The Challenges of Assessing In-the-Moment Ethical Decision-Making

Ms. Tori N. Wagner, University of Connecticut

Tori Wagner is a doctoral student at the University of Connecticut studying Engineering Education. She has a background in secondary science education, playful learning, and digital game design.

Dr. Daniel D. Burkey, University of Connecticut

Daniel Burkey is the Associate Dean of Undergraduate Programs and the Castleman Term Professor in Engineering Innovation in the College of Engineering at the University of Connecticut. He earned his B.S. in Chemical Engineering from Lehigh University in 1998, his M.S.C.E.P and Ph.D., both in Chemical Engineering, from the Massachusetts Institute of Technology in 2000 an 2003, respectively, and his M.A.Ed with a focus in Research Methods, Measurement, and Evaluation from the University of Connecticut in 2023.

Dr. Richard Tyler Cimino, New Jersey Institute of Technology

Dr. Richard T. Cimino is a Senior Lecturer in the Otto H. York Department of Chemical and Materials Engineering at New Jersey Institute of Technology. His research interests include the intersection of engineering ethics and process safety, and broadening inclusion in engineering, with a focus on the LGBTQ+ community.

Dr. Scott Streiner, University of Pittsburgh

Scott Streiner is an Assistant Professor in the Industrial Engineering Department, teaches in the First-Year Engineering Program and works in the Engineering Education Research Center (EERC) in the Swanson School of Engineering at the University of Pittsburgh. Scott has received funding through NSF to conduct research on the impact of game-based learning on the development of first-year students' ethical reasoning, as well as research on the development of culturally responsive ethics education in global contexts. He is an active member of the Kern Engineering Entrepreneurship Network (KEEN), the American Society for Engineering Education (ASEE), and the Institute of Industrial and Systems Engineers (IISE)

Dr. Kevin D. Dahm, Rowan University

Kevin Dahm is a Professor of Chemical Engineering at Rowan University. He earned his BS from Worcester Polytechnic Institute (92) and his PhD from Massachusetts Institute of Technology (98). He has published two books, "Fundamentals of Chemical Engineer

Dr. Jennifer Pascal, University of Connecticut

Jennifer Pascal is an Assistant Professor in Residence at the University of Connecticut. She earned her PhD from Tennessee Technological University in 2011 and was then an NIH Academic Science Education and Research Training (ASERT) Postdoctoral Fellow at

The Challenges of Assessing In-The-Moment Ethical Decision-Making

The engineering education community lacks a consensus on an effective assessment tool to gauge the growth of undergraduate students' ethical reasoning throughout a course or program. The Engineering Ethics Reasoning Instrument (EERI) was developed by a team at Purdue and is based on the NSPE Code of Ethics [1,2]. Previous research has shown that the EERI failed to detect significant growth in ethical reasoning during a single-semester course, which contained substantial ethics content [3].

We hypothesized that perhaps the EERI could detect a significant change in students' ethical reasoning over the course of a four-year undergraduate program, during which students are typically exposed to many engineering-contextualized ethical dilemmas, both via coursework as well as potential work experiences.

Using a quasi-experimental design, we used the EERI to measure changes in the ethical reasoning of 178 undergraduates at a Public R1 university in the Northeast across multiple engineering disciplines. Analysis of EERI data typically focuses on two outputs - a student's P score and N2 score. The P score measures the extent to which students employ Kohlbergian postconventional thinking, which is characterized by ethical reasoning based on universal good [1,4]. The N2 score takes into account how much postconventional thinking is used and preconventional (self-interested) thinking is absent [1,4]. We found that over the course of the four-year program, the EERI did not indicate any change in N2 score ($n = 178$, $p = 0.65$), but showed a decrease of -3.38 in P score ($n = 178$, $p = 0.017$). This suggests that over four years, there is a reduction in students prioritizing decisions that were altruistic and based on universal good. It is challenging to predict why this occurs, but we tentatively suggest that it may reflect a more accurate representation of students' thoughts on these ethical dilemmas. Additionally, it might indicate a deeper consideration of the complex factors typically involved in real ethical decisions, rather than merely an abstract evaluation of what a reasonable engineer should do.

Given these results and to gain a fuller understanding of students' changes in ethical reasoning throughout their undergraduate programs, we contend that qualitative measures should also be employed. Ethical reasoning can be ill-defined and multidimensional, making quantification of a student's ethical reasoning challenging and difficult to interpret. A qualitative instrument designed to be 1st person, situated, contextually-rich, and playful might more accurately capture students' in-the-moment ethical decision-making.

