Engaging Everyone in Research Ethics: Assessment of a Workshop for Engineering and Computer Science Graduate Students

SUSAN B. WAINSCOTT
MOHAMED B. TRABIA
DAVID E. JAMES
University of Nevada, Las Vegas
Las Vegas, NV

ABSTRACT

How can we engage new engineering and computer science graduate students in meaningful conversations about research and publication ethics without establishing a common understanding of the issues and expectations? Most universities offer extensive responsible conduct of research (RCR) training programs, which are usually a semester-long. Absent a requirement, it is unlikely that engineering and computer graduate students and their advisors would prioritize a lengthy training during the student’s first semester. Recognizing this, the University of Nevada, Las Vegas’ Howard R. Hughes College of Engineering designed and implemented an introductory research ethics workshop for all graduate students entering engineering and computer science. We engaged an interdisciplinary team of faculty and staff in the workshop’s design and implementation, and approached our design within the sensemaking framework for ethical decision-making. Each workshop included lecture content in four priority topic areas identified by the college faculty: research design and data ethics, publication ethics, computer coding ethics, and intellectual property. The workshops also included a face-to-face panel discussion with experts including engineering, computer science, and law professors; librarians; and technical writers. Our assessment showed that after completing the workshop, students demonstrated increased content knowledge, and their self-assessed expertise ratings were better aligned with their content knowledge.

Key words: Research Ethics; Graduate Education; Knowledge Gain

INTRODUCTION

Engineering and computer science graduate students start their advanced degree programs with a variety of cultural and educational backgrounds and professional experience. Although student
participation in research and publication of research results is a key component of graduate school in the United States (National Academy of Engineering 1995), many students may have little experience conducting research or be lack knowledge about expectations for ethical conduct in research and publication. While engineering ethics have been required in ABET-accredited undergraduate curricula for more than two decades (Rabins 1998; Hess and Fore 2018), the subject matter in most of these courses does not emphasize research or publishing ethics. Instead, the emphasis is on engineering industrial practice associated with the engineering professional societies' codes of ethics, such as the National Society of Professional Engineers (NSPE) Code of Ethics (Hess and Fore 2018; NSPE 2019). For example, the NSPE Code of Ethics covers topics that reflect engineering practices, such as protection of proprietary information and lawful practice of engineering, but does not mention research conduct (NSPE 2019). Thus, training on research ethics topics would likely be of benefit to matriculating engineering and computer science graduate students.

One of the frameworks for ethical decision-making shows promise for research ethics training participants who have a wide variety of experience and existing knowledge: sensemaking. In the sensemaking framework for ethical decision-making, there are three considerations: (1) recognition, (2) acknowledgement of complexity, and (3) forecasting outcomes of ethical dilemmas and decisions (Mumford et al. 2008). Notably, the sensemaking framework accommodates the subjective nature of complexity characterizations, which are based upon each person's knowledge, experiences, and current situation (MacDougall, Martin, Bagdasarov, & Mumford 2014). Use of sensemaking implies that a person must be able to identify (recognize) an ethical issue in order to employ an ethical response. Again, while undergraduate engineering ethics courses may cover some related topics, it is unlikely that transfer of engineering ethics knowledge would address all areas of research and publishing ethics. Graduate students' skills in recognizing and then analyzing ethical dilemmas would benefit from research ethics training that orients them to a common set of expectations as well as resources for navigating ethical dilemmas within a research setting. These considerations led us to seek existing research ethics training models that addressed these needs.

Training in ethical and responsible conduct of research (RCR) is often available to students upon entering a graduate degree program, and is required for those working on research projects funded by certain agencies. Some U.S. federal agencies, such as the National Science Foundation (NSF) and National Institutes of Health (NIH), require training in RCR for graduate and post-doctoral student researchers who receive certain federal research funding (Cordova 2018; National Institutes of Health 2009). These agencies have also issued recommendations for the instructional methods, minimum duration, and content of RCR training. According to the National Academies of Sciences, Engineering, and Medicine (2017) several mechanisms can be used to provide RCR training including stand-alone courses, workshops, or online training. The training is commonly longer than eight hours and
often includes group discussions of cases in which researchers grapple with ethical considerations. Only a few universities have made RCR training mandatory for all students and/or post-doctoral researchers (Resnik and Dinse 2012).

Many universities have centralized the design and delivery of RCR training (Resnik and Dinse 2012), and will include many topics in order to address the wide variety in research methods used across disciplines. Common RCR topics are research misconduct, the protection of human subjects, the welfare of laboratory animals, conflicts of interest, data management practices, mentor and trainee responsibilities, collaborative research, authorship and publication, and peer review (Steneck 2007). Centralized RCR training is also likely to include either generalized examples of ethical cases, or examples from many disciplines. Topics and examples that are specific to other disciplines may appear irrelevant to engineering and computer science students (Bulger and Heitman 2007). Thus, completion of centralized RCR training might be delayed or avoided by some students.

