AC 2012-3569: MICROETHICS AND MACROETHICS IN GRADUATE EDUCATION FOR SCIENTISTS AND ENGINEERS: DEVELOPING AND ASSESSING INSTRUCTIONAL MODELS

Dr. Heather E. Canary, University of Utah

Heather E. Canary (Ph.D., Arizona State University, 2007) is Assistant Professor in the Department of Communication at the University of Utah. Her work appears in The International Encyclopedia of Communication and Communication and Organizational Knowledge: Contemporary Issues for Theory and Practice. She has published articles in the American Journal of Public Health, Communication Education, Health Communication, the Journal of Applied Communication Research, the Journal of Business Ethics, and Management Communication Quarterly, among other scholarly journals. Canary is Co-principal Investigator for two inter-disciplinary projects of graduate ethics education, funded by the National Science Foundation. Her other research foci include organizational and family communication, particularly as those processes co-influence each other in contexts of disability, health, and public policies.

Dr. Joseph R. Herkert, Arizona State University

Joseph R. Herkert, D.Sc., P.E., is Lincoln Associate Professor of ethics and technology in the School of Letters and Sciences and the Consortium for Science, Policy & Outcomes at Arizona State University. He has taught engineering ethics and related courses for nearly 25 years. His work on engineering ethics has appeared in engineering, law, social science, and applied ethics journals. Herkert is the Past Editor of IEEE Technology & Society and a founding Associate Editor of Engineering Studies. He is the current Chair of the ASEE Liberal Education/Engineering and Society Division and a Distinguished Life Member of the Executive Board of the National Institute for Engineering Ethics.

Dr. Karin Ellison, Arizona State University
Dr. Jameson M. Wetmore, Arizona State University

©American Society for Engineering Education, 2012

Microethics and Macroethics in Graduate Education for Scientists and Engineers: Developing and Assessing Instructional Models

Abstract[1]

While the government and the public look to universities to educate students in research ethics, those who teach ethics to science and engineering graduate students still struggle to find the most effective models for ensuring that their students internalize professional values and make them part of their scientific and technical practices. (1) This paper will report on a four year research project to develop and assess four different instructional models that introduce and educate science and engineering graduate students to the micro- and macroethical issues in their work. Efforts at integrating micro- and macroethics in graduate education of engineers and scientists have been few. To be effective such efforts require incorporation of interdisciplinary concepts and methods drawn from such fields as science and technology studies and applied ethics. The four models included in the project are: 1) a standalone course on societal implications of science and engineering; 2) micro- and macroethics material embedded in a required science course; 3) a hybrid online/face-to-face course on responsible conduct of research; and 4) engaging ethics in the lab. In the paper we discuss development of the course models and assessment results of students' knowledge of relevant standards, ethical sensitivity, and ethical reasoning, as well as student-instructor communication.

Introduction

In an earlier paper we reported on the beginning phases of a four year research project to develop and assess four different instructional models that introduce and educate science and engineering graduate students to the micro- and macroethical issues in their work, including a detailed summary of the project's initial Coordination Workshop (1). As noted in the earlier paper, "microethics" refers to moral dilemmas and issues confronting individual researchers or practitioners, whereas "macroethics" refers to moral dilemmas and issues that collectively confront the scientific enterprise or the engineering profession, as well as societal decisions about science and technology, decisions that scientists and engineers may participate in as individuals or collectively. The importance of macroethics in science and engineering has gained wider appreciation in recent years, especially in fields such as science & engineering ethics and science & technology studies. Efforts at integrating micro- and macroethics in graduate education of engineers and scientists have to date been few in number.

The project's goals are to: 1) formulate educational outcomes for the integration of micro- and macroethics in graduate science and engineering education; 2) develop and pilot different models for teaching micro- and macroethics to graduate students in science and engineering; 3) assess the comparative effectiveness of the instructional models; 4) facilitate adoption of the instructional models and assessment methods at other academic institutions; and 5) provide for widespread dissemination of course materials and assessment results in the engineering, science, and ethics education communities.

[1] Portions of the text of this paper draw on reference (1).

The project team includes ethics, science, engineering, science & technology studies, and communication faculty at Arizona State University (including the PI Joseph Herkert, Co-PI Karin Ellison and Co-PI Jameson Wetmore) and the University of Utah (Co-PI Heather Canary, who conducted the project assessment), an Advisory Council of faculty from four other universities, three consultants with national reputations in science and engineering ethics education, and a number of "unofficial" project contributors from various fields

The purpose of the present paper is to update the descriptions of the project's instructional models and report on the quantitative portion of the project assessment.

Instructional Models

An important premise of the project is that scientists and engineers cannot simply learn about ethics, but must learn how to incorporate new perspectives into their daily practice and professional behavior. Effective ethics instruction thus requires integration of interdisciplinary concepts and methods drawn both from applied ethics and such fields as science and technology studies. The four instructional models included in the project, chosen to provide a range of approaches to achieving such integration, are: 1) a standalone course on social and ethical implications of science and engineering; 2) micro- and macroethics material embedded in a required science course; 3) a hybrid online/face-to-face course on responsible conduct of research; and 4) engaging ethics in the lab.

