Developing the Preparation in STEM Leadership Programs for Undergraduate Academic Peer Leaders

Stacey Blackwell, Sari Katzen, Nipa Patel, Yan Sun, Mary Emenike
Rutgers, The State University of New Jersey

Author Notes

Stacey Blackwell, Sari Katzen, Nipa Patel, Learning Centers (New Brunswick Campus); Dr. Mary Emenike, Department of Chemistry and Chemical Biology; Yan Sun, Graduate School of Education

Correspondence concerning this article should be addressed to Stacey Blackwell, Learning Centers, Rutgers University, 151 College Ave., New Brunswick, NJ 08901. Email: stacey.blackwell@rutgers.edu; Phone: (848) 445-0984.

Abstract

The authors introduce the Preparation in STEM Leadership Program at Rutgers, The State University of New Jersey. This NSF-Funded program and research study creates a centralized training program for peer leaders that includes a battery of assessments to evaluate peer leaders' content knowledge, pedagogical knowledge, communication skills, and leadership practices over time. This article introduces readers to the program and its theoretical background, explains some lessons learned regarding the program design and implementation, and briefly describes preliminary findings on a broad-scale from peer leaders' assessments intended to measure skills and content development.

Introduction

In any given college or university, there often exists a variety of academic peer leadership opportunities available to students. Academic peer leadership positions for undergraduate students range from tutoring in one-on-one or group settings to facilitating small-group study sessions, leading large review sessions, or providing in-class learning support. Because each program has its own place within a campus' academic community, faculty and staff interested in implementing academic support programs will choose one or more programs to address the diverse needs and wants of their students and instructors, while considering their access to spaces and budgetary requirements. All of these factors result in significant variation in the training, evaluation, and instructional methods of the peer leaders.

Although Rutgers University employs many peer leaders through its various academic support programs, there are no standard expectations for the amount, type, or frequency of training programs for peer leaders. We propose that employment in peer leader positions provides peer leaders with the opportunity to develop their content knowledge and prepare for success in the STEM workforce; however, we also suppose that the lack of consistency across programs leads to variability in the value of leadership positions to students' professional development. Through the Preparation in STEM Leadership (PSL) Program, we have created an opportunity for peer leaders from various programs to participate in a standard training program and have developed a series of assessment measures to study their content knowledge, pedagogical knowledge, leadership practices, and communication skills. Through the implementation and assessment of this program, we intend to identify the elements of peer leader training that are particularly impactful and the specific areas of skill and knowledge that are enhanced through employment as a peer leader.

If successful, the benefits of this program will reach beyond impacting the career opportunities of participating peer leaders. Because peer leaders in the PSL Program will receive increased training in content knowledge, communication skills, and research-based instructional strategies, a logical outcome would be that the students they serve benefit by experiencing increased learning gains (Figure 1).

In this article, we introduce the PSL program model and provide examples of assessments that can be integrated into the practices of peer leader training programs. We provide some initial results as samples of the type of data that can be collected from these assessment practices. We also discuss lessons learned and remaining questions. Future publications will include extended analyses of the assessment results as well as discussions of how such measures are being used to inform decisions about the structure, content, and format of our training program.

Relevant Literature

Research has uncovered "a number of positive effects of peer education on student success," such as "increased satisfaction, persistence and retention, social development, and academic performance" (Ganser & Kennedy, 2012, p. 17). Kuh (2008) highlighted a set of "high impact practices" that result in "substantial educational benefits" (p. 1), and the positive results of these practices are particularly notable for underprepared and historically underserved populations. Kuh proposes teaching and learning practices that "have been widely tested and have been shown to be beneficial for college students from many backgrounds" (p. 9) are effective because they involve one or more of the following characteristics:

- demand considerable time and effort,
- place students in situations that demand interaction with faculty and with one another,
- increase the likelihood of contact with people from diverse backgrounds,
- involve frequent feedback,
- provide opportunities for higher level learning, and
- deepen the undergraduate experience through increased self-awareness (p. 14-17).

When part of a well-designed program, peer leader positions involve all of these characteristics. They demand considerable time and effort through training, lesson planning, and instructional time. The peer leaders meet regularly with peers and instructors of diverse backgrounds and receive frequent feedback from their supervisors, peers, and instructors. They are also taught to engage their students in higher-level learning and metacognitive practices, which requires the peer leaders themselves to engage in such practices as they prepare lessons and discuss course activities with their fellow peer leaders and course instructors.

As universities focus on educating students for success in the STEM workforce, they are increasingly searching for ways to help students develop 21st century skills because individuals in STEM fields "must be able to adapt to new work environments, communicate using a variety of mediums, and interact effectively with others from diverse cultures" (Koenig, 2011, p. viii). In *Education for life and work: Developing transferable knowledge and skills in the 21st Century*, Pellegrino & Hilton (2012) define 21st century skills as fitting into three primary categories: cognitive, intrapersonal, and interpersonal (p. 21). Peer leadership promotes development in all three of these categories.

Peer leadership training, when implemented according to best practices, emphasizes the "application of knowledge, skills, and responsibilities to new settings and complex problems" (Shook & Keup, 2012, p. 10). The emphasis of these applications leads to the application and development of "skills and capabilities such as self-direction, leadership, oral communication, intercultural skills, civic engagement, teamwork, and critical thinking," all of which are "identified as twenty-first century learning objectives for college and that are also highly desirable skills among employers" (Shook & Keup, 2012, p. 10). Students who undergo such training and serve in peer leader positions have reported "increased confidence in their ability to manage group dynamics, facilitate learning, and empathize with their students" (Shook & Keup, 2012, p. 10), skills that would equip them to become leaders in the STEM workforce. Peer leaders also learn to address "real-world," ill-defined problems that "require multiple areas of knowledge and multiple modes of inquiry" (Shook & Keup, 2012, p. 11). Moreover, Tingson-Gatuz (2009) found that several studies have demonstrated growth in critical thinking, problem solving, and group processing and that peer mentoring opportunities have the potential to "increase leadership capacity among students-of-color" (p. 3) and "after graduation, these peer leaders can engage in higher levels of leadership both in professional and community capacities" (p. 87). We anticipate that peer leadership positions specifically focused on academic discourse would have similar impacts on critical thinking, problem solving, and group processing for students-of-color, who have historically been under-represented in STEM majors and, therefore, in leadership positions within the STEM workforce.

