Sci Eng Ethics (2016) 22:1649–1667
DOI 10.1007/s11948-015-9725-2

ORIGINAL PAPER

Navigating Bioethical Waters: Two Pilot Projects in Problem-Based Learning for Future Bioscience and Biotechnology Professionals

Roberta M. Berry[1] · Aaron D. Levine[2] · Robert Kirkman[2] · Laura Palucki Blake[3] · Matthew Drake[4]

Received: 2 July 2015 / Accepted: 3 November 2015 / Published online: 12 November 2015
© Springer Science+Business Media Dordrecht 2015

Abstract We believe that the professional responsibility of bioscience and biotechnology professionals includes a social responsibility to contribute to the resolution of ethically fraught policy problems generated by their work. It follows that educators have a professional responsibility to prepare future professionals to discharge this responsibility. This essay discusses two pilot projects in ethics pedagogy focused on particularly challenging policy problems, which we call "fractious problems". The projects aimed to advance future professionals' acquisition of

Electronic supplementary material The online version of this article (doi:10.1007/s11948-015-9725-2) contains supplementary material, which is available to authorized users.

✉ Roberta M. Berry
robertaberry@gatech.edu

Aaron D. Levine
aaron.levine@pubpolicy.gatech.edu

Robert Kirkman
robert.kirkman@pubpolicy.gatech.edu

Laura Palucki Blake
lpblake@hmc.edu

Matthew Drake
drake987@duq.edu

1 Georgia Tech Honors Program, Georgia Institute of Technology, Armstrong, Room 005, Atlanta, GA 30332-0755, USA

2 School of Public Policy, Georgia Institute of Technology, 685 Cherry Street, Atlanta, GA 30332-0345, USA

3 Institutional Research and Effectiveness, Office of Institutional Research, Harvey Mudd College, 301 Platt Boulevard, Claremont, CA 91711-5901, USA

4 Palumbo Donahue School of Business, Duquesne University, 600 Forbes Avenue, Pittsburgh, PA 15282-3016, USA

"fractious problem navigational" skills, a set of skills designed to enable broad and deep understanding of fractious problems and the design of good policy resolutions for them. A secondary objective was to enhance future professionals' motivation to apply these skills to help their communities resolve these problems. The projects employed "problem based learning" courses to advance these learning objectives. A new assessment instrument, "Skills for Science/Engineering Ethics Test" (SkillSET), was designed and administered to measure the success of the courses in doing so. This essay first discusses the rationale for the pilot projects, and then describes the design of the pilot courses and presents the results of our assessment using SkillSET in the first pilot project and the revised SkillSET 2.0 in the second pilot project. The essay concludes with discussion of observations and results.

Keywords Fractious problems · Navigational approach · Ethics education · Problem-based learning (PBL) · Ethics assessment

Introduction

The scope of professional responsibility of bioscience and biotechnology professionals, we believe, includes a responsibility to contribute to the resolution of ethically fraught policy problems generated by their work. Those who educate future bioscience and biotechnology professionals, accordingly, should prepare them to discharge this responsibility. We discuss here two pilot projects in ethics education in which we aimed to do so, and in which we administered a new assessment instrument designed to measure the success of our effort.

The main learning objective of our pilot courses was to cultivate skills for understanding and resolving ethically fraught policy problems. We also aspired to a related secondary objective, buoyed by observations of students enrolled in our first pilot courses: to advance students' motivation to apply these skills to help their communities understand and resolve these problems.

Our courses focused on the study of particularly challenging policy problems generated by advances in bioscience and biotechnology, what we call "fractious problems" (Berry 2007; see Appendix 1 of ESM). Fractious problems are notable for the contentiousness of the debate and policymaking processes that surround them. Recent and prospective examples include policy problems associated with human embryonic stem cell research, human reproductive cloning, human enhancement technologies, neuroimaging for lie detection or to predict violence, and the creation of DNA databases for forensic identification. These problems share five characteristics that appear to contribute to the challenges they pose; in brief, they are novel, complex, ethically fraught, unavoidably of public rather than purely private concern, and unavoidably divisive.

The skills cultivated in our pilot courses were designed to support an effective policymaking approach for fractious problems, what we call a "navigational approach" (Berry 2007). Our six "fractious problem navigational" (FPN) skills are designed to address the characteristics of fractious problems; in brief, the FPN skills include: bringing multiple perspectives to bear, drawing on similar or related precedent, predicting the consequences of possible policy resolutions, generating an expansive array of possible policy resolutions, incorporating persistence in policy resolutions to allow for iterative response to change, and identifying and applying consensus principles to guide the crafting of policy resolutions.

To advance our main learning objective and our secondary aspiration for enhanced motivation, we employed problem-based learning (PBL) pedagogy, and we devoted significant effort to cultivating PBL skills for team problem solving as well. PBL skills and the PBL process they support—engaging teams of learners in actively and collaboratively striving to understand and resolve problems—are congruent with FPN skills and the navigational approach they support (see Berry et al. 2013). In addition, we intentionally composed our PBL teams of diverse, multi-disciplinary learners to support the acquisition of FPN skills, which aim to enable broad and deep understanding of fractious problems and the design of problem resolutions that effectively address their challenging characteristics.

To measure the success of our effort, we developed a new assessment instrument, "Skills for Science/Engineering Ethics Test" (SkillSET). The first version of SkillSET, administered in the first pilot project, was designed to measure gains in FPN skills. The revised SkillSET 2.0, administered in the second pilot project, was designed to measure gains both in FPN skills and in the motivation to apply them (see Appendix 2 of ESM).

We discuss, first, the rationale for the pilot projects, including the asserted scope of professional responsibility, our main and secondary learning objectives, the PBL pedagogy we employed, and the assessment instrument we designed and administered. We then describe the design of the pilot project courses and present the results of our assessments with SkillSET and SkillSET 2.0. We conclude with a discussion of our observations and results, including suggestions for future pedagogical experimentation. Appendix 1 of ESM (Four Fractious Problems) and Appendix 2 of ESM (SkillSET 2.0) are included to supplement our discussion and support future experimentation; additional course materials appear in two online ethics education repositories (Berry 2012; Beckford et al. 2013).

