Debunking the Fallacy of the Individual Decision-maker: An Experiential Pedagogy for Sustainability Ethics

Thomas P. Seager, Evan Selinger, Daniel Whiddon, David Schwartz, Susan Spierre, Andrew Berardy

***Abstract* – Existing pedagogical approaches to ethics education in engineering and science reinforce what this paper terms "the fallacy of the individual decision-maker" by suggesting an oversimplified, individualistic model of ethical decision-making, rather than recognizing the organizational, cultural, or group deliberative context of an ethical dilemma. Consequently, students fail to develop the group deliberative and ethical reasoning skills necessary to properly recognize and resolve ethical questions. This paper critiques existing approaches and presents an alternative pedagogy that emphasizes active, participatory, and experiential learning that is intended to more deeply immerse students in questions of fairness, justice, and equity in the context of sustainability by playing the *Externalities Game.* Preliminary testing supports the hypotheses that game play results in deeper consideration of ethical issues, more emotionally engaged students, fosters greater deliberative discourse, and encourages experimentation with different ethical strategies. The *Externalities Game* may be an appropriate piece of a larger course in sustainability ethics when combined with traditional reading and pedagogical strategies.**

***Index Terms*— Educational Games, Environmental Ethics, Sustainability. Moral Philosophy**

I. INTRODUCTION

WHEN ethics is formally introduced in graduate science & engineering curricula, it is typically through an individualistic, professional ethics approach that supplements indoctrination to rules of conduct with case study readings about the catastrophic consequences of ethical "failures." Although these case studies are often abstracted from real events, they fall short of cultivating an emotional resonance among students with limited real world experience, and give the misleading impression that impropriety occurs only when *bad apples* act without due regard for others. Consequently, science & engineering students often feel disconnected from a traditional ethics pedagogy, which typically lacks the active, experimental and collaborative learning exercises that have been proven effective in other science & engineering subject areas. This is particularly problematic for questions of sustainability that demand a greater level of complexity and broader scope than professional ethics. Incorporation of *sustainable development* into the fundamental ethical canons of engineering has presented at least two serious challenges to science & engineering educators. First, sustainability is not reducible to a tractable definition that is amenable to traditional problem-solving strategies. Second, there is a paucity of exemplary course modules and/or case studies that illustrate how to put the principles of sustainability as an ethical concept *into practice*.

The dominant approaches to science and engineering ethics education share a common focus on the *proper conduct that individuals, defined as individual members of a profession, should exhibit.* Theorists have identified several limitations to this educational approach.

- Many of the dominant "catastrophe/disaster" case studies and thought experiments are narrated through "thin" descriptions that occlude relevant detail, especially context
- Students often do not identify with the people who are discussed in case studies and thought experiments. They overestimate their own moral fortitude, and perceive these case studies and thought experiments to concern people who are fundamentally not like them.
- Students often have trouble connecting case studies that focus on historical episodes of conflicts that occurred in the past with current and future cases of conflict.

As we see it, a consistent problem in ethics education for engineering and science students is that they remain within a *humanities horizon*. They are constrained by conventional pedagogical tools typical of the humanities: students read and discuss new books and articles, view new dramatic perfor-

This work was funded in part by National Science Foundation grant # 0932490 and the RIT Office of the Vice President for Research.

T.P. Seager is an Associate Professor of Sustainability at Rochester Institute of Technology, Rochester NY 14623 USA (585-475-6612; email: thomas.seager@rit.edu
E. Selinger is an Associate Professor of Philosophy at Rochester Institute of Technology, Rochester, NY 14623, USA (email: eselinger@rit.edu).
D. Whiddon is an undergraduate Interactive Games & Media major at Rochester Institute of Technology, Rochester NY 14623 USA (email: dlw4736@rit.edu).
D. Schwarz is an Assistant Professor of Interactive Games & Media at Rochester Institute of Technology, Rochester NY 14623 USA (email: disvks@rit.edu).
S. Spierre and A. Berardy are Ph.D. – Sustainability students at Rochester Institute of Technology, Rochester, NY 14623, USA (email: sgs3458@rit.edu).

mances (e.g., movies and plays), and write critical, individual essays. By changing only the learning objectives without substantially altering the learning activities, three problems remain.

