GAME-BASED ETHICAL INSTRUCTION IN UNDERGRADUATE ENGINEERING

COLTER MOOS
UNIVERSITY OF CONNECTICUT NEAG SCHOOL OF EDUCATION

LAUREN P. DOUGHER
UNIVERSITY OF CONNECTICUT NEAG SCHOOL OF EDUCATION

LANDON BASSETT
UNIVERSITY OF CONNECTICUT SCHOOL OF ENGINEERING

DR. MICHAEL F. YOUNG
UNIVERSITY OF CONNECTICUT NEAG SCHOOL OF EDUCATION

DR. DANIEL BURKEY
UNIVERSITY OF CONNECTICUT SCHOOL OF ENGINEERING

ANALYSIS

Ethics is a topic in undergraduate engineering curricula believed to frequently fall short in terms of professional application. In this paper we examine game-based learning approaches to engineering ethics in first year undergraduate students evaluated through the Engineering Ethical Reasoning Instrument (EERI). The EERI builds on the standard engineering ethics instrument, the DIT-2, and is focused on moral reasoning, while the game-based learning is based on situated decision making. Results showed no statistically significant interaction between participants who had game-based learning and lecture-based instruction over the time of the intervention, though ethical reasoning improved across all treatments. One conclusion we draw is that the lack of statistically significant interaction suggests that the EERI may not be the correct measure for these engineering ethics games, and further studies should research the creation of new instruments to incorporate this type of ethics instruction.

Keywords: Engineering, ethics, game-based learning, situated decision making, moral reasoning, assessment

GAME-BASED ETHICAL INSTRUCTION IN UNDERGRADUATE ENGINEERING

When visiting a typical undergraduate engineering classroom, one may expect to find a professor lecturing at the front of the room about formulae, equations, and theories in front of a lecture hall full of students. While these concepts are an important part of future engineers' education, so too is ethics, especially when you consider decisions that many engineers will be faced with when working in the field. In the United States, the Accreditation Board for Engineering and Technology (2022) guidelines for 2022-2023 require that students have: "an ability to recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts" (Criterion 3: Student outcomes). In other words, engineers not only need to understand how to create within their field, but they must also be able to consider the implications of their work.

MORAL REASONING AND SITUATED DECISION MAKING

In this paper, we focus on two key terms in relation to ethics: moral reasoning and situated decision making. Moral reasoning is based on how people think about what is right and wrong, such as whether they prioritize self-interest, or look toward justice for all people (Kohlberg & Hersh, 1977). We define situated decision making as grounded in how contextual factors, such as authentic details, and people's past experiences in dealing with ethical problems, influence their current decision making (Bagdasarov et al., 2013). People respond and act differently when immersed in a situation as opposed to when presented with more "disengaged" scenarios. In the context of this study, we predict that an engineer will react and respond differently to ethical dilemmas at their workplace than to a predetermined scenario in compliance training or a classroom. Their moral reasoning, or how they think about what is right and wrong, may be the same in both contexts, but the decision they make may differ as a result of the aforementioned contextual factors.

ETHICS AND ENGINEERING

Ethics can be an obscure topic to teach and difficult for many students to learn. As a result, some undergraduate engineering programs have employed game-based learning toward enhancing student interest and achievement. While game-based learning can be defined in a variety of ways, most definitions agree that game-based learning is a type of gameplay with defined learning outcomes (Shaffer et al., 2005, as cited in Plass et al., 2015). Studies on game-based learning have yielded varying results, including significant improvement in learning outcomes for participants who underwent the learning intervention (Chee & Tan, 2012; Franciosi, 2017) as well as non-significant differences between groups' learning outcomes, specifically in content knowledge and critical thinking skills (Cicchino, 2015). Despite these mixed results, all of these authors see the utility of game-based learning and call on others in the field to continue these efforts and associated research.

In doing so, we follow the recommendations of Marklund and Taylor (2016) who call on researchers to ensure that the game-based learning being utilized aligns with best practices and is implemented by those with expertise in this technique. Specifically, the professor was trained to utilize a variety of targeted games they constructed collaboratively with co-authors of this study who have expertise in game-based learning. As such, we could place greater focus on the learning outcomes from the game-based learning, rather than their implementation. Additionally, this approach allowed us to view game-based learning in the niche, yet expansive, topic of engineering ethics for insight into this game-based approach. If successful, this approach would have great implications for engineering schools and, in turn, engineering practice across the nation.

Our study was guided by the following research question: How did various arrangements of three engineering ethics games and modalities of play across three years differ from lecture-based instruction in their effect on students' ethics?

With a better understanding of the key terms used in this study and our research question, we next review some broader topics of interest, including engineering ethics and game-based learning. Next, our methods explore the participants, the three games used in the intervention, the lectures, as well as the outcome tool, the Engineering Ethical Response Inventory, or EERI. Using the EERI output, we will discuss the analysis and results of the two-way mixed-design ANOVAs across each year. We conclude with a discussion of our results and final conclusions regarding the role and importance of assessment in game-based ethical instruction.

LITERATURE REVIEW

To set the context for this study, we begin with a broad discussion of ethics education. We then focus on what ethics looks like in engineering education to better understand how lecture and game-based instruction are leveraged in this study. Next, we will consider "games" throughout the literature and an examination of what other studies have found when using games in the engineering ethics space.

