Investigating the Epistemological Development of Academic Peer Leaders Across STEM Disciplines: Exploring Changes Over Time, By Gender, and by Discipline

Mary Emenike, Thomas Cuthbert, & Stacey Blackwell

Rutgers, The State University of New Jersey

Author Notes

Dr. Mary Emenike, Department of Chemistry and Chemical Biology

Thomas Cuthbert and Stacey Blackwell, Learning Centers (New Brunswick Campus)

Abstract

Academic peer leadership positions provide opportunities for undergraduate students to develop content knowledge, Twenty-first Century Skills, and their beliefs about teaching and learning. To explore peer leaders' (PLs') epistemological development, the Epistemological Beliefs about Physical Sciences (EBAPS) survey was administered to 135 PLs three times a year, over a three-year period. This instrument was demonstrated to be valid and reliable for use with STEM PLs. Although the majority of positive shifts within instrument constructors occurred during the fall semester, responses did not regress by the end of the spring semester, demonstrating a retention of the new or altered beliefs over time. Implications for the design of PL development programs are discussed.

Investigating the Epistemological Development of Academic Peer Leaders Across STEM Disciplines: Exploring Changes Over Time, By Gender, and by Discipline

Introduction

As universities focus on educating students for success in the STEM workforce, they are increasingly searching for ways to help students develop Twenty-first 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). Twenty-first Century Skills are defined as fitting into three primary categories: cognitive, intrapersonal, and interpersonal (Pellegrino & Hilton, 2012). While many programs provide opportunities for students to develop these skills, not all appear to be equally emphasized among STEM disciplines. Undergraduate research experiences, for example, are highly valued and highly recommended for STEM students. Yet, within some departments at any given institution, there may be a number of impediments to participation, including the competitiveness of UG research positions, limitations in the number of opportunities available, time and travel constraints, and similar challenges that limit students' opportunities for UG research. Academic peer leadership positions provide an alternative or complementary opportunity for students because they promote development in all three of these categories described above by Pellegrino and Hilton (2012). These positions are offered on-campus and can provide meaningful experiential learning opportunities that may be otherwise inaccessible to students with travel restrictions or other barriers to accessing experiential learning opportunities that are off-campus or unpaid. The research completed through our Preparation in STEM Leadership (PSL) Program investigated the Twenty-first Century Skills and competencies developed through academic peer leadership to provide practitioners with evidence that can be used to evaluate the value of these programs as experiential learning opportunities in STEM disciplines.

For our work, we use the term "academic peer leader" (PL) to include undergraduate students who tutor (e.g., walk-in, group tutoring, one-on-one tutoring, or holding "office hours"); lead study groups or review sessions; or facilitate active and collaborative learning with (or without) course instructors or TAs in lectures, recitations, workshops, or laboratory sessions. We include students whose PL positions involve an emphasis on supporting students with course content. We do not include students whose peer leadership roles are solely associated with general mentoring, such as in residential life, general academic coaching, or extracurricular organizations. Within a large, public, R1 institution, there are, of course, many PL programs providing support in STEM disciplines, such as Learning Assistants (LAs; Otero, 2015), Supplemental Instructors (SI; Martin & Arendale, 1992), and Teaching Interns (TIs, Atieh & York, 2020). These programs have varying requirements for formal training or professional development.

Quality PL programs lead to improved learning gains (Barrasso & Spilios, 2021; White et al., 2016) and "increased satisfaction, persistence and retention, social development, and academic performance" of the students served (Ganser & Kennedy, 2012, p. 17). These programs provide the opportunity to broadly impact the teaching and learning community of STEM courses. However, because most of the research literature involves assessing or evaluating the impacts on students served by PL programs, assessing the development of the PLs themselves and evaluating PL programs remain areas needing attention (Table 1). While there is evidence that PL positions contribute to enhanced professional identity for learning assistants (Close et al., 2013; Close et al., 2016; Nadelson & Fannigan, 2014), a robust understanding of the skills and competencies developed through various PL experiences is still needed.

Because research has found that "neither success nor sustainability can be attained in a peer leadership program without thoughtful and intentional planning, management, and training" (Esplin et al., 2012, p. 85), the overarching goal of our PSL program was to develop a standard for PL training that could be assessed rigorously and implemented across multiple programs (Blackwell et al., 2017). Through the PSL program, PLs in STEM disciplines earned a scholarship to participate in advanced training and professional development workshops, including enrollment in a 3-credit, 300-level pedagogy course that was typically only available to first semester LAs. Part of the PSL project involved identifying instruments that could be used to assess the development of content knowledge, pedagogical knowledge, leadership styles, and communication skills.

