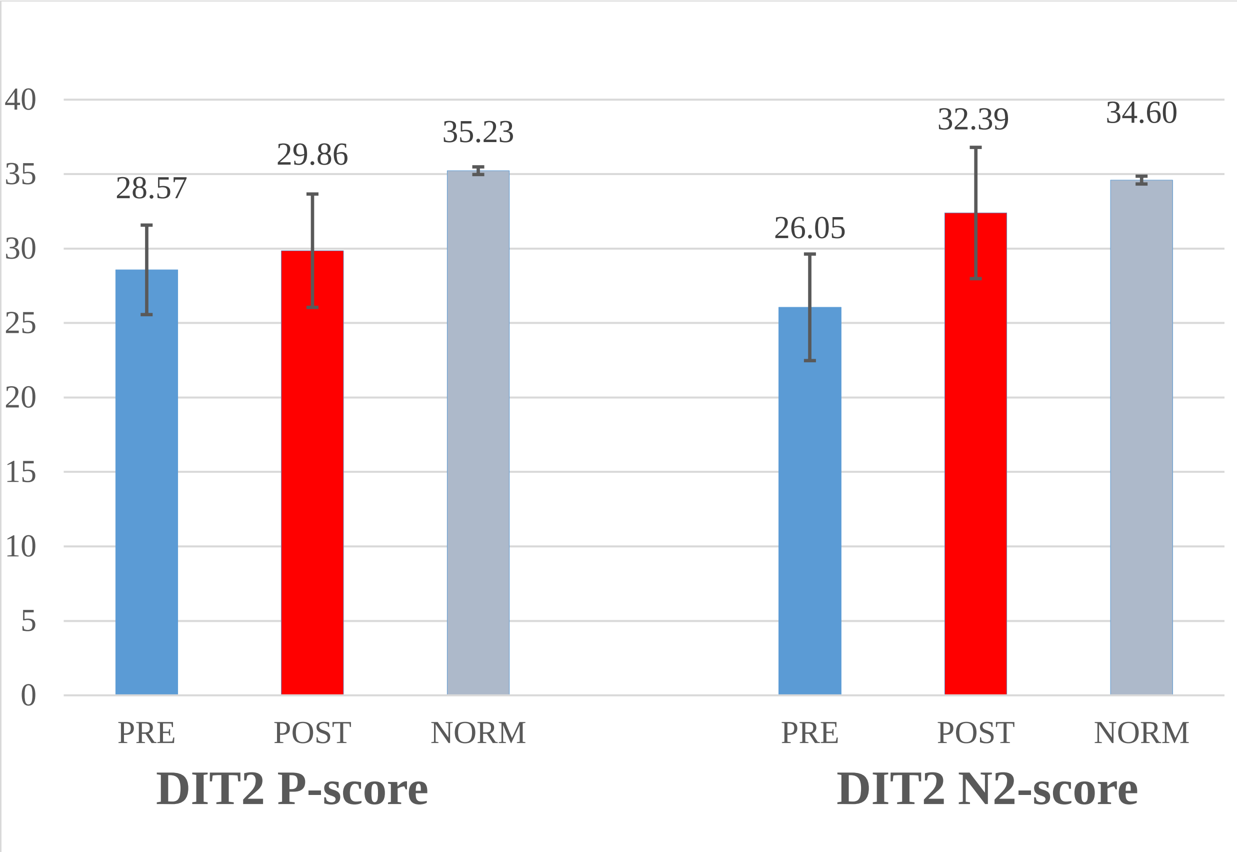

Figure 1: DIT2 pre and post scores (N=14) with comparison to 2005-09 national norms [12] for sophomores (N =3542). (Mean +- S.E.)