In Western culture, many undergraduate universities and colleges offer philosophy courses focused on teaching philosophical ethics. These courses provide a theoretical background of topics such as virtue ethics, feminist ethics, and consequentialist ethics, among others (Paulson & Kretz, 2018). As Callahan (1980) explains, such courses should be included in undergraduate curriculum because "morality is part of any reflective personal life, and because ethical perspectives and specific moral rules are part of any cultural and civic life… ethical problems are inescapable…it is difficult to think of any aspect of personal or public life that will not be determined or conditioned by moral values" (p. 62). For students, such courses, at the very least, can provide clear evidence that there are ethical problems in all aspects of life, and how one understands and responds to these problems can make a difference in the lives of others, both positively and negatively (Callahan, 1980).

Ethics education at this stage is appropriate because ethical problems can happen at any time in a person's life and are also a part of jobs and professions (Callahan, 1980). For this reason, not only do we see introductory philosophy/ethics education as important in undergraduate studies, but also as an integral part of many different disciplines' education. Some examples of these fields include counseling (Lamb, 1991), psychology (Plante & Pistoresi, 2017), and medicine (Wong et al., 2022). Engineering is yet another example of a field heavily intertwined with ethics, especially with the reach of its many branches including chemical, civil, electrical, industrial, mechanical, and so on. Considering a slice of ethics education, via the engineering discipline, may help improve ethics and engineering education broadly.

History shows us what can happen when ethical problems are encountered but poorly attended to, such as with the case Challenger disaster in 1986 where seven space shuttle members died due to known flaws in the spacecraft ("Space Shuttle Challenger Disaster", n.d.). But in engineering there is still a "disconnect between the ethics education of contemporary engineering students…and the ethics realities of contemporary engineering practice" (McGinn, 2018, p. 3). Some students and educators may even consider ethics as a simple box to check off rather than an essential part of the engineering design process (Lloyd & van de Poel, 2008).

To help curb these deficiencies some institutions have turned to game-based learning for their engineering ethics education, with Barab and Dede (2007) noting that, over the past few decades, game-based learning methodologies emerged as a type of curricula in science education. Lau et al. (2012) provided one example of game implementation in the engineering classroom. In this study, students were tasked with creating a design using colored paper in accordance with a list of constraints. Each group of students was set up with roles and could not talk to one another about their constraints. They were then given a "briefing" shortly before the end of the session that posed an ethical dilemma. The different roles within each team were then pressured in different ways to either move forward with the product or halt it due to concerns regarding its ability to function properly. The scenario was created in such a way as to mimic the Challenger disaster and the decisions/ lack of communication that led to it (Lau et al., 2012). Between the implementation of the game and viewing of the real Challenger disaster video, students' opinions on the importance of ethical statements in engineering practice increased for all but one statement (Lau et al., 2012).

The Challenger disaster case study is also often used in traditional engineering ethics instruction during lectures. Some methods used in this teaching format include providing students with "do's and don'ts" lists related to ethical engineering practice, having students use basic scenarios to apply said principles, and the use of case studies where students must analyze and provide an approach for resolving the case (Alfred & Chung, 2012). The latter is the most effective approach used in classroom lectures (Whitbeck, 1996, as cited in Alfred & Chung, 2012), but Drew (2011) described how engineering courses taught through lectures often fail to interest students, leading to reduced engagement and shallow learning. Therefore, other approaches, like games, have been introduced to increase the usefulness and engagement of engineering ethics instruction.

The example above describes Lau et al.'s (2012) study as implementing a "game," but there are additional terms with more distinct meanings used in education. Games, gamification, and game-based learning are terms that are frequently used interchangeably in educational settings. While there is some disagreement on exact definitions, Plass

(2017) differentiated between them. Specifically, game-based learning is distinct in that the original learning task is transformed into a game with a design grounded in discipline-specific applications (Plass, 2017). Bodnar et al. (2016) went into further detail, defining gamification as "the application of game design elements to nongame scenarios" (p.148). A common and simple form of gamification is the use of points, badges, and leaderboards where participants earn points for completing tasks (Bodnar et al., 2016). Further, they continued to explain that game-based learning has many benefits including that it provides immediate feedback, informs participants they are making progress, and motivates them (Bodnar et al., 2016). Bodnar et al.'s (2016) systematic review of the games meant to teach undergraduate engineering students, provided support for the conclusion that the implementation of games in undergraduate engineering classrooms improved student learning and attitudes.

For this paper, we define game-based learning as it aligns with Plass (2017) and Shaffer et al. (2005) in that an original learning task has been transformed into a game with learning outcomes. In the original lecture-based instruction scenario at the study site, students would engage in discussions of historic engineering and philosophical ethics problems, with the learning task being that they would contribute to the discussion about, and listen to, the problems. For the larger National Science Foundation (NSF) study from which this paper is derived, three games were created to transform the role of the student in different ways, as an individual (1) voting on an ethical response to a problem, (2) choosing a potentially ethically dubious card option due to the nature of the situation, or (3) ordering other engineers' views on ethical responses to issues in the field. All these games stemmed from the same situated engineering ethics problems of the lecture, but had the player take on a role with richer context and details, more agency, and from a different perspective on the ethical problem, leading to the varied learning outcomes from each game.