Table 1

Non-Inclusive List of Studies Investigating Various Aspects of Peer Leader Programs

Category	Sources
Assessing Students Served by PL Programs	Allenbaugh & Herrera (2014); Alzen et al. (2017); Alzen et al. (2018); Blanc et al. (1983); Eren-Sisman et al. (2018); Fayowski & MacMillan (2008); Gok (2012); Hockings et al. (2008); Knight et al. (2015); Lewis (2011); Martin & Arendale (1992); Moore & LeDee (2006); Mutanyatta-Comar & Mooring (2019); Parkinson (2009); Peterfreund et al. (2008); Rath et al. (2007); Rath et al. (2012); Talbot et al. (2015); Van Dusen & Nissen (2017); Van Dusen & Nissen (2020); Weidler-Lewis et al. (2013); White et al. (2016)
Exploring the Development of PLs	Atieh & York (2020); Barrasso & Spilios (2021); Becker et al. (2016); Bourne et al. (2021); Cao et al. (2018); Close et al. (2016); Gray & Otero (2009); Hite et al. (2021); Lockie & Van Lanen (2008); Nadelson & Fannigan (2014); Top et al. (2018)
Evaluating PL Programs	Arco-Tirado et al. (2011); Colver & Fry (2016); Sabella et al. (2016); Wilson & Varma-Nelson (2016)

This article summarizes our findings from investigating the use of the Epistemological Beliefs about Physical Science (EBAPS) survey to assess the epistemological development of PLs over the course of one semester and one year. We included epistemological development as part of our category on content knowledge development, although it is not necessarily discipline-specific. Moreover, epistemological development includes more than simply understanding or applying the foundational concepts within a discipline; rather, it involves an understanding of the nature of knowing. The epistemological views of PLs are important to be able to assess because PLs will be interacting with, and likely influencing, the views and beliefs of the students whom they are helping to learn. We would argue that PL program directors and coordinators are aware that PLs are students who are still developing their core skills, competencies, and content knowledge in their own right; PLs are still developing as students, as leaders, and as professionals. It is important for PL programs to be able to assess this construct in order to demonstrate this outcome as a benefit to participation in PL positions and to evaluate whether or not any implemented (or modified) training and professional development opportunities resulted in a shift in this construct for the PLs.

Research Questions

In this article, we present findings to answer three research questions:

1. Is the Epistemological Beliefs About Physical Science (EBAPS) survey instrument reliable with a population of academic peer leaders (PLs)?
2. Do PLs' epistemological beliefs (as measured by EBPAS) shift after serving in their position for one or more semesters?
3. If epistemological beliefs shift over time, are there differences between groups of academic peer leaders by gender or STEM discipline?

Relevant Literature

Peer leadership training, when implemented according to best practices, emphasizes the "application of knowledge, skills, and responsibilities to new settings and complex problems" and 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" (Shook & Keup, 2012, p. 10). Students who undergo such training and serve as PL have reported "increased confidence in their ability to manage group dynamics, facilitate learning, and empathize with their students," as well as learning to address "real-word," ill-defined problems that "require multiple areas of knowledge and multiple modes of inquiry" (Shook & Keup, 2012, p. 11). Several studies have demonstrated growth in critical thinking, problem solving, and group processing (Table 1) and that peer mentoring opportunities "increase leadership capacity among students-of-color" (Tingson-Gatuz, 2009, p. 3).

The development of 21st century skills described by Shook and Keup (2012) and Tingson-Gatuz (2009) are multifaceted and overlap with other developmental constructs such as views and attitudes towards learning (Gray & Otero, 208), beliefs about the nature of knowledge and learning, and scientific identity development (Close et al., 2013; Close et al., 2016; Hite et al., 2021). To the best of our knowledge, epistemological development has not been investigated based on participation in STEM academic peer leader programs. These programs, however, have the potential to impact the development of epistemological beliefs because PLs approach understanding and learning the content in a new way when they assume an instructional role.

While many instruments are available to measure undergraduate students' attitudes, beliefs, and expectations about learning in science or in specific STEM disciplines, fewer instruments have been developed to measure epistemological beliefs within the STEM disciplines at the post-secondary level (Appendix A). Duell (2001) reported on a variety of instruments to measure epistemological beliefs; at that time, discipline-specific instruments to measure epistemological development were relatively new to the field. Although survey instruments such as CLASS-Phys (Perkins et al., 2005), CLASS-Chem (Barbera et al., 2008), CLASS-Bio (Semsar et al., 2011), MPEX (Redish et al., 2000), and CHEMX (Grove & Bretz, 2007) were developed between 1998-2011, these surveys were designed to measure attitudes, beliefs, and expectations about learning in chemistry, physics, or biology courses, which do not necessarily include epistemological beliefs. Duell's report, however, provided a useful list of instruments (from inventories to interviews to vignettes), organizing them by uni- or multidimensional constructs and providing theoretical frameworks. DeBacker *et al.*(2008) analyzed the factor structure and internal consistency of three instruments (Epistemological Questionnaire (Schommer, 1990), Epistemic Beliefs Inventory (Schraw, Bendixen, & Dunkle, 2002), Epistemological Beliefs Survey (Wood & Kardash, 2002). In all cases, DeBacker *et al.* uncovered psychometric issues that likely result from the conceptualization and specificity of epistemic beliefs. Ultimately, all of these instruments were focused on measuring epistemic beliefs of students and not PLs.