Embedded Model

Embedding ethics education in a required core course provides for presenting ethics concepts and technical concepts in comparable ways. During the fall 2008 semester, the research team piloted the embedded model of delivering integrated macro and micro ethics education to science and engineering graduate students. Ellison and two other project senior personnel provided ethics instruction in the Core Course for first-year graduate students in the Ph.D. program in Biological Design. Thirteen students enrolled in the course which met three hours a day, five days a week. The two senior personnel attended all class sessions and one of them, an expert in science policy education, introduced macro-ethics topics for discussion as opportunities arose. He also served as a resource for students on social, ethical and political aspects of science. For example, discussion of biotechnology was related to public acceptance of genetically modified organisms, control of the cell cycle lead to dilemmas in use of stem cells, and tissue engineering was joined to transhumanism questions. Ellison served as one of the invited speakers for the course and delivered a five-session module on research ethics. The pilot study suggested that providing macro- and micro-ethics content together had synergistic effects. Students sometimes raised macro ethical questions with Ellison and microethical questions with the "embedded" faculty member. Also, informal indicators such as quality of submitted student assignments suggested that embedding the ethical content in a technical core course led students to engage the material at a high level.

During the fall 2009 semester, the research team used the embedded model to deliver integrated micro- and macroethics education to science and engineering graduate students with assessment conducted as part of this project. The same course was used and the same instructors participated as in the earlier pilot. Again, thirteen students enrolled in the course, which generally met three

hours a day, five days a week. In addition to the faculty roles discussed for the pilot, a second-year PhD student in the Biological Design program assisted Ellison with the ethics module by leading one class session and serving as a resource for students as they developed cases for presentation. The research ethics sessions were three hours each. The experience with this module continued to suggest that providing macro- and micro-ethics content together had synergistic effects.

During the Spring 2011 semester, the research team used and assessed the embedded model of delivering integrated macro and micro ethics education to science and engineering graduate students for the second time. While the faculty participants were the same as before, the embedded module was implemented in the second semester of the Biodesign core seminar, rather than the first as we did in fall 2009. This changed the context of the ethics instruction somewhat as the course has fewer contact hours, a more applied curriculum, and spread the research ethics component over more days. Eleven students enrolled in the course which met for ninety minutes three days a week. The course focused on integrative areas of bioinformatics, systems biology and synthetic biology, and their application in biomedical research. Due to shorter class meeting time, Ellison presented the RCR materials over 9 class sessions rather than 5. On the whole, these changes are believed to have strengthened the integrated nature of ethics content delivery for these students. More applied content gave more direct opportunities to raise ethical questions. Delivering the RCR content over more sessions provided more sustained reflection on these topics.

Hybrid Model

Based upon feedback from the Coordination Workshop (1) and the extensive availability of online course materials for RCR, the PIs decided to change the online modules envisioned in the project proposal to a hybrid course model. This permitted maximum use of existing materials for the online portion of the course and allowed for better integration of micro- and macroethics in the face-to-face portion of the course.

Ellison had taught a traditional one-credit face-to-face course in the Responsible Conduct of Research (RCR) to life science graduate students for several years; she and Herkert worked together in converting the course to a hybrid format with integrated macroethical content. The revised course has three pedagogical goals:

- Identifying ethical dilemmas in science
- Identifying norms, policies, and regulations that speak to those ethical dilemmas
- Using problem solving skills to generate and select options for action in the face of such dilemmas

Ellison (with some assistance from Herkert) taught the hybrid course in both fall 2009 and spring 2010. In the fall semester, 15 students enrolled; 14 students enrolled in the spring.

The course introduced students to the ethical and regulatory issues in the conduct of research. It covered the nine core areas of responsible conduct of research (RCR), as identified by the National Institutes of Health (NIH). The areas are:

1. Animal welfare
2. Collaborative science
3. Conflict of interest and commitment
4. Data acquisition, management, sharing, and ownership
5. Human subjects
6. Mentor/trainee responsibilities
7. Peer review
8. Publication practices and responsible authorship
9. Research misconduct

In addition, students explored two macroethical topics: sustainability and military funding of university research. Discussion of real and hypothetical cases tied broad ethical considerations to current concerns of scientists and helped students develop ethical reasoning skills.