Rationale for Pilot Projects

The Scope of Professional Responsibility and Ethics Education

The responsibility of science and engineering professionals to adhere to professional ethical codes and rules for the responsible conduct of research is uncontroversial (see Lee et al. 2013). The time-honored virtues of professional practice and instrumental values coincide in support: the breach of a fiduciary obligation to a client is disloyal and bad for business, fraud and falsification are dishonest and undercut progress in research, the abuse of human subjects is maleficent and undermines the social support essential to the research enterprise.

Despite ample evidence of widespread ethical lapses by professionals and the energetic efforts of many educators to address these problems through better educational preparation, explicit instruction in the ethics of professional conduct was late to arrive and slow to expand (see Eisen and Berry 2002; Herkert 1999). These efforts were significantly assisted by the adoption of ethics education requirements for grantees by the National Institutes of Health (NIH) in 1989 and the National Science Foundation (NSF) in 1999, and for accredited engineering programs by ABET, Inc., effective in 2001. The institutional response to these requirements has resulted in the first extensive diffusion of ethics education across science and engineering curricula in the U.S. (see Berry et al. 2013; Eisen and Berry 2002).

Our pilot projects centered on the ethical implications of professional practice—what has been referred to as "macroethics" (Herkert 2005)—rather than the ethics of professional conduct. We believe that the professional responsibility of scientists and engineers extends to macroethics and includes responsibility for contributing to the understanding and resolution of fractious problems. The virtues of professional practice and instrumental values once again coincide in support.

A virtue of social responsibility in addressing the ethical implications of professional practice arises from the nature of the modern, socially embedded science and engineering enterprise. Science and engineering professionals possess both distinctive power to impact the society in which they practice and distinctive expertise essential to addressing the implications. The fractious problems generated by the enterprise often prompt significant social discord and, on occasion, paralyze the policymaking process; the expertise of bioscience and biotechnology professionals is essential to understanding these problems and crafting well-informed policy resolutions.

Instrumental values support this professional responsibility as well. Fractious problems threaten both the policymaking process and the bioscience and biotechnology enterprise that gives rise to the problems. Public trust in and support for the enterprise—including the provision of public funding—are premised on an expectation of social benefit; contentious and unproductive policy debates and accompanying policy gridlock may undermine that expectation and erode public trust and support.

Ethics education for scientists and engineers, thus, should acknowledge the full scope of professional responsibility and prepare future professionals to contribute to the resolution of fractious problems. Our pilot projects were one effort to prepare future professionals through a skills-based curriculum employing PBL pedagogy.

Learning Objectives: FPN Skills and the Motivation to Apply Them

Our pilot projects aimed, first, to advance the acquisition of FPN skills. The design of the pilot projects around this learning outcome was anchored in earlier work of co-author (Berry), who served as principal investigator for both projects. Berry observed that certain policy problems generated by advances in bioscience and biotechnology stimulate contentious and sometimes fruitless policy debates accompanied by policymaking dysfunction in diverse, pluralistic communities. She noted five shared characteristics of these problems—they are novel, complex, ethically fraught, unavoidably public, and unavoidably divisive—that appeared to render them "fractious" in this way. Berry proposed a "navigational approach" to

Table 1 Description of the six fractious problem navigational skills

Skill		Description
P1	Perspectives	Engage multiple perspectives to help understand the ethical, social, scientific, and technological dimensions of the problem and possible policy resolutions
P2	Precedent	Draw on analogies to similar or related problems to help understand the problem or aspects of the problem and possible policy resolutions
P3	Prediction	Consider the consequences of possible policy resolutions from the perspectives of all who might be affected
P4	Possibilities	Employ imagination and flexibility to expand the range of possible policy resolutions considered
P5	Persistence	Design policy resolutions that allow for ongoing, iterative response to emerging and evolving aspects of the problem
P6	Principles	Identify and apply consensus principles to guide the crafting of a policy resolution

addressing these problems drawing on other problem-solving approaches that have proved effective in addressing problems sharing one or more of these characteristics, including common law decision making, contextualized approaches to bioethical decision making, and negotiation (see Berry 2007, 2011a, b, Berry et al. 2013).

This navigational approach was operationalized for the pilot projects as the six FPN skills, and the focus of the educational interventions in both pilot projects was cultivation of these skills. The six FPN skills, which we number P1 through P6, are briefly summarized in Table 1.[1]

The FPN skills are designed to enable problem solvers to gain a broad and deep integrative understanding of the characteristics of fractious problems and to design effective policy resolutions for them. Consider, for example, a practice problem presented to our students at the outset of both pilot projects, "Bringing a Neanderthal to Life," (Appendix 1 of ESM, p. 1). The problem asks how a state legislature should respond to the prospect that a research team, including researchers located in the state, might clone and bring to life a Neanderthal to advance scientific understanding of differences between Neanderthals and modern humans (see Appendix 1 of ESM, p. 1; Berry 2011a, b).

Examination of diverse perspectives (P1) enables understanding of the novel, complex, ethically fraught issues and the crafting of a policy resolution that addresses the issues of unavoidably public and divisive concern, for example, regarding reproductive cloning, freedom of scientific research, experimentation involving a hominid subject, and the restoration of extinct species. Examination of precedent (P2)—such as current policies permitting cloning of non-hominid animals but not homo sapiens—also helps address these characteristics and initiate consideration of potential policy resolutions. Predicting the consequences of potential policy resolutions (P3) is also necessary: what would be the consequences for these researchers, the research enterprise, the potential Neanderthal newborn, and the community in which the Neanderthal is reared, considered from the

[1] The summary titles of the six FPN skills appear in Tables 3 and 5, below.

perspectives of those with the diverse worldviews and interests that contribute to the problem characteristics? Generating a wide range of possible resolutions (P4) that incorporate persistence (P5)—iterative response to future developments—also helps cope with the problem characteristics. For example, a policy resolution might permit researchers to proceed only to a certain point, avoiding problematic consequences while yielding valuable scientific results, and this resolution might require continuing oversight and re-evaluation as the consequences and the scientific value of continuing research come to be better understood. And anchoring a policy resolution in consensus principles (P6) that span diverse worldviews and interests will be essential to designing a good policy resolution that addresses the problem characteristics.

If our pilot courses could succeed in developing these FPN skills, would the future professionals enrolled in our pilot courses be motivated to one day deploy these skills to help their communities address fractious problems? Discharge of the professional responsibility that we posit here requires both skills and the motivation to use them. Accordingly, our secondary learning objective was to cultivate students' motivation.