Because our PLs spanned many STEM disciplines, we were interested in administering a survey that was not constrained to a single discipline but was more specific to science or STEM than a general epistemological instrument. We also needed an instrument that could be administered to, and scored relatively easily with, a large number of students. For these reasons, we chose to investigate the potential of the Epistemological Beliefs about Physical Sciences (EBAPS) survey to assess PL epistemological development, even though it was designed to measure students' epistemological beliefs within the context of physical science examples (Elby, 2006b; Otero & Gray, 2008) and some of our PLs were assigned to mathematics, natural sciences, life sciences, computer science, and engineering disciplines.

The EBAPS Survey contains 30 items organized into three sections based on the format of the items. Part 1 contains 17 statements with a 5-point Likert scale from strongly disagree to strongly agree, with "neutral" listed as the mid-point option. Part 2 contains 6 fixed-response items. Part 3 contains 7 items written as a short discussion between two students, and the fixed-response choices asks for the degree to which the participant agrees with only one student or with both students. All EBAPS items are scored using a non-linear scoring scheme that ranges from 0 (least sophisticated) to 4 (most sophisticated) that takes into account whether a "neutral" response is more or less sophisticated than other options; this scoring scheme differentiates the EBAPS survey design from traditional attitude instruments that contain only Likert-scale items.

According to the EBAPS designers, 26 of the 30 items map onto one of five Axes (Table 2): (1) *Structure of scientific knowledge*, (2) *nature of knowing and learning*, (3) *real life applicability*, (4) *evolving knowledge*, and (5) *sources of ability to learn*. Two of the 30 items map onto two Axes and two items do not map onto any Axis but are included in the overall score. Recently, Johnson and Willoughby (2018a) explored the underlying structure of the EBAPS items using exploratory factor analysis on data collected over five years from 1,258 students at the end of an introductory astronomy course. Johnson and Willoughby reported on both a 3-factor and a 5-factor model--where the 5-factor model partially overlaps with the original Axes described by Elby--but around half of the items did not fit either model. Because Johnson and Willoughby's study focused on the epistemologies of introductory astronomy students, the analysis reported herein will utilize the original five Axes described by the EBAPS designers (Appendix B).

Reports in the literature regarding the use of the EBAPS survey have primarily been conducted with students enrolled in introductory college-level physics (Otero, 2008; Warren, 2018 & 2020) and astronomy (Johnson & Willoughby, 2018a & 2018b) courses. Typically, the EBAPS instrument has been administered to students to study the effects of specific curricula or instructional interventions (e.g., Physics & Everyday Thinking (PET) and Physical Science & Everyday Thinking, Bayesian activities (PSET; Otero, 2009)). Elby (2001), for example, observed that students were only likely to develop more sophisticated epistemologies when the curriculum had an explicit focus on epistemological development (5.3% gains on overall instrument). Otero and Gray (2008) reported that students from courses using the PET and PSET curriculum had 14%-25% higher scores than those in traditional courses. Johnson and Willoughby (2018a) administered the EBAPS to investigate changes in epistemological development among students in an introductory astronomy course that had been modified to include a focus on the *nature of science*. Johnson and Willoughby (2018a) observed some differences in epistemological beliefs (and the degree to which these beliefs decayed over time) among students enrolled in different colleges (e.g., Letters of Science, Business, Education) and by gender.

Methods

Overview

PLs completed assessments at the beginning of the fall semester (pre), end of the fall semester (mid), and end of the spring semester (post). All data was collected and analyzed under a research protocol approved by Rutgers University's Institutional Review Board. PLs who participated in this research study could earn human subjects' payments in the form of a gift card for each assessment completed.

Dedicated time to complete assessments was provided during new peer leader orientation sessions (beginning of fall semester) and the pedagogy course (beginning and end of fall semester). In addition, PLs could arrange a time to complete assessments in a supervised setting. Paper-and-pencil responses were transcribed into electronic format. PLs could also complete some surveys (including EBAPS) online using Qualtrics (*Qualtrics*, 2020).