Slota and Young (2014) described the importance of implementing game-based learning beyond simply taking a lesson and adding game elements like rewards or points. Rather, incorporating principles where games can change to sustain player interest, have game narratives, and include opportunities for players to explore, expand, or build within the game, can result in a more effective game-based learning environment (Slota & Young, 2014). These game elements lead to more richly situated game-based learning environments, and the need for similarly richly situated assessments of engineering ethics. It is with this idea of richly detailed, and highly contextual, games in mind that we circle back to the current study, where we aim to examine whether there are differences between such engineering ethics games and more traditional scenario-based lectures.

METHODS

As part of a larger NSF-funded study on the use of researcher-created engineering ethics games on situated decision making, data was collected over the course of three years from an undergraduate Introduction to Engineering course at a public northeastern state research university. Throughout this study, various games were employed through different modalities across each of the three years of the study. The games played by the experimental groups included Mars: An Ethical Expedition (MAEE), Cards Against

Engineering Ethics (CAEE), and Toxic Workplaces (TW), with more detailed descriptions of each game below. Some key commonalities during the three-year period: the course always ran during spring semester and had no other connection to ethics, every year the Tuesday section had game-based learning and the Thursday section had lecture-based classroom instruction. These lectures occurred during the two full days set aside for CAEE, and later TW, in the middle of the semester. The lecture-based group had no substitute for MAEE, only receiving two lectures; first on the Challenger disaster discussed earlier, and second on the classic philosophical ethics trolley problem. In this problem, a bystander observes that a trolley will move forward and kill 5 people unless the bystander switches the track away, the result being that one person on the other track is killed. During each year of this study, participants took the Engineering Ethical Reasoning Instrument (EERI), our outcome measure, pre- and post- intervention.

Outside of these commonalities, the implementation differed across the years, both in terms of which games students played and via which modality (see Table 1 for further explanation). In 2019, MAEE and CAEE were run in person, with CAEE played for two weeks in the middle of the semester, while TW was not played. For 2020, MAEE was run half in person for the first six weeks, and half online for the last six weeks due to the COVID-19 pandemic. CAEE was not played in 2020, instead, for the first week, the game-based group received the same lecture as the lecture-based group. For what would be the second CAEE session, the game-based group instead discussed the prompts for TW used to generate ranked ethical choices for the responses to case studies in the game as will be discussed later. All three games were played online in 2021, with the COVID-19 pandemic still interfering with in person instruction.

PARTICIPANTS

Each year participants were students enrolled in a shared Introduction to Engineering class at a public northeastern state research university in the spring semester of their first year, a course required for all pursuing a bachelor's in engineering. The following guidelines were used as exclusion criteria for analysis: did not provide consent, did not answer all the questions for either the pre or post EERI, failure to complete either a pre or post EERI, or if a student reported they switched conditions (game vs. lecture) between the pre and post. Additionally, anyone who completed the EERI under 10 minutes was excluded as they were believed to have not taken enough time to review and respond to questions based on the EERI length. Last, if a participant submitted multiple times only their first complete submission was used and the rest excluded.

As seen in Table 2, after these factors for exclusion were implemented, there were roughly similar sample sizes across all conditions except the lecture-based instruction in 2021. We believe this difference was not due to the elimination criterion above, as the individuals removed were roughly equal across the classes, but instead due to participant self-selection. As the lecture-based instruction class was always held on Thursday, and the class of 2021 was fully online due to the COVID-19 pandemic, we believe that fewer people enrolled in that class for unknown personal reasons, resulting in an uneven sample size which we account for later.

Table 2 highlights the percentages of students across gender identity and race/ethnicity. Across all years, individuals who identified as male were predominant within each condition, while those identifying as Caucasian followed by Asian American or Pacific Islander were the highest and second-highest percentages across every condition, respectively, aligning with the larger demographics of engineering undergraduates (American Society for Engineering Education, 2022). From this, we felt comfortable comparing the participants across conditions within their respective years as there were no statistically significant differences.

GAMES

As part of the aforementioned NSF funded study, the research team designed the following three engineering ethics games with multiple rounds of playtesting and feedback before incorporation into the study, to teach students about engineering ethics through gameplay. Mars: An Ethical Expedition and Cards Against Engineering Ethics were created by some of the authors on this paper, while Toxic Workplaces was created by researchers at another institution working under this NSF study and supplemented and tested by some authors on this paper.

MARS: AN ETHICAL EXPEDITION

In this multiplayer choose-your-own-adventure game, the students collaborated as a class to make ethical decisions to survive on a Mars Colony. At the beginning of the semester the students were given a narrative introduction to the overall game, and then each week consisted of a specific ethical scenario. For each weekly scenario, the students would have a choice of two or three decisions they could choose from in handling the ethical situation. The students then voted on which decision they wanted to make, and the choice with the most votes was taken. Based upon their choice the class received a different situation prompt the next week, continuing for 12 weeks with branching events based on their choices. At the end of the 12 weeks, the students would learn whether they were successful in surviving/saving the Mars Colony.

The learning outcome from MAEE is for students to be able to take various perspectives in making situated ethical engineering decisions. Through playing the game they needed to vote and make an ethical decision. The goal is for them to make this decision with consideration of the impact their decisions would have on the Mars Colony from various characters' perspectives and determine not only whether those impacts and decisions are ethical, but which they would advocate for through their vote.