We pursued this objective in the second pilot project with some expectation that we might achieve measurable success due, in part, to observations by research personnel in the first pilot project of positive motivational effects on some of the participating students (see Berry et al. 2013; Berry 2011b). The students in the first pilot project, however, were self-selected participants recruited through notices distributed across four campuses participating in the project, and the participants were advanced learners, enrolled in graduate and professional programs. Perhaps these students were predisposed to be motivated.

In the second pilot project, the student participants were undergraduates who enrolled in the course without prior knowledge of the pilot project. The vast majority were science and engineering majors seeking to satisfy distributional humanities and ethics requirements. These students were to a limited extent self-selected: they were able to select out of the course by dropping it, and those who completed the assessment instrument self-selected to participate as human subjects, receiving extra credit points in exchange for their completion of pre- and post-course assessment instruments. On balance, the participation of these students likely reflected less selection bias for a predisposition to motivation than the graduate and professional students in the first pilot project.

Would these undergraduates display positive motivational effects? We sought to advance and measure three elements of motivation: a sense of engagement with ethical and policy issues resulting from the active-learning pedagogy and the acquisition of FPN skills, an increased willingness to participate in policy problem solving due to this engagement, and increased confidence in the capacity of the policymaking process due to experience with the capacity of PBL teams to address fractious problems.

Pedagogical Approach: Problem-Based Learning

A number of innovative efforts in ethics education employing active-learning modalities, including PBL, have shown promising results (see e.g. Chang and Wang

2011; Jones et al. 2010a, b; Jonassen et al. 2009). The pedagogy employed in both of our pilot projects was facilitated PBL.

In the first pilot, we conducted two small courses, each divided into two PBL teams of four to six graduate and professional students drawn from four Atlanta-area institutions. Each team was facilitated by one and, on occasion, two faculty members from our research team. In the second pilot, we conducted a 135-student undergraduate course, divided into 20 PBL teams of six or seven Georgia Institute of Technology (Georgia Tech) undergraduate students each, with five Masters students in public policy facilitating four PBL teams each.

Our PBL pedagogy was adapted from the approach developed in the Department of Biomedical Engineering (BME) at Georgia Tech (Newstetter 2005, 2006). In Georgia Tech's BME PBL curriculum, PBL teams are presented with authentic BME problems together with an engineering model for problem solving. Following a "cognitive apprenticeship" approach to learning, facilitators provide support for the teams as they engage in the problem-solving process, and gradually recede as the novice learners on their teams become more skillful in applying the problem-solving model.

In similar fashion, we presented our pilot course students with fractious problems designed to be plausible current or imminent policy problems (see Appendix 1 of ESM; Beckford et al. 2013; Berry 2012). To enhance authenticity, PBL teams were asked to present and defend their analyses and problem resolutions before panels that included subject-matter experts, stakeholders, and those with policymaking expertise or experience (see Berry et al. 2013; Berry 2012, 2011a, 2011b). The FPN skills, P1 through P6, served as our model for problem solving. Our facilitators supported the problem-solving process, gradually receding as the novice learners on their teams advanced in their capacity to apply the FPN skills.

Our choice to employ PBL pedagogy, despite the logistical challenges of coordinating inter-institutional team meetings in the first pilot project and of coordinating facilitation of 20 PBL teams in the second, was based on our judgment of its promise for advancing our learning objectives. PBL has been found more effective than traditional approaches in advancing problem-solving skills (see Hmelo 1998; Patel et al. 1991, 1993), a fit with our emphasis on cultivating FPN skills rather than imparting a pre-specified range of subject matter knowledge. Active engagement with authentic problems also appeared well suited to developing our students' capacity to integrate perspectives, precedent, and principles spanning multiple disciplinary domains (see Newstetter 2005, 2006), a capacity demanded by the characteristics of fractious problems. And PBL has been shown effective in developing cognitive flexibility (Spiro et al. 1991); flexibility was essential if our students were to bring multiple perspectives to bear and generate multiple possible problem resolutions allowing for persistent, iterative responses to fractious problems.

We also anticipated that PBL pedagogy would be well suited to enhancing future professionals' motivation to apply their skills to help their communities cope with fractious problems. We noted that PBL has been shown to be effective in fostering the ability to work with others (see Hmelo 1998; Patel et al. 1991, 1993). We believed that team problem solving, under conditions of considerable freedom but also responsibility to arrive at and defend resolutions (see Barrows and Tamblyn

1980; Newstetter 2005, 2006), held promise for developing a sense of engagement, willingness, and confidence in the possibility of achieving good policy resolutions.

Our first pilot project entailed significant devotion of faculty effort in planning the courses, designing the course materials, and facilitating PBL team meetings, and we intentionally enrolled very few students in the two pilot courses. Our second, scaled-up pilot project benefited from the planning and course materials from the first pilot, and incorporated graduate students as facilitators of a 135-student class of undergraduates divided into 20 PBL teams, offering the promise of a scalable and sustainable model for delivering undergraduate PBL ethics education.

Assessment: SkillSET Free Response and Likert Scale Questions

In both pilot projects, we aimed to assess whether our efforts would yield measurable results in advancing our learning objectives. Our learning objectives differed from those amenable to measurement by existing assessment instruments, such as versions of the widely used Defining Issues Test (DIT), which aim to measure outcomes anchored in the work on moral development of Lawrence Kohlberg and successors (see e.g. Drake et al. 2005; Borenstein et al. 2010). For that reason, members of the first pilot project research team developed a new assessment instrument—"Skills for Science/Engineering Ethics Test" (SkillSET)—designed to capture the acquisition of FPN skills. In the second iteration of the instrument, SkillSET 2.0, employed in the second pilot project, three Likert Scale questions were included to capture improvement in students' motivation. In both pilot projects, these assessment instruments were administered to students enrolled in the pilot courses and to control students enrolled in other courses.

Both SkillSET and SkillSET 2.0 (see Appendix 2 of ESM) presented a detailed scenario involving a biomedical engineer facing a choice whether to accept a position with a biotech startup firm engaged in controversial research involving the use of human embryonic stem cells (hESCs). In the scenario, as the biomedical engineer is considering the firm's offer of employment, the work of the firm comes to light, contributing to the public controversy surrounding hESC research in the state. The scenario also indicates that legislative committee hearings have been scheduled to consider possible state legislation addressing the ethical and policy issues.