Introduction

The assessment of ethical reasoning is of paramount importance in engineering education. As future engineers are poised to face increasingly complex ethical dilemmas, amplified by rapid technological advancements, it becomes essential to ensure they are well-equipped with robust ethical reasoning skills. These skills are not just a cornerstone of academic accreditation, such as that stipulated by ABET, but also a critical component of engineers' success and responsibility in their professional lives [5].

In response to this need, various assessment tools have been developed to evaluate the ethical reasoning abilities of engineering students. One such tool is the Engineering Ethics Reasoning Instrument (EERI) [1]. The EERI, rooted in the NSPE Code of Ethics for Engineers, aims to offer a standardized measure to assess the ethical reasoning skills of engineering students [1,2]. It is structured similarly to the DIT-2, but is situated in engineering [6]. The development of the EERI draws upon Kohlberg's theory of moral development [4]. This theory delineates the stages of moral evolution, from pre-conventional levels focused on self-interest and external rewards to post-conventional levels where abstract principles and universal values predominate [4]. The EERI uses this framework to evaluate where students stand in their moral development and how they apply these principles to real-world engineering scenarios. Typically, the evaluation of data from the EERI concentrates on two key metrics: the P score and the N2 score. The P score is an indicator of the degree to which students engage in Kohlbergian postconventional thinking [1,4]. This form of thinking is marked by ethical reasoning grounded in the notion of universal good. The N2 score reflects the extent of post-conventional thinking employed by the students, while also considering the absence of pre-conventional thinking, which is oriented around self-centeredness [1,4].

There is a notable gap in current research using the EERI. Most previous studies employing the EERI have been limited in scope - either involving small sample sizes, lacking a longitudinal perspective, or focusing primarily on graduate students rather than undergraduates [7,8]. Our study aimed to address this gap by conducting a larger-scale, longitudinal analysis specifically targeting undergraduate students.

In our previous research, we found that students' ethical reasoning abilities, as measured by the EERI, exhibit minimal change over the course of a semester [3]. This raises questions about whether our current ethics curriculum is effectively fostering moral reasoning development or if the EERI might be insufficient in capturing the subtleties of students' situated understanding and ability to reason and act ethically in authentic scenarios. In response to these findings, we broadened the scope of our study to encompass the full duration of students' undergraduate careers. This expansion was driven by the hypothesis that a single semester of ethics education within the curriculum might not be sufficient to effect significant changes in students' ethical reasoning. However, we speculated that over the entirety of their undergraduate experience, a more notable change in their ethical reasoning might become evident.

Methodology

This longitudinal study employed a quasi-experimental design to investigate the ethical reasoning progression of undergraduates across all engineering disciplines excluding computer science (the computer science students take a separate first-year course different from the other engineering majors). The investigation was conducted at a Public R1 university located in the Northeast United States. The participants consisted of 178 undergraduate students enrolled in engineering programs (See Table 1). The selection criterion included students who were in their first year during the Spring semester of 2020 and who subsequently progressed to their fourth year by the Spring semester of 2023.

Table 1 - Demographic Information

Gender		
Female	47	(26.40%)
Male	127	(71.35%)
I Prefer to Identify as… (short response)	0	(0.00%)
Prefer Not to Answer	4	(2.25%)
Race		
African American or Black	6	(3.37%)
Asian or Pacific Islander	26	(14.61%)
Hispanic	23	(12.92%)
American Indian/ Other Native American	0	(0.00%)
Caucasian (Other than Hispanic)	132	(74.16%)
Other	4	(2.25%)
Prefer not to answer	5	(2.81%)
Concentration		
Chemical	24	(13.48%)
Civil	23	(12.92%)
Environmental	8	(4.49%)
Mechanical	84	(47.19%)
Other	1	(0.56%)
Biomedical	30	(16.85%)
Electrical	7	(3.93%)
Computer	5	(2.81%)

The undergraduate engineering curriculum begins with a common first-year course where initial instruction in ethics is introduced. As students progress through their respective programs, the extent and nature of additional ethics instruction vary significantly across different engineering disciplines. Each discipline incorporates ethics education in its own unique way, adapting it to fit the specific context and demands of the field. In their final year, all students complete a capstone course, which allows students to apply their accumulated knowledge in a practical, project-based context. Students are required to not only focus on the technical aspects of design, but also the ethical and societal implications of their decisions, as per ABET requirements.