Research findings have revealed that "neither success nor sustainability can be attained in a peer leadership program without thoughtful and intentional planning, management, and training" (Esplin, Seabold, & Pinnegar, 2012, p. 85). Therefore, we believe that creating a standard for peer leader training that can be assessed rigorously and implemented across multiple programs will have far-reaching effects for peer leaders, the students they serve, and academic support services as a whole. Successful training programs involve intensive initial preparation with ongoing training and utilize methods that are "as hands-on, applicable, and engaging as possible" and incorporate "pedagogical and leadership theories, models, approaches, and research findings" (Esplin et al., 2012, p.94), which are already part of our pedagogy course and our ongoing trainings. Both the course and the training workshops engage students in practical applications of learning theory specific to their individual roles and fields of study, as well as introduce them to educational research through readings and group discussions. Our training programs are structured with an iterative cycle that involves continuous reflection on teaching and learning, as well as gradually increased complexity. It can be envisioned as a type of spiral that involves revisiting and expanding on previous learning and experience using theory and educational research, as demonstrated in Figure 2. Steps 1-3 take place during our initial, intensive Peer Leader Training program (two days) and are revisited with more advanced topics throughout the pedagogy course and ongoing workshops, which both expand on the initial training by including steps 4-7.

Two national models of academic support programs that have specifically addressed peer leader training in their theoretical design are Supplemental Instruction (SI; Stone & Jacobs, 2006) and the Learning Assistant (LA) Model (Otero, Pollock, & Finkelstein, 2010). While both the SI and LA models include pedagogical and content components, only the LA model explicitly requires these components. In SI implementations, one or both of these components might not be included, which is evidenced by SI developer Martin's quote: "the single most common reason for the failure of an SI program is the lack of consistent training and supervision for the leaders" (Burmeister, 1996, p. 33). While the SI model allows for interpretation and modification of the model, the LA model is more restrictive, requiring specific elements such as participation in a sustained pedagogy course or seminar for at least one semester while serving as a learning assistant and consistent content meetings with the course instructor(s).

Institutional Context

Rutgers University is a land-grant R1 institution that serves both New Jersey residents and international students. The university currently enrolls more than 65,000 students from all 50 states and more than 115 countries. More than half of Rutgers University students identify as non-Caucasian and more than 80% receive financial aid, making Rutgers University a diverse campus both culturally and socioeconomically. Because of its diverse student population, Rutgers University provides many programs that promote the retention and success of students from underrepresented minority groups in STEM fields, and the Learning Centers (LCs) have longstanding partnerships with these programs. The Rutgers -- New Brunswick campus offers STEM degrees through multiple schools within the university. The research opportunities, the large and increasing number of academic support programs, and the many leadership opportunities available to students at Rutgers University make this institution an ideal site for the assessment of peer leader development opportunities.

Rutgers University's Learning Centers

The LCs provide four core service programs, including academic coaching, writing coaching, walk-in group tutoring, and the Learning Assistant (LA) Program, which follows the LA Model (Otero et al., 2010). Peer Leader Training, our fifth core program, is provided to our own student employees and to other programs that utilize peer leaders, such as group and one-on-one tutors, supplemental instruction leaders, peer mentors, teaching interns, and study group leaders. Training requirements vary among the different peer leader positions on campus. Several programs require their peer leaders to attend the New Peer Leader Training sessions offered by the LCs at the beginning of the semester and a few programs request specific training from the LCs for their peer leaders during the semester, while other programs do not utilize any of the trainings offered at the LCs.

Pedagogy Courses

The LCs' LA Program includes a required 300-level, 3-credit pedagogy course for all first-semester learning assistants. The course focuses on effective methods of college teaching and instructional strategies. Students participate in activities designed to increase their understanding of the role of a peer educator. The instructional strategies taught in this class are grounded in principles of student-centered, active, cooperative learning environments and differentiated classroom instruction. Through this course, students learn how to work with course instructors and teaching assistants to facilitate small-group learning among students in their lectures and recitations, lead study groups (which involves designing student-centered activities and facilitating small group learning), and help individuals and small groups of students during tutoring sessions by engaging in dialogic discourse and utilizing effective questioning techniques.

Peer leaders teaching in Rutgers University's First-Year Interest Group Seminars (FIGS) are required to take a 300-level, 3-credit pedagogy course that is related to their role as peer instructors. Although the LA and FIGS pedagogy courses share a common name and similar course number, they are distinct courses that are structured differently to meet the unique needs of the peer leaders in each program. Undergraduate teaching assistants in general chemistry laboratory courses are required to co-enroll in a 400-level, 3-credit course associated with their position. The course's goals for students are to develop teaching, supervisory, organizational, and communication skills by teaching in the laboratory ("Chemistry 499," n.d., para. 2).

The general chemistry Teaching Interns (TI) Program includes an optional pedagogy course as part of a larger program that culminates in a "Certificate of Chemistry Education." The chemistry interns' 300-level, 3-credit pedagogy course was modeled after the

LA Program's pedagogy course and has some similar structure and assignments (flipped classroom structure, in-class activities, peer observation, midterm exam, teaching philosophy paper). However, the chemistry course is specific to the discipline of chemistry (mostly general chemistry topics) and includes TIs' time working with students as part of the students' time spent in class.

PSL Program Design

Any undergraduate student who had at least completed his/her first year of college and who had secured a peer leader position related to one or more STEM disciplines was eligible to apply to participate in the PSL program. The term "peer leader" includes, but is not limited to, tutors; study group leaders, supplemental instruction, or review sessions; undergraduate teaching assistants (TAs), teaching interns, or peer mentors; and learning assistants.

Eligible peer leader positions required the facilitation of learning of STEM content because the program included a strong pedagogical component to the training, and we were interested in assessing learning gains in disciplinary content. Consequently, students whose positions were solely associated with general mentoring, counseling, or academic coaching were not eligible to participate in the PSL program. Students applied over the summer and participants were chosen from among those who applied based on year in school, program affiliation (to ensure representation from as many peer leader programs as possible), and short answer responses in the application form that asked the candidate to describe his/her interest in peer instruction, interest in leadership, career goals, a statement of teaching philosophy, and potential for growth through participation in the PSL program. The general timeline of activities is provided in Table 1.

All first-year participants in the PSL program were required to attend the New Peer Leader Training days at the beginning of the fall semester and co-enroll in the 300-level, 3-credit pedagogy course that, up until this point, was only offered to LAs. By expanding this course to include participants in the PSL program, the LCs provided an opportunity for peer leaders across programs to receive in-depth training in pedagogy and research-based instructional strategies.

Those peer leaders who enrolled in the PSL program were also required to participate in a minimum of four 80-minute training or professional development workshops related to pedagogy and leadership each semester (example topics provided in Figure 2). They were also encouraged to attend a seminar each semester offered by a professional in the STEM industry, but attendance was not mandatory for those who had a scheduling conflict. All first-year PSL Program participants who successfully completed the minimum requirements earned a $500 scholarship in both the fall and spring semesters.