In the first SkillSET instrument, students are asked to place themselves in the shoes of the biomedical engineer in answering a series of five free-response questions. In questions 1 through 4, students are asked to summarize their job choice and situation (summarize decision), list questions they would like answered before deciding whether to accept the job (identify unknowns), indicate the most important considerations in making their choice (key considerations), and articulate the reasons why a friend might disagree with that choice (alternative views). The fifth and final question asks what the biomedical engineer would tell the state legislative committee if invited to testify (legislative testimony). The final question, in particular, is meant to elicit students' application of the FPN skills in the post-course administration of the instrument.[2]

[2] The parentheticals provide the summary titles of these questions as they appear in Table 2, below.

The second iteration of the instrument, SkillSET 2.0, includes a slightly revised version of the scenario and substitutes two free-response questions for the earlier set of five. Both questions are designed to elicit application of the FPN skills. The first asks what input the biomedical engineer would provide if testifying to the legislative committee (legislative testimony). The second asks what assessment criteria the biomedical engineer would include in an editorial assessing an imagined legislative resolution enacted in the previous month (assessment criteria).[3]

For both pilot projects, student responses were scored by teams of two scorers (for the first pilot project, one of the teams included three scorers). Each scorer independently scored each student response to each question for evidence of application of each of the FPN skills, P1 through P6, on a scale from 0 to 3. (0 = No evidence of application of skill in answering question; 1 = Minimal application of skill; 2 = Moderate application of skill; 3 = Significant application of skill). Scorers then met to reconcile any differences in scoring either by agreeing on a consensus score, or if they were only able to agree within one on the 0–3 scale, by a consensus score averaging the independent scores. All scorers were blinded to the pre-post, control-experimental conditions. In the first pilot project, all scorers were members of the research team. In the second pilot project, the scorers were graduate students in public policy who were not involved in the research project but were trained by project personnel.

In the second pilot project, three Likert Scale questions were included in SkillSET 2.0 to assess improvement in motivation. The questions aimed to capture students' engagement with science- and technology-related ethical and policy issues, their willingness to participate in policymaking efforts surrounding these issues, and their confidence in the possibility of resolving policy problems involving these issues.

Pilot Project Course Design and Results

Design of Pilot Courses

In the first pilot project, we conducted our pilot courses in the fall semesters of 2009 and 2010, with support from the National Science Foundation (NSF) Ethics Education in Science and Engineering (EESE) program (National Science Foundation 2010). In each semester, two small interdisciplinary and inter-institutional teams of four to six graduate and professional students engaged in problem-based learning (PBL), tasked with analyzing and proposing policy resolutions for two fractious problems. The PBL teams met once per week with one or two faculty facilitators drawn from our research team. Total enrollment in the two courses was 19 students.

The disciplinary domains of the students spanned science, engineering, medicine, law, the social sciences, and the humanities, and their institutional homes spanned Georgia Tech, Emory University, Georgia State University College of Law, and the

[3] The parentheticals provide the summary titles of these questions as they appear in Table 4, below.

Morehouse School of Medicine. Each PBL team was composed of students who were diverse in gender, ethnicity, disciplinary domain, and institutional home. The research personnel on the project included 24 faculty also spanning these disciplinary domains and institutions, plus faculty drawn from three additional institutions.

In the second pilot project, we conducted one pilot course in the spring semester of 2013 in a regular offering of a Georgia Tech undergraduate philosophy course, "Science, Technology, and Human Values." The course satisfies a humanities distribution requirement for Georgia Tech students and also satisfies the ethics requirements for several science and engineering majors. The course enrolled 135 students who were arrayed in 20 interdisciplinary PBL teams of six or seven students each. Students once again were tasked with analyzing and proposing policy resolutions for two fractious problems.

The students enrolled in the second pilot course spanned most science and engineering majors at Georgia Tech and several other majors as well. Each PBL team was composed of students who were diverse in gender, ethnicity, and major, including one or two non-science and engineering majors per team. Five Masters students in public policy served as facilitators, each Masters student facilitating four PBL teams meeting twice per week for 1-h sessions. An additional Masters student in public policy, who was a member of the research team, assumed primary responsibility for training the five facilitators, both in PBL and in the FPN skills. Another member of the research team, a Ph.D. student in mechanical engineering, assumed primary responsibility for an additional weekly 1-h 135-student lecture; these lectures incorporated additional exercises in the theory and application of the FPN skills. The research personnel on the second pilot project included two members of the research personnel for the first project plus three graduate students spanning public policy, mechanical engineering, and electrical and computer engineering.

In both pilot projects, students were introduced to the FPN skills by discussing the practice fractious problem, "Bringing a Neanderthal to Life," (Appendix 1 of ESM, p. 1). Each PBL team then spent 6–7 weeks preparing presentations and written reports of their analyses and policy proposals for each of two assigned fractious problems. Examples of fractious problems undertaken by PBL teams in the pilot courses include:

- Neuroimaging and Violence in the Educational Setting.
- Forensic DNA Identification.
- Sports Enhancement in U.S. Professional Sports Leagues.

(See Appendix 1 of ESM, pp. 1–3.)

Course documents, including syllabi and additional fractious problems assigned in the pilot courses, are available at two online ethics education repositories (Beckford et al. 2013; Berry 2012). More detailed discussions of the pilot project courses, fractious problems, the application of the FPN skills to the problems, student problem resolutions, and observations by research personnel appear elsewhere (see Berry et al. 2013; Berry 2011a, b).

Table 2 SkillSET results by question for pilot 1 courses

	Description	Average	Δ Pilot	Δ Control	Estimated effect
Q1	Summarize decision	0.34	0.02	−0.15*	0.16
Q2	Identify unknowns	0.41	0.08	−0.09	0.17
Q3	Key considerations	0.37	0.18	−0.08	0.26
Q4	Alternative views	0.48	0.03	−0.12*	0.15
Q5	Legislative testimony	0.46	0.03	−0.12	0.14

N = 48 students (16 pilot, 32 control)

For each question, the average score across all six FPN skills is shown. Paired t-tests were used to assess the statistical significance of the difference between pre- and post-tests for both the pilot and control groups. For the estimated effect, a difference-in-difference estimate (Δ Pilot − Δ Control) was calculated and its statistical significance assessed using an ordinary least squares (OLS) regression with an interaction term to capture the combined effect of being in the pilot group and the post-test

† $P < 0.1$; * $P < 0.05$; ** $P < 0.01$

Both pilot projects were approved and overseen by Institutional Review Boards (IRBs) at the participating institutions. Informed consent was obtained from all individual participants included in the pilot studies.