1. Scientific and engineering students are placed in a learning environment that significantly differs from the experimental settings that typify the arenas where their major studies are conducted.
2. The material that is offered continues to place students in the roles of spectators and readers. Because spectators and readers are limited to discussing decisions that other people have made, students are not inclined to identify themselves as candidates for future ethical conflict. Moreover, it is difficult for them to place themselves empathetically in the metaphorical shoes of the people who are faced with ethically problematic choices.
3. Priority is not given to teaching students deliberative skills—skills that are crucial in a range of collaborative projects involving scientists and engineers.

To partly remedy these difficulties, we propose a game-based approach to ethics education in which, through repeated game play, students can experiment with different ethical strategies and reflect upon the outcomes in a manner that is analogous to teaching strategies that are already made familiar in science and engineering laboratory courses.

The Coase (1960) theorem states that when transaction costs are low and property rights are determined, optimal allocation of resources is achievable without government intervention. The polluting and damaged parties will negotiate a transfer of payments between them to either accept damage or reduce pollution on the basis of which was more profitable. The *Externalities Game* allows participants to experiment with no transaction cost, determined property rights, and no third party intervention and therefore test the Coase theorem.

II. HYPOTHESES

Educational games have previously been found effective at delivering technical content in a collaborative, engaging process that captures the attention of the Millennial Generation (Isaacs et al 2008, Mayo 2007). Games have long been used in a team setting in business education, although ethical implications are rarely direct learning objectives (Wolfe & Fritzsche 1998). The overarching goal is to create an active, deliberative, and experiential setting that has historically been absent from ethics education for science and engineering students. In so doing, we hope to make progress on the impasse that honest and reflective instructors readily admit in private conversation: the current approaches to teaching engineering and research ethics primarily enable universities to fulfill accreditation requirements, but fail to fully engage students in ethical thinking. Understood within this context, bored students who look for transformative experiences elsewhere are deemed acceptable sacrifices.

One of the critical advantages of a carefully designed game is that we can avoid what we term the *fallacy of the individual decision maker*. By framing sustainability problems in terms of non-cooperative game theory, in which the outcomes that impact each player are determined partly by the actions of other players, we position students in explicitly social settings that require coordination of decision processes to ensure group success.

We claim that the experiential, game-based pedagogy has at least four advantages that are lacking in the humanities alternatives. First, because students experiment with decision-making strategies in situations with unknown outcomes, they are cognitively oriented towards forwards-looking hypothesis formation, in contrast to the retrospective orientation that is to be expected when students judge decisions made by the people discussed in historical case studies or thought experiments. Students who engage in debate through the various game models will not only experience firsthand problems and theories discussed previously (such as in assigned readings), but will also have the opportunity to formulate and test new potential solutions to complex problems in a collaborative setting. Knowing that decisions of each individual will impact the entire group creates moral implications that are felt on a personal level. Bargaining and agreements that are made may or may not be upheld, and future rounds may be played differently based on the actions of other players in earlier rounds.

Second, because students experiment with decision-making strategies in the context of point-accumulating games while being observed by peers and an instructor, they will be motivated to test and develop *successful* hypotheses with regard to resolution of the games, such as voluntary agreements, appeals to altruism, or censure of pathological or maladaptive behavior. Moreover, they can reasonably be expected to feel a sense of pride at discovering successful strategies, resulting in an emotional investment that is necessary to progress from novice to expert stages of competence.

Third, students will not only be emotionally invested in strategies, but also in outcomes. While initially students will be motivated to win a given game, they can be expected to alter their outlooks in response to the emotional cues that other students express through body language and the articulation of propositional claims. This emotional sensitivity will enliven the affective associations and intellectual links that students develop between the strategies that they select, strategies that classmates select, limitations imposed by game structures, and just and unjust outcomes that arise when certain strategies are used in certain games.

Fourth, the emotional investment in strategy and outcome will incentivize students to develop increasingly better ways to communicate with their peers to improve communication – especially for discussion of issues from multiple perspectives.

III. METHODS

The focus of this paper is a game designed to teach students the economic concept of environmental externalities. In a classroom setting, students are presented with a simultaneous, non-cooperative game theory problem where there are three types of producers (luxury, intermediate and subsistence) with varying production rates for externalities and profitability. The externalities are calculated using an exponential function, whereas the points from production are produced with a linear function. The students' goal is to maximize grade points (calculated as private profits minus their share of social costs generated by the entire class). Sub-optimal production levels may lead to failure of a portion or the entire class to obtain passing grades.