Data Analysis

Data was analyzed using SPSS (*IBM SPSS Statistics 28 for Windows*, 2021). A repeated-measures t-test (for the pre & mid data) and repeated-measures ANOVAs (for the pre, mid, & post data) were used to examine the main effect of time. Mixed-model ANOVAs were used to examine the effects of all demographic variables. Fisher's least significant difference procedure was used as the post-hoc test where applicable. Due to the exploratory nature of this study, we chose to not use alpha correction to ensure our tests were as sensitive as possible to potential differences in order to illuminate avenues for future research. The full descriptive statistics for all measures described herein can be accessed online at: https://doi.org/10.6084/m9.figshare.21097921.v1

Institutional Context

Rutgers, the State University of New Jersey is a land-grant R-1 institution that serves both New Jersey residents and students from around the world. The New Brunswick (NB) campus currently enrolls more than 33,000 undergraduate students from all 50 states and more than 115 countries. More than half of Rutgers-NB students identify as non-Caucasian and more than 80% receive financial aid, making Rutgers-NB a diverse campus both culturally and socioeconomically. On the New Brunswick campus, the majority of STEM majors and programs are concentrated in three schools—the School of Arts and Sciences (SAS), the School of Engineering (SOE), and the School of Environmental and Biological Sciences (SEBS).

Participants

Overall, 165 PLs provided informed consent for this study. Of these PLs, 14% participated in the PSL program, 84% were learning assistants, and 52% self-identified as female. Participants' school enrollment and self-identified race/ethnicity is provided in

Axis	Label	Description	Items*
1	**Structure of scientific knowledge**	Students' view physics and chemistry knowledge as "a bunch of weakly connected pieces without much structure and consisting mainly of facts and formulas" or as "a coherent, conceptual, highly-structured, unified whole"	10
2	**Nature of knowing and learning**	Students' view learning science as "consist[ing] mainly of absorbing information" or "constructing one's own understanding" through active engagement, experiences, and reflection	8
3	**Real-life applicability**	Students' view scientific knowledge and ways of thinking as limited only to specific academic/scientific settings or more broadly to real life settings. "These items tease out students' views of the applicability of scientific knowledge *as distinct from* the student's own desire to apply science to real life, which depends on the student's interests, goals, and other non-epistemological factors."	4†
4	**Evolving knowledge**	Students' view scientific knowledge along a continuum from absolutism (e.g., "all scientific	3‡

		knowledge is set in stone") to extreme relativism (e.g., "no distinctions between evidence-based reasoning and mere opinion").	
5	**Source of ability to learn**	Students' views of being "good" at science exist along a continuum from fixed natural ability to the result of hard work and effective study strategies. Note: these views are intended to be "*distinct from* [students'] self-confidence and other beliefs about themselves."	5†‡

* Two items do not map onto any Axis.

† One item is included in both Axis 3 and Axis 5.

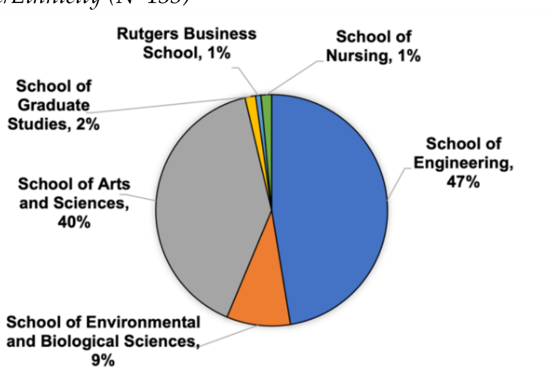
‡ One item is included in both Axis 4 and Axis 5.

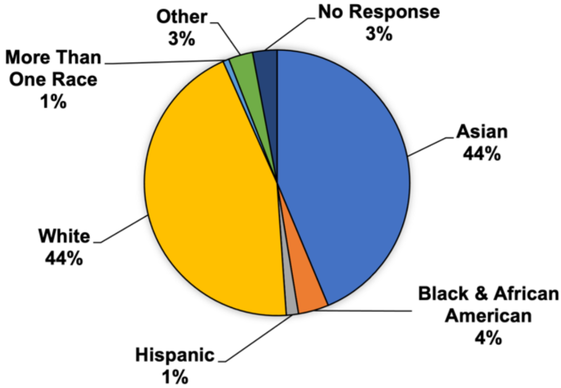
. Our sample's racial and ethnic demographics were largely representative of the university as a whole, with the exception of Asian and Hispanic students: Asian students were overrepresented in our sample, while Hispanic students were underrepresented.

Figure 1.

Study Participants' Enrolled in School and Self-Identified

Race/Ethnicity (N=135)