First Pilot Project Results

With respect to FPN skill acquisition captured in SkillSET, Tables 2 and 3 present the combined results of our fall 2009 and fall 2010 pilot course offerings. These tables show measured changes in SkillSET scores from pre- to post-course for students in our pilot courses and control courses (ethics-focused graduate classes taught in the same semesters at our four participating institutions). As described above, SkillSet scores range from a low score of 0 (no evidence of application of skill) to a high score of 3 (significant application of skill).

Table 3 SkillSET results by FPN skill for pilot 1 courses

	Description	Average	Δ Pilot	Δ Control	Estimated effect
P1	Perspectives	0.87	0.28	−0.30**	0.57*
P2	Precedent	0.16	−0.11	−0.05	−0.06
P3	Prediction	0.34	0.18	−0.04	0.22
P4	Possibilities	0.22	0.01	−0.10†	0.11
P5	Persistence	0.13	0.01	−0.08†	0.1
P6	Principles	0.75	0.03	−0.1	0.13

N = 48 students (16 pilot, 32 control)

For each FPN skill, the average score across all five questions is shown. Paired t-tests were used to assess the statistical significance of the difference between pre- and post-tests for both the pilot and control groups. For the estimated effect, a difference-in-difference estimate (Δ Pilot − Δ Control) was calculated and its statistical significance assessed using an ordinary least squares (OLS) regression with an interaction term to capture the combined effect of being in the pilot group and the post-test

† $P < 0.1$; * $P < 0.05$; ** $P < 0.01$

Table 2 presents results for each of the five SkillSET questions (Q1 through Q5). For each question, the average improvement across all six FPN skills for pilot students was positive but not statistically significant. Question 3, asking students as individuals contemplating accepting a job in an ethically contentious field "what would be the most important considerations in making your decision [whether to take the job]?" elicited the largest measured gains. Question 5, on the other hand, which was intended to capture the application of the FPN skills to the task of achieving a policy-level resolution—"what would you tell [the legislative committee]?"—elicited relatively little measured gain.

Table 3 presents results for the application of each of the six FPN skills (P1 though P6), calculated by averaging individual scores for each FPN skill across all five questions. Similar to the analysis by question above, there were no statistically significant gains, although all were positive except P2 (Precedent). Relatively larger suggestive, but not significant, gains were observed for P1 (Perspectives) and, to a lesser extent, P3 (Prediction).

We note that the very small number of observations in this first pilot project—involving a total of 19 students in the two pilot course offerings and 16 completed assessments—raised the bar on achieving statistically significant results on these individual pre- and post-comparisons. To help address this concern, we conducted a second analysis, focusing on the direction of changes in the SkillSET results between the pre- and post-tests. Specifically, we recorded the movement for each of P1 through P6 for each of Q1 through Q5 for pilot and control students. We then classified each of these thirty measurements as positive (if post-test $>$ pre-test) or non-positive (if post-test $\leq$ pre-test). In the pilot group, 19 of 30 measurements were positive, compared with 4 of 30 in the control group, providing suggestive evidence that the pilot course may have had a modest effect.

The results in Tables 2 and 3 show no measured gains among control students on any of Q1 through Q5 or P1 through P6 and only four of the 30 individual Q and P combinations were classified as positive. We would not expect to see gains in control classes, given that they were not designed to cultivate FPN skills. The prevalence of declines, including statistically significant declines, we believe is likely due to control students encountering lengthy assessment instruments near the end of the semester. For students whose only engagement with the exercise was the offer of an extra-credit point for completing both pre- and post-course assessments, there would be minimal motivation to write well-developed free response answers that might display the application of FPN skills.

Finally, we note that statistically significant results with respect to P1 (Perspectives) appear when we calculate an estimated effect of the pilot course, using a difference in differences estimate (see Table 3). This is due to the relatively large suggestive pilot student gains for P1 and the relatively large and statistically significant reductions among control students for P1.[4]

[4] We repeated these analyses with a larger group of control classes, including those with some ethical content in addition to those focused specifically on ethics instruction and found similar results (analysis not shown). The results were very similar, with no evidence of meaningful differences in results.

Second Pilot Project Results

With respect to our first learning objective—improved FPN skills—a revamped SkillSET 2.0 in our second pilot project asked students to answer only two free-response questions. Q1 asked students to indicate advice they would give to a legislative committee about arriving at a policy resolution, and Q2 asked students to imagine that the legislature had already arrived at a policy resolution and to indicate what assessment criteria they would include in an editorial assessing the legislature's effort. The goal for this revision was to ensure that students addressed the problem in the context for which the FPN skills were designed: designing a policy resolution.

Tables 4 and 5 present the measured changes in SkillSET 2.0 scores from pre- to post-course for pilot students and control students. Table 4 shows the results for each of the two SkillSET 2.0 questions. Statistically significant results were obtained for both Q1 ($P < 0.05$) and Q2 ($P < 0.01$). Table 5 presents the results for each of the six FPN skills. Statistically significant improvements were seen for three of the FPN skills: P1 (Perspectives) ($P < 0.1$), P4 (Possibilities) ($P < 0.01$), and P5 (Persistence) ($P < 0.01$). Non-statistically significant gains appear for the other FPN skills. In addition, the estimated effect of the pilot course, using a difference in differences estimate, was statistically significant for P2 (Precedent) ($P < 0.1$), P4 (Possibilities) ($P < 0.05$), and P5 (Persistence) ($P < 0.01$). And examination of the directional changes for the 12 measurements, P1 through P6 for Q1 and Q2, found that 11 out of 12 of these measurements showed improvement from pre- to post-tests in the pilot group, compared with 4 of 12 in the control group.