The Engineering Ethics Reasoning Instrument (EERI) was assigned as a homework assignment for data collection purposes. This instrument measures two key metrics: the P score and the N2 score. The EERI was administered as a mandatory component of required courses in both the first and fourth years of the undergraduate engineering curriculum. The EERI was first administered in the Spring semester of the first academic year (2020) and then again in the Spring semester of the fourth academic year (2023). The analysis exclusively focused on paired data, meaning that only data from participants who completed the EERI in both their first and fourth years were included. Unpaired data, representing students who completed the EERI only in their first or fourth year, were excluded from the analysis. Furthermore, any incomplete EERI submissions were discarded to maintain the integrity of the dataset. To ensure the quality and sincerity of responses, submissions completed in less than 1000 seconds (~17 minutes) were also excluded, under the presumption that these responses did not involve a thorough engagement with the instrument. An expert in the field was tasked with taking the assessment to establish a baseline for the amount of time reasonably required to complete the EERI thoughtfully. The change in score was calculated for each individual, and histograms, Q-Q Plots and the Shapiro-Wilks test were used to evaluate the normality of this data [9]. A paired pre/post *t*-test was employed to evaluate the differences in the EERI scores from the first to the fourth year. This test was chosen for its effectiveness in comparing two related samples.

Results

Initially, the P score, which measures postconventional thinking based on universal good, had a mean pre-score of 60.62 (SD = 17.59). Over four years, it decreased to a mean post-score of 57.24 (SD = 19.37), resulting in a mean difference of -3.38 (SD = 18.71). The N2 score, reflecting the balance of postconventional and preconventional (self-interested) thinking, showed a mean pre-score of 57.49 (SD = 16.99) and a slightly increased post-score of 58.08 (SD = 18.28), with a mean difference of 0.59 (SD = 17.30). (See Table 2)

Table 2 - Descriptive Statistics

Group	Statistic	Prescore		Postscore		Difference	
		M	*SD*	*M*	*SD*	*M*	*SD*
N = 178	P	60.62	17.59	57.24	19.37	-3.38	18.71
	N2	57.49	16.99	58.08	18.28	0.59	17.30

Histograms and Q-Q plots of individuals' change in P and N2 scores suggested normality (Appendix A). The Shapiro-Wilks test *p*-value exceeded the significance level of 0.05, also suggesting that the data are normally distributed. This normality assumption supports the appropriateness of using the *t*-test for our analysis.

The change in the P score was statistically significant with a t-statistic of -2.41 and a *p*-value of 0.017, indicating a significant decrease in postconventional ethical reasoning. The effect size (Cohen's d) was -0.18, suggesting a small effect. The change in the N2 score was not statistically significant, with a t-statistic of 0.46 and a *p*-value of 0.65. This indicates that there was no significant change in the balance between postconventional and preconventional thinking among the students. The effect size was 0.03, indicating a negligible effect (Table 3).

Table 3 - EERI Paired *t*-test Results

Group	Statistic	Normality Tests		95% CI		*t*-test results		
		W	*p*	Lower	Upper	*t*-stat	*p*	*d*
N = 178	P	0.99	0.56	-6.14	-0.61	-2.41	0.017*	-0.18
	N2	0.99	0.44	-1.97	3.15	0.46	0.65	0.03

*Significant at the 0.95 level

The significant decrease in the P score suggests a reduction in students' propensity for ethical reasoning based on universal good, potentially indicating a shift towards more pragmatic or context-dependent ethical considerations. This might reflect a maturation in their ethical reasoning, as they encounter more complex, real-world scenarios in their engineering education. The lack of significant change in the N2 score indicates a stable level of ethical reasoning in terms of balancing self-interested and postconventional thinking. This stability might suggest that while the nature of their ethical reasoning shifts (as indicated by the P score), the overall balance between self-interested and universal considerations remains constant.

These findings are significant in the context of engineering education, as they highlight a potential shift in ethical reasoning among students over the course of their undergraduate studies. The reduction in postconventional ethical reasoning (P score) could reflect a more nuanced and realistic approach to ethical decision-making, influenced by the complexities of real-world engineering problems. Understanding these shifts is crucial for developing curricula that effectively support the ethical development of future engineers, ensuring they are prepared to make decisions that are not only technically sound but also ethically responsible.