When played in person, the professor would stand at the front of the class and share their screen with the narrative introduction and weekly scenario as appropriate, while the students would use polling software on an electronic device to vote on the decision they personally would make. When played online the professor would similarly share their screen in the video call and collect votes from the polling software using the same method. In both cases, after the votes were collected, the professor would share their screen with the results of the vote and then continue with the rest of the lesson for that week unrelated to the gameplay.

Implementation of Games and Lectures Across Years

Table 1

Implementation of Games and Lectures Across Years

	2019		2020		2021	
	Games	Lecture	Games	Lecture	Games	Lectur e
Lecture Modality	N/A	In Person	Online	Online	N/A	Online
Number of Lectures	0	2	1	2	0	2
MAEE Modality	In Person	N/A	Hybrid	N/A	Online	N/A
CAEE Modality	In Person	N/A	N/A	N/A	Online	N/A
TW Modality	N/A	N/A	Prompt Online	N/A	Online	N/A

Note. For TW Modality in 2020 "Prompt - Online" is used because while the students did not play Toxic Workplaces, they did discuss the prompts online to generate the ranked ethical choices, which differs from the Lecture condition, and the play in 2021.

CARDS AGAINST ENGINEERING ETHICS

This game is based on the similar party games Apples to Apples and Cards Against Humanity, in which players match a card with a word or phrase on it in their hand to a shared prompt. If a player's card is chosen as the best fit by the judge of that round, the player earns points. In this version, cards are focused on situations that can occur as an engineer, with some prompts and phrase cards based on real-world engineering ethics situations. Students played this game in groups of four, with the judge rotating between the players throughout their play. After 30 to 45 minutes the groups stopped playing and the whole class discussed some of the real-world ethical situations on the cards and considered their goals in playing the game. After this discussion, gameplay resumed for another 30 to 45 minutes before ending with a debrief on the decisions players made and a discussion of what they would do in similar real-life circumstances.

The learning outcome from CAEE is for players to recognize how the context of a situation modifies their situated decision making. Through both playing combinations of cards that are unethical and discussing why the context of playing the game allowed them to play such unethical combinations, students would be able to better recognize the importance of context in modifying their situated decision making. In this case, the learning came not from the playing of CAEE directly, but from the discussions and analysis of how the students were playing the game.

When played in person, the players would use physical playing cards and gather in groups with those sitting near them. When played online, the players would be randomly grouped together with others in their class through video conference breakout groups to play a digital version of the game. The virtual card deck contains nearly all the cards from the physical version, with some omissions due to copyright purposes. Therefore, the largest changes between the in person and virtual implementations were the use of slightly fewer cards in the digital version, the randomization of groups, and using video conferencing to communicate in which not all players would turn on their cameras or

DEMOGRAPHICS BREAKDOWN BY YEAR AND CONDITION

Table 2

Demographics Breakdown by Year and Condition

Demographics	2019		2020		2021	
	Games	Lecture	Games	Lecture	Games	Lectur e
Sample Size	128	138	125	120	108	52
Gender Identity						
% Male	65.6	58.7	71.2	65	53.7	53.8
% Female	34.4	40.6	28.8	35	45.4	46.2
% Other	0	0.8	0	0	0	0
% Prefer Not To Answer	0	0	0	0	0.9	0
Race/Ethnicity						
% African American	3.9	4.3	3.2	3.3	6.5	3.8
% AAPI	10.2	18.8	11.2	15	16.7	11.5
% Hispanic	3.9	2.2	10.4	9.2	5.6	3.8
% Native American	0	0	0	0	0	0
% Caucasian	71.9	62.3	64	64.2	57.4	61.5
% Other	0.8	1.4	1.6	0.8	0	0
% Prefer Not to Answer	0	0	0	0	1.9	7.7
% Mixed	9.4	10.9	9.6	7.5	12	11.5

Note. The percentages in this table are all relative to their condition within their year. For example, the 2019 Games column has a sample size of 128 students, of which 65.6 identify as male, and 34.4 identify as female. This same interpretation applies to race/ethnicity.

keep their microphones open. While it is possible that those seated near each other are strangers, there is a higher chance that players would know at least one or two others in their group of four, while in the online random groups it is less likely that the players were familiar with each other.

TOXIC WORKPLACES

Toxic Workplaces uses a Family Feud style of play to discuss situated decision making. In this version however, instead of students blindly guessing what responses others have made, they are presented with responses to an ethical case study and have to rank order those responses based on their belief of how many people would choose that course of action. The response prompts were created by students from another university who voted both on how ethical each response was for each prompt, and how many of them would choose that course of action in the case study. This survey data was collected in 2020 and used to create the correct rank order for play in 2021. The gameplay occurred instead of the second week of CAEE.

The learning outcome from Toxic Workplaces is for players to be able to recognize how others within the engineering community perceive engineering ethics. This is achieved through not only the students' ranked ordering, but also through their scores which would be higher the more accurately their order aligned with the rank order of their peers from the previous year. Through this, they were directly able to get feedback on their alignment with how ethical others within the engineering community perceive various choices.