Assessment and Evaluation

To measure the effects of the PSL Program on the peer leaders' development across several dimensions, pre- and post-semester data was collected by asking the peer leaders to respond to several assessment tools measuring content development, leadership practices, pedagogical knowledge, and communication skills. Peer leaders were also asked to submit a teaching philosophy statement, and to be formally observed in their position. The assessment instruments are provided in Table 2. Whenever possible, discipline-specific assessments (e.g., concept inventories or concept mapping) were used to measure content development. Teaching philosophy statements were collected upon application to the PSL program. These statements were also collected as a summative assessment measure within the pedagogy course. All protocols were approved by Rutgers' Institutional Review Board, and only data from students providing informed consent are included in this report.

During training sessions, the pre-assessments were administered in the following order: 1) teaching and learning concept map, 2) discipline-specific content map, 3) concept inventory, and 4) beliefs about physical sciences survey. Specific assessment tools with references can be found in Table 2. Unfinished surveys were completed independently under the supervision of the Program Coordinator by the third week of September. The beliefs about physical sciences survey, the Student Leadership Practices Inventory (LPI), COMSA-R2 communication survey, and the submission of the teaching philosophy statements were all completed online. This combination of assessments was designed to use two methods to evaluate students' content knowledge (concept maps and inventories) and ideas about teaching (concept maps and teaching philosophy statements). The multiple assessment tools complement one another and enable students to demonstrate their knowledge using different approaches to account for variance in learning styles. The LPI and COMSA were administered to evaluate changes in 21st Century skills (group management, communication, and interpersonal skills) needed in increasingly collaborative work environments within the STEM fields.

The peer leaders' formal evaluation occurred between the 5th and 15th week of the semester in both fall and spring semester (Tables 1 & 3). By the 5th week of the semester, most courses for which peer leaders provide a service have administered the first exam to the students, and peer leaders have had time to develop a rapport with the students they serve.

Data was collected from a control group of students serving as peer leaders who were not enrolled in the PSL program. This data was collected to help distinguish between the impact of serving as a peer leader and the impact of the training provided for peer leaders. We assumed that some learning gains would result from working in a peer leader position, regardless of whether or not the peer leader participated in regular training and professional development practices. We intended to explore 1) the learning gains experienced through service as a peer leader regardless of training, 2) the learning gains experienced as a result of rigorous and structured training and professional development practices, and 3) the difference in outcomes between those peer leaders participating in regular training and those not participating. The control group for this study was larger than the treatment group because assessments were incorporated into the regular training practices of the LCs, and, consequently, any peer leader could choose to participate in the research study. Because of the resources required to implement the PSL program (e.g., scholarships, pedagogy instructors), the PSL program was limited to 30 participants each semester. However, the cohorts consisted of fewer than 30 students because of the significant time commitments from students; most peer leaders are full-time students and have at least one other job on campus

Cohort Participants

We invited other peer leaders employed at our LCs and in other programs to participate in the research study. These control group peer leaders were compensated through human subjects payments for their time completing assessments if they agreed to participate in the research study. Of the 324 non-PSL Peer Leaders for whom we collected one or more assessment measures during the past two years, 208 (64%) provided informed consent. Of the 37 PSL participants, 34 (92%) provided informed consent for the research study. Only those students providing informed consent are included in the data and analysis discussed herein.

In the first year of the PSL Program, 24 students participated in the PSL Program, eight of whom were learning assistants employed at the LCs. The other participants had positions as tutors, supplemental instructors, study group leaders, calculus peer mentors, undergraduate teaching assistants, and chemistry teaching interns. Both our first and second cohorts of PSL participants included a range of disciplinary assignments, not all of which aligned with each participant's major. Of the 34 PSL participants, 56% self-identified as female, while 41% self-identified as male and 2% chose not to identify, 30% were second-year students, 40% were third-year students, and 30% were students in their senior year (see Table 3).

Data and Results

Overview of Data Collected

All PSL participants completed assessments and were observed. However, participants could choose whether or not they would allow their data to be included in analysis for public dissemination. Not all types of data resulted in the same response rate.

As an example, during the fall 2016 pre-assessment period, the highest assessment responses from all peer leaders were for the teaching and learning concept map (N=198) and disciplinary concept maps (N=121). The concept inventories (N=90) and beliefs about learning science survey (N=76) had moderate response rates, but it is important to remember that these surveys were not available or applicable for all disciplines. The online leadership practices inventory (N=25) and communications survey (N=47) had relatively low response rates, while the teaching philosophy statement response rate was high for peer leaders in the pedagogy course because it was a required assignment (N=120), but low for all other peer leaders (N=0).

Due to the variance within peer leader positions at the university, some are observed on a regular basis as part of the program's policy and others are not. Of the peer leaders included in this study (both PSL and control group), most were observed (PSL N=7). For those PSL participants who were not observed (N=6), the observation did not occur because of the nature of the position (e.g., no interaction between peer leader and students) or because of scheduling constraints. Anecdotally, PSL participants who normally would not be observed in their position expressed appreciation for the opportunity to receive feedback.

Preliminary Findings

We are in the process of collecting data from year 2 participants and analyzing the data we have collected from both year 1 and year 2 cohorts. We will disseminate our findings through several publications once the program pilot has been completed. Below are some sample findings to demonstrate the type of information we are collecting and how that information is used to inform the direction of our existing training program.

Content Knowledge. One example of evaluating content knowledge involves administering concept inventories. Best practices for using concept inventories include administering the entire test with the same order of questions in their original wording both at the beginning of the semester (or before direct instruction on the topic) and again at the end of the semester (or after students have been formally tested on the topic through a course exam) (Madsen et al., 2016). Physport's data explorer was used to analyze physics concept inventory results for matched student data (www.physport.org/DataExplorer/Preview.cfm). In the first year cohort, 15 peer leaders completed Thornton and Sokoloff's (1998) Force and Motion Conceptual Evaluation (FMCE) concept inventory at both the beginning and end of the fall semester. The FMCE measures students' conceptual understanding of Newtonian mechanics, including kinematics, forces, energy, and graphing. These students' average score increased from 65% ± 6% on the pre-test to 76% ± 6% on the post-test, with an average normalized gain of 0.40 ± 0.10. A students' normalized gain score is calculated by dividing the difference in points on the post and the pre-test by the potential points the student could improve upon after the pre-test (i.e., 100% - pre score). Seven peer leaders completed the FMCE at the beginning of the fall semester and also at the end of the spring semester. For this year-long matched sample, the students' average score increased from 65% ± 6% on the pre-test to 82% ± 8% on the post-test, with an average normalized gain of 0.24 ± 0.15. We are encouraged by these results for first-semester physics concepts, which indicate students are improving their content knowledge during the time they are participating as peer leaders.