Discussion

A significant drop in P scores alongside stable N2 scores among engineering undergraduates over a four-year period indicates a complex evolution in ethical reasoning. This pattern suggests that as students mature through their engineering education, there's a notable decrease in their postconventional thinking, which is focused on universal principles, as evidenced by the declining P scores. Simultaneously, the stability in the N2 scores implies that their preconventional thinking, characterized by self-centered perspectives, also diminishes. This dual shift suggests a nuanced development in ethical reasoning: while students move away from idealistic, universal principles, as might be expected of novices, they also show a reduction in self-interested thinking. Such a trend could be the result of increased exposure to real-world complexities and pragmatic problem-solving, leading to a more balanced, contextually sensitive approach to ethical decision-making. This balanced approach seems to replace both the idealism of postconventional thinking and the self-focused nature of preconventional reasoning with a mature, nuanced perspective that recognizes the intricacies and practical implications of ethical decisions in the field of engineering.

The EERI relies on Kohlberg's theory of moral development [1,4]. Our study's results prompt us to think about the degree and manner in which students' moral thinking, particularly their progression towards higher-level reasoning, changes over the course of a four-year undergraduate program. The typical duration of undergraduate studies may not always be sufficient for significant Kohlbergian moral development, which is often a gradual and multifaceted process. It's influenced not just by academic instruction but also by a range of other factors, including personal experiences and environmental influences. As such, the rate and nature of moral development can vary significantly among individuals. Some students may naturally progress to higher stages of moral reasoning, while others may require longer periods or different experiences to exhibit similar changes.

Kohlberg's theory of moral development has also been subject to several critiques that question its comprehensiveness and applicability. A primary criticism centers on its perceived overemphasis on justice and rationality, potentially overlooking other moral dimensions such as care and compassion, as highlighted by Carol Gilligan's ethics of care theory [10]. Gilligan criticizes its limited scope in addressing the complexity of moral behavior, as it primarily focuses on moral reasoning and judgment, without adequately considering factors like emotions, social environment, and individual experiences that significantly influence moral actions [10]. This focus has led to allegations of a gender bias, suggesting that the theory favors a moral reasoning style more commonly associated with males [10]. Additionally, critics argue that Kohlberg's stages are culturally biased, reflecting a Western, individualistic perspective on morality that may not be universally applicable or relevant in collectivist societies [11]. There is also skepticism about the rigidity of the stage progression, with critics contending that moral development is more fluid and context-dependent than the theory suggests [10,11,12]. The limitations and criticisms of Kohlbergian theory points to the potential need for a more holistic educational approach, emphasizing the importance of richly authentic ethical education in undergraduate engineering programs.

In assessing ethical reasoning through our study, we have come to critique the reliance on quantitative methods for their limited scope in capturing the multifaceted nature of ethics. While the EERI is a valuable tool for assessing ethical reasoning, it has limitations. The EERI, like any standardized instrument, may oversimplify the complexity of real-world ethical dilemmas. This simplification could lead to misunderstandings of students' ability to navigate complex ethical scenarios. Our experience with the changes in P scores and the stagnant N2 scores has highlighted the shortcomings of purely quantitative data. These metrics, while offering some level of insight, largely leave us in a position of conjecture regarding their deeper significance. Ethics is inherently multidimensional and not easily confined to a singular definition. It poses a significant challenge when attempting to be quantified into a singular numeric value. This realization has led us to advocate for the inclusion of qualitative measures in the study of ethical reasoning. We believe that qualitative methods, with an emphasis on detailed narratives and subjective experiences, are likely to provide a more accurate representation of student reasoning. These methods may enable a richer exploration of the complexities and subtleties in ethical decision-making, thereby offering a more comprehensive and nuanced understanding than what quantitative measures can reveal on their own. This approach is in line with our view that ethical reasoning is a complex, subjective process that requires a broad method of study to be fully understood.