As this game was only played online, there was one implementation method used which utilized video conferencing and breakout groups, and Google Slides as the basis for playing the game. As mentioned before, after reviewing the case study players would discuss amongst themselves how to rank order the responses by dragging the response slides into the order from what they thought was most popular to least. After the responses were ordered, everyone in the group would take a screenshot of the order and record their choices for the end of the game. Groups would then be formed randomly two more times for a total of three rounds of play before scoring would occur. For every response in the correct rank order, the group would score a point. Individuals won if their teams scored the most points correctly throughout the three rounds.

LECTURES

Students in the lecture-based course received two lectures on ethics, one each over the course of two classes. In the first lecture, the professor would discuss the Challenger Shuttle disaster, situating the ethical problem for students in the engineers' knowledge about the potential explosion, and their decisions around moving forward with the shuttle launch anyway.

After going through the context of the ethical problem, the professor then led a discussion about alternative approaches the engineers could have taken and the ethics behind them, as well as the potential impacts of those decisions. The learning objective from this lecture was to have students take the perspective of an engineer in that ethical situation and consider what they might have done differently.

For the second lecture, the professor discussed the philosophical trolley problem described earlier. After setting this premise, the lecturer opened the class up for discussion about what they would do and the ethical implications of these actions. As the students presented different solutions, the professor lectured on the impact of different choices on others. The learning objective was to have students take the perspective of someone who had only bad options and determine what they would choose to do in an ethically complicated situation.

OUTCOME TOOL: ENGINEERING ETHICAL REASONINGS INSTRUMENT

Educators at Purdue University and the Illinois Institute of Technology recognized the gap between engineering ethics education and students' ability to apply these principles in practice. Together, they created an instrument for "individual ethical decision-making in a project-based design" (e.g., team-based, problem-solving) called the Engineering Ethical Reasoning Instrument (EERI) based off the leading ethics assessment of the time, the Defining Issues Test-2 (DIT-2) (Zhu et al., 2014). This instrument was chosen for the present study due to the DIT-2 being the standard ethics assessment in the engineering ethics literature, and the EERI improving upon the DIT-2 through situating the scenarios in engineering.

In the EERI the participants were given six ethically complex scenarios, asked to decide on performing an ethically questionable action, and were then assessed on how important each factor was to them, and which factors were most important, in making their decision. In one such scenario, the participant's student project team is tasked with rating the overall quality of buildings in an impoverished section of the community to improve the quality of housing, some of which are thought to be unsafe to live in. The participant has good reason to believe that the data will be used to raze the least safe of those buildings, forcing residents to relocate with some residents stating that they may end up homeless if their residence is demolished. Thus, the ethical dilemma is to 1) report the unsafe dwellings so that people aren't exposed to the danger of the building but possibly leaving them unhoused or 2) not report the dwellings accurately, allowing residents to stay in their unsafe homes but off the streets. The participant must then decide whether to rate the homes, and then rate the importance of factors that could influence the decision. Some factors included whether the participant was friends with an affected resident, whether it could help the participant's future career, whether the unintended social ramifications of this work should be a concern, etc. Then after the importance of these factors were chosen, the participant would rank the four factors that were most important to them from most important to fourth most important.

From these scenarios, the EERI produces two scores: P and N2. The P score represents the extent to which participants' four most important factors were based on looking beyond self-interest (e.g., whether the unintended social ramifications of this work should be a concern). The N2 score is statistically derived from the P score, while also factoring in whether participants' rated factors based on self-interest (e.g., whether it could help the participant's future career) as less important than factors based on looking beyond self-interest. For the purposes of this paper, we are focusing on the N2 score, as it is the more comprehensive and interpretable score from the EERI, and is derived from the P score.

ANALYSIS

For participants who remained after exclusion criteria, we distilled their EERI results into pre and post P, and pre and post N2 scores. We had already split these individuals into groups based on their year of taking the course and whether they were in the game- or lecture-based instruction version of the course. Given our question's focus on determining if the conditions significantly differ in each year, we conducted an interaction analysis to determine how they may differ from pre to post, and if there are other underlying factors such as their groups, or time, that lead to perceived significant pre-post changes. We used two-way mixed-design ANOVAs for each year independently to compare the pre and post EERI N2 scores.

Before conducting the two-way mixed-design ANOVA we also analyzed whether there were statistically significant differences first between the pre EERI N2 scores for each group in their respective year to ensure the groups were comparable at the outset. We found that, within each year, there was no statistically significant difference between the pre EERI N2 scores for the groups. In running the multiple two-way mixed-design ANOVAs not all assumption tests on normality indicated a normal distribution. The smallest sample size of a potentially non-normal distribution was 108 which can cause the assumption tests to be less accurate, as was the case here upon visual inspection of the distributions. For the assumption tests on the homogeneity of variance, there were multiple instances where the variances were below 0.05, but above 0.01. Thus, we adopted 0.01 as our threshold for these tests, as in all cases where the samples were identical or nearly identical, Hartley's Test resulted in a variance ratio under 1:2. The assumption tests identified a few outliers, however, removing them did not significantly change anything so they were kept in the model to reflect the data more accurately. Last, there was no assumption of sphericity as the within-subject factor of time only had two levels, pre, and post. From the above, we believe the two-way mixed-design ANOVAs used to analyze the EERI N2 data can be interpreted normally.