Students completed online modules, wrote brief case analyses prior to each classroom session, and actively participated in class discussions. In addition to the Collaborative Institutional Training Initiative (CITI) Program online modules in the Responsible Conduct of Research for biomedical researchers (https://www.citiprogram.org/), the students completed the Southeast Regional Center of Excellence for Biodefense and Emerging Infections' Online Training in Ethical and Legal Issues in Biological Research (http://www.serceb.org/dualuse/welcome.htm) and the NIH Office of Extramural Research online module "Protecting Human Research Participants"(http://phrp.nihtraining.com/users/login.php). Using publically available online resources created some challenges. For examples, updating of resources meant that the instructors could not rely on the content or location of the assigned materials to remain stable between syllabus development prior to the start of term and the date students completed or tried to complete assignments. Some students also had trouble accessing materials, especially video clips. On the positive side, students received their CITI Program RCR training certificates as part of the course. Starting in June 2010, our institution has required all graduate students supported on extramural funding to complete CITI training.

Ellison continues to teach the course at least once a year and continues to use a hybrid format and integrated macroethics content, modifications to the course made for this research project.

Stand-Alone Model

Many undergraduate science and engineering programs offer stand-alone courses in technical ethics, but they are rare at the graduate level. Because of this we developed a number of strategies to attract graduate students to a course covering micro- and macroethics (1). Initially we had hoped to offer a three credit course that would engage science and engineering students. We found, however, is that it is difficult for a large number of students to justify, either to themselves or to their advisors, a time consuming course that will not likely help them advance toward their degree. Consequently, we decided to make sure that a larger number could participate by creating a one credit course. We were also able to increase enrollment by negotiating with the Chemistry department to allow the course to meet its mandatory CHM 501 (current topics in chemistry) requirement.

Because the course is offered nearly every semester we decided to rotate the topics presented. In one semester we might do a broad overview of science policy and the macroethical issues it raises while in another semester we might focus on the responsibility and skill necessary to communicate science to the public. To further make sure that we are covering topics of interest to the students, and thereby motivating them to more fully participate in class and reflect on the issues, we let the students play an important role by choosing some of the readings and projects we engage in.

Thus far through the stand alone course model we have offered five one-credit courses. Wetmore and two other project senior personnel collaboratively taught the "Science Policy for Scientists and Engineers" course in fall 2008, spring 2009, fall 2009, spring 2010, and spring 2011. (Project assessment was undertaken during the fall 2009 and spring 2010 semesters.) The class is taught as both a Chemistry requirement and Engineering elective and routinely fills to capacity at 20 students. Students explore the social and ethical implications of their work and science and engineering in general through weekly readings, class discussions, and public outreach projects including the team development of museum cart demonstrations created with the aim of fostering conversations on the macroethical issues raised by cutting edge technologies. In spring 2011 the course was based on a series of mock Congressional hearings where students had to prepare, present, and defend testimony before a panel pretending to be members of Congress. A number of faculty, staff, and graduate students helped to staff the panels in this exercise. These courses have inspired the creation of two additional one credit courses that are not necessarily focused on ethics, but have some ethical content. One was taught in spring 2009 on "Energy Policy." The second, entitled "Science, Technology & Developing Areas," was taught in fall 2009. There are currently discussions to expand this program into the School of Sustainability in the future.

Lab-Engagement Model

This model is based on the idea that scientists and engineers sometimes disregard traditional ethics training in the classroom because they don't see how the lessons could pertain to their daily work or how the ethics instructor could understand their situation. Holding these sessions in laboratories where the students are comfortable helps convey the message that the ethics instructors acknowledge that science and engineering ethics is a cooperative endeavor. The focus of this pedagogy is not to teach the graduate students something completely new, but rather to help them to think about what they already know, analyze it with new tools and perspectives, and reflect on the impact of their daily decisions. This idea of being reflective – the ability to explore where scientific and social values come from, what they mean, and how they may be related to decisions about science and engineering – is a key component of the process to get scientists to engage with ethical and social issues as they conduct their technical practice. (1,2)

The laboratory engagement group completed its first module in spring 2010. A neuro-scientist/engineer volunteered his lab for three months of the program. Wetmore and one other member of the project team, a philosopher, met with this group of seven graduate students, seven undergraduate students, a post doc, and visiting faculty about every other week for a semester during their normally scheduled lab meetings. The module was deemed by all involved as a success. Students and faculty in the lab were very much engaged in the discussions and looked forward to our regular meetings. The instructors thought that the group would shy away from issues with laboratory animals, but on the contrary the lab group was very much interested in

exploring the ethical implications of their research on monkeys. Lab participants raised a number of issues and understood them with such nuance that it was clear to the researchers that the most productive use of their skills was to provide a space and structure for these conversations to happen rather than to lead the discussions themselves. The only issue the researchers had was that the lab participants changed regularly during the program and it was not possible to do an adequate assessment of the success of the program. The second lab engagement was conducted in 2010-2011 in a solar engineering lab with similar results.