RCR training that emphasizes compliance expectations for participants’ fields of study and examples from their discipline might be desirable and more effective (Mulhearn et al. 2017). While only a small percentage of RCR training at research-intensive universities includes content tailored to disciplines (Philips et al. 2018), several examples of RCR training designed for engineering or computer science graduate students appear in the literature. These include using case-study examples in engineering and computer science courses (Barry and Ohland 2009; Brock et al. 2008; Kligyte et al. 2008; Newberry et al. 2011; O’Neill-Carrillo et al. 2008; Schneider et al. 2016; Valdes et al. 2009). Other ethics courses focused on engineering and computer sciences cases (Canary et al. 2012; 2014; Pennock and O’Rourke 2017; Phillips et al. 2018; Schneider et al. 2016; Tractenberg et al. 2015; Witten 1992). Some RCR-focused lessons were developed and added to existing subject courses (Canary et al. 2012; 2014; Keefer et al. 2014; May and Luth 2013; Walter 2012). These examples are all tied to credit-bearing courses, which require instructor time to develop and a considerable time commitment for a graduate student. Unless these courses are required, graduate students may not elect to take them, or may enroll after their first year in the degree program.

In the sensemaking framework, recognition and awareness of potential ethical dilemmas is foremost, and early training, before matriculating students encounter such issues in their degree programs, is useful. At institutions where all entering graduate students are not required to take the NSF and NIH-mandated RCR training, additional research ethics training has been offered that can provide this foundation. Some universities offer shorter duration research ethics training to introduce a few key RCR concepts. This training may be part of a graduate student orientation program (Poock 2004), while other training occurs at different points within the degree program. In some instances, this training is decentralized and focuses on a few disciplines, such as engineering and computer science research. For instance, Pennock and O’Rourke (2017) described a set of several
facilitated 50- to 90-minute workshops for science and engineering graduate students that covered one or two “science virtue” ethical topics. Buffington et al. (2018) described a 30-minute discussion on research integrity that was led by the College of Engineering’s associate dean as a part of a three-day retreat for graduate students. Another workshop included an instructor-led introductory discussion and a group case-study discussion (Vallero 2007). Other formats include a multi-player card game focused on learning and discussing RCR concepts within engineering contexts (Briggle et al. 2016) or adding a standing agenda item for a discussion of ethical issues within existing, regularly scheduled meetings of a research group (Canary et al. 2012; 2014).

Ultimately, all RCR training aims to decrease the occurrence of irresponsible conduct and to promote ethical decision-making (Herkert 2000; Steneck and Bulger 2007), but this goal may be difficult to measure directly. For example, Marusic et al. (2016) found only a few published studies that assessed actual or intended behavioral changes, including some evidence for improved reasoning skills and knowledge or awareness of ethical situations. Some argue that educators should instead measure intermediate outcomes, such as content knowledge and skill building (Resnik 2014). It is also challenging to achieve large shifts in content knowledge after a single training. Antes et al. (2010) found that RCR training at 21 U.S. research institutions demonstrated limited effectiveness in increasing the knowledge or awareness of ethical situations or participants’ self-reported behaviors in hypothetical situations. Meta-analysis of several RCR training efforts’ assessment data has shown moderate improvements in participants’ content knowledge and ethical-analysis skill building (Antes et al. 2009; Watts et al. 2017). With such modest expectations in mind, we focused our efforts on building awareness and content knowledge for the most salient research and publishing ethics topics for our college.

In this work, we present a foundational engineering and computer science research ethics workshop and the associated assessment. We employed the sensemaking framework for ethical decision-making (Mumford et al. 2008) in our design, delivery, and assessment of the workshop. The proposed approach may apply to other institutions where the NIH- and NSF- mandated RCR training is not required of all graduate students.

METHODS

Background

The Howard R. Hughes College of Engineering at the University of Nevada, Las Vegas (UNLV) offers M.S. and Ph.D. degrees in four departments: Civil and Environmental Engineering and Construction, Computer Science, Electrical and Computer Engineering, and Mechanical Engineering. Year-to-year graduate student enrollment during the period of this study ranged between 235–265; nearly half were international students.
UNLV’s Office of Research Integrity provides centralized training programs for research ethics involving human subjects, as well as NSF and NIH-compliant training in RCR. During this study period, faculty and students had two options for completing NSF and NIH-compliant RCR training: either an in-person credit-bearing course, or a set of online training modules plus participation in a series of in-person seminar sessions. This RCR training is not mandatory for all graduate students. The college does receive research funding from a variety of government and industry sources, but most of the funded projects do not require the RCR training. Therefore, it was not surprising that approximately 10% of engineering and computer science graduate students had completed the online-only training course for research involving human subjects, and 1% had completed the RCR training (personal communication, C. Shafer, UNLV Research and Economic Development 2017). We do not know how many of our incoming graduate students have completed RCR training elsewhere.

In 2015 the College of Engineering designed a new, foundational research ethics training that all incoming graduate students within the college should complete during the first semester after matriculation. We designed a workshop grounded in the sensemaking framework of ethical decision-making, which requires that individuals must first recognize that an ethical dilemma exists in order to then consider potential ethical expectations and common issues (Mumford et al. 2008). To inform the workshop content, we identified patterns in common ethical issues reported by college faculty and staff who regularly help students navigate best practices in publishing and research. We sought to build students’ awareness of salient ethical dilemmas that they were likely to encounter as researchers during their graduate program. We engaged experts across the university to include specific engineering and computer science topics and examples that would augment, but not replace, the central RCR training available at our university. The workshop content also acknowledged that many ethical questions do not have simple answers, and we aimed to encourage our students to initiate analysis and discussion with their mentors, advisors, and other experts.