In response to the limitations of quantitative methods in assessing ethical reasoning, we propose the development of a new, qualitative instrument that adopts a digital game-based approach. This tool would feature immersive, first-person scenarios that place participants directly in contextually rich, realistic situations, mirroring the complexities they might encounter in professional settings. The use of game mechanics, such as branching storylines and interactive problem-solving, is intended to actively engage participants, making the process more dynamic and reflective of actual decision-making processes [13]. The game-based tool should be rooted firmly in the learning theory of situated cognition [14,15]. Situated cognition posits that knowledge and learning are integrally tied to the context and situation in which they occur [14]. The instrument would be designed to not only assess but also foster ethical reasoning skills. By engaging participants in these narrative-driven scenarios, it would encourage them to consider the implications of their decisions in real-time, thereby providing deeper insights into their ethical reasoning processes. In these richly authentic scenarios, students may be more likely to transfer their developing ethical reasoning skills to situations outside the classroom and into their engineering careers [14]. The inclusion of reflective elements, such as prompts for introspection and feedback mechanisms, would further enhance this learning process. Participants would be encouraged to articulate and reflect on their decision-making rationale, offering valuable insights into their thought processes and ethical perspectives. Additionally, the digital infrastructure of this game-based tool would allow for the collection of detailed data on student interactions, decisions, and reasoning processes within the game. This data could include metrics such as decision times, choice patterns, and pathways taken through the game's narrative. This qualitative, game-based tool would offer significant benefits over traditional assessment methods. It could be an innovative and effective approach to understanding and cultivating ethical reasoning in a manner that aligns closely with the multifaceted nature of real-world ethical decision-making.

In our ongoing research, we are developing a narrative game titled *Mars: An Ethical Expedition* (*Mars*) to foster ethical decision-making skills in undergraduate engineering students [16]. This game is rooted in situated cognition, providing an authentic, contextualized, and playful environment for students to apply and reflect upon their ethical reasoning abilities [14,16]. Our objective is to create an engaging, immersive tool that facilitates the development of engineering ethics within the narrative context of realistic decision-making scenarios. Looking ahead, we aim to enhance *Mars* by incorporating open-ended responses for key decisions made by the students. These will not only prompt reflection and deeper engagement with ethical dilemmas but also serve as instructive tools for educators. Furthermore, we plan to utilize text classification techniques to assess and provide feedback on the quality of students' ethical reasoning. By analyzing and categorizing student responses based on a scoring rubric, we aim to evaluate the game's effectiveness in developing ethical decision-making skills.

Conclusion

We found no detectable growth in students' ethical reasoning across a four-year undergraduate engineering program, as measured by the EERI. The observed decrease in P scores and stability of N2 scores may suggest a nuanced shift in ethical reasoning, though we can only speculate as to the nature of that shift. We suspect that the reduction in naivety commonly experienced by students during their undergraduate years might have led to the observed decrease in adherence to idealistic, universal principles. This change was accompanied by a concurrent reduction in self-centered thinking, causing the change in EERI scores. These results indicate that current methods of assessing ethics may not sufficiently capture students' ethical development and emphasize the need for a more dynamic, contextually-rich, and interactive approach to assessing engineering ethics.

References

[1] Q. Zhu, C. B. Zoltowski, M. K. Feister, P. M. Buzzanell, W. C. Oakes, and A. D. Mead, "The Development of an Instrument for Assessing Individual Ethical Decisionmaking in Project-based Design Teams: Integrating Quantitative and Qualitative Methods." Presented at ASEE Annual Conference & Exposition, Indianapolis, IN, USA, June, 2014. 10.18260/1-2--23130

[2] National Society of Professional Engineers. "NSPE Code of Ethics for Engineers." NSPE.org. https://www.nspe.org/resources/ethics/code-ethics [accessed Jan. 4, 2023].

[3] C. Moos, L. Dougher, L. Bassett, M. Young, and D. D. Burkey, "Game-Based Ethical Instruction in Undergraduate Engineering," *NEAG Journal*, no. 1, pp. 20–37, Mar. 2023, doi: 10.59198/8259gnir7.

[4] L. Kohlberg, *The psychology of moral development: The nature and validity of moral stages*. San Francisco, CA: Harper & Row, 1984.

[5] Accreditation Board for Engineering and Technology. "Criteria for accrediting engineering programs, 2022-2023". ABET. https://www.abet.org/accreditation/accreditation-criteria/accreditation-policy-and-procedure-manual-appm-2023-2024/ (accessed Feb. 7, 2024).

[6] Rest, James; Narvaez, D.; Bebeau, M.; Thoma, S. (1999). "DIT-2: Devising and testing a new instrument of moral judgment," *Journal of Educational Psychology*. vol. 91, no. 4, pp. 644-659. 1999.

[7] L. G. Kisselburgh, J. L. Hess, C. B. Zoltowski, J. Beever, and A. O. Brightman, "Assessing a Scaffolded, Interactive, and Reflective Analysis Framework for Developing Ethical Reasoning in Engineering Students," *ASEE Annual Conference and Exposition*, Jul. 2016, doi: 10.18260/p.26288.