Assessment Methods

This study used a nonequivalent control-group quasi-experimental design. The four instructional models constituted four experimental groups. Students enrolled in graduate engineering and science courses that did not include ethics instruction constituted the control group. This study design is common in educational settings when random assignment is difficult or impossible. Additionally, this study design provides adequate controls for threats to internal validity when recruitment is standard for all groups and participants across groups do not differ on relevant variables, which for this study include demographics and pretest scores (3). The institution's Institutional Review Board (IRB) approved the study and participants were provided cover letters before consenting to participate.

Participants

Students in experimental and control groups were invited to participate at the beginning of each instructional experience. For courses, this was the beginning of the semester for which they were enrolled. For the lab engagement model, this was early in the academic year when new lab members had joined and the lab director had agreed to participate in the project. Eighty-one participants completed all sections of the pretest and posttest from fall 2009 through spring 2011: 21 students in the embedded model, 20 students in the hybrid model, 14 students in the stand-alone model, and 26 students in the control group. Due to changes in lab membership and laboratory time constraints, only two participants in the lab engagement model completed both the pretest and the posttest; accordingly, the lab engagement model was not included in results analysis.

Participants were primarily graduate students, with 50 doctoral students and 20 master's students. The remaining 11 participants were either undergraduate students or in a transitional program between undergraduate and graduate degrees. Participant ages ranged from 20 – 34 years old ($M = 24.23$, $SD = 2.72$), with 55 males and 26 females. Several academic programs were represented in the sample: engineering ($N = 30$), biodesign ($N = 21$), biology ($N = 12$), chemistry/biochemistry ($N = 9$), and several others with a few participants each. Most participants were native English speakers ($N = 54$). Other native languages reported by more than two participants each include Chinese ($N = 10$) and Indian languages ($N = 8$). Several other native languages were reported by only one or two participants each. Similarly, most participants reported they were from North America ($N = 51$), with Asia/China Region reported by 13 participants and Asia/India Region reported by 11 participants.

Procedures
Pretests were administered in the first week of each semester. Surveys took approximately 30 minutes to complete in paper-and-pen format. Posttests were administered approximately one week before the end of the course and also took about 30 minutes to complete in paper format. Participants completed study materials during their regularly-scheduled instruction time and were provided snacks each session in appreciation of their participation. Data collection procedures were repeated over two semesters for all groups, except the lab engagement group, in order to obtain an adequate sample size.

Instrumentation
Several measures were used to assess the three learning outcomes identified early in the project as relevant and realistic for the project scope, which were (1) increased knowledge of relevant standards, (2) increased ethical sensitivity, and (3) improved ethical reasoning. These measures included three study-specific scales to assess learning outcomes for four general topic areas included in all instructional models: data management, conflict of interest, military research, and sustainability. The study-specific measures were developed in a three-stage process over several months using a national sample of graduate students in sciences and engineering. First, the project team developed a pool of potential items for assessment of the learning outcomes in the four areas. Next, the item pool was sent to a panel of five experts in ethics education for feedback regarding readability, exhaustiveness, and clarity. The final stage included refining items with panel feedback, testing the pool with an online pilot survey of graduate students, and further refining the measures based on statistical tests of reliability from the pilot survey.

Open-ended items on surveys assessed participant perceptions of their roles in society and their instructional experiences. These responses, which were analyzed using qualitative data analysis techniques, are not included in this paper.

Study-specific measures. Knowledge of relevant standards was assessed with 8 true/false/don't know items, with a possible score range of 0 (no knowledge) to 16 (high level of knowledge). Sensitivity to ethical issues was measured with 14 Likert-type items with choices ranging from 1 (strongly disagree) to 5 (strongly agree). The ethical sensitivity score is an average of all 14 items. Additionally, we developed a six-item study-specific measure of ethical reasoning but results indicated that the control group scored significantly lower than all experimental groups in both pretest and posttest, with all experimental groups scoring exactly the same on pretest and posttest. Accordingly, results of this measure are not included in our report of findings.

Ethical reasoning. Ethical reasoning was assessed using two existing measures of moral judgment. Through discussion, the research team determined that ethical reasoning and moral judgment are essentially the same, or very similar, processes such that previously developed and tested measures of moral judgment would be appropriate to measure this outcome. We used a short form of the Engineering and Sciences Issues Test (ESIT) (4), which reflects James Rest's Defining Issues Test (5) and was specifically designed for science and engineering students. We deleted two sections of the measure that are least applicable to a graduate student sample in order to shorten the measure. The measure generates two scores: the P-score represents the percentage of post-conventional reasoning present, and the N2-score uses the P-score but also accounts for the absence of pre-conventional thinking. As noted by Borenstein et al. (10) post-conventional

reasoning is characterized by "the search for moral ideals on which a social order ideally ought to be based," whereas pre-conventional reasoning is characterized by "narrow personal interest" (p. 391). The Moral Judgment Test (MJT) (6) is another existing measure of moral reasoning that is more general in scope. There were no statistically significant gains on this measure between pretest and posttest and there were no significant group differences on either the pretest or posttest; accordingly this report only includes results from the ESIT instrument.