Outline of the Ethics Workshop

As shown in Figure 1, the workshop consisted of two components: a lecture and a panel discussion. The lecture covered four topics: research design and data ethics, publishing ethics, computer

<table>
<thead>
<tr>
<th>5 minutes</th>
<th>10 minutes</th>
<th>40 minutes</th>
<th>10 minutes</th>
<th>50 minutes</th>
<th>5 minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start of Workshop: Pre-Workshop Survey</td>
<td>Introduction by the Associate Dean for Research and Graduate Studies</td>
<td>Lecture on four ethics topics: • Research Design &amp; Data • Publishing • Computing &amp; Information Technology • Intellectual Property</td>
<td>Short break. Participants hand in anonymous questions on notecards</td>
<td>Panel discussion. Panelists from: • Engineering College • Law School • Library</td>
<td>End of workshop: Post-Workshop Survey</td>
</tr>
</tbody>
</table>

*Figure 1. Duration and format of the UNLV College of Engineering ethics workshop for engineering and computer science graduate students.*
Engaging Everyone in Research Ethics: Assessment of a Workshop for Engineering and Computer Science Graduate Students

The lecture focused on the identified common ethical dilemmas, grouped into four topic areas:

- **Research Design and Data Management (research):** We briefly addressed foundational concepts of the generalized scientific method and data quality. We also encouraged participants to think about what to do when errors are detected in their research methods and data analysis.

- **Publishing:** We described plagiarism avoidance, citation practices, and discussed what publishers expect from authors and co-authors. Publication retractions and errata notices were also discussed.

- **Computing and Information Technology (computing):** We discussed the different expectations for authorship, licensing, and citation of computer code within course assignments, research projects, and professional publications.

- **Intellectual Property (IP):** We discussed copyrights, patents, potential industry trade secrets, and innovation considerations. We also discussed copyright considerations for reuse of images. While not legal advice, the presented content was periodically reviewed by members of the UNLV's William S. Boyd School of Law faculty to ensure it correctly reflected current interpretations of law.

The second consideration in the sensemaking framework, after recognizing an ethical dilemma, is for individuals to analyze the dilemma for potential complexity, rather than expecting a quick and simple answer (Mumford et al. 2008). To address this sensemaking consideration, content knowledge and information technology ethics, and intellectual property. Our learning outcomes and knowledge assessment measures (Table 1) focused on building awareness of those ethical dilemmas common for our engineering and computer science graduate students (Table 1).

<table>
<thead>
<tr>
<th>Workshop Learning Outcome</th>
<th>Workshop Component</th>
<th>Lecture Topic</th>
<th>Sensemaking Consideration</th>
<th>Assessment Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Be aware that ethical research conduct includes understanding the scientific method,</td>
<td>lecture</td>
<td>research</td>
<td>awareness</td>
<td>content knowledge</td>
</tr>
<tr>
<td>reporting errors and failures, calibrating equipment, and data management.</td>
<td></td>
<td></td>
<td></td>
<td>question 1</td>
</tr>
<tr>
<td>Be aware that reuse of images and other figures require both citation and permission.</td>
<td>lecture</td>
<td>publication and IP</td>
<td>awareness</td>
<td>content knowledge</td>
</tr>
<tr>
<td>Be aware that software code must be cited and reuse or modification may also require</td>
<td>lecture</td>
<td>computing and IP</td>
<td>awareness</td>
<td>question 2</td>
</tr>
<tr>
<td>permission and licensing.</td>
<td></td>
<td></td>
<td></td>
<td>question 3</td>
</tr>
<tr>
<td>Be aware that nations may have different intellectual property (IP) laws.</td>
<td>lecture</td>
<td>IP</td>
<td>awareness</td>
<td>content knowledge</td>
</tr>
<tr>
<td>Be aware that coauthors need to communicate about shared responsibility.</td>
<td>lecture</td>
<td>publication</td>
<td>awareness</td>
<td>question 4</td>
</tr>
<tr>
<td>Understand that analysis of ethical dilemmas can include consideration of multiple</td>
<td>panel</td>
<td>all</td>
<td>perspective taking</td>
<td>none</td>
</tr>
<tr>
<td>perspectives.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Workshop learning outcomes, workshop component, lecture topic, sensemaking framework consideration, and assessment measurement.
question 5 addressed the topic of collaboration in research. We also concluded each workshop with a panel discussion with faculty from the College of Engineering, the UNLV Libraries, and UNLV’s William S. Boyd School of Law to address complex topics such as grant-funded intellectual property ownership. Additionally, Law School faculty discussed U.S. and international copyright and intellectual property law and emphasized that some ethical dilemmas have both legal ramifications and professional employment expectations. We used questions submitted anonymously by the students to moderate each workshop panel discussion.

The third and final consideration in the sensemaking framework is for individuals to consider what might be the likely outcomes of each alternative decision for themselves and how their decision might affect others (Mumford et al. 2008). While addressing this type of analysis was not an explicit learning outcome for this two-hour workshop, panelists were asked to discuss accounts and consequences of both good and poor ethical decisions, emphasizing cases that have affected graduate students (Table 1).