A round of play consists of two simultaneous decisions. First, students are informed of the basic premise of the game and assigned roles as luxury, intermediate or subsistence goods producers. They are told what their production limits are and told they must choose a positive (or zero) production level, but not greater than the limit. Players are encouraged to discuss production decisions among each other and explore the relationship between production levels and social costs either mathematically (through exact equations) or experimentally (by using a spreadsheet program that automates calculations). After deciding on production levels, grade points are calculated and reported back confidentially to each player. In the second decision phase of the round, students are able to donate points to classmates. Again, students are encouraged to discuss how and if to distribute points between players before finalizing their decision regarding point sharing. Agreements are not binding; participants may agree to one thing among themselves, but hand in a different decision to the Instructor. As it has been tested thus far, the game consists of three rounds, and players are told when it is the final round. Final points may be determined based on an average of scores for each round, or just the score acquired in the final round. This has implications for how participants will likely distribute or not distribute points at the end.

The game is calibrated such that the social optimum (maximum class average grade) is found at a level of production lower than the private optimum. As such, each student has an incentive to unilaterally improve their own position by overproducing. However, if all (or a preponderance of) students produce more than the social optimum, the entire class will fail, proverbially choking on their own pollution. Therefore this is a fundamentally collaborative game. The maximum obtainable grade for luxury goods producers is well in excess of the 90 points needed to earn an 'A' grade, whereas intermediate and subsistence goods producers have progressively lower theoretical maximum grade point potential. Nonetheless, unilateral overproduction by any *group* of players of a single type (acting in concert) would be sufficient to ensure failure of the entire class. Consequently, the game forces student to confront several moral issues:

- On what basis should production be allocated among the three different types of producers?
- Are moral arguments alone sufficient to prevent over production by all players?
- To what extent can cooperative strategies emerge both among players of a common type and across types?
- How can agreements between players be enforced, and what should the consequences be for not upholding them?

The exact equations that determine the grades of the luxury, intermediate, and subsistence players are given below:

$$u_{i,n} = \pi_i \times q_{i,n} - \frac{\left[\sum_{i=1}^{3} e_i \sum_{n=1}^{N_i} q_{i,n}\right]^D}{\sum_{i=1}^{3} N_i}. \qquad (1)$$

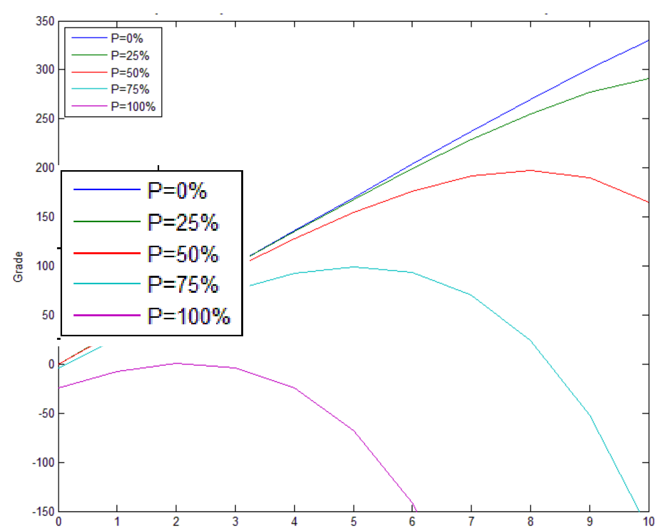
$$D = \log_x(0.2111)$$

$$x = 30\sum_{i=1}^{3} N_i$$

$$e_i = \frac{.19}{N_L}, \frac{.0204}{N_M}, \frac{.007}{N_S}$$

Variables are defined as follows: $u_{i,n}$ = the grade points of a player n of producer type i. π_i = the marginal profitability of producers of type i. $q_{i,n}$ = the quantity produced by player i,n. e_i = marginal emissions. N_i = number of producers of type i, where the subscript L indicates the Luxury type, M indicates Intermediate, and S indicates Subsistence. The constants appearing in (1) and subsequent equations have been chosen to calibrate the eventual grade points to a scale in which 90 points (or more) is considered an 'A' grade, 80 points is considered a 'B', etc. Intermediate and subsistence production quantities are limited to 55 and 160, respectively, so that the overall class average can not exceed 99 points. However, *individual* grades *can* exceed 100 (at the expense of other players).

To facilitate understanding of the mathematical model, students are given graphs to show the optimum levels of production for each type of player, given certain assumptions about what others players do. Figure 1 depicts an optimization problem for a luxury player as a function of production by all other players, expressed as a percentage of maximum production.