Leadership Practices. The Student Leadership Practices Inventory (Kouzes & Posner, 1998) results revealed some notable trends in peer leaders' self-perceptions of their leadership practices. This inventory requires students to rank the frequency of their use of specific leadership behaviors on a Likert scale of 1-5. These behaviors correspond to five leadership practices. The "Model the Way" practice is to "clarify values by finding your voice and affirming shared values" (The Student Leadership Challenge, 2017). It also refers to practices that "set the example by aligning actions with shared values." "Inspire a Shared Vision," refers to practices that "envision the future by imagining exciting and ennobling possibilities" or "enlist others in a common vision by appealing to shared aspirations." "Encourage the Heart," is to "recognize contributions by showing appreciation for individual excellence" or to "celebrate the values and victories by creating a spirit of community." "Enable Others to Act," refers to practices that "foster collaboration by building trust and facilitating relationships" and "strengthen others by increasing self-determination and developing competence." Finally, "Challenge the Process" refers to practices that "search for opportunities by seizing the initiative and by looking outward for innovative ways to improve" or "experiment and take risks by constantly generating small wins and learning from experience." The total points for each practice are calculated as the student's score for each practice, which can range from 6 (responding with a ranking of one for each of the six items associated with that category) to 30 (responding with a ranking of five for each of the six items associated with that category). Peer leaders' scores for each practice were collected, as well as their rankings for each behavior. Here we address the general trends of the group as a whole; detailed analyses by practice and behavior will be discussed in a subsequent publication. This inventory was not intended to be evaluative, in that no one category is necessarily more desirable than another; rather, our intention is to identify trends in how peer leaders' practices may change in response to their experiences as peer leaders, and the training they receive, to better understand how such positions may impact the type of leadership practices these students value and utilize.

Figure 4 shows the mean scores from all students for the five practices of the LPI across three-waves of data (beginning of fall, end of fall, and beginning of spring semesters). The plot shows a general pattern that students achieved a higher score in the mid-test than the first wave and then decreased score at the post-test. Splitting data by group, we can track pattern differences across groups (Figure 5). The patterns of the PSL group and control group are not identical; for example, for "Model the way", the control group (solid line) tended to decrease at the end of the school year, while PSL group (dash line) tended to increase. The same pattern was observed in the "Inspire a shared vision" and "Challenge the process" practices. The decrease in scores for the post-test has led to additional research questions. For example, because this inventory relies on self-report, the decrease could indicate increased self-awareness of one's practices or a change in how peer leaders understand each behavior to which they responded. If, through their training and practice as a peer leader, they are becoming increasingly aware of improvements they can make in their practices, they may report lower scores on the post-test than the pre-test simply because their expectation for the particular behavior has been increased or their understanding of the behavior has been refined. Additional qualitative data such as interviews or follow-up surveys would be needed to better understand the changes in self-reported behaviors.

Gender had a marginally significant effect on "Model the way;" females generally reported higher than males. Two significant effects were found for "Challenge the process," the main effect of wave and the interaction of wave by group. The relationship of the two effects has been delineated in Figure 4; although at the beginning, control groups had a higher average score than the PSL group, at the post-test, the PSL group outperformed the control group. For the remaining two practices, no significant effect was found (Table 4) via an analysis of variance (ANOVA) conducted using SPSS (version 24).

Pedagogy. Concept maps and teaching philosophy statements were used to study changes in peer leaders' ideas about teaching and learning before and after a semester of peer leader training. As a preliminary assessment of pedagogical change over time, we coded thesis statements within the teaching philosophies by topic and compared the frequency of each topic in pre-semester and post-semester drafts. Identifying the changes that the peer leaders made to their thesis statements over the course of the semester could indicate changes in how they value each topic and find it relevant to their work as peer leaders. All coding was managed using QSR International's NVivo 10 Software.

We were able to collect both pre-semester and post-semester teaching philosophy statements from seventeen PSL participants and from three control group peer leaders. All participants except for one in the control group had completed the pedagogy course. Table 5 contains the number of instances for each topic in the pre-semester and post-semester statements. Overall, specific learning theories, methods, and practices were more frequently mentioned in post-semester thesis statements than pre-semester statements. There was more variety of learning theory, methods, and practice included in post-semester thesis statements; also, peer leaders were more specific in the learning methods and practices identified. For example, they discussed dialogic discourse rather than just discourse and collaborative learning in place of more general group work. The most common learning theories identified throughout the statements were collaborative or cooperative learning, dialogic discourse, and meaningful learning. There was increased discussion of the learning environment in the post-semester drafts, as peer leaders often used descriptors such as "positive", "safe", and "student-centered."

Because the classroom environment became a prominent topic in post-semester drafts, we looked more closely at the descriptors used to discuss classroom environment (Table 6). In pre-semester drafts, classroom environment was mentioned in 6 statements (20% of total statements), and a total of 8 descriptors were used. In post-semester drafts, it was mentioned in 22 statements (73% of total statements) and 32 descriptors were used. General descriptors "effective", "optimal", "best", "ideal" were mentioned with equal frequency in both drafts, but the specific descriptors "safe" and "comfortable" increased from 0 to 6 instances and "student-centered" increased from 0 to 4 instances. These trends suggest an increase in peer leaders' valuation of the classroom environment and that they developed both an increasingly specific understanding of the characteristics of an effective classroom environment and a vocabulary to express these views. As we continue to evaluate these statements for additional trends, we will be modifying the pedagogy course to more fully address concepts that do not appear prominently in the philosophy statements.

Peer Leaders' Perceptions of Program Elements. From the qualitative data, we observed that peer leaders describe the pedagogy course as the most critical component of their training and professional development.

> "The pedagogy course helped me a lot more than the other [PD trainings] because I would have it every week, and I was exposed to it more. Like it was literally learning how to learn every week, which sticks with you more than you just going to four workshops every semester because you could easily just get those done and get them out of the way, and not think about them ever again. But with the pedagogy course it sticks with you every week, and I think that helps- that helped me learn everything and apply it, too."

Fewer students described the importance and value of the peer leader trainings, but most peer leaders mentioned the trainings in their interviews. Interestingly, the interview data revealed that peer leaders do not all define training and professional development the same way, and the definition of these terms can lead to confusion about the relevance or usefulness of the peer leader training workshops offered as part of the PSL Program. For example, two peer leaders believed the terms were synonyms; these peer leaders both viewed the workshops as a supplement to the pedagogy course. One described training as “needed in order for peer leaders to do their job” and professional development as “general skills or concepts that could be applied to any profession and might not necessarily help the peer leader perform his/her job.” For this student, the workshops were mostly considered professional development; whereas, the pedagogy course was described as training because it specifically related to being an academic peer leader. On the other hand, another peer leader described training in a negative tone—as “teaching someone to do a specific task”—and described professional development in a positive tone—as “helping someone develop skills that lead to professionalism.” One student viewed training as a subset of professional development, but not the other way around. Yet, other students did not feel a need to distinguish between these two terms. Exploring how peer leaders viewed the terms *training* and *professional development* provides insight to program directors and enables them to frame these experiences positively within the program requirements. For example, by framing the workshops as activities that contribute positively to students’ professional development, as opposed to framing them as job training, program directors might garner more participation and engagement related to this requirement.