[8] J. L. Hess, J. Beever, C. B. Zoltowski, L. G. Kisselburgh, and A. O. Brightman, "Enhancing engineering students' ethical reasoning: Situating reflexive principlism within the SIRA framework," *Journal of Engineering Education*, vol. 108, no. 1, pp. 82–102, Jan. 2019, doi: 10.1002/jee.20249.

[9] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (Complete samples)," *Biometrika*, vol. 52, no. 3/4, p. 591, Dec. 1965, doi: 10.2307/2333709.

[10] C. Gilligan. "In a Different Voice: Women's Conceptions of Self and of Morality." *Harvard Educational Review*, vol. 47, no. 4, 1977, pp. 481–517, sfonline.barnard.edu/sfxxx/documents/gilligan.pdf, https://doi.org/10.17763/haer.47.4.g6167429416hg5l0.

[11] J.R. Snarey. "Cross-Cultural Universality of Social-Moral Development: A Critical Review of Kohlbergian Research" *Psychological Bulletin*, vol. 97, no. 2, pp. 202–232, Jan. 1985, doi: 10.1037/0033-2909.97.2.202.

[12] R. A. Shweder, "Thinking through Cultures: Expeditions in Cultural Psychology," *Philosophy East and West*, vol. 43, no. 2, p. 326, Apr. 1993, doi: 10.2307/1399622.

[13] J. P. Gee, *What video games have to teach us about learning and literacy*. New York, NY: Palgrave Macmillan, 2003.

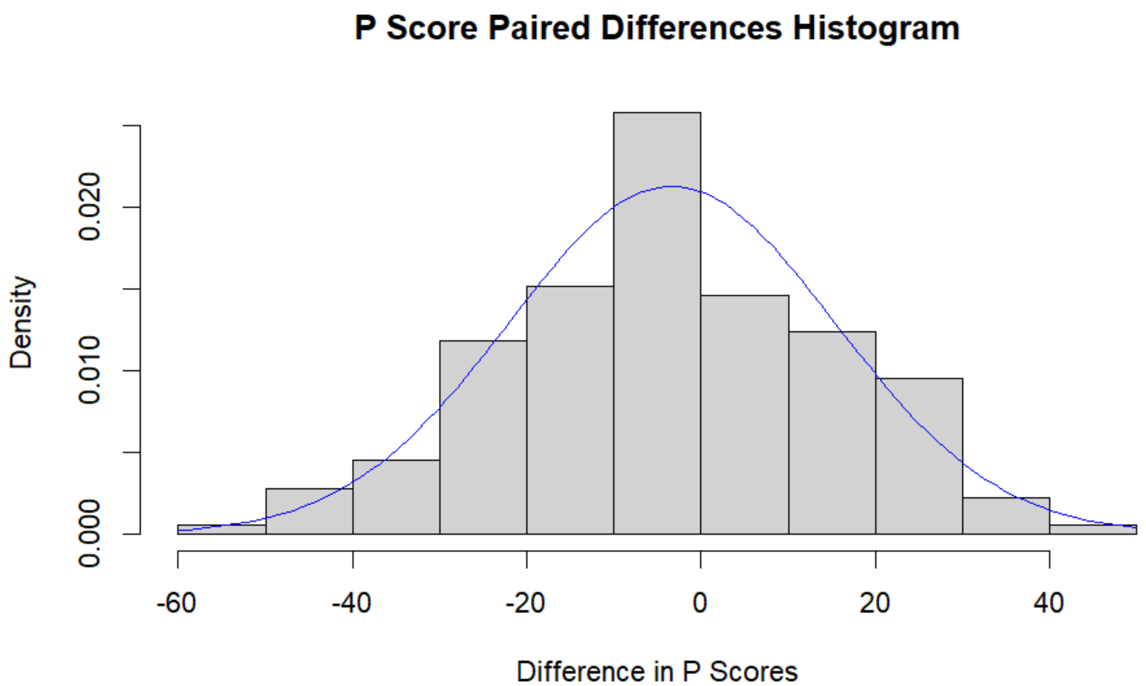
[14] J.S. Brown, A. Collins, and P. Duguid, "Situated cognition and the culture of learning." *Educational Researcher*, 1989, vol. 18, no. 1, pp. 32-42.