We have offered at least one workshop in each Spring and Fall semester since Spring of 2015. The workshop lecture content and overall format were fairly consistent from the first workshop phase covering Spring 2015 - Spring 2018. The content of the lecture was adjusted slightly over this 3-year period in response to feedback from students, panelists, and engineering faculty, and new examples or materials were incorporated.

In the second phase, starting Fall 2018, we added a new lecturer and made changes to the lecture format. In Spring 2020, the COVID-19 pandemic required a shift to a third fully remote delivery phase. This paper presents findings from the first workshop phase. Overall, 187 students attended the workshops during the first phase. This includes 173 individuals who attended at least one workshop, 13 who attended twice, and one student who attended three sessions. Additional details about the workshop design and content can be found in Trabia et al. (2016) and in Appendix A.

**Workshop Assessment**

In light of our unique learning outcomes, and the limited gains in ethical knowledge and intention reported by many RCR educational interventions, we aimed for an assessment that would measure short-term knowledge gained in the topics we addressed, as well as provide insights that would improve future workshop delivery. Our efforts began as classroom assessments for internal purposes. From the workshop’s inception, we have used short surveys to assess student awareness of the learning outcomes by measuring student content knowledge before and after the workshop. We also collected student questions submitted for use in moderation of the panel discussions. The questions’ content and frequency were reviewed for evidence of student engagement with the topics.
the lecture addressed. In Fall 2017 and onward, we also collected students' self-assessment of their expertise in research ethics before and after the workshop through an additional survey question. In advance of the Fall 2017 workshop, we decided to more formally assess student workshop knowledge gains and received UNLV Institutional Review Board (IRB) approval (Exempt #1096005-1), including approval to add a control (non-participants) group for survey questions. This IRB approval also applied retroactively to earlier assessment data. Below, we first describe the assessment measure collection tools and the types of measurement data.

Assessment Measure Collection Tool: Surveys

Pre-workshop surveys were collected before the start of the lecture, and post-workshop standardized surveys were collected after the panel discussion. Students were not provided the correct answers to the pre-workshop survey. Students also were not informed that they would receive a post-workshop survey until it was distributed at the end of the workshop. None of the surveys requested the respondent’s identity, and any unsolicited identifying information was removed prior to data archiving and analysis. All pre- and post-workshop surveys included five ethics content knowledge questions derived from information covered in the lecture, described in detail below. Respondents were informed that there might be more than one correct answer to each multiple-choice question. After IRB approval, starting in Fall 2017, we also included a consent form and a separate five-point Likert scale ethical expertise self-assessment question to the pre- and post-workshop surveys. All pre-workshop survey forms also included a free text response field for students to pose questions for the panelists. Our survey forms and distribution methods are further described in Appendix A.

In Spring 2018, to establish a control group, we surveyed 123 graduate students enrolled in the College who had not previously attended any of the workshops. The online control group Qualtrics® survey content was similar to the questions used in post-workshop survey forms used for the Fall 2017 and Spring 2018 workshops and included the IRB-approved consent form and the Likert self-assessment of expertise. The control group received three email requests to complete a Qualtrics® online survey.

Survey response rates varied from semester to semester, as shown in Table 2. In order to increase the response rate, the survey administration method changed over time (see Appendix A). We did not administer surveys in Fall 2016 to allow design time to address survey distribution methods and increase response rates. We achieved an overall increase in response rates for the three most recent semesters. For the 154 surveyed workshop attendees, 113 (73%) responded to the pre-workshop survey and 77 (50%) responded to the post-workshop survey. For comparison, the response rate for the control group of 123 students was 15 (12%).
Assessment Measure Collection Tool: Notecards

In addition to the pre-workshop survey, students could submit additional questions for the panel discussion immediately after the lecture. These questions were submitted anonymously via paper notecards. The panelists’ identities and departmental affiliations were announced at the start of the panel discussion, so student questions were not biased by this knowledge. Students could also ask questions during the panel discussion, but these were not recorded. During survey data archiving, participants’ free-text note card questions from the Spring and Fall 2016 workshops were combined and could not be disaggregated for analysis.

Measurement: Content Knowledge (Awareness)

To measure student awareness of potential ethical issues, we used five multiple-choice content knowledge questions to assess the five awareness learning outcomes (Figure 2 and Table 1). The questions were examined for face validity by the lecture and assessment designers.

The maximum score for any pre- or post-workshop content knowledge assessment was five (5) correct answers, with the exception of Spring and Fall 2015. Due to survey form design issues in the Spring and Fall 2015 workshops, answers to questions 2 and 3 (Figure 2) could not be correctly entered. These data were excluded from analysis, leaving a possible maximum score of 3 for these semesters (Table 2). Thus, the 2015 scores were analyzed separately from the Spring 2016 and later workshops. Response data from the Spring 2017 pre-workshop survey were also excluded because one student responded twice to the survey, and totals of correct responses