Student-instructor interaction. Because student gains in knowledge and development may be influenced by instructor characteristics and classroom dynamics, several student-instructor interaction variables were measured for experimental groups. These variables include: out-of-classroom communication with instructor (7), classroom climate (8), instructor verbal aggressiveness (9), and instructor verbal assertiveness (9). Additionally, two items asked participants how many conversations about research ethics they had in the previous three months outside of the classroom/lab with (a) peers and (b) lab directors/principal investigators. Response choices were 0, 1, 2, and 3 or more.

Assessment Results

Statistical analyses of difference (ANOVAs) and paired samples *t*-tests were conducted to compare outcomes between the three instructional models as well as between instructional models and the control group. Table 1 presents a summary of outcome measures by study group. Additionally, previous ethics education research indicates that there might be differences in outcomes between native English speakers and non-native English speakers (10) as well as between males and females (11). Accordingly, we analyzed results in terms of these demographic variables as well. The following sections present results of assessment and demographic analyses.

Knowledge of Relevant Standards

Using paired samples *t*-tests, results indicate a statistically significant increase for experimental groups from pretest ($M = 11.89$, $SD = 2.15$) to posttest ($M = 13.31$, $SD = 2.04$), $t\ (54) = 5.02$, $p < .001$. There was one significant group difference in pretest scores. The hybrid group scored significantly higher ($M = 12.55$, $SD = 2.01$) than the control group ($M = 10.42$, $SD = 2.27$), $F\ (3,78) = 4.06$, $p = .01$. There were also significant group differences in posttest scores, $F\ (3, 78) = 11.03$, $p < .001$. An LSD *post hoc* test revealed that the control group scored significantly lower on the posttest ($M = 10.52$, $SD = 3.06$) than all experimental groups. Also, the hybrid group scored significantly higher ($M = 14.40$, $SD = 1.43$) on the posttest than the stand-alone ($M = 12.36$, $SD = 2.13$) and embedded ($M = 12.90$, $SD = 2.07$) groups.

Ethical Sensitivity

There were no significant group differences determined by an omnibus ANOVA test for pretest scores, although a *post hoc* LSD test revealed that the Embedded group ($M = 3.44$, $SD = .37$) scored higher than the Control group ($M = 3.21$, $SD = .35$) at the .05 significance level. There were significant group differences on posttest scores of ethical sensitivity ($F(3,78) = 3.99$, $p = .01$). An LSD *post hoc* test revealed that the Control group scored significantly lower ($M = 3.21$, $SD = .41$) than all experimental groups and that there were no significant differences between experimental groups.

Table 1

Outcomes by Study Group

Outcome Measure	Embedded Mean (S.D.)	Stand-Alone Mean (S. D.)	Hybrid Mean (S.D.)	Control Mean (S.D.)
Pretest N2-Score	8.11 (3.66)	7.62 (4.39)	8.39 (4.24)	6.64 (3.53)
Posttest N2-Score	8.70* (4.16)	8.76* (3.53)	10.14* (3.93)	5.18 (4.47)
Pretest Knowledge	11.57 (2.16)	11.43 (2.24)	12.55* (2.01)	10.42 (2.27)
Posttest Knowledge	12.90* (2.07)	12.36* (2.13)	14.40* (1.43)	10.62 (3.07)
Pretest Ethical Sensitivity	3.44* (.37)	3.28 (.46)	3.36 (.34)	3.21 (.35)
Posttest Ethical Sensitivity	3.48* (.45)	3.51* (.47)	3.60* (.37)	3.21 (.41)

Note: * indicates significantly higher than Control Group at $p < .05$ level.