For example, if all intermediate and subsistence players produce at 100% of their allowable production levels, the optimum grade for the luxury producer will be obtained by producing 2 units, resulting in a grade of minus 10 (before donations of points from other players). The luxury player clearly has an incentive to convince the intermediate and subsistence players to curtail production, because when these other types of players produce at only 50% of maximum, the optimal grade for the luxury producer is obtained at 8 units, resulting in a grade of 180. Because only 90 points are required to earn an 'A' grade, in this latter scenario, the luxury producer may offer to share excess points with other players, as an incentive for them to forgo their own production.

IV. RESULTS

The game has been play-tested in several forms. The most sophisticated of these, in which all three producer types were represented, was tested with students who had already completed several core courses (including, at minimum, *Fundamentals of Sustainability Science* and *Economics of Sustainability*) in the PhD – Sustainability curriculum at Rochester Institute of Technology.

During a play-test there was some moderate communication in the first round regarding strategy between some participants, but largely independent decision making without group collaboration. One student claimed a moral obligation to not harm his classmates required him to choose a zero production point. Most students significantly overproduced, seemingly with the expectation that if they produced enough for themselves they would get a passing grade. When the first round results came back, the class average was a negative number – although luxury producers fared better than others and one luxury player obtained a passing grade (over 80). Subsistence and intermediate producers asked luxury producers to share points with them. Because the class average number of points was so low, sharing would have meant that the luxury goods producers would have sacrificed their own grades, without being able to obtain passing grades for other. Nonetheless, the intermediate and subsistence players invoked a principle of equity that required luxury producers to "share the wealth." The luxury producers elected *not* to share.

In negotiating the second round, intermediate producers realized that in the absence of cooperation from luxury goods producers, the intermediate and subsistence producers could sabotage the luxury goods producers by maximizing all of their production so that *all* players failed. This provided a sufficiently credible threat that allowed the intermediate and subsistence goods producers to convince luxury goods producers to decrease production in the second round. There was debate over which producers should cut back and how much. Participants generally agreed that in the second round, producers of a certain type would all produce about the same amount. All three groups would produce about half of their respective capacities. Consequently, all students agreed to reduce production levels in the second round. Nevertheless, only luxury producers obtained passing grades. To the extent that luxury players had accumulated points in excess of 80, they were willing to keep their promise to share. However, even after sharing, none of the other types of producers obtained better than a 'D' grade.

Before the final round, a participant proposed a solution to assign production values based on efficiency that would allow every student in the class to get an 'A'. A model was run that showed that by having the luxury group stop production entirely, and maximizing all other production, the class average could be a 99. By agreeing to share points with luxury producers, the intermediate and subsistence producers could assure the luxury producers of 'A' grades, providing they entirely forgo production. However, the luxury group in this case did not completely trust the intermediate and subsistence groups to honor the point sharing part of the agreement. This was partly because they feared that intermediate and subsistence producers would feel entitled to keeping points for themselves to make up for the fact that luxury producers scored comparatively better in previous rounds. As one student said, "In the absence of transparency, no trust can be given." The luxury producers also expressed their feeling that their group should not shut down entirely. Because of this, they elected a token amount of production which ended up reducing the class average in the final round to a 'B'.

Intermediate and subsistence players were "shocked" that the luxury goods producers had not entirely complied with the model, to the detriment of the entire class. Feelings were mixed in the intermediate and subsistence groups, both of which followed the proposed plan to maximize the score for all, about whether agreements to share points with the luxury producers should be honored at all. While some felt honor-bound to live up to their agreement, others felt that since the luxury group produced needlessly, they should be punished by receiving no points from the other groups. Most students were somewhere in-between. One student felt he was entitled to keep his points because he earned them, claiming "this is capitalism, not socialism." This provided a teaching moment in which the instructor could clarify the difference. The decision to share points for services is capitalism, whereas in socialism the government mandates involuntary sharing.

After the final round of game play, students were asked to reflect on their experiences and the lessons or limitations of the game. Almost all the students reported that they had "fun", which is consistent with the literature of serious game play and supports the goal of engaging students in active learning about ethics. One student remarked that through the game they were actually able to experience the theories they previously only learned about through readings. Students realized the difficulty of applying the Coase theorem in practice even in ideal conditions.