### Table 2. Response Rates to Pre- and Post-Workshop Standardized Surveys.

<table>
<thead>
<tr>
<th>Semester</th>
<th>Survey Method</th>
<th>Number of Content Knowledge Questions</th>
<th>Number of Participants</th>
<th>Survey Response Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring 2015</td>
<td>Email</td>
<td>3</td>
<td>11</td>
<td>55%</td>
</tr>
<tr>
<td>Fall 2015 (two sessions)</td>
<td>Email</td>
<td>3</td>
<td>71</td>
<td>69% 8%</td>
</tr>
<tr>
<td>Spring 2016</td>
<td>Qualtrics®</td>
<td>5</td>
<td>13</td>
<td>38% 54%</td>
</tr>
<tr>
<td>Fall 2016 (two sessions)</td>
<td>Survey not administered</td>
<td>NA</td>
<td>33</td>
<td>NA NA</td>
</tr>
<tr>
<td>Spring 2017</td>
<td>Email pre- Paper post-</td>
<td>5</td>
<td>15</td>
<td>107% 93%</td>
</tr>
<tr>
<td>Fall 2017</td>
<td>Email pre- Paper post-</td>
<td>5</td>
<td>26</td>
<td>88% 100%</td>
</tr>
<tr>
<td>Spring 2018</td>
<td>Email pre- Paper post-</td>
<td>5</td>
<td>18</td>
<td>83% 100%</td>
</tr>
<tr>
<td>Spring 2018 Control</td>
<td>Qualtrics®</td>
<td>5</td>
<td>123</td>
<td>12% NA1</td>
</tr>
</tbody>
</table>

1 One student submitted two pre-workshop survey forms.
2 Control group completed only one survey form.

---

**Table 2. Response Rates to Pre- and Post-Workshop Standardized Surveys.**
could not be computed for that survey. Survey administration methods, participant numbers, and number of content knowledge questions for each semester are provided in Table 2 with additional details provided in Appendix A.

For each content knowledge question, we calculated the average and standard deviation of the number of correct responses. We used response rate and sample size data to compute the sampling Margin of Error (MOE), defined as a statistical estimate of the uncertainty around the sample mean (Scheuren 2004), using 95% confidence limits ($\alpha = 0.05$) for each computed average for the pre- and post-workshop surveys for each semester. The sampling MOE was computed using Equation (1) from James, Schraw, and Kuch (2015) and converted to a percentage of the range.

To reduce sampling MOE, responding participant sample sizes, $n$, and workshop participant population sizes, $N$, were pooled across the Spring 2015 – Fall 2015 workshops and separately for the Spring 2016 through Spring 2018 workshops for each question using methods described in James, Schraw, and Kuch (2019). We set a MOE threshold of 5% (Dillman et al 2009; James, Schraw and Kuch 2015, 2019; Royal 2016). Any resulting computed MOE less than 5% of the maximum number of correct scores was considered small enough for the measured knowledge change from pre-workshop to post-workshop to be informative.

Figure 2. Content knowledge questions. Post-completion correct answers depicted in bold font.
Measurement: Self-assessment of Expertise (Confidence)

The pre- and post-workshop surveys in Fall 2017 and Spring 2018, as well as the Spring 2018 control group survey, included a question for the respondents to self-assess their expertise in research ethics, based on a five-point ranked Likert scale (see Appendix A). These data allowed us to calculate values of the Gamma coefficient (Sirkin 2011) to describe any association between respondents’ self-assessed Likert scale expertise ratings and their number of correct content knowledge answers to detect possible over- and under-confidence (Kruger and Dunning 1999) compared to their content knowledge. We also compared workshop participants’ Gamma values before and after the workshop to a control group’s (non-participants) responses and confidence. Additional information about the Gamma coefficient is provided in Appendix B.

Measurement: Student Questions

We examined the questions that students submitted for the panel discussion to measure their engagement with each topic area. Questions could be submitted via each pre-workshop survey form, or on notecards after the lecture. Two of the study authors participated in a content analysis of the questions and counted the number of questions per topic area for each workshop. If no students asked questions about a topic across several workshops, this may indicate that the topic was not relevant to the students’ experiences or their understanding of their future research projects. We classified student questions across five predefined themes: Intellectual Property, Publishing, Research, Software (Computer Coding and Information Technology), and Other. Depending on the question content, a question could be classified into more than one category. We independently coded the data, compared results, and discussed differences to determine a consensus coding using methods in Taylor-Powell and Renner (2003).

Additional details about the workshop assessment design and data collection can be found in the Appendices.

RESULTS

Content Knowledge (Awareness)

From the content knowledge questions, we calculated the mean number of correct scores, standard deviations (SD) and computed sampling margins-of-error (MOE) for the content knowledge questions for each semester’s workshop(s) and the control group. Mean scores with MOE values less than 5% of range for $\alpha = 0.05$ are indicated in bold font in Table 3. Also shown are the pooled numbers of surveyed workshop participants ($N$) and pooled numbers of respondents with a completed survey ($n$). Standard deviations were computed using the pooled data and responding sample sizes.
We were able to establish a pre-workshop measurement of awareness for Spring-Fall 2015, but not for the post-workshop because of a large MOE. The post-workshop MOE of 1.57 was 52% of the possible range of scores (3). Survey design issues eliminated 2 of the 5 content knowledge questions from analysis, and we also experienced a low response rate for the post-workshop surveys. For these two early workshops, we are unable to determine if there was any change in participants’ awareness of potential ethical dilemmas. However, after we addressed the survey design and administration starting in Spring 2016, we achieved measurements that yielded MOE’s well within the 5% of the maximum score range. The pooled Spring 2016-Spring 2018 mean scores increased from 2.31 pre-workshop with a MOE of 4% to 3.73 post-workshop with a MOE of 1.5%.