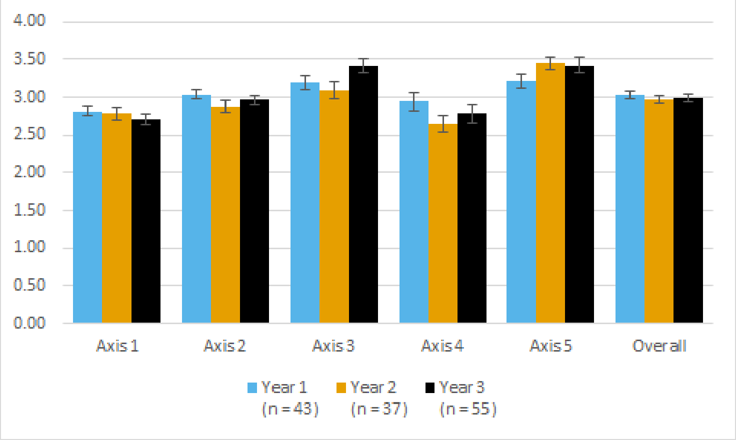


Data and Results

General Descriptive Statistics

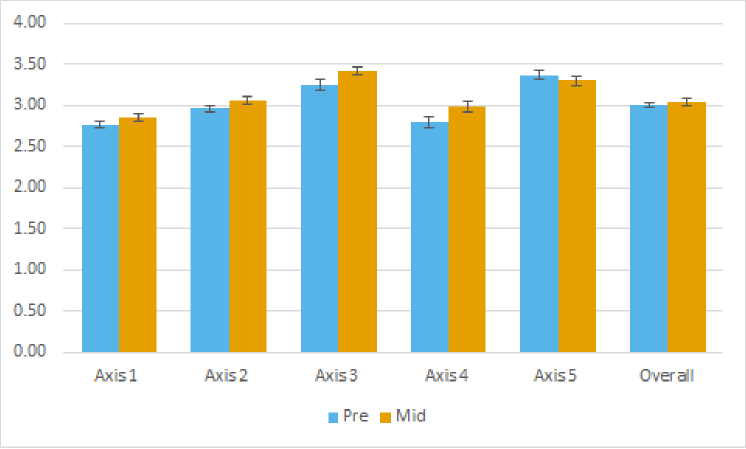
Over a three-year period, 135 PLs completed the EBAPS survey at the beginning of the fall semester (43 in the first year, 37 in the second year, and 55 in the third year; 82% response rate overall). There were no statistically significant differences among the three cohorts for Axes 1, 2, 4, and 5. On Axis 3 (*real-life applicability*) cohorts 2 and 3 differed significantly using a one-way ANOVA, F (2, 132) = 3.31, p = .040, η_p^2 = .05, with students in cohort 3 (m = 3.42) scoring higher than students in cohort 2 (m = 3.08, p = .015). Because there was only the one difference among two cohorts on a single Axis, the cohorts were combined into one dataset to explore changes in their responses over time and among subpopulations (Figures 3, 4, 5, and 6).

Figure 2.
Scores on each Axis and the overall EBAPS instrument for three cohorts of PLs at the beginning of the fall semester. Error bars represent 1 SEM.



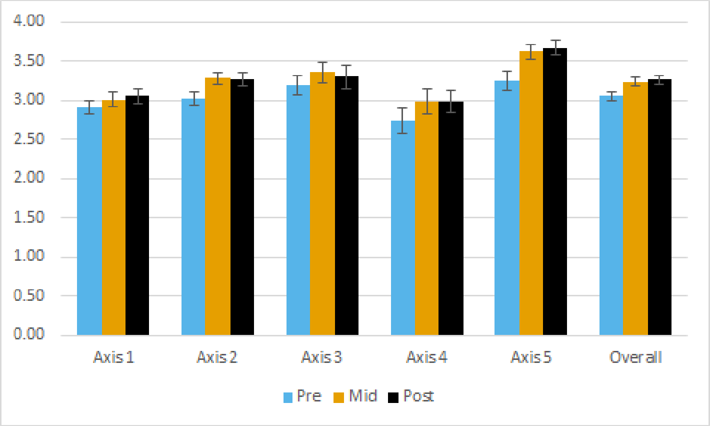
The overall average pre- scores on the five Axes range from 2.76 ± 0.46 to 3.36 ± 0.62 (Figure 3), with students scoring highest on Axis 5 (*source of ability to learn*, *m* = 3.36) and lowest on Axis 1 (*structure of scientific knowledge*, *m* = 2.76). Because scores range from 1 to 4, Axes 3 and 5 have the potential for observed ceiling effects with average pre-scores of above 3.2.

Figure 3.
Scores on EBAPS subscales and overall instrument for PLs (N = 129) who completed the survey at the beginning of fall semester (pre) and end of fall semester (mid). Error bars represent 1 SEM.



Of the 135 PLs who completed the pre-survey, 129 PLs also completed the survey at the end of the fall semester (mid-survey, Figure 3), and 20 of those PLs completed the survey for a third time at the end of the spring semester (post-survey, Figure 4). Survey participation rates were high for the pre- and mid-administration periods because the PLs were assessed during orientation at the start of the semester and at the beginning and end of a pedagogy course that is required for the Learning Assistant PL position. Survey participation rates were low for the post-administration period because we did not have direct access to the PL at the end of the spring semester; we invited PLs to come to the office to complete the assessments, but many PLs choose not to complete the assessments on their own time, despite the financial compensation provided for it. Only data from PLs with matched responses are included in the investigation of changes over time (RQs 2 & 3).

Figure 4.
Scores on EBAPS subscales and overall instrument for PLs (N = 20) who completed the survey at the beginning of fall semester (pre) and end of fall semester (mid) and end of the spring semester (post). Error bars represent 1 SEM.