Ethical Reasoning

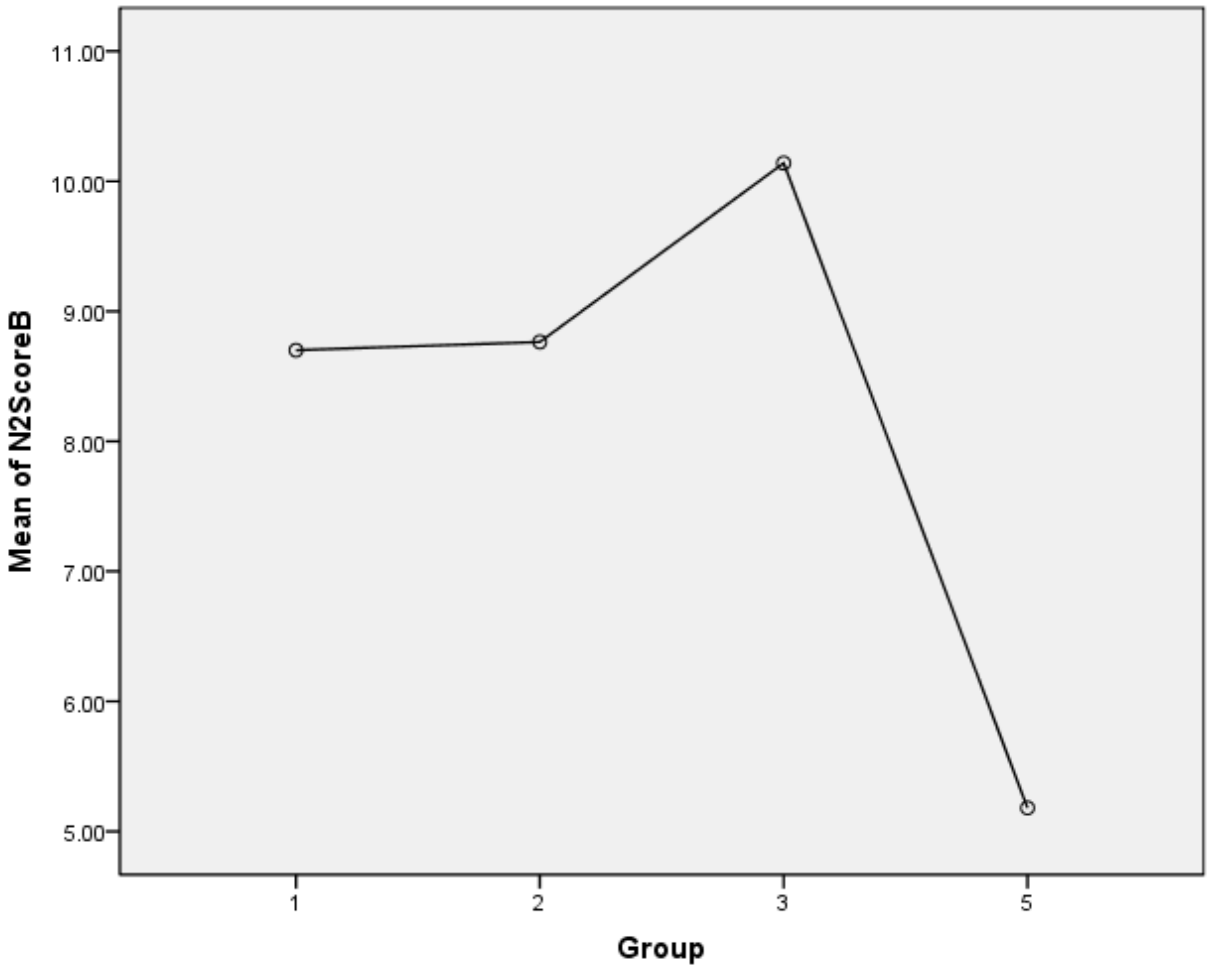
There were no significant gains in P-scores for experimental or control groups from pretest (M = .52, SD = .17) to posttest (M = .53, SD = .15). However, there were significant gains in N2-scores for experimental groups. Accordingly, we focus on the N2-scores. There were no statistically significant group differences in pretest N2-scores between study groups. For experimental groups, N2-scores increased from M = 8.22, SD = 3.92 in the pretest to M = 9.25, SD = 4.37 in the posttest, $t(53) = 2.54$, $p < .05$. There was a statistically significant decrease in control group scores from pretest to posttest (M = 6.64, SD = 3.53 to M = 5.18, SD = 4.47, $t(26) = -2.13$, $p < .05$). An ANOVA indicated statistically significant group differences on posttest scores ($F(3,77) = 5,36$, $p = .002$). An LSD *post hoc* test revealed that the control group scored significantly lower than all experimental groups but there were no significant differences among experimental groups (see Table 1). Figure 1 presents a plot of posttest N2-scores.

Demographic Differences

To probe for language differences, we transformed native language responses to reflect two groups: native English speakers and non-native English speakers. There were differences between these two groups on all outcome measures, including pretest scores. In all cases, native English speakers scored higher than non-native English speakers (Table 2).

Figure 1

N2 Scores by Study Group



Note: Group 1 = Embedded Group; Group 2 = Stand-Alone Group; Group 3 = Hybrid Group; Group 5 = Control Group

Differences between male and female participants were also probed (Table 3). Borenstein et al. (10) reported no sex differences in ESIT scores so we were interested in whether that result was consistent in our sample. It was not. Although there was no statistically significant difference in pretest N2-scores between males and females, there was a significant difference in posttest scores (t (78) = -2.40, $p < .05$). Females scored higher in the posttest than males (M = 9.72, SD = 3.59 and M = 7.06, SD = 5.06, respectively). Additionally, females had a higher difference score than males (M = 1.42, SD = 3.07 and M = -.36, SD = 3.33, t(78) = -2.30, $p < .05$). There was only one sex difference in the other outcome measures, which was that females had a significantly higher posttest knowledge score (M = 13.35, SD = 1.94) than males (M = 12.02, SD = 2.43), t(79) = -2.10, $p < .05$.