We next examined each of the content-knowledge question responses for the Spring 2016-Spring 2018 group (Table 4). The percentage of correct responses were evaluated for each of the five questions, and there was an increase post-workshop in correct responses for all questions. The largest increase was for Question 1, “Research Ethics involves …”, and the smallest increase was for Question 5, “When there is more than one collaborator…”.

**Table 3. Pooled Pre- and Post-workshop and Control Mean Content Knowledge Scores.**

<table>
<thead>
<tr>
<th>Pooled workshops (range of number of correct answers)</th>
<th>Pre-workshop</th>
<th>Post-workshop</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>n</td>
</tr>
<tr>
<td>Spring and Fall 2015 (3)</td>
<td>82</td>
<td>54</td>
</tr>
<tr>
<td>Spring 2016-Spring 2018 (5)</td>
<td>72&lt;sup&gt;2&lt;/sup&gt;</td>
<td>43</td>
</tr>
<tr>
<td>Spring 2018 Control (5)</td>
<td>123</td>
<td>15</td>
</tr>
</tbody>
</table>

1 Mean scores with MOE’s less than 5% of range for that group are shown in **bold font**
2 Spring 2016, Spring 2017, Fall 2017 and Spring 2018 attendees
3 Spring 2017 pre-workshop survey responses were excluded from the pre-workshop analysis due to one participant completing two surveys.

**Self-assessment of Expertise (Confidence) Association with Content Knowledge (Awareness)**

Gamma coefficient values increased after workshop participation for both Fall 2017 and Spring 2018 as shown in Figure 3 and Table 5. Figure 3 shows contingency tables for pre- and post-workshop content knowledge (awareness) and self-assessed expertise (confidence) for both semesters. Scores falling inside the gray-shaded cells are concordant pairs in which individuals scored their expertise level in a manner consistent with their number of correct content knowledge answers. Responding
### Table 4. Pre-workshop to Post-workshop change in Correct Pooled Responses for each content Knowledge Question.

<table>
<thead>
<tr>
<th>Content Knowledge Questions (Spring 2016, Fall 2017 and Spring 2018 pooled)</th>
<th>Percent Correct (%)</th>
<th>Pre to post Increase (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Research ethics involves…</td>
<td>21.4</td>
<td>83.3</td>
</tr>
<tr>
<td>2. To use an exact copy of a picture…</td>
<td>42.9</td>
<td>73.5</td>
</tr>
<tr>
<td>3. When re-using code…</td>
<td>33.3</td>
<td>59.2</td>
</tr>
<tr>
<td>4. Are laws for copyrights and patents…</td>
<td>69.0</td>
<td>83.3</td>
</tr>
<tr>
<td>5. When there is more than one collaborator…</td>
<td>65.9</td>
<td>73.5</td>
</tr>
</tbody>
</table>

### Figure 3. Contingency tables and Gamma coefficient values for the Fall 2017 and Spring 2018 workshops. Gray cells indicate the possible perfectly concordant scores for each table.
sample sizes were too small to generate statistical significance at the 0.05 level, so Gamma coefficient results must be considered preliminary. Table 5 summarizes the Gamma values for Fall 2017 and Spring 2018 workshop groups and the Spring 2018 control group.

The Fall 2017 pre-workshop participant responses' negative Gamma (-0.326; Figure 3a, Table 5) showed no concordant association between content knowledge (awareness) and self-assessed expertise (confidence). Post-workshop, this group’s Gamma coefficient had increased by 0.354 to a mildly concordant value of 0.028 (Figure 3b and Table 5). While the Spring 2018 workshop participants’ Gamma pre-workshop was mildly concordant (0.111, Figure 3c), an even larger increase in Gamma (0.468) was achieved post-workshop (Table 5) to a value of 0.579 (moderate concordance).

**Student Questions**

A total of 83 participant questions, collected from pre-workshop surveys and notecards, were classified into five categories using methods described in Taylor-Powell and Renner (2003). A few sample questions are shown in Table 6.

As shown in Table 7, all workshops generated participant questions in all four topical categories with one exception. In Spring 2018, no software-related questions were present in the collected data; although, it is possible that questions about this topic were posed during the panel discussion. The most frequent topic category was Publishing (60%), followed by Intellectual Property (42%).

<table>
<thead>
<tr>
<th>Table 6. Sample student questions for each category.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Publishing</strong></td>
</tr>
<tr>
<td>Why is self-plagiarism bad?</td>
</tr>
<tr>
<td>Does it just need to be cited properly?</td>
</tr>
<tr>
<td><strong>Intellectual Property</strong></td>
</tr>
<tr>
<td>Are the data published in a paper copyrighted?</td>
</tr>
<tr>
<td>Does it belong to me, my advisor, UNLV or the journal?</td>
</tr>
<tr>
<td><strong>Research</strong></td>
</tr>
<tr>
<td>How and when should I report my failure in a paper or an external report?</td>
</tr>
<tr>
<td><strong>Software</strong></td>
</tr>
<tr>
<td>How to deal with codes shared on open blogs?</td>
</tr>
<tr>
<td><strong>Other</strong></td>
</tr>
<tr>
<td>Can you help us refine our research question?</td>
</tr>
</tbody>
</table>
Our assessment of a short and early training intervention for engineering and computer science graduate students shows promise. Our learning outcomes (Table 1) focused on improving participants’ knowledge and awareness of several ethical issues that were commonly encountered by our population of graduate students, addressing the first consideration in the sensemaking framework for ethical decision-making (Mumford et al. 2008). We detected a short-term increase in participants’ knowledge of key ethical topics addressed in the workshop. This was shown in increases in both mean overall scores (Table 3) and increases in the percentage correct answers for each question (Tables 4 and 5, Figure 3). As the workshop content and questions were designed to focus on topics and expectations that our faculty had observed to be salient issues for our students, this was a key component of our assessment, and demonstrated student learning improvements met our learning outcomes. Our data suggest that students were more aware of these key ethical expectations after completion of the workshop.