RQ1: Reliability of the EBAPS survey with academic peer leaders

Cronbach's alpha was calculated to establish internal consistency for the EBAPS survey with PLs. The designers of the EBAPS survey argued against this measure because "the assessment items were designed so that students were allowed to disagree with themselves within a subscale" and because "epistemological beliefs may be triggered depending on context" (Elby, 2006b). However, other studies reported Cronbach alpha of 0.7 for students in chemistry courses (Keen-Rocha, 2008; Lekhi, 2018).

For the 129 students who completed the pre- and mid-surveys, Cronbach's alpha was 0.61 for the pre-testing and 0.79 for the mid-testing with no items removed. Additionally, the EBAPS pre-scores were consistent across cohorts of PLs over three years, with the previously mentioned exception on Axis 3. Because we are exploring this instrument with a new, targeted population of STEM PLs who have demonstrated success in previous STEM courses, and given the instrument has 30 items, we are limited to observing the internal consistency of the items as a proxy for instrument reliability. Despite the Cronbach's alpha scores falling in the *questionable to acceptable* range, the consistent pre-scores across cohorts provides supporting evidence that this instrument is reliable with a population of undergraduate academic PLs in STEM disciplines.

RQ2: Exploring epistemological development over time

For the 129 PLs who completed the EBAPS surveys at the beginning (pre) and end (mid) of the fall semester, positive shifts were observed on four of the five axes, but not on Axis 5 (source of ability to learn) and not on the instrument overall (Figure 3).

- For Axis 1 - *structure of scientific knowledge*, there was a marginally significant increase from students' pre (m = 2.76) to mid (m = 2.85) scores, t (128) = 1.97, p = .051, d = 0.17.
- For Axis 2 - *nature of knowing and learning*, there was a significant increase from students' pre (m = 2.96) to mid (m = 3.06) scores, t (128) = 2.27, p = .025, d = 0.20.
- For Axis 3 - *real-life applicability*, there was a significant increase from the students' pre (m = 3.25) to mid (m = 3.42) scores, t (128) = 3.22, p = .002, d = 0.28.
- For Axis 4 - *evolving knowledge*, there was a significant difference with a small effect size between the pre (m = 2.79) and mid (m = 2.98) scores, t (128) = 2.66, p = .009, d = 0.23.
- For Axis 5 - *source of ability to learn*, the difference between the pre (m = 3.37) and mid (m = 3.31) scores was not significant, t (128) = -1.03, p = .30, d = -0.09.
- For the instrument overall, the difference between the pre (m = 3.00) and mid (m = 3.04) scores was not significant, t (128) = 1.40, p = .17, d = 0.12.

For the 20 PLs who completed the EBAPS surveys both at the beginning (pre) and end (mid) of the fall semester, and at the end of the spring semester (post), positive shifts were observed on Axis 2 (nature of knowing and learning), Axis 5 (source of ability to learn), and on the instrument overall (Figure 4). For all three cases where there was a statistically significant difference, an LSD post-hoc test revealed that the mid and post scores were significantly higher than the pre scores, but the mid and post scores were not significantly different from each other.

- For Axis 1 - *structure of scientific knowledge*, the main effect for time was not significant, F (2, 38) = 2.23, p = .12, η_p^2 = .11.
- For Axis 2 - *nature of knowing and learning*, there was a significant main effect for time, F (2, 38) = 7.03, p = .003, η_p^2 = .27, such that scores increased from pre (m = 3.02) to mid (m = 3.28, p = .007), and from pre to post (m = 3.27, p = .005) but the mid scores did not differ significantly from the post scores (p = .90).
- For Axis 3 - *real-life applicability*, the main effect for time was not significant, F (2, 38) = 0.91, p = .41, η_p^2 = .05.
- For Axis 4 - *evolving knowledge*, the main effect for time was not significant, F (2, 38) = 1.37, p = .27, η_p^2 = .07.
- For Axis 5 - *source of ability to learn*, there was a significant main effect for time, F (2, 38) = 7.57, p = .002, η_p^2 = .29, such that scores increased from pre (m = 3.25) to mid (m = 3.62, p = .010) and from pre to post (m = 3.67, p = .008), but the mid

scores did not differ significantly from the post scores (p = .48).

- Overall - There was a significant main effect for time, F (2, 38) = 12.89, $p < .001$, $\eta_p^2 = .40$, such that scores increased from pre (m = 3.05) to mid (m = 3.24, p = .002), and from pre to post (m = 3.26, $p < .001$) but the mid scores did not differ significantly from the post scores (p = .43).

RQ3: Exploring epistemological development among groups

Observed differences by gender

In our sample of PLs, 47.29% self-identified as men (n = 61) and 52.71% self-identified as women (n = 68). Significant main effects for *gender* were observed on Axis 2, 5, and the instrument overall, and a marginally significant main effect for *gender* was observed on Axis 4. In each case the women scored higher than men (Figure 5). There was a significant interaction between *time* and *gender* on Axis 5, and marginally significant interactions between time and gender on Axis 4 and the instrument overall; in each case, the women scored higher than men on the mid survey.