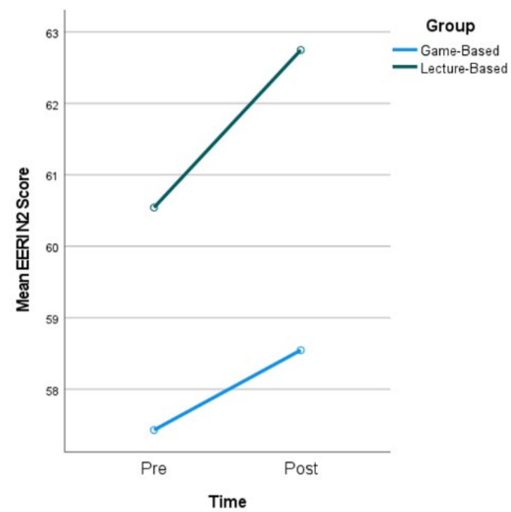
RESULTS

2019 RESULTS

In 2019, there were no statistically significant effects for the N2 score, though multiple effects neared significance. Specifically, our main interest was in the lack of statistically significant interaction effect for the N2 score ($F(1, 264) = .358$, $p = .550$, $\eta^2 = .001$), suggesting that students experienced equivalent changes in their moral reasoning regardless of treatment group. Figure 1 shows how the slopes of the lines do not significantly differ, supporting the non-significant interaction between group and time.

Figure 1

Estimated Marginal Means of EERI N2 for 2019

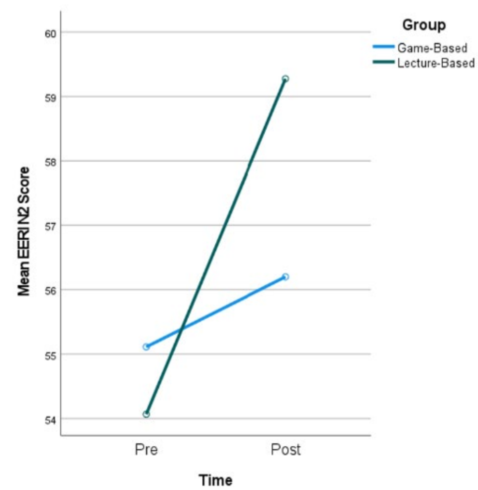


2020 RESULTS

The results of the two-way mixed-design ANOVA for 2020 indicate that the only finding of statistical significance was the main effect difference between the pre and post-EERI N2 score ($F(1, 243) = 8.453$, $p = .004$, $\eta^2 = .034$). Thus, in this case, we can understand students in either condition ended their time in the course with a significantly higher EERI N2 score. Figure 2 highlights the direction of this change, as in both groups there was a positive difference

Figure 2

Estimated Marginal Means of EERI N2 for 2020



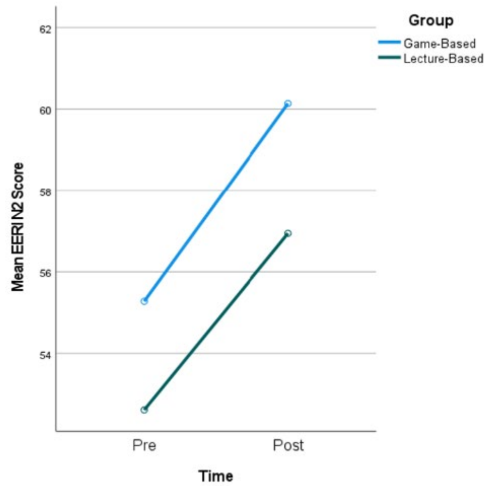
in their EERI N2 score from pre to post. We again see that the interaction between time and group was not statistically significant for the N2 score ($F(1, 243) = 3.620$, $p = .058$, $\eta^2 = .015$). This continues the trend that, regardless of the condition participants were in, there was an increase in N2 score. Figure 2 also highlights how the slopes of the lines do not significantly differ with the proportional difference over time in EERI N2 score being similar across groups.

2021 RESULTS

Again, the only finding of statistical significance was for the main effect difference between the pre and post N2 EERI score ($F(1, 158) = 11.727$, $p = .001$, $\eta^2 = .069$). Figure 3 highlights this difference in EERI N2 scores across time in both groups. As well, for the third time, the interaction between time and group was not statistically significant for the N2 score ($F(1, 158) = .040$, $p = .842$, $\eta^2 = .000$), such that students' scores were proportional over time regardless of treatment. Figure 3 also highlights how the slopes of the lines do not significantly differ, supporting the non-significant interaction between group and time.

Figure 3

Estimated Marginal Means of EERI N2 for 2021



DISCUSSION

From the results above it is clear that, as assessed through the EERI, there is a lack of evidence that within any year, the game-based learning condition significantly differed from the lecture-based instruction. This is a potentially encouraging finding in its own right, as a lack of significant difference between conditions, but significant increases in N2 scores over time in 2020 and 2021 may indicate that game-based learning is as impactful as traditional lectures. While other analyses should be performed to support this claim, there are other possible explanations for these results. For instance, it is possible that other shared elements of this engineering class led to increases over time, despite there being no other direct discussion or assessment of ethics in the course. Regardless of this positive change in moral reasoning over time, the focus of this study was on whether game-based learning and lecture-based instruction differed in their effect on ethics. Given this focus, we are more interested in how such drastically different methodologies resulted in no statistically significant difference.