In this second pilot project, the control class was a 35-student section of the same course, with 18 control students completing the pre- and post-course SkillSET 2.0. Again, the control course offering did not aim to cultivate FPN skills, so we would not expect to see statistically significant gains, and Tables 4 and 5 indicate that there were none.

To assess the secondary objective pursued in the second pilot project—improved motivation—we included a set of three Likert Scale questions in the revamped SkillSET 2.0. These questions were designed to assess students' sense of engagement (L1), willingness to participate in addressing policy problems (L2),

Table 4 SkillSET 2.0 results by question for pilot 2 course

	Description	Average	Δ Pilot	Δ Control	Estimated effect
Q1	Legislative testimony	0.63	0.11*	−0.04	0.15
Q2	Assessment criteria	0.53	0.12**	−0.002	0.12

N = 93 students (75 pilot, 18 control)

For each question, the average score across all six FPN skills is shown. Paired t-tests were used to assess the statistical significance of the difference between pre- and post-tests for both the pilot and control groups. For the estimated effect, a difference-in-difference estimate (Δ Pilot − Δ Control) was calculated and its statistical significance assessed using an ordinary least squares (OLS) regression with an interaction term to capture the combined effect of being in the pilot group and the post-test

† $P < 0.1$; * $P < 0.05$; ** $P < 0.01$

Table 5 SkillSET 2.0 results by FPN skill for pilot 2 course

	Description	Average	Δ Pilot	Δ Control	Estimated effect
P1	Perspectives	0.76	0.12†	0.07	0.05
P2	Precedent	0.40	0.08	−0.15	0.23†
P3	Prediction	0.61	0.05	0.11	−0.06
P4	Possibilities	0.43	0.17**	−0.08	0.25*
P5	Persistence	0.47	0.22**	−0.03	0.24*
P6	Principles	0.79	0.06	−0.05	0.11

N = 93 students (75 pilot, 18 control)

For each FPN skill, the average score across all five questions is shown. Paired t-tests were used to assess the statistical significance of the difference between pre- and post-tests for both the pilot and control groups. For the estimated effect, a difference-in-difference estimate (Δ Pilot − Δ Control) was calculated and its statistical significance assessed using an ordinary least squares (OLS) regression with an interaction term to capture the combined effect of being in the pilot group and the post-test

† $P < 0.1$; * $P < 0.05$; ** $P < 0.01$

Table 6 Summary of Likert Scale questions for pilot 2 course

	Description	Mean	Δ Pilot	Δ Control	Estimated effect
L1	Importance of ethics and policy	3.76	0.11	0.06	0.05
L2	Willingness to testify	3.82	0.23	−0.11	0.34
L3	Confidence in legislature	2.63	0.15	0.0	0.15

N = 93 students (75 pilot, 18 control)

Reponses on L1 ranged from 1 (Not at all important) to 5 (Extremely Important), responses on L2 ranged from 1 (Definitely not willing) to 5 (Definitely willing) and responses on L3 ranged from 1 (Very unlikely) to 5 (Extremely likely). Paired t-tests were used to assess the statistical significance of the difference between pre- and post-tests for both the pilot and control groups. For the estimated effect, a difference-in-difference estimate (Δ Pilot − Δ Control) was calculated and its statistical significance assessed using an ordinary least squares (OLS) regression with an interaction term to capture the combined effect of being in the pilot group and the post-test

and confidence in the capacity of the policymaking process to arrive at good policy resolutions (L3). L1 asked students to rate the importance of ethical and policy issues in making their individual decision whether to take the job. This question does not go directly to motivation, but offers a baseline measure, pre- and post-course, of students' engagement with the ethical and policy issues associated with bioscience and biotechnology, a likely pre-condition to motivation. L2 asks students to rate their willingness to provide input on the ethical and policy issues to the legislative committee if asked to do so. And L3 asks students, given the circumstances surrounding the ethical and policy issues, how likely they thought it was that the legislature would succeed in "effectively analyzing the ethical and policy issues and developing a policy resolution."

Table 6 presents the results for these three Likert Scale questions and indicates modest gains for pilot students on L1, L2, and L3, with the greatest suggestive

Table 7 Assessment of inter-rater and inter-team reliability

	SkillSet (intra-team)	SkillSet 2.0 (intra-team)	SkillSet 2.0 (inter-team)
Exact match	83.9 % (7879)	65.6 % (2929)	59.7 % (1333)
Within 1	98.4 % (9237)	96.7 % (4316)	96.3 % (2150)
N	9390	4464	2232

Each cell shows the percent and number of times the two scorers on each team (intra-team) or the two teams (inter-team) reported the same score (or scores within 1) for a given measurement

evidence of gains in L2: willingness to participate in the policymaking process. Control students show a modest gain in L1 (engagement), a reduction in L2 (willingness), and no change in L3 (confidence). The difference in differences between pilot and control students shows non-statistically significant relative gains for pilot students on all three questions, with the greatest difference again in L2 (willingness).

Measurement Reliability: SkillSET and SkillSET 2.0

The value of our analysis of the efficacy of our two pilot projects with respect to the first learning objective is dependent upon the reliability of our scoring of the free-response questions in SkillSET and SkillSET 2.0. It is important to ensure that when scorers read the same student answer, both evaluate that answer as indicating the same level of competency in the application of the skills that our educational effort aims to cultivate (see Hammersley 1992). To assess the reliability of our novel instruments, we calculated both intra- and inter-team agreements. Table 7 shows the results of intra-team matches for SkillSET scoring in our first pilot project (fall 2010 course only) and for SkillSET 2.0, the instrument used in our second pilot project. Table 7 also shows the inter-team matches for SkillSET 2.0.

Within-one "matches" for intra-team scoring on the 0 to 3 scoring scale were very high, exceeding 95 % in all cases. Exact matches for intra-team scoring were more variable, however: for SkillSET nearly 84 %; for SkillSET 2.0 (averaging the intra-team matches for both teams) almost 66 %. Inter-team reliability was only assessed for SkillSet 2.0—within-one "matches" were very high, over 96 %, but exact matches between teams were just under 60 %. In general, we believe these reliability levels, particularly the strong within-one "matches," are appropriate for pilot studies of new qualitative instruments. We attribute this to reasonably specific SkillSET and SkillSET 2.0 questions, well-defined FPN skills, and the guidance of a member of the research team with expertise in assessment in designing our instrument and establishing our scoring procedures.