We also took notice of the overall lower absolute scores of the small engineering class to the norms. This may be a function of the relatively non-selective home institution compared to other 4-year institutions, or possibly influenced by the specific class composition of perhaps more technically-oriented mentalities of engineering students (as compared to general undergraduate populations with all majors). Or perhaps one factor is gender skew as males generally score lower than females (and the class was predominantly male). These arguments may find some basis in the norms which show lower Tech and Junior College scores (N2-score of 27.67 for males, N=535) than for undergraduates.

The EERI scores are compared to the results of Cimino and Streiner [10] in Table 2. Once again, the home institution absolute scores were lower, especially the pretest scores. This may be a function of institutional and/or demographic differences. At the same time, the post-intervention scores were improved in our case compared to the downshift present in the reference. Acknowledging the limitations in sample size and high variability, several possible explanations come to mind. First, the identical set of students assessed may have facilitated more consistency in the pre/post comparison than the independent tests. Another possible factor may be the class size: smaller, discussion-based course vs. an assumed larger class. Or perhaps timing was a possibility: 2nd semester sophomore course vs. 1st semester freshmen course. Lastly, liberal arts institutions, for whatever reason(s), are the only institutional type with consistently large effect size with respect to DIT gains [13], which may possibly be transferable to consistency in EERI gains.

Table 2: Comparison of pre and post EERI scores.

	Home Institution Study		Reference Study [10]	
	Pre-test (N=14)	Post-test (N=14)	Pre-test (N=34)	Post-test (N=28)
EERI P-score	45.14	51.00	56.56	51.71
EERI N2-score	39.86	49.93	54.32	53.57

The comparison of the absolute scores between the two instruments is worth mentioning. Similar to Cimino and Streiner [10], the DIT2 indices in our case were also much lower than those of the EERI. From our perspective, it is not necessarily an expectation that the absolute values of the two tests to be similar; they are different tests with different scenarios and different responses and may have different scoring scales. At the same time, the EERI is based on the framework of the DIT2 with items assigned according to the NeoKohlbergian schema theory and the difference in absolute scores is somewhat substantial. Now this difference may simply be a result of a different scale or offset. However, it may be possible that engineering students are better at choosing the assigned “principled” response items in the EERI than the DIT2 due to the discipline-specific dilemmas. This offset may or may not hold true for general undergraduate respondents, who may score lower on the EERI but potentially higher on the DIT2 than engineering students. On the other hand, given the high pre-test EERI scores of incoming freshmen [10], which is probably more representative of a general population than of trained engineering professionals, it would seem that the “principled” items are more frequently selected (in general) in the EERI due to its dilemmas and response items.

While the scores on the EERI and DIT2 were different in absolute value, we endeavored to understand if there was any association between students’ scores on the respective exams. That is, we compared the association between:

- the DIT2 Pre-test P-score with the EERI Pre-test P-score,
- the DIT2 Pre-test N2-score with the EERI Pre-test N2-score,
- the DIT2 Post-test P-score with the EERI Post-test P-score,
- the DIT2 Post-test N2-score with the EERI Post-test N2-score,

We compared all associations using Spearman's rank-order correlation coefficient test (Spearman’s rho). As the pairwise data between the scoring systems on each exam did not show a clear linear correspondence, this non-parametric test was chosen.

The association was weakest between the DIT2 Pre-test P-score and EERI Pre-test P-score. However, the N2-scores in both the pre-test and post-test showed a stronger positive association, indicating that students scoring higher on the DIT2 post-test with N2-scoring also scored higher on the EERI Post-test with N2-scoring. Due to the nature of the statistical test, we can only interpret the relative ranking of the scores earned. From the limited sample, we conjecture that the N2-scoring is a more consistent indicator between the exams and think it is worthy of future examination in a larger population to compare the same students taking the different exams.

Table 3: Pairwise comparison of scores using Spearman's rho.