Table 2

Outcomes by Language Group

Outcome Measure	Native English Speakers Mean (S.D.) N = 54	Non-Native English Speakers Mean (S. D.) N = 27
Pretest N2-Score*	8.53 (3.85)	5.82 (3.38)
Posttest N2-Score*	9.28 (4.73)	5.12 (3.53)
Pretest Knowledge*	11.83 (2.28)	10.59 (2.08)
Posttest Knowledge*	13.30 (2.11)	10.74 (3.01)
Pretest Ethical Sensitivity*	3.40 (.33)	3.16 (.43)
Posttest Ethical Sensitivity*	3.61 (.35)	3.08 (.40)

Note: * indicates significant group differences at the $p < .05$ level.

Table 3

Outcome by Sex Group

Outcome Measure	Male (N = 55) Mean (S.D.)	Female (N = 26) Mean (S. D.)
Pretest N2-Score	7.31 (3.80)	8.30 (4.08)
Posttest N2-Score*	7.06 (5.06)	9.72 (3.59)
Pretest Knowledge	11.18 (2.21)	11.92 (2.38)
Posttest Knowledge*	12.02 (2.93)	13.35 (1.94)
Pretest Ethical Sensitivity	3.32 (.35)	3.31 (.45)
Posttest Ethical Sensitivity	3.42 (.44)	3.45 (.45)

Note: * indicates significant difference at the $p < .05$ level.

Student-Instructor Interaction

Student-instructor interaction variables were analyzed for any group differences and for associations with outcome measures. The ANOVA indicated only one difference between experimental groups, which was verbal aggressiveness ($F(2, 49) = 3.49$, $p < .05$). The LSD post hoc analysis revealed that the Embedded group reported higher instructor aggressiveness ($M = 18.33$, $SD = 1.22$) than the Hybrid group ($M = 14.21$, $SD = 4.69$). Interestingly, one of the instructors participated in both of these groups. However, regression analyses and correlational analyses did not indicate that instructor aggressiveness was associated with any outcome measures. Furthermore, there were no other group differences in student-instructor interaction variables, indicating that classroom dynamics were fairly similar across the different instructional models.

Student-instructor out-of-class communication was associated with two outcome measures. There was a significant negative correlation between out-of-class communication and posttest ethical sensitivity scores ($r = -.35$, $p < .01$). Additionally, out-of-class communication was positively correlated with posttest reports of research ethics conversations with lab directors/principal investigators ($r = .34$, $p < .05$).

Frequency of ethics conversations outside of class/lab increased for experimental groups from pretest to posttest. A paired samples *t*-test indicated a significant increase from pretest ($M = 1.10$, $SD = 1.19$) to posttest ($M = 2.12$, $SD = .89$) in the number of ethics conversations with peers ($t(41) = 5.53$, $p < .001$). The number of ethics conversations with PI/lab directors also increased from pretest ($M = .38$, $SD = .76$) to posttest ($M = .60$, $SD = .86$), but the increase was not statistically significant.

Discussion of Assessment Results

The purpose of this study was to develop and assess efficacy of different instructional models for integrating micro- and macroethics into graduate science and engineering education. Our efforts contribute to research and practice of ethics education in several ways.

One interesting, and unexpected, finding was the differences in knowledge scores between the Hybrid group and other experimental groups. The Hybrid model included both online and face-to-face instruction, with existing public material and training modules comprising the online component. Although online education often is the target of scathing criticism, it could be that including the online component for delivering material about ethics standards, cases, and other knowledge-building content was a more effective way for teaching knowledge of relevant standards. The other instructional models covered relevant standards for ethical conduct, but not in as great detail nor with the online channel that was used in the Hybrid model. This finding is consistent with findings of a meta-analysis of online learning conducted by the US Department of Education (12), which found that combining online learning with face-to-face instruction led to better learning outcomes than either mode on its own.

Differences in outcome scores between native English speakers and non-native English speakers also deserve attention. These results indicate that caution should be used when using these measures in non-native English speaking samples. As found in an earlier study using ESIT (10),

these measures might reflect cultural knowledge that foreign students studying in the United States do not have. On the other hand, these differences also might reflect cultural differences in ethical interpretations. A third potential reason for score differences is that language used to describe scenarios and response choices on the ESIT is difficult for non-native English speakers to access, leading to answers that do not reflect actual levels of moral reasoning. These differences clearly deserve further investigation.

Relatedly, sex differences in outcome measures should be interpreted with caution. The difference could be an artifact of the language difference phenomenon. Of the 54 male participants, 33 were native English speakers and 22 were non-native English speakers. Of the 26 female participants, 21 were native English speakers and only 5 were non-English speakers. Accordingly, the sex differences could actually be a function of the language differences more than a function of differences between males and females. Peterson et al. (11) argued that because research regarding sex differences in ethics education outcomes is inconsistent, it is highly likely that other factors play a larger role in differences than gender alone.