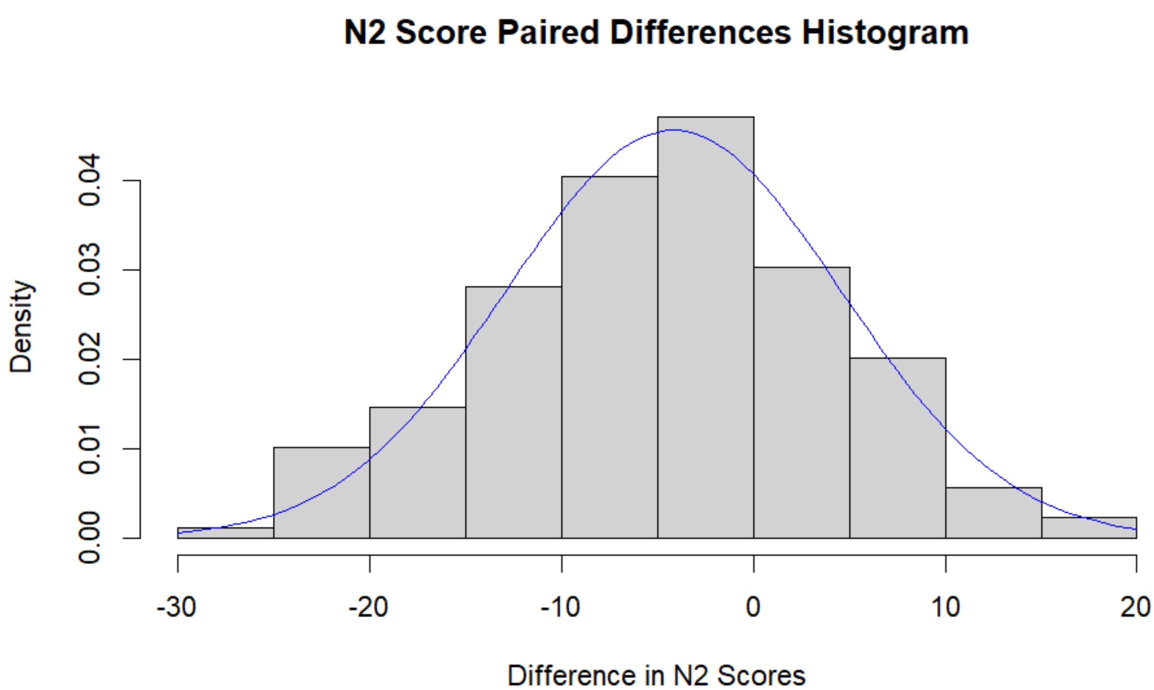
[15] J. Lave and E. Wenger, *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press, 1991.

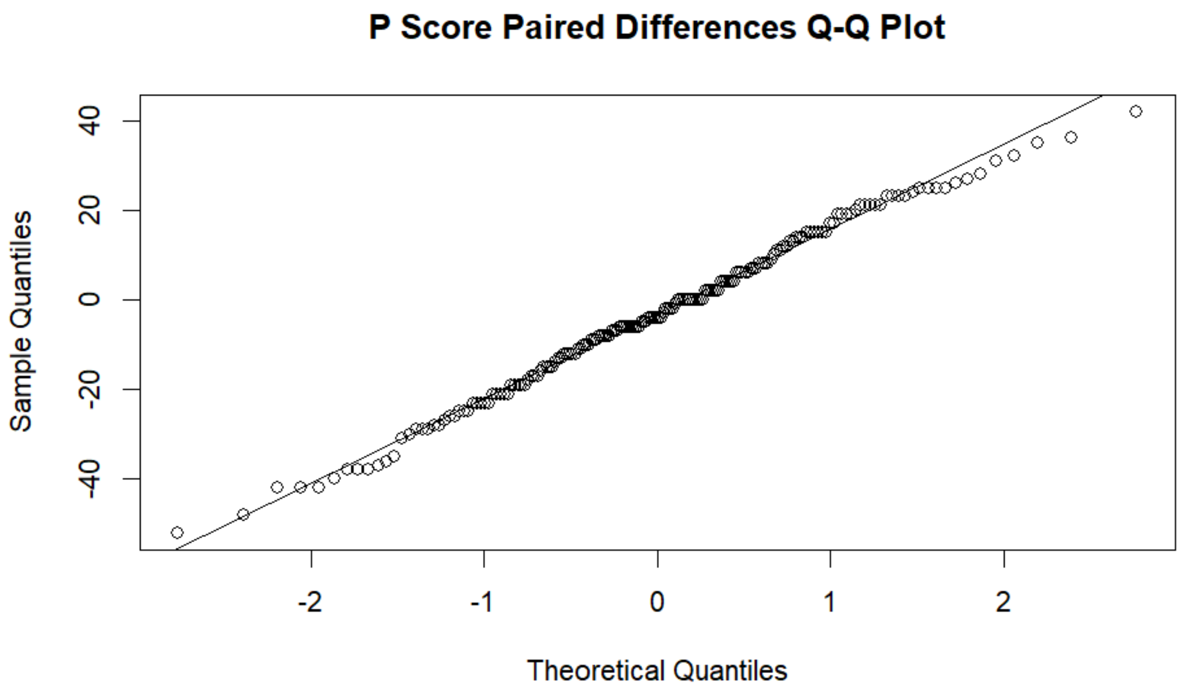
[16] T. Wagner, L. Bassett, M. F. Young, D. D. Burkey, S. Streiner, and J. Pascal. “The Power of Playful Learning – Ethical Decision Making in a Narrative-Driven, Fictional, Choose-Your-Own Adventure”, *2023 ASEE Annual Conference and Exposition*, Baltimore, MD, June 2023, work-in-progress paper.