Comparison	Spearman's rho	p-value
DIT2 Pre-test P, EERI Pre-test P	0.1078	0.7388
DIT2 Pre-test N2, EERI Pre-test N2	0.5754	0.0503
DIT2 Post-test P, EERI Post-test P	0.5643	0.0559
DIT2 Post-test N2, EERI Post-test N2	0.4315	0.1612

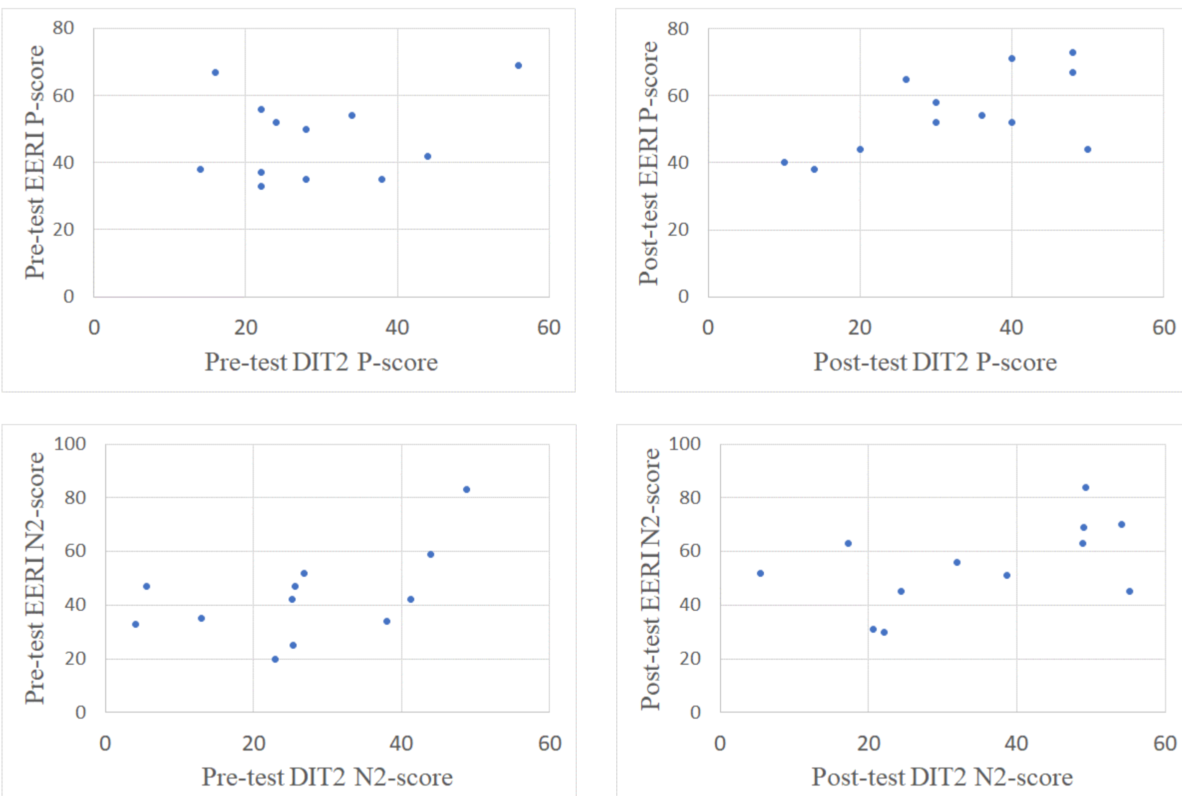

Figure 2: Graphical comparisons of EERI score vs. DIT score for each of the four associations.

A Partial Confirmatory Factor Analysis [14] suggests that the items used in the EERI do not map to the three neo-Kohlbergian schemas, which supports the possibility that DIT2 and EERI are measuring different things. This doesn't mean that changes in the EERI scores (P or N2) cannot be used as a proxy for changes in the DIT2 scores. Indeed, the Spearman's rho test shows some evidence that EERI N2 score gains might be a proxy for DIT2 N2 gains.