For the first cohort of participants, the invited seminars seemed to have a minimal impact on the peer leaders. The seminar component of the program revealed one challenge in that it is a one-time activity each semester, and if the peer leaders have scheduling conflicts, they were excused from the activity. While the PSL program directors observed that the peer leaders who attended these seminars enjoyed hearing about how leadership, communication skills, and teamwork are valued in industry, no participants referenced this programmatic component in their interviews when asked about essential aspects of an academic peer leader program.

Some peer leaders could see an obvious connection between their work as a peer leader and their future careers; other peer leaders had to think about this connection when asked about it in an interview. An example of the latter occurred when a peer leader who was a tutor was asked whether or not any of the peer leader training workshops would help her in her future career. After describing how the training workshops helped her be a tutor, she thought more about the workshops she had attended and replied, "This semester I went to *Thinking Critically* and then *Managing Group Sessions*. That actually helped me a lot for tutoring, but, I mean, if I was a boss in my future career, then I guess it would help me manage groups and stuff." By answering this interview question, this peer leader reasoned out that the training workshop was in fact related to her future career in industry, especially if she were to secure a managerial position.

Lessons Learned

As expected, it was challenging to collect assessment data from peer leaders via online surveys, emailed written responses, or even through a scheduled meeting to complete paper-and-pencil concept maps and concept inventories. Response rates were highest when we included assessments as part of training workshops and the pedagogy course. We asked everyone to submit assessments as part of the training session, and then only used the responses from peer leaders who provided consent. We also wanted to demonstrate to students that we valued the assessments and were willing to use time during workshops for this activity. While we expected that the 10-15 minute online surveys would be the most attractive to students, we found that many would forget, or were not willing, to take the survey on their own time, even with monetary incentives in the form of human subjects payments. Similarly, it was difficult to collect baseline data the year prior to the first PSL cohort. One challenge was that we did not include human subjects payments, so there was no incentive for students to take the assessments outside of their normal peer leader position activities. This was one reason we decided to incorporate the assessments into our training and pedagogy course sessions during the first year of the program.

When considering the use of online assessments that are administered through a third-party (e.g., LPI & COMSA-R2), it is important to consider validity and reliability measures, the associated costs of each survey, and the format in which data is provided back to the program staff. For example, some third-party survey developers will not provide validity and reliability measures. We considered this as a warning sign and were only interested in utilizing surveys in which validity and reliability was established and shared with users. Cost structures for surveys vary and some provided more flexibility in distribution than others. The format of the survey results is another important consideration because, for example, some third party developers will not provide raw data, only aggregate or summarized results. While these reports can be useful, if the user intends to look closer at the data, perform his/her own statistical analyses, or view results at the participant level, the raw data is needed. We chose the LPI and COMSA-R2 surveys because they met our standards for all of these criteria related to validity and reliability, cost, and availability of data.

Another challenge to the PSL program design was that not all programs hire their peer leaders before the fall semester starts. Unfortunately, these students are not identified as being eligible for PSL participation early enough to be recruited to, or to apply for, the PSL program. While we allowed two peer leaders to join the program in the second week of the semester, joining in the third week or later in the semester would be difficult because the participant would have missed two weeks of the pedagogy course, new peer leader training workshops, and the window to complete pre-survey assessments. We are considering an alternative model of the pedagogy course that would start in week three or four of the semester and be taught at an accelerated pace in order to provide the same amount and level of content and class activities. We are also offering the pedagogy course in the spring semester during the second year of the program in the hopes that we can attract students who were hired in the fall semester for year-long positions.

The final challenge when working with peer leaders from other programs is that the activities of the peer leader position might not align with the types of activities we promote at the LCs: creating a positive learning environment that fosters collaborative learning to develop independence. The peer leader trainings and pedagogy course are both designed with opportunities for peer leaders to learn about theory and best practices related to active and collaborative learning and for the peer leaders to reflect on their interactions with students. For PSL participants who are in positions that might include more traditional "teaching" activities—leading instructor-centered review sessions, assisting a teaching assistant with administrative duties, grading, directly answering questions—the content and activities in the pedagogy course and peer leader trainings do not align with what they do in their positions. Moreover, the formal job observation is not always possible or effective because the peer leader might not be interacting with students enough for the observer to be able to judge the peer leaders' performance.

Limitations of the Study

While we collected data from 71 peer leaders who were not in the PSL program, the majority of this "control group" population consisted of LC learning assistants. Our Learning Assistant (LA) program follows the LA program model developed at UC Boulder (Otero et al., 2010). The program requires LAs to participate in staff meetings with the course instructors and to co-enroll in a pedagogy course. Our PSL program extends these requirements to include participation in additional training and professional development workshops and provides the opportunity for such development to peer leaders working in other positions, such as tutors and study group leaders. Consequently, our control group represents a biased sample of peer leaders who already participate in a large amount of training and professional development within their program. We hope to find better ways to include non-LA peer leaders in our study to observe the effects of a wider variety of required training and professional development activities across institutional programs.

Implications for Future Research

While it does not fall within the scope of this particular project, future studies should investigate the actual career outcomes of students in peer leader positions. We are investigating effective methods for tracking students post-graduation to identify a method for evaluating the extent to which the skills developed through peer leader positions translate to success in the STEM workforce. In addition, a study of former peer leaders' perceptions of the value of their positions to their future careers would add significantly to this body of knowledge.

We are collaborating with faculty in computer science to extend this research and compare the assessment of computer science learning assistants' content knowledge gains to that of their peers who are not academic peer leaders, but who are taking similar upper-level courses. We expect that this comparison will help determine whether gains in content knowledge can be attributed to the peer leader position (and subsequent re-learning and teaching of the content in the lower level courses) and not only to the fact that these students are taking upper level courses which might review or build upon the concepts taught at the lower level.

Conclusions

As we begin the last semester of our two-year pilot program, we believe the evidence suggests that such a program has potential to lead to gains in peer leaders' skill and knowledge development, though further analysis is required to identify the specific elements of the program that have the highest impact on these developments and to compare peer leaders to non-peer leaders in the same major. We understand that centralized programs such as this may not be feasible for all institutions and programs because constraints such as hiring timelines, peer leaders' availability, and financial resources may limit the degree of training and assessment that can be enacted.