Finally, results regarding out-of-class communication have implications for longer-term effects. The significant negative correlation between out-of-class communication and ethical sensitivity might indicate that students who struggle the most with identifying ethical issues seek additional dialogue with instructors. This finding deserves further study to unpack the nature of out-of-class communication between engineering/science graduate students and ethics educators. Furthermore, results indicate that fostering informal ethics dialogues seems to have a spillover effect. Students who reported more out-of-class communication with project instructors also reported more out-of-class/lab ethics conversations with their lab directors/principal investigators by the end of the project (but not at the beginning). This could mean that students who become comfortable with talking informally with ethics educators become more comfortable with bringing up ethics issues in their professional settings. This finding is also consistent with previous research of long-term effects of ethics education on professionals' abilities to identify ethical breaches in their workplaces (13). Although this project was not longitudinal in design, results concerning ethics conversations represent a positive indicator for long-term effects.

Students also reported having more ethics conversations outside of the classroom or laboratory with both their peers and their PIs by the end of their ethics instruction. Although only the peer conversation difference was statistically significant, the general upward trend is encouraging. These results, interpreted in light of the results discussed above, indicate that ethics education has a positive impact on students' propensity to engage in conversations about ethical issues.

Summary and Conclusions

Overall, results of the quantitative assessment conducted for this project can be summarized as follows (qualitative assessment results will be reported on in a subsequent paper):

- Students in all experimental groups showed gains in desired learning outcomes; gains were significantly higher than control group gains
- All instructional models improve students' ethical knowledge, sensitivity, and reasoning

- Study-specific measures of knowledge of standards and ethical sensitivity tap changes and correlate with existing measures
- Incorporating online components to ethics education seems to improve knowledge of relevant ethical standards
- ESIT is more appropriate for this population than MJT
- Measures of instructor-student interactions and classroom dynamics are similar across different instructional models
- Ethics education has a positive impact on student engagement in ethics conversations outside of the classroom

Our experience in conducting this project has led us to the following conclusions:

- Graduate students in engineering and science can effectively engage in material that integrates micro- and macroethical issues and perspectives

- A variety of instructional models and pedagogies (embedded/stand-alone, face-to-face/online, classroom instruction/lab engagement) can be successfully employed in these endeavors

- Student and advisor flexibility in scheduling courses; dedicated, interdisciplinary teams of ethics instructors; and commitment and cooperation of engineering and science faculty and departments are critical to the success of these endeavors.

Acknowledgements

The work in this Project is funded by National Science Foundation Award Number 0832944. The authors, who are PI and Co-PIs on the project, acknowledge the contributions of other project senior personnel, the project consultants, our Advisory Panel, and other participants at the Coordination Workshop.

References

1. Herkert, J., Wetmore, J., Canary, H., & Ellison, K. (2009). Integrating microethics and macroethics in graduate science and engineering education: Developing instructional models, Proceedings of the American Society for Engineering Education Annual Conference, http://soa.asee.org/paper/conference/paper-view.cfm?id=11085.
2. McGregor, J., & Wetmore, J. (2009). Researching and Teaching the Ethics and Social Implications of Emerging Technologies in the Laboratory. *NanoEthics*, *3*(1), 17-30.
3. Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Boston: Houghton Mifflin.
4. Borenstein, J., Kirkman, R., & Swann, J. (2005). *Engineering and Sciences Issues Test* version 1.0p. Georgia Institute of Technology.
5. Rest, J., & Narvaez, D. (1998). *DIT-2: Defining issues test*. St. Paul, MN: University of Minnesota.
6. Lind, G. (2009). *The Moral Judgment Test*.

7. Myers, S. A., Martin, M. M., & Knapp, J. L. (2005). Perceived instructor in-class communicative behaviors as a predictor of student participation in out of class communication. *Communication Quarterly, 53*, 437-450.
8. Rosenfeld, L. R. (1983). Communication climate and coping mechanisms in the college classroom. *Communication Education, 32*, 167-174.
9. Myers, S. A., & Rocca, K. A. (2000). The relationship between perceived instructor communicator style, argumentativeness, and verbal aggressiveness. *Communication Research Reports, 17*, 1-12.
10. Borenstein, J., Drake, M. J., Kirkman, R., & Swann, J. L. (2010). The Engineering and Science Issuees Test (ESIT): A discipline-specific approach to assessing moral judgment. *Science and Engineering Ethics, 16*, 387-407.
11. Peterson, D., Rhoads, A., & Vaught, B. C. (2001). Ethical beliefs of business professionals: A study of gender, age and external factors. *Journal of Business Ethics, 31*(3), 225-232.
12. Department of Education, Office of Planning, Evaluation, and Policy Development. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies.* Washington, D. C.: Department of Education.
13. Gale, K., & Bunton, K. (2005). Assessing the impact of ethics instruction on advertising and public relations graduates. *Journalism & Mass Communication Educator, 60,* 272-285.