There are multiple possible explanations for this lack of significant difference between the ethics games and traditional lectures. One explanation is that both interventions were beneficial themselves and simply did not differ significantly in how effective they were. Another reason for these findings could be that the lectures were particularly excellent, and that other types of ethics lectures by another professor would not have these same results. We can also speculate that the games may require more time to engage students than anticipated, so the games may have been less effective than theoretically possible due to constraints such as class length, time to get students settled and focused, and other classroom priorities. Lastly, it is possible that switching the modalities (in person vs. online) and games played between years, and the effects of COVID-19, did not allow the instructors and students to engage with the games to the best extent possible. All of these possible explanations are worth further investigation, however, we believe there is another explanation that we view as the most compelling.

We believe that the best possible explanation for no statistically significant interaction terms lies with a potential misalignment between the EERI's measure of ethics through moral reasoning and the ethics games focus on exploring students' situated decision making. The EERI is focused solely on the underlying moral reasons, defined as their thoughts on right and wrong. The ethics games were focused on situated decision making, which is grounded in contextual factors, authentic details, and people's past experiences. This explanation is not rooted in data from the EERI, but in recognizing how the games were designed based on situated decision making, in contrast to the moral reasoning design of the EERI. These findings of no significant difference between the forms of pedagogy indicate to us that the difference between the games' focus on situated decision making as distinct from the lectures, was not captured by the EERI's measure of moral reasoning.

As the EERI is focused solely on the underlying moral reasons that students find most important (e.g,. self-interest versus justice), games seeking to change this underlying moral reasoning would potentially show some difference from traditional lectures as assessed through the EERI. For games that differ from lectures in their authentic, specific, and contextually rich scenarios, like the games used in this study, the EERI is unlikely to detect differences between the lecture and game-based pedagogies due to the EERI's format. The EERI's scenarios are antithetical to those from our games and therefore cannot accurately detect the changes in students' ethics. For these types of games, such as the ones in this study, research needs to be conducted to create new instruments that are targeted to assess these principles, rather than try to apply existing ethics measures that are not aligned.

CONCLUSION

Given this believed misalignment between the EERI's measurement of moral reasoning and the situated decision-making design of the engineering ethics games used in this study, some questions remain; how should assessments of ethics be designed to better capture situated decision making, and how can these assessments of ethics be better aligned with the interventions on ethics instruction? These are questions for both further designers of assessment and for researchers in choosing which assessments are the most appropriate for your research questions when studying situated decision making, or moral reasoning, in engineering instruction. There is likely no single answer to any of these, as various assessments are built from different moral and ethical philosophies and frameworks, and the various playful and other interventions being studied may be designed incongruously. This result occurred within this study, as the EERI was the best measure we could choose, being based on the standard ethics assessment in the literature and further focused on engineering, but was not the correct measure for our engineering ethics games. Recognizing the importance of varied assessments of situated decision making and moral reasoning in engineering, how the design of ethics games aligns with these theories, and the impact alignment has on measuring the impact of ethics games on situated decision making, can generate further research on how ethics games may help shape the ethics of future engineers.

LIMITATIONS

There are several limitations to this study, most of which have been discussed already: the uneven sample size in 2021, and the COVID-19 pandemic occurring in the middle of 2020. The uneven sample sizes in 2021 were a minor limitation in the analysis of the EERI P and N2 scores for that year. While the analysis was conducted due to the use of assumption tests that account for uneven sample sizes, it is still limiting that one group had slightly over twice the number of participants as the other. It is impossible to control which section students sign up to take and whether they properly finish and take seriously the EERI, there could be better incentive structures increasing retention so that more students complete the EERI both times.

The other limitation of note is the COVID-19 pandemic that occurred halfway through 2020 and fundamentally changed the instruction in both conditions for 2020 and 2021, also causing a change in the interpretations. It is impossible to control major pandemics and their impact on students, all that can be done in these cases is to acknowledge their presence, and how they may impact interpretations. In this case, it is possible that some of the non-significant results such as interactions could have been the impact of the pandemic. While the findings are interpreted as the result of a misalignment between the EERI and games, it is possible that for 2020 and 2021, there were effects of the pandemic that impacted these results, unrelated to the problem of alignment.

ACKNOWLEDGMENTS

This research is being funded by the National Science Foundation, "Collaborative Research: Learning Engineering Ethics Through High-Impact Collaborative and Competitive Scenarios" (IUSE – 1934702).

REFERENCES

Accreditation Board for Engineering and Technology. (2022). Criteria for accrediting engineering programs, 2022-2023. ABET. https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2022-2023/

Alfred, M., & Chung, C. A. (2012). Design, development, and evaluation of a second generation interactive Simulator for Engineering Ethics Education (SEEE2). Science and engineering ethics, 18(4), 689-697.

American Society for Engineering Education. (2022). Profiles of Engineering and Engineering Technology, 2021. https://ira.asee.org/wp-content/uploads/2022/11/Engineering-and-Engineering-Technology-by-the-Numbers-2021.pdf

Bagdasarov, Z., Thiel, C. E., Johnson, J. F., Connelly, S., Harkrider, L. N., Devenport, L. D., & Mumford, M. D. (2013). Case-based ethics instruction: The influence of contextual and individual factors in case content on ethical decision-making. Science and Engineering Ethics, 19(3), 1305-1322. https://doi.org/10.1007/s11948-012-9414-3

Barab, S., & Dede, C. (2007). Games and Immersive Participatory Simulations for Science Education: An Emerging Type of Curricula. Journal of Science Education and Technology, 16(1), 1–3. https://doi.org/10.1007/s10956-007-9043-9

Bodnar, C. A., Anastasio, D., Enszer, J. A., & Burkey, D. D. (2016). Engineers at play: Games as teaching tools for undergraduate engineering students. Journal of Engineering Education, 105(1), 147-200.