Additionally, the Gamma values analysis suggest that our workshop may have reduced over-confidence in some participants, which may lead to more willingness to consider the complexity of ethical dilemmas and to consider ethical issues from a variety of perspectives. This would address the second and third components of the sensemaking framework. in the post-workshop survey, individuals rated their own expertise (confidence) in closer alignment with their measured knowledge (awareness), yielding increases in Gamma values (Table 5 and Figure 3). This suggests that the workshop reduced both under- and over-confidence among some students. While these data are preliminary, if any students had remained over-confident (Kruger and Dunning 1999) with a lower

---

**Table 7. Numbers of Attendees (N), Number of questions Submitted, and Percentage categorization of Questions by Topical Categories for Each Workshop Group.**

<table>
<thead>
<tr>
<th>Number of Workshop Attendees (N)</th>
<th>Questions Submitted</th>
<th>Publishing*</th>
<th>Intellectual Property*</th>
<th>Research*</th>
<th>Software*</th>
<th>Other*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring 2015 (11)</td>
<td>9</td>
<td>67%</td>
<td>11%</td>
<td>44%</td>
<td>11%</td>
<td>0%</td>
</tr>
<tr>
<td>Fall 2015 (71)</td>
<td>20</td>
<td>60%</td>
<td>45%</td>
<td>30%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Spring and Fall 2016 (46)</td>
<td>18</td>
<td>72%</td>
<td>50%</td>
<td>6%</td>
<td>6%</td>
<td>6%</td>
</tr>
<tr>
<td>Spring 2017 (15)</td>
<td>5</td>
<td>40%</td>
<td>40%</td>
<td>20%</td>
<td>20%</td>
<td>0%</td>
</tr>
<tr>
<td>Fall 2017 (26)</td>
<td>15</td>
<td>53%</td>
<td>20%</td>
<td>20%</td>
<td>7%</td>
<td>33%</td>
</tr>
<tr>
<td>Spring 2018 (18)</td>
<td>16</td>
<td>56%</td>
<td>69%</td>
<td>31%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Aggregate (187)</td>
<td>83</td>
<td>60%</td>
<td>42%</td>
<td>24%</td>
<td>7%</td>
<td>10%</td>
</tr>
</tbody>
</table>

*Some questions were coded to multiple categories based on their content, so topic percentages reading across a row will sum to values exceeding 100%.
awareness of these salient ethical issues, it may have deterred those students from detecting and exploring potential ethical dilemmas during their early research experiences.

While the control group awareness and confidence data (Table 5) are limited by the small responding sample size, the discordant Gamma value is intriguing. The results may be explained by the possibility that the control group respondents included a higher proportion of experienced students who might have previously completed ethics training elsewhere, learned from experience, and/or received mentoring about ethical issues during their graduate studies. Such information was not collected.

We also found support for our decision to address the four selected research ethics topics (Figure 1). With one exception (Spring 2018), students in each workshop posed questions to the panelists in all four topic areas (Table 7), indicating engagement by workshop participants across all topics in nearly all workshops. While student motivation to ask any particular question is unknown, posing a question for the panelists implies that the student wanted the panelists to discuss the question. The lower aggregate frequency of questions in the Software category (7%, Table 7) compared to the other three topics suggests that participants found the software topic to be the least relevant, yet in all but one semester, at least one question related to this topic was submitted. This result supports our decisions to include these four topic areas and to engage colleagues from UNLV’s William S. Boyd School of Law and University Libraries as workshop designers and panelists. While not directly measured by our assessment, this suggests that the workshop design supports the second and third considerations of the sensemaking framework of ethical decision-making (Mumford et al. 2008).

We also succeeded in increasing the overall number of students exposed to those key research ethics topics and expectations. By the end of our 2015–2018 study period, 62% of the graduate students enrolled in the College had attended at least one of the workshops. This is a much higher proportion than the 1% of the College’s students completing the university provided NSF- and NIH-compliant RCR training and the 10% of students who completed the human subjects of research training during the same time period. For universities where research funding comes from a variety of sources and few research participants have funder-mandates to complete RCR training programs, this type of short intervention should be considered. Shorter and earlier interventions such as this may also be a useful step for programs that aim to eventually require the more-extensive RCR training for all enrolled graduate students. Alternatively, an introductory research ethics module could be included in college or departmental-level graduate student orientation programs (Poock 2004).