4.0 CONCLUSION & FUTURE WORK

In the context of the specific engineering course and population, the EERI showed better discrimination in the pre and post test results. This is of course likely due to the context-specific nature and synergy of the engineering ethics course and the survey itself. We would recommend the EERI for assessment purposes of engineering programs or interventional classes. The DIT2, especially the N2-score, may be considered a somewhat less desirable, but acceptable, surrogate for longer duration assessment efforts.

This single course study provides some encouraging evidence on ethics assessment approaches in the context of a liberal arts engineering program. However, this study and previous literature seem to raise additional research questions, particularly with respect to repeatability and reproducibility in relation to different institutional types. We are currently conducting additional studies to better understand the impact of a liberal arts environment on engineering ethics.

ACKNOWLEDGEMENTS

We would like to especially thank Dr. Carla Zoltowski at Purdue University for Qualtrics administration and data reporting of the EERI scores.

REFERENCES

[1] *Code of Ethics for Engineers. Publication #1102*. National Society of Professional Engineers. Revised July 2019.

[2] *Criteria for Accrediting Engineering Programs, 2023 – 2024,* Criterion 3. Student Outcomes, ABET, 2023.

[3] L. Kohlberg, "The Cognitive-Developmental Approach to Moral Education." *The Phi Delta Kappan*, 56(10), 670–677. 1975.

[4] J. Rest, D. Narvaez, M.J. Bebeau, and S.J. Thoma. "A Neo-Kohlbergian Approach: the DIT and Schema Theory." *Educ. Psychol. Rev.* 11(4): 291-324, 1999.

[5] J. Rest, D. Narvaez, M.J. Bebeau, and S.J. Thoma. *Postconventional Moral Thinking: A neo-Kohlbergian Approach*. Mahwah, NJ: Erlbaum Publishers, 1999

[6] S.J. Thoma. "Research on the defining issues test" *in Handbook of moral development,* M. Killen, & J. G. Smetana, Mahwah, NJ: Erlbaum Publishers. 2006, pp. 67-91.

[7] M.J. Bebeau and S.J. Thoma. *Guide for DIT-2*. Center for the study of Ethical Development. University of Alabama. 2003.

[8] J. Borenstein, M.J. Drake, R. Kirkman, and J.L. Swann. "The Engineering and Science Issues Test (ESIT): a discipline-specific approach to assessing moral judgment." *Sci Eng Ethics.* 16(2):387-407. 2010. doi:10.1007/s11948-009-9148-z

[9] Q. Zhu, C.B. Zoltowski, M.K. Feister, P.M. Buzzanell, and W.C. Oakes. "The Development of an Instrument for Assessing Individual Ethical Decision making in Project-based Design Teams: Integrating Quantitative and Qualitative Methods." *121st ASEE Annual Conference*. Indianapolis, IN. Paper ID #10060. 2014.

[10] R.T. Cimino and S.C. Streiner. "Effectiveness of Ethical Interventions in a First-Year Engineering Course: A Pilot Study." *2018 FYEE Conference.* Glassboro, NJ. Paper ID #24421. 2018.

[11] S. Streiner, D.D. Burkey, M.F. Young, R.T. Cimino, and J. Pascal. "Engineering Ethics Through High-Impact Collaborative/Competitive Scenarios (E-ETHICCS)." *2021 ASEE Virtual Annual Conference.* Paper ID #33860. 2021.

[12] Y. Dong, "Norms for DIT2: From 2005 to 2009." Center for the Study of Ethical Development. University of Alabama. Available at: https://ethicaldevelopment.ua.edu/uploads/8/4/9/8/84986096/norms-for-dit2__05-09_.pdf Accessed September 1, 2022.

[13] 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.), Springer Netherlands, 2004, pp. 375–440.

[14] P.W. Odom and C.B. Zoltowski. "Statistical Analysis and Report on Scale Validation Results for the Engineering Ethical Reasoning Instrument (EERI)." *2019 ASEE Annual Conference*. Paper ID #26859. 2019.