Callahan, D. (1980). Goals in the teaching of ethics. In: Callahan, D., Bok, S. (Eds.) Ethics Teaching in Higher Education. (pp. 61-80). Springer, Boston, MA. https://doi.org/10.1007/978-1-4613-3138-4_2

Chee, Y. S., & Tan, K. C. (2012). Becoming chemists through game-based inquiry learning: The case of legends of Alkhimia. Electronic Journal of e-Learning, 10(2), 185-198.

Cicchino, M. I. (2015). Using game-based learning to foster critical thinking in student discourse. Interdisciplinary Journal of Problem-Based Learning, 9(2). http://dx.doi.org/10.7771/1541-5015.1481

Drew, C. (2011, November 4). Why science majors change their minds. New York Times. http://www.nytimes.com/2011/11/06/education/edlife/why-science-majors-change-theirmind-its-just-so-darn-hard.html

Franciosi, S. J. (2017). The effect of computer game-based learning on FL vocabulary transferability. Educational Technology & Society, 20(1), 123-133.

Kohlberg, L., & Hersh, R. H. (1977). Moral development: A review of the theory. Theory Into Practice, 16(2), 53–59. https://doi.org/10.1080/00405847709542675

Lamb, C. S. (1991). Teaching professional ethics to undergraduate counseling students.

Psychological Reports, 69(3_suppl), 1215–1223. https://doi.org/10.2466/pr0.1991.69.3f.1215

Lau, S. W., Tan, T. P. L., & Goh, S. M. (2012). Teaching engineering ethics using BLOCKS game. Science and Engineering Ethics, 19(3), 1357–1373. https://doi.org/10.1007/s11948-012-9406-3

Lloyd, P., & van de Poel, I. (2008). Designing games to teach ethics. Science and Engineering Ethics, (14), 433–447. https://doi.org/10.1007/s11948-008-9077-2

Marklund, B. B., & Taylor, A. A. (2016). Educational Games in Practice: The challenges involved in conducting a game-based curriculum. The Electronic Journal of e-Learning, 14(2), 122-135.

McGinn, R. (2018). The ethical engineer: Contemporary concepts and cases. Princeton University Press.

Paulson, J., & Kretz, L. (2018). Exploring the potential contributions of mindfulness and compassion-based practices for enhancing the teaching of undergraduate ethics courses in philosophy. The Social Science Journal, 55(3), 323-331. https://doi.org/10.1016/j.soscij.2017.12.003

Plante, T., & Pistoresi, S. (2017). A survey of ethics training in undergraduate psychology programs at Jesuit universities. Pastoral Psychology, 66(3), 353–358. https://doi-org.ezproxy.lib.uconn.edu/10.1007/s11089-017-0755-3

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. Educational psychologist, 50(4), 258-283.

Plass, J. L. (2017, March 15). Gamification v. Game-based Learning v. Playful Learning. LinkedIn. https://www.linkedin.com/pulse/gamification-v-game-based-learning-playful-jan-l-plass/

Shaffer, D. W., Squire, K. R., Halverson, R., & Gee, J. P. (2005). Video games and the future of learning. Phi delta kappan, 87(2), 105-111.

Slota, S. T., Young, M. F. (2014). Think games on the fly, not gamify: Issues in game-based learning research. Journal of Graduate Medical Education, 6(4), 628-630. https://doi-org.ezproxy.lib.uconn.edu/10.4300%2FJGME-D-14-00483.1

Space Shuttle Challenger Disaster. (2022, September 22). In Wikipedia. https://en.wikipedia.org/w/index.php?title=Space_Shuttle_Challenger_disaster&oldid=1111721759

Whitbeck, C. (1996). Ethics as design: Doing justice to moral problems, The Hastings Center Report, 26(3), 9-16. https://doi.org/10.2307/3527925

Wong, M. K., Hong, D. Z. H., Wu, J., Ting, J. J. Q., Goh, J. L., Ong, Z. Y., Toh, R. Q. E., Chiang, C. L. L., Ng, C. W. H., Ng, J. C. K., Cheong, C. W. S., Tay, K. T., Tan, L. H. S., Ong, Y. T., Chiam, M., Chin, A. M. C., Mason, S., & Radha Krishna, L. K. (2022). A systematic scoping review of undergraduate medical ethics education programs from 1990 to 2020. Medical Teacher, 44(2), 167-186. https://doi.org/10.1080/0142159X.2021.1970729

Zhu, Q., Zoltowski, C. B., Feister, M. K., Buzzanell, P. M., Oakes, W. C., & Mead, A. D. (2014, June), The Development of an Instrument for Assessing Individual Ethical Decisionmaking in Project-based design teams: Integrating quantitative and qualitative methods Paper presented at 2014 ASEE Annual Conference & Exposition, Indianapolis, Indiana. https://doi.org/10.18260/1-2--23130