LIMITATIONS AND PROPOSED FUTURE WORK

Our workshop consisted of a single group pre-test and post-test research design that may have some risks to the internal validity of the responses (Marsden and Torgeson 2012). Respondents may
have matured in their studies of research and publishing ethics between the pre- and post-workshop surveys, thus workshop participation as an intervention may not have been entirely responsible for demonstrated learning gains and increasing Gamma values. However, the time between the pre- and post-workshop surveys ranged from 2 hours to 16 days (Table 2 and Appendix A). Thus, the matura-
tion risk for this time is low. To further minimize this risk, future post-workshop surveys and control surveys will include questions evaluating participants’ outside-workshop program or self-directed ethical learning or experiences that have occurred during the workshop period.

Participants’ memory of the pre-workshop survey questions may have also affected the internal validity of the assessment. This test effect bias (Marsden and Torgerson 2012) may have influenced what content participants paid attention to and learned during the workshop. Although this risk was reduced somewhat by not sharing the correct survey answers and not notifying students that they would be taking a post-
workshop survey, this potential bias remains. To further reduce the risk of test effect bias, a more robust experimental design will be implemented that includes continued use of a comparison (non-workshop) group as well as a revised post-workshop survey with different questions in the same content areas.

The increase in Gamma coefficient values seen after completion of the workshops in Fall 2017 and Spring 2018 is promising and worthy of further assessment. These preliminary data suggest that the workshop might address the second and third considerations of the sensemaking framework. As sample sizes increase, we can pool pre- and post-workshop data to increase sample size to determine if demonstrated levels of concordance or changes in concordance were statistically significant.

We will explore students’ willingness to discuss ethical issues with others and their intentions to learn more about research ethics using existing validated surveys, such as the Engineering and Science Issues Test (Borenstein, Drake, Kirkman and Swann 2010). Use of such instruments will also allow for comparisons with other RCR related training.

We will also further explore the student-submitted questions data using emergent coding to detect any additional themes in the student questions. In particular, we will look for evidence that students expressed disagreement with our lecture content, such as our focus on the scientific method, the inclusion of data management, and citation of computer code.

The changes we have made to the workshop beginning in summer 2018 can now be compared to the workshop design presented in this paper. These changes include our efforts to provide this workshop in a fully remote, online format starting in Fall 2020 due to the COVID-19 pandemic shift to remote learning. The lecture portion was recorded in five modules (one welcome module and one module per topic area) and completed asynchronously, followed by a live remote panel discussion.

We will use the insights from this assessment and a discussion with colleagues to refresh our learning outcomes and lecture content. We anticipate moving to a synchronous, hybrid online/in-person workshop design for maximum flexibility in the future.
CONCLUSION

We found that the short workshop design and implementation were successful in improving new engineering graduate students’ awareness of ethical research issues and conduct expectations. While our use of a classroom assessment approach and the nature of our learning outcomes make direct comparisons with other training difficult, this type of short, foundational training may be useful to other institutions providing foundational research ethics training to new graduate students. We believe that this workshop design provides a model that could be used by any graduate program seeking to provide early research ethics training to establish key expectations for responsible research conduct. We strongly recommend including lecture designers and panelists from multiple campus units where practical. Their inclusion demonstrates that many campus colleagues have expertise and bring valuable perspectives to discussions of responsible conduct of research.

ACKNOWLEDGEMENTS

We offer our sincere appreciation to all of our panelists, especially those who helped create the content for the workshop lecture; to the faculty of the Howard R. Hughes College of Engineering who supported the workshop development; to Carisa Shafer, Kari Locke, and Cheryl Bella for providing data; and to the journal’s manuscript reviewers. We also thank Shelby Smith for her improvements to the manuscript. Julie A. Longo, Technical Writer (retired) for the college, was instrumental in the design and implementation of this workshop. This work would not have been possible without the help of those at UNLV who believed that research ethics training for engineering and computer science graduate students was both needed and wanted.

REFERENCES


Engaging Everyone in Research Ethics: Assessment of a Workshop for Engineering and Computer Science Graduate Students


Engaging Everyone in Research Ethics: Assessment of a Workshop for Engineering and Computer Science Graduate Students

AUTHORS

**Susan Wainscott** is Associate Professor & Engineering Librarian at the University of Nevada, Las Vegas University Libraries. She holds a MLIS from San Jose State University and a MS in Biological Sciences from Illinois State University. As liaison librarian to several departments at UNLV, she teaches information literacy, provides reference assistance to the campus and community, and maintains the collection in assigned subject areas. Her research interests include information literacy instruction and assessment, the notion of threshold concepts, and improving access to technical literature.

**Dr. Mohamed Trabia** is a Professor of Mechanical Engineering since 2000 at UNLV. His research interests include design and optimization of mechanical systems, characterization of material properties under dynamic loading, biomaterials, system identification and control of smart actuators. Dr. Trabia has been the author of more than 200 technical journal and conference papers. He was involved in multiple funded research projects. He is a Fellow of the American Society of Mechanical Engineers (ASME).

**Dave James**, Ph.D. PE, F. NSPE, is Associate Professor, & Director of Solar and Renewable Energy Programs at the University of Nevada, Las Vegas (UNLV). His published research expertise is in water quality, dust emissions, and teaching evaluations. He teaches engineering design, research methods, and has previously taught engineering ethics, water and wastewater treatment, and air pollution control in UNLV’s Department of Civil and Environmental Engineering and Construction.