Overall, we are optimistic about the model and, with some revisions to the timeline and methods of survey distribution, intend to build upon our Peer Leader Training Program using the knowledge gained from evaluating peer leaders' development over time. Our training program is being completely redesigned this coming year and we will be using the assessment results we gather through these practices to inform decisions about program change. We will be providing more instruction provided by faculty in each discipline in course content for all peer leader roles and will be providing more opportunities for assessment and feedback than we have previously. We intend to build assessment practices into the regular activities of the Peer Leader Training Program as a way to continue the evaluation of our training methods without creating assessment fatigue. We hope to see similar assessment measures replicated nationally so that benchmarks can be defined and the value of peer leader positions to students' skill development and content knowledge gains become widely publicized. We believe that the dissemination of such research will aid academic support services in gaining funding and institutional support for programs utilizing peer-led and near-peer instruction as the role of these positions in students' 21st century skills development is better understood.

Acknowledgements

The authors gratefully acknowledge funding support from the National Science Foundation's (NSF) Directorate of Undergraduate Education's (DUE) and Directorate for Education and Human Resources' (HER) Improving Undergraduate STEM Education (IUSE) program for this work (Award #1432394). Any opinions, findings, conclusions or recommendations expressed in this report are solely those of the authors and do not necessarily reflect those of the National Science Foundation.

References

Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and Evaluation of the Conceptual Inventory of Natural Selection, *Journal of Research in Science Teaching, 39*(10), 952-978.

Burmeister, S. L. (1996). Supplemental Instruction: An interview with Deanna Martin, *Journal of Developmental Education, 20*(1), 22-24, 26.

Chemistry 499 Introduction to Teaching Chemistry Lab. (n.d.). Retrieved from http://chem.rutgers.edu/sites/default/files/coursefiles/courses_sp15/Chem_499_Sp_15_Syllabus.pdf

Ding, L., Chabay, R., Sherwood, B., & Beichner, R. (2006). Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment, *Physical Review Special Topics - Physics Education Research, 2*, 010105

Dorn, B., & Tew, A. E. (2015). Empirical Validation and Application of the Computing Attitudes Survey, *Computer Science Education, 25*(1), 1-36.

Elby, A., Frederiksen, J., Schwarz, C., & White, B. Ideas behind EBAPS, Retrieved May 23, 2012, from http://www2.physics.umd.edu/~elby/EBAPS/idea.htm

Esplin, P., Seabold, J., & Pinnegar, F. (2012). The architecture of a high-impact and sustainable peer leader program: A blueprint for success. *New Directions for Higher Education, 157*, 85-100.

Ganser, S. R., & Kennedy, T. L. (2012). Where it all began: Peer education and leadership in student services. *New Directions for Higher Education, 157*, 17-29.

Hurley, M., Jacobs, G., & Gilbert, M. (2006). The Basic SI Model. In M. E. Stone and G. Jacobs (eds.), *Supplemental Instruction: New Visions for Empowering Student Learning* (pp. 11 – 22). San Francisco, CA: Josey Bass.

Klymkowsky, M. W., Underwood, S. M., & Garvin-Doxas, R. K. (2010). Biological Concepts Instrument (BCI): A diagnostic tool for revealing student thinking.

Koenig, J. A. (2011). *Assessing 21st century skills: Summary of a workshop*. Washington DC: National Academies Press.

Kouzes, J. M., & Posner, B. Z. (1988). *Leadership Practices Inventory (LPI): A self-assessment and analysis*. San Diego, CA: University Associates

Kuh, G. D. (2008). Brief summary of High-impact educational practices: What they are, who has access to them, and why they matter. Washington, DC: Association of American Colleges and Universities.

Lawson, A. E. (2006). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching, 15*(1) 11-24.

Madsen, A., McKagan, S. B., & Sayre, E. C. (2014). Best Practices for Administering Concept Inventories.

McClary, K., & Bretz, S. L. (2012). Development and assessment of a diagnostic tool to identify organic chemistry students' alternative conceptions related to acid strength. *International Journal of Science Education, 34*(5), 2317-2341.

Mulford, D. R., & Robinson, W. R. (2002). An Inventory for Alternate Conceptions among First-Semester General Chemistry Students. *Journal of Chemical Education, 79*(6), 739-744.

Ngothai, Y., & Davis, M. C. (2012). Implementation and analysis of a Chemical Engineering Fundamentals Concept inventory (CEFCI). *Education of Chemical Engineers, 7*(1), e32-e40.

Otero, V. K., & Gray, K. E. (2008). Attitudinal gains across multiple universities using the Physics and Everyday Thinking curriculum, Physical Review Special Topics - Physics Education Research, 4, 020104

Otero, V., Pollock, S., & Finkelstein, N. (2010). A physics department's role in preparing physics teachers: The Colorado learning assistant model. *American Journal of Physics, 78*(11), 1218-1224.

Parker, M. C., Guzdial, M, & Engleman, S. (2016). Replication, validation, and use of a language independent CS1 knowledge assessment. *ICER 2016 - Proceedings Of The 2016 ACM Conference On International Computing Education Research*, (ICER 2016 - Proceedings of the 2016 ACM Conference on International Computing Education Research), 93-101. doi:10.1145/2960310.2960316

Pellegrino, J. W., & Hilton, M. L. (Eds.). (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century.* Washington, D.C.: National Academies Press.

Psychtests. COMSA-R2 [Measurement Instrument]. Retrieved from http://corporate.psychtests.com/tests/?sec=tests

Shi, J., Wood, W. B., Martin, J. M., Guild, N. A., Vicens, Q., & Knight, J. (2010). A diagnostic assessment for introductory and molecular and cell biology. *CBE Life Science Education, 9*(4), 453–461.

Shook, J. L., & Keup, J. R. (2012). The benefits of Peer Leader programs: An overview from the literature. *New Directions for Higher Education*, (157), 5-16.

Smith, M. K., Wood, W. B., & Knight J. K. (2008). The Genetics Concept Assessment: A new concept map for gauging student understanding of genetics. *CBE Life Sciences Education, 7*(1), 422-430.

Stone, M. E., & Jacobs, G. (Eds.), (2006). *Supplemental Instruction: New visions for empowering student learning.* San Francisco, CA: Josey Bass.

The Student Leadership Challenge. (2017). The Student Leadership Practices Inventory 360 Online. Retrieved March 25, 2017, from http://www.studentleadershipchallenge.com/Assessment/assessment-studentLPI-online.aspx

Thornton, R. K., & Sokoloff, D. R. (1998). Assessing Student Learning of Newton's Laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula. *American Journal of Physics, 66*(4), 338-352.

Tingson-Gatuz, C. (2009). *Mentoring the leader: The role of peer mentoring in the leadership development of students-of-color in higher education.*

(Dissertation). Retrieved from ProQuest Dissertations and Theses, (Order No. 3381412)

Zelik, M., Schau, C., & Mattern, N. (1998). Misconceptions and their change in university-level astronomy courses. *The Physics Teacher, 36*(2), 104-107.