- For Axis 1 - *structure of scientific knowledge*, the main effect for gender was not significant, F (1, 127) = 0.70, p = .41, $\eta_p^2 = .01$, nor was the interaction between gender and time, F (1, 127) = 0.45, p = .50, $\eta_p^2 = .004$.
- For Axis 2 - *nature of knowing and learning*, there was a significant main effect for gender, F (1, 127) = 7.07, p = .009,

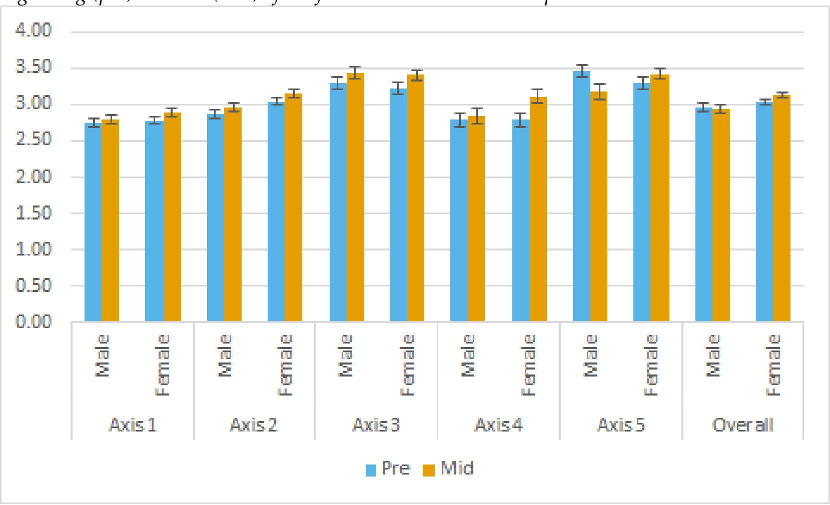
η_p^2 = .05, such that women (*m* = 3.09) scored higher than men (*m* = 2.92). The interaction between time and gender was not significant, *F* (1, 127) = 0.01, *p* = .93, η_p^2 = .0001.

- For Axis 3 - *real-life applicability*, neither the main effect for gender, *F* (1, 127) = 0.28, *p* = .60, η_p^2 = .002, nor the interaction between gender and time, *F* (1, 127) = 0.14, *p* = .71, η_p^2 = .001, were significant.
- For Axis 4 - *evolving knowledge*, the interaction between gender and time was marginally significant, *F* (1, 127) = 3.36, *p* = .069, η_p^2 = .03, such that women (*m* = 3.11) scored higher than men (*m* = 2.84, *p* = .043) at the mid testing. The main effect for gender was not significant, *F* (1, 127) = 1.44, *p* = .23, η_p^2 = .01.
- For Axis 5 - *source of ability to learn*, there was a significant interaction between time and gender, *F* (1, 127) = 10.41, *p* = .002, η_p^2 = .08, such that women (*m* = 3.42) scored higher than men (*m* = 3.18, *p* = .049) at the mid testing. Scores for men at the mid-testing were also significantly lower than their scores at the pre-testing (*m* = 3.46, *p* = .003). The main effect for gender was not significant, *F* (1, 127) = 0.19, *p* = .67η_p^2 = .001.
- Overall – The main effect for gender was significant, *F* (1, 127) = 4.89, *p* = .029, η_p^2 = .04, with women (*m* = 3.08) scoring higher than men (*m* = 2.95). There was a marginally

significant interaction between time and gender, F (1, 127) = 3.57, p = .061, η_p^2 = .03, with women (m = 3.13) scoring higher than men (m = 2.94, p = .011) at the mid testing. The women's scores also increased from the pre-testing (m = 3.03) to the mid testing (p = .022).

Figure 5.

Scores on EBAPS subscales and overall instrument for men (n = 61) and women (n = 68) at the beginning (pre) and end (mid) of the fall semester. Error bars represent 1 SEM.



Observed differences among PLs' discipline

Because students can secure academic PL positions in courses outside their enrolled school, participants were grouped using two different categorization schemes. First, PLs were grouped by whether or not their assigned course was within the discipline of their major (i.e., matched major-discipline). Of the 129 PLs for whom a major was identified, 43 were assigned to a course that matched the discipline of their major (12 Computer Science; 10 Life Sciences and Environmental Sciences, 12 Engineering; 5

Mathematics, Economics, and Logic and 4 Physics and Astronomy), while 86 were assigned to a course that did not match the discipline of their major (21 Chemistry; 4 Computer Science; 7 Life Sciences and Environmental Sciences, 1 Engineering; 4 Mathematics, Economics, and Logic and 49 Physics and Astronomy).