V. CONCLUSIONS

These preliminary experiences reinforce the hypothesis that the game-based approach fosters a sense of emotional resonance and emphasizes group deliberation. To the extent that discussions regarding the fair distribution of resources, what constitutes ethical behavior in the context of the game, and the relevance of hypothetical games to real-world behaviors emerged spontaneously (i.e., without prompting from the Instructor), the game as play-tested achieved several instructional goals. Students certainly were more engaged in discussions and active learning. The extent to which disagreements and frustrations emerged is indicative of the emotional investment that students were willing to make in the exercise. Moreover, a willingness to adapt strategy to feedback indicates that students were behaving experimentally, as they tried different forms of agreement or persuasion. In short, the trial game play provides qualitative evidence that supports each of the four hypotheses originally described herein.

Nonetheless, these results are not entirely consistent with the thesis originally termed the "fallacy" of the individual decision-maker. In the first two rounds, where the class averages were poorest, the students correctly interpreted failure as the result of group action, without assigning disproportionate blame to any player or group. However, in the case of the final round, when the class average was the best, students from intermediate and subsistence groups clearly blamed individuals in the luxury group for what they felt was a unilateral defection from the agreement. In this case, the ethical failing could (in the minds of the aggrieved players, at least) be pinned directly on a small group of individuals.

Finally, it as a result of play-testing in other settings, it has been pointed out that students should be called upon to reflect upon the game itself, and not just their experience playing it. That is, students should discuss what the game does and does not represent. For example, one limitation is that accumulating points in the game is not necessarily indicative of learning. Therefore, grading should be based only partially (the Instructors recommend 10%) on game performance, but mostly on other assessment tools, such as essays, term papers or class discussion, that test the knowledge gained through participation in the game. Ultimately, it may be appropriate for students to modify the design of the game (or design their own games) to incorporate additional elements that have yet to be conceived by the Instructors.

Lastly, it is important for science and engineering graduate students to think about how the lessons learned at such a simplistic level could be applied to complex problems in the real world. For example, classroom discussions of externalities in the context of climate change typically center on basic propositions of justice and equity. Agreement on general principles (such as an egalitarian per capita distribution of carbon dioxide emissions allowances) is often quickly achieved in the classroom, but elusive in real negotiations (e.g., Copenhagen). However, in game play, the stakes are more real to students. Issues that emerge in climate change (such as rectification of past inequities) also emerge in the context of the game. Similarly, in a classroom setting, when students are asked to consider whether the US should cut back on CO_2 emissions so that developing countries can emit more, students readily agree because the proposition is abstract. But in game play, luxury students were considerably more circumspect about putting such a principle into practice – a situation that more closely represents real climate negotiations.

Therefore, the *Externalities Game* will likely be most helpful in conjunction with other readings, classroom discussion, and/or pedagogical strategies that help students extract and understand the salient moral or ethical questions, as well as the hypotheses that might lead to their resolution. In this way, the game can be used as a tool for raising the question of correlation: What real world events, actions, and players does the game correspond to? What is a luxury good? Who produces luxury goods? What does a given end state "solution" represent? Asking these questions focuses more attention on the constraints of the model than typical discussion of a philosophical text does. The overall learning exercise would then consist of three phases: simple discussions based on simply modeled moral texts, complex discussions based on playing mathematically simple games, and, eventually, complex discussions based on relating the texts and game experience to provocative and more open-ended questions of moral philosophy.

ACKNOWLEDGMENT

Ron Kline, Park Doing and Stephen Hilgartner at Cornell University offered several helpful observations regarding existing models of ethics education and the effectiveness of the pilot game module. Jacqueline Isaacs of Northeastern University provided helpful insights with regard to the use of educational games as an effective tool for teaching twenty-something graduate students known as the Digital Generation.

REFERENCES

Coase, Ronald H. 1960. The Problem of Social Cost. *Journal of Law and Economics. Chicago IL, 1-23. October 1960.*

Isaacs JA, Laird JT, Seager TP. 2008. Engineering students game to green the automobile supply chain. *IEEE Symposium on Energy and Environment. San Francisco CA, 19-21 May 2008.*

Mayo MJ. 2007. Games for science and engineering education. Communications of the ACM. 50(7)31-35.

Seager TP, Selinger E. 2009. Experiential teaching strategies for developing ethical reasoning skills relevant to sustainability. *Proceedings of the 2009 IEEE International Symposium on Sustainable Systems and Technology, Phoenix AZ, 18-22 May 2009*.

Wolfe J, Fritzsche DJ. 1998. Teaching business ethics with management and marketing games. Simulation & Gaming. 29(1):44-59.