Discussion

Our two pilot projects yielded several interesting and promising observations and results, which, we believe, suggest the value of our PBL approach to ethics education aimed at cultivating FPN skills, and the motivation to use them, in future bioscience and biotechnology professionals. We found both suggestive and statistically significant evidence that our pilot courses fostered the ability to bring multiple perspectives to bear in addressing fractious problems and the ability to develop an expansive set of possible policy resolutions that allow for persistent, iterative response to change. We also found suggestive evidence that our approach might advance future professionals' motivation to apply the FNP skills to help their communities understand and resolve fractious problems. We developed a reliable qualitative assessment instrument for measurement of FPN skill acquisition. And we succeeded in delivering a scaled-up and, potentially, sustainable PBL approach to undergraduate ethics education using graduate student facilitators.

The most promising skill improvement suggested by the evidence in the first pilot project was in P1 (Perspectives). While PBL teams in both pilot projects were diverse in ethnicity and gender—similar to the diversity in the institutions from which they were drawn—and this may well have contributed to the improvement, we note also the very broad disciplinary diversity of students enrolled in the first pilot project courses, spanning science, engineering, medicine, law, the social sciences, and the humanities. The students in our second pilot project worked in PBL teams with less disciplinary diversity—the overwhelmingly majority were drawn from science and engineering majors—but with more direct engagement with disciplinary diversity on their PBL teams than typical either in disciplinary or ethics coursework for these majors, and these students showed modest but statistically significant improvement in P1 (Perspectives). These results suggest that collaborative learning in PBL teams composed of peer learners from diverse disciplines may be important to advancing the capacity of future professionals to bring multiple perspectives to bear in addressing fractious problems. We believe that this possibility merits further educational experimentation.

In our second, scaled-up pilot project, the statistically significant improvements in P1 (Perspectives), P4 (Possibilities), and P5 (Persistence) are the most noteworthy results. While more traditional approaches to the study of issues at the interface of science, technology, and society, might be expected to increase students' awareness of the importance of considering the predicted consequences of different policy choices (P3) and of principles in guiding the formulation of policy (P6), the distinctiveness of the navigational approach operationalized in the FPN skills appears to be reflected in these results. To the extent that the theoretical case for the value of incorporating multi-perspectival, flexible, and iterative approaches to problem solving in ethics education is persuasive, the results suggest the value of the PBL modality, which has been found particularly well suited to promoting cognitive flexibility in addressing complex problems (see Spiro et al. 1991).

The value of any pedagogical effort to prepare future bioscience and biotechnology professionals to discharge their professional responsibility is ultimately dependent upon their motivation to do so. Whether educators can exert an influence on this motivation, and, if so, whether we can hope to measure the effects of that influence are difficult to know. To the extent our Likert Scale questions in SkillSET 2.0 succeed in measuring that influence, we found only suggestive evidence for the efficacy of our approach, including especially encouraging suggestive evidence with respect to willingness to engage with bioethical policymaking. Given the importance of motivation—both to the genuine integration of professionalism into the lives of future professionals and to the realization of the benefits of this by their communities—further research on pedagogical efficacy in influencing motivation is clearly needed.

Success in cultivating the FPN skills, our first learning objective, was not amenable to assessment by versions of the DIT, which aim to measure improvement in learning outcomes different from ours. Accordingly, we developed a new assessment instrument, SkillSET and SkillSET 2.0, and we found this new instrument to be reliable. We recognize the important role of continuing pedagogical efforts aimed at learning outcomes measurable by the DIT, and we also encourage experimentation and assessment anchored in other conceptions of the aims of ethics education. More extensive discussion of the value of the FPN skills, and the navigational approach they are intended to operationalize, appears elsewhere (see Berry et al. 2013; Berry 2007, 2011a, b).

Finally, we observe that it is possible, even if sometimes challenging, to deliver bioethics education employing PBL, both in the small post-graduate classroom and in the large, scaled-up undergraduate class setting. Scaling up with the efficiency essential to sustainability, we note, is possible with the engagement of graduate students as facilitators. Given the potential of active-learning modalities such as PBL to advance learning outcomes of the kind we and others have pursued (see Chang and Wang 2011; Jones et al. 2010a, b; Jonassen et al. 2009), we hope that experimentation in both small and scaled-up PBL ethics education as well as other active-learning modalities will continue.

In sum, we believe the results of our two pilot projects offer good reason for continued experimentation with and refinements of our approach and similar active-learning approaches to ethics pedagogy, with the reasonable expectation that these approaches might advance the problem-solving skills and, we hope, the motivation of future bioscience and biotechnology professionals. In light of the professional responsibility of these professionals to contribute to understanding and resolving the challenging policy problems generated by their enterprise, it is incumbent on educators to discover and design effective pedagogies to prepare them to do so, and effective assessment instruments to measure our success.

Acknowledgments We acknowledge support in conducting these pilot projects and preparing this article from a National Science Foundation (NSF) Ethics Education in Science and Engineering (EESE) grant, NSF EESE Award ID 0832912 (PI: R.M.B.) and additional support in preparing the article from a NSF Faculty Early Career Development (CAREER) grant, NSF DRL Award ID 1150114 (PI: A.D.L.). We acknowledge additional support for the second pilot project from the School of Public Policy, Georgia Institute of Technology. Any opinions, findings, and conclusions or recommendations set forth in this article are those of the co-authors and do not necessarily reflect the views of these funders. We acknowledge the efforts of the NSF EESE Research Team for the first pilot project: John D. Banja, Gillian Hue Beckford, Roberta M. Berry, Laura Palucki Blake, Jason Borenstein, Robert J. Butera, Lara Denis, Matthew Drake, Martha L. Elks, Kathy Kinlaw, Robert J. Kirkman, Michelle Lampl, Aaron D. Levine, Paul A. Lombardo, Patricia Marstellar, Robert M. Nerem, Wendy C. Newstetter, Douglas F. Paulsen, Edward L. Queen II, Mark Risjord, Charity Scott, Lisa A. Tedesco, Keith D. Wilkinson, Leslie E. Wolf. We thank Wendy C. Newstetter, a member of the NSF EESE Research Team, in particular, for her assistance in operationalizing the "navigational approach" in the form of the "fractious problem navigational" (FPN) skills. We also extend thanks to the following members of the NSF EESE research team for their efforts in scoring the results of the first pilot project: Roberta M. Berry, Laura Palucki