Appendices

Appendix A: Figures

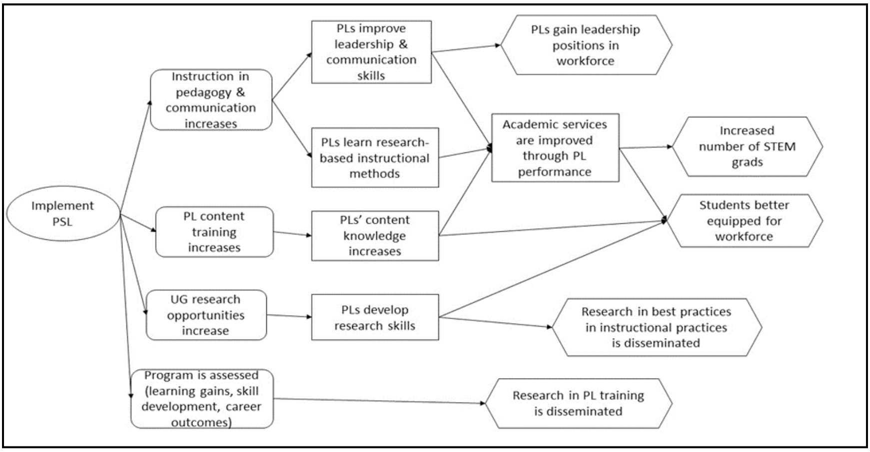

Figure 1. Expected outcomes (hexagons) of the PSL program based on program activities (rounded rectangles) and associated skill and content development (rectangles)

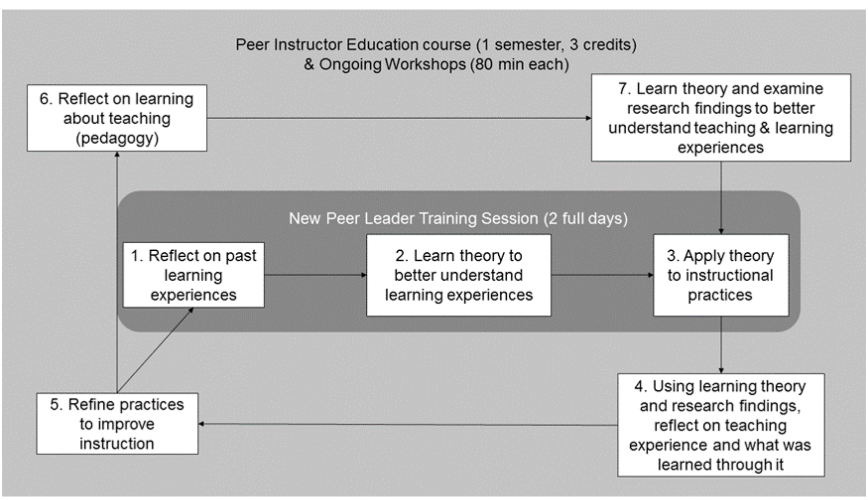

Figure 2. Training program iterative cycle

Access and Equity

- Working with students with disabilities
- Diversity and microaggressions
- Understanding challenges faced by underserved populations

Leadership

- Academic integrity & professionalism
- Leadership practices
- Managing groups
- Mentor training and conflict resolution

Communication

- Crucial conversations: Managing conflict
- Effective questioning
- Structuring the conversation
- Communication skills
- Student dependency & how to address it
- Communicating your peer leader position to potential employers

Learning Theory

- Understanding and utilizing mental models
- Understanding and implementing metacognition: Reflection on instructional methods
- Critical thinking skills and strategies
- Formative assessment
- Domains of knowledge in STEM fields

Figure 3. Examples of training and professional development workshops

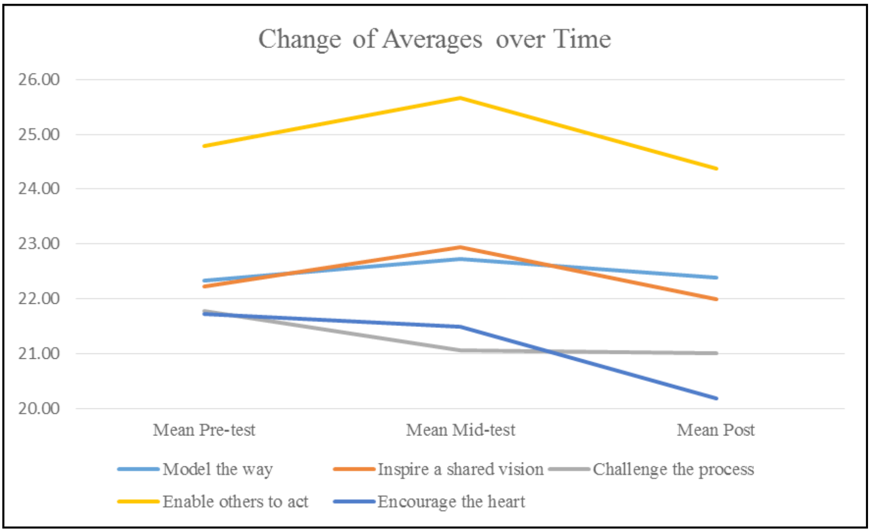

Figure 4. LPI changes of average over time
Note. Total scores range from 6-30 per category.

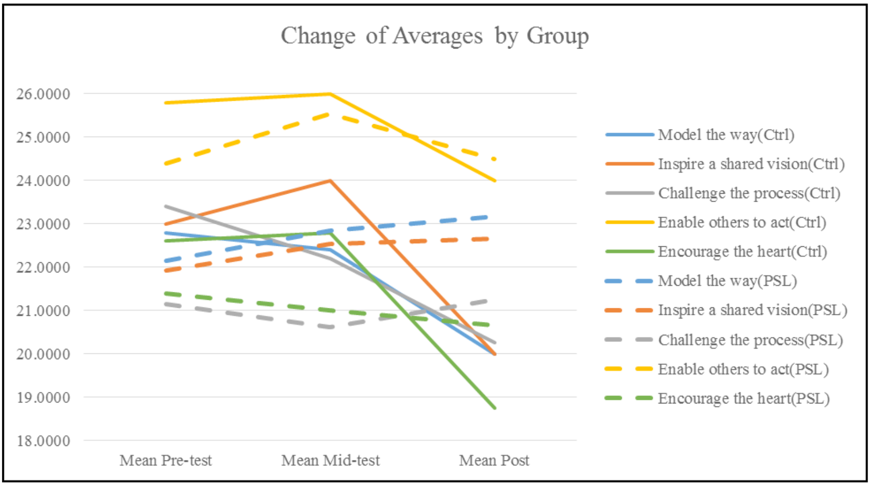

Figure 5. LPI changes of averages by group

Appendix B: Tables

Table 1

Timeline of Activities for PSL Program Administration

Month	Activity
June – August	Accept applications and invite participants; New Peer Leader Trainings (pre-semester assessment time included)
September	Co-enrollment in Pedagogy course (pre-semester assessment time included)
October – December	Observations; training and professional development workshops; invited seminar speaker
December – January	Post-fall semester assessments
January	New Peer Leader Trainings (pre-semester assessment time included); Co-enrollment in pedagogy course
February – April	Observations; training and professional development workshops; invited seminar speaker
April – May	Post-spring semester assessments