APPENDIX A

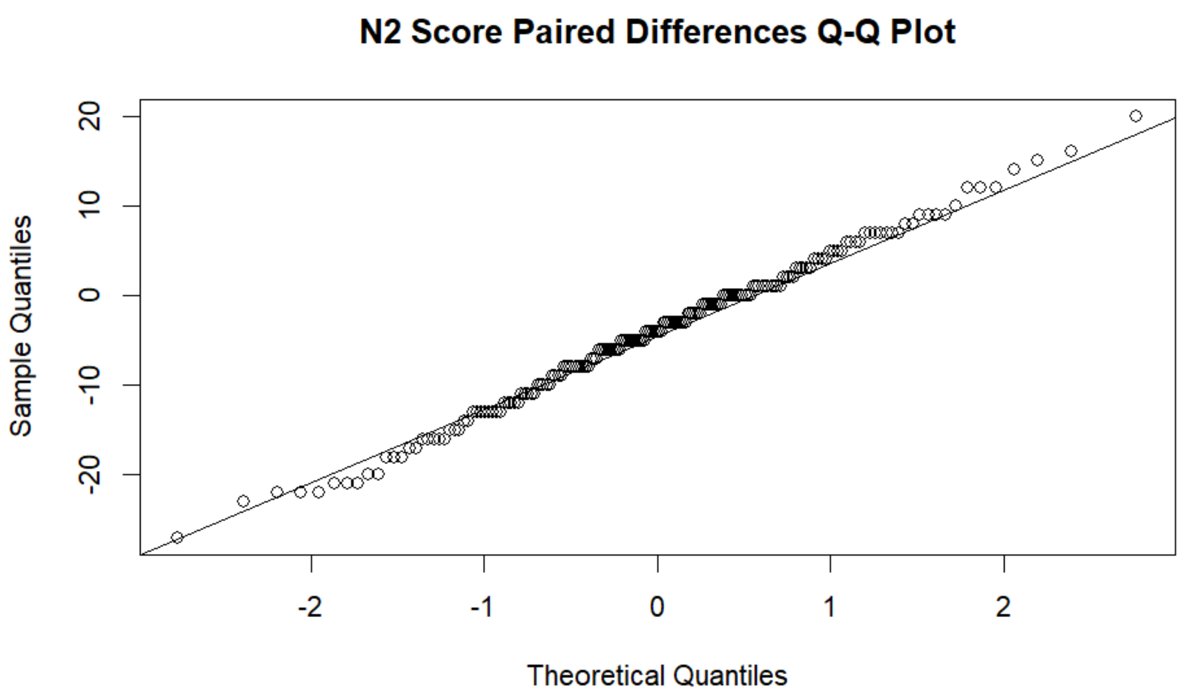
HISTOGRAMS AND Q-Q PLOTS OF INDIVIDUALS' CHANGES IN EERI SCORES







P Score Paired Differences Q-Q Plot
Sample Quantiles
Theoretical Quantiles
40
20
0
-20
-40
-2
-1
0
1
2



N2 Score Paired Differences Q-Q Plot
Sample Quantiles
Theoretical Quantiles
20
10
0
-10
-20
-2
-1
0
1
2