Blake, Jason Borenstein, Robert J. Butera, Kathy Kinlaw, Robert J. Kirkman, Aaron D. Levine, Wendy C. Newstetter, Edward L. Queen II, Leslie E. Wolf. We thank Ethan Butler for his extensive and invaluable research assistance in support of the work reported here for both pilot projects. We also thank both of the following for their very helpful efforts as graduate research assistants for the NSF EESE pilot project: Alexander M. Smith, William Edward Staley, Jr. We acknowledge the efforts of the Science, Technology, and Human Values Research Team for the second pilot project: Roberta M. Berry, Laura Palucki Blake, Ruchir N. Karmali, Sharon E. Norman, Jason L. Wang. We extend thanks to the following for their efforts in scoring the results of the second pilot project: Caroline H. Appleton, Kathryn J. Kline, Stephanie C. Noble, Carter M. Parker. We also thank the following for their efforts in facilitating the Science, Technology, and Human Values PBL teams: Christopher Y. Bellew, Bekim Haliti, Susan Kibler, Hannah T. Santoro, Claire E. Woodring.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

Barrows, H. S., & Tamblyn, R. (1980). *Problem-based learning: An approach to medical education*. Springfield, IL: Problem-based Learning Institute.

Beckford, G. H., Berry, R. M., Queen, E. L., II, Kinlaw, K., Newstetter, W. C., & Wolf, L. E. (2013). *Problem based learning (PBL) course addressing 'fractious problems' in science and technology*. National Academy of Engineering: Online Ethics Center for Engineering and Science. http://www.onlineethics.org/Resources/TeachingTools/Modules/27534.aspx.

Berry, R. M. (2007). *The ethics of genetic engineering*. New York: Routledge.

Berry, R. M. (2011a). A small bioethical world? *HealthCare Ethics Committee Forum, 23*(1), 1–14.

Berry, R. M. (2011b). Teaching health law: Problem-based learning regarding "fractious problems" in health law: Reflections on an educational experiment. *Journal of Law, Medicine and Ethics, 39*(4), 694–703.

Berry, R. M. (2012). *NSF EESE interdisciplinary PBL course on "fractious problems"—course materials*. National Center for Professional and Research Ethics: EthicsCORE-Resources. https://nationalethicscenter.org/resources/808.

Berry, R. M., Borenstein, J., & Butera, R. J. (2013). Contentious problems in bioscience and biotechnology: A pilot study of an approach to ethics education. *Science and Engineering Ethics, 19*(2), 653–668.

Borenstein, J., Drake, M. J., Kirkman, R., & Swann, J. L. (2010). The engineering and science issues test (ESIT): A Discipline-specific approach to assessing moral judgment. *Science and Engineering Ethics, 16*(2), 387–407.

Chang, P.-F., & Wang, D.-C. (2011). Cultivating engineering ethics and critical thinking: A systematic and cross-cultural education approach using problem-based learning. *European Journal of Engineering Education, 36*(4), 377–390.

Drake, M. J., Griffin, P. M., Kirkman, R., & Swann, J. L. (2005). Engineering ethical curricula: Assessment and comparison of two approaches. *Journal of Engineering Education, 94*(2), 223–231.

Eisen, A., & Berry, R. M. (2002). The absent professor: Why we don't teach research ethics and what to do about it. *American Journal of Bioethics, 2*(4), 38–49.

Hammersley, M. (1992). *What's wrong with ethnography?: Methodological explorations*. London; New York: Routledge.

Herkert, J. R. (1999). *ABET's engineering criteria 2000 and engineering ethics: Where do we go from here?* Paper presented at the OEC International Conference on Ethics in Engineering and Computer Science, March 1999. National Academy of Engineering: Online Ethics Center for Engineering and Science. http://www.onlineethics.org/Education/instructessays/herkert2.aspx.

Herkert, J. R. (2005). Ways of thinking about and teaching ethical problem solving: Microethics and macroethics in engineering. *Science and Engineering Ethics, 11*(3), 373–385.

Hmelo, C. E. (1998). Problem-based learning: Effects on the early acquisition of cognitive skill in medicine. *Journal of Learning Sciences, 7*(2), 173–208.

Jonassen, D. H., Shen, D., Marra, R. M., Cho, Y. H., Lo, J. L., & Lohani, V. K. (2009). Engaging and supporting problem solving in engineering ethics. *Journal of Engineering Education, 98*(3), 235–254.

Jones, N. L., Peiffer, A. M., Lambros, A., & Eldridge, J. C. (2010a). Problem-based learning for professionalism and scientific integrity training of biomedical graduate students: Process evaluation. *Journal of Medical Ethics, 36*(10), 620–626.

Jones, N. L., Peiffer, A. M., Lambros, A., Guthold, M., Johnson, A. D., Tytell, M., et al. (2010b). Developing a problem-based learning (PBL) curriculum for professionalism and scientific integrity training for biomedical graduate students. *Journal of Medical Ethics, 36*(10), 614–619.

Lee, L. M., Viers, H. W., & Anderson, M. A. (2013). The Presidential Bioethics Commission: Pedagogical materials and bioethics education. *Hastings Center Report, 43*(5), 16–19.

National Science Foundation. (2010). *Ethics education in science and engineering (EESE)*. http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=13338.

Newstetter, W. C. (2005). Designing cognitive apprenticeships for biomedical engineering. *Journal of Engineering Education, 94*(2), 207–213.

Newstetter, W. C. (2006). Fostering integrative problem solving in biomedical engineering: The PBL approach. *Annals of Biomedical Engineering, 34*(2), 217–225.

Patel, V. L., Groen, G. J., & Norman, G. R. (1991). Effects of conventional and problem-based medical curricula on problem solving. *Academic Medicine, 66*(7), 380–389.

Patel, V. L., Groen, G. J., & Norman, G. R. (1993). Reasoning and instruction in medical curricula. *Cognition and Instruction, 10*(4), 335–337.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1991). Knowledge representation, content specification, and the development of skill in situation-specific knowledge assembly: Some constructivist issues as they relate to cognitive flexibility theory and hypertext. *Educational Technology, 31*(9), 22–25.