Table 2

Assessment Instruments

Area	**Purpose**	**Assessment Tools**	**Administration**
Pedagogy	Evaluate ideas about teaching and learning	Teaching Philosophy Teaching Concept Map	App, Dec., May InTr, Dec., May
All Content Area Disciplines	Evaluate complexity and extent of disciplinary content knowledge	Discipline Concept Map	InTr, Dec., May
Leadership Practices	Identify types of leadership practice and behaviors	LPI	Sept., Dec., May

Table 2 Continued

Area	Purpose	Assessment Tools	Administration
Communication Skills	Evaluate communication skills	COMSA-R2	Sept., Dec., May
Instructional Techniques & Group Management	Evaluate teaching practice and application of training topics	Observation	Wk5-15
Physics and astronomy	Evaluate accuracy of disciplinary content knowledge	Astronomy Pre/ Post-test Questions (Zelik et al., 2010)	InTr, Dec., May
		Force and Motion Conceptual Evaluation (Thornton & Sokoloff, 1998)	InTr, Dec., May
		Brief Electricity and Magnetism Assessment (Ding, 2006)	InTr, Dec., May
Mathematics	Evaluate accuracy of mathematics content knowledge	Lawson Test of Formal Reasoning (Lawson, 2006)	InTr, Dec., May
Chemistry & Biochemistry	Evaluate accuracy of disciplinary content knowledge	Chemical Concept Inventory (Mulford & Robinson, 2002)	InTr, Dec., May
		Concept Inventory on Acid Strength (McClary & Bretz, 2012)	InTr, Dec., May
Biological & Life Sciences	Evaluate accuracy of disciplinary content knowledge	Biological Concepts Instrument (Klymkowsky et al., 2010)	InTr, Dec., May
		Concept Inventory of Natural Selection (Anderson et al., 2002)	InTr, Dec., May

Table 2 Continued			
Area	**Purpose**	**Assessment Tools**	**Administration**
Biological & Life Sciences	Evaluate accuracy of disciplinary content knowledge	Genetics Concept Assessment (Smith et al., 2008)	InTr, Dec., May
		Introductory Molecular and Cell Biology Assessment (Shi et al., 2010)	InTr, Dec., May
Engineering & Computer Science	Evaluate accuracy of disciplinary content knowledge	Chemical Engineering Fundamentals Concept Inventory (Ngothai & Davis, 2012)	InTr, Dec., May
		Secondary assessment of Computer Science 1 knowledge (SCS1; Parker et al., 2016)	InTr, Dec., May
Attitudes & Beliefs	Evaluate attitudes related to learning in computer science	Computing Attitudes Survey (Dorn & Tew, 2015)	InTr, Dec., May
	Evaluate epistemological beliefs about physical sciences	Epistemological Beliefs about Physical Science Survey (Elby, 2012; Otero & Gray, 2008)	InTr, Dec., May

Note. App = submitted upon application to program; InTr = completed during initial training; Wk5-15 = observed during the 5th – 15th week of the semester

Table 3

Demographic Data for PSL Participants in Cohorts 1 & 2

	PSL Participants (N=34)	Rutgers University SAS, SEBS, SOE (N=16,068)
Sex		
Male	14	49%
Female	19	51%
Unknown or prefer not to answer	1	NA
Race and Ethnicity		
African American	1	9.1%
American Indian		0%
Asian	14	24.7%
Native Hawaiian[a] or Pacific Islander		0.2%
Latino		14.1%
White	9	39.2%
Two or more	3	3.7%
Black, Non-Hispanic		NA
Hispanic, or Hispanic Non-Puerto Rican	5	NA
Foreign		7.4%
Puerto Rican		NA
Unknown or prefer not to answer	2	1.6%
Class Year		
Second year	10	NA
Third year	14	NA
Fourth year	10	NA

Notes. SAS = School of Arts and Sciences; SEBS = School of Environmental and Biological Sciences; SOE = School of Engineering

[a] This is not the official label from the DOE, but is the label used within our university's student data system, which is the source of the data gathered here.

Table 4

Result of Repeated Measures for LPI

Source	Model the way	Inspire a shared vision	Challenge the process	Enable others to act	Encourage the heart
Within Subject					
Wave	0.14	0.3	11.46***	1.61	1.75
Wave x Gender	2.16	0.6	1.48	1.41	1.9
Wave x Group	1.4	0.81	7.09**	0.63	1.73
Between Subject					
Gender	4.18*	2.25	0.72	0.14	0.01
Group	0	0.09	0.64	0.14	0.93

Note. * $p < .10$. ** $p < .05$. ***$p < .01$

Table 5

Number of Instances for Each Topic in Teaching Philosophy Thesis Statements (Fall 2015)

Category	Pre–Fall	Post–Fall
Learning/Classroom Environment	6	22
Collaborative/ Cooperative Learning	5	16
Learning Types/Styles	3	2
Meaningful Learning	3	12
Motivation	3	4
Discourse	2	11
Cognition/Metacognition	2	1
Active Learning	2	0
Constructivism	1	0
Group (work, environment)	1	5

Table 5 Continued		
Category	**Pre–Fall**	**Post–Fall**
Independent Learning	1	5
Investigative Science Learning Environment (ISLE)	1	1
Mental Models	1	1
Multiple Intelligences	1	7
Problem-Solving	1	2
Reflection	1	3
Dialogic Discourse		8
Higher orders/levels of learning and thinking		5
Teacher-Student Relationships		3
Bloom's Taxonomy		2
Content Literacy		2
Critical Thinking		2
Univocal Discourse		2
Assessment		1
Communication		1
Gesturing		1
Group Processing		1
Memory		1
Questioning		1
Sensory Registering		1

Table 6

Instances of Descriptor use in Teaching Philosophy Thesis Statements

Descriptor of Classroom Environment	Pre Fall (N=6)	Post Fall (N=22)
“Effective”, “Optimal”, “Best”, “Ideal”	4	4
“Positive”	2	3
“Cooperative” or “Collaborative”	1	3
“Diversity”	1	1
“Challenging”	0	1
“Safe” or “Comfortable”	0	6
“Student-centered”	0	4