Second, and separately, PLs were grouped by the discipline of their assigned course: Chemistry; Computer Science; Engineering; Life Sciences and Environmental Sciences; Mathematics, Economics, and Logic; and Physics and Astronomy. There is particular mixing of students from different schools in SAS courses because students in SAS, SOE, and SEBS often have prerequisite courses offered by SAS (e.g., introductory math, biology, chemistry, and physics courses) or because these courses fulfill “core requirements” from SAS. In our sample, for example, 30 of the 53 PLs in physics and astronomy were students from SOE who were learning assistants in calculus-based physics courses. Of the remaining physics and astronomy PLs, 18 were from SAS and four were from SEBS; these PLs supported an introductory astronomy course for non-science majors and two different algebra-based physics courses that are primarily taken by life sciences majors in SEBS. Of the 17 PLs in Life Sciences and Environmental Sciences disciplines, seven were from SEBS and eight were from SAS. The exception, however, is that PLs in engineering disciplines were all enrolled in either SOE or the School of Graduate Studies (SGS). For the purposes of this study, we excluded the students enrolled in the School of Nursing (n = 2), SGS (n = 1), and Rutgers Business School (n = 2) due to their small group sizes.

There were a few significant differences between PLs assigned to courses that matched their major and PLs assigned to those not matched to their major. On Axis 2, Axis 5, and the Overall instrument, PLs assigned to courses that matched major scored lower than those assigned to courses that did not match their major. However, for Axis 2, PLs assigned to courses that matched their major scored higher at the end of the fall semester than at the beginning of the semester, indicating a positive shift in epistemological development for the *nature of knowing and learning*.

- For Axis 1 - *structure of scientific knowledge*, neither the main effect for matched major-discipline, F (1, 127) = 1.57, p = .21, η_p^2 = .01, nor the interaction between matched major-discipline and time, F (1, 127) = 0.77, p = .39, η_p^2 = .01, were significant.
- For Axis 2 - *nature of knowing and learning*, the main effect for matched major-discipline was marginally significant, F (1, 127) = 2.92, p = .090, η_p^2 = .02, such that PLs assigned to a course outside of their major (m = 3.05) scored higher than PLs assigned to a course inside of their major (m = 2.93). The interaction between matched major-discipline and time was significant, F (1, 127) = 4.21, p = .042, η_p^2 = .03, such that PLs

assigned to a course inside of their major scored higher at the mid testing (*m* = 3.04) than at the pre testing (*m* = 2.81).

- For Axis 3 - *real-life applicability*, neither the main effect for matched major-discipline, F (1, 127) = 0.72, p = .40, η_p^2 = .01, nor the interaction between matched major-discipline and time, F (1, 127) = 2.26, p = .14, η_p^2 = .02, were significant.
- For Axis 4 - *evolving knowledge*, neither the main effect for matched major-discipline, F (1, 127) = 1.76, p = .19, η_p^2 = .01, nor the interaction between matched major-discipline and time, F (1, 127) = 0.58, p = .45, η_p^2 = .01, were significant.
- For Axis 5 - *Source of ability to learn*, the main effect for matched major-discipline was marginally significant, F (1, 127) = 2.85, p = .094, η_p^2 = .02, such that PLs assigned to a course outside of their major (*m* = 3.40) scored higher than did PLs assigned to a course inside of their major (*m* = 3.22). The interaction between matched major-discipline and time was not significant, F (1, 127) = 0.01, p = .93, η_p^2 = .0001.
- Overall - There was a significant main effect for matched major-discipline, F (1, 127) = 4.10, p = .045, η_p^2 = .03, such that PLs assigned to a course outside of their major (*m* = 3.06) scored higher than did PLs assigned to a course inside their major (*m* = 2.93). The interaction between matched major-discipline and time was not significant, F (1, 127) = 0.30, p = .59, η_p^2 = .002.

Because of the extent of mixing of PLs assigned to courses from various majors, the findings above could be dependent on which courses and disciplines were in each group. Grouping PLs by their assigned course would involve groups with too few participants to compare. Consequently, we compared PLs by assigned discipline, even though we know that experiences across the courses within each discipline are likely to vary. We found a statistically significant main effect for discipline on the instrument overall with PLs assigned to engineering courses scoring lower than PLs assigned to chemistry and physics and astronomy courses (Figure 6). There was a marginally significant main effect for discipline on Axis 5, such that PLs assigned to engineering courses scored lower than PLs assigned to chemistry, computer science, and physics and astronomy courses. There was a significant interaction between time and discipline on Axis 3, such that PLs assigned to physics and astronomy courses scored higher on the mid-survey than the pre-survey, and PLs assigned to engineering courses scored lower at the pre-survey than PLs assigned to all other courses. There was a marginally significant interaction between discipline and time on Axis 5, such that PLs assigned to engineering courses scored lower than PLs assigned to all other disciplines on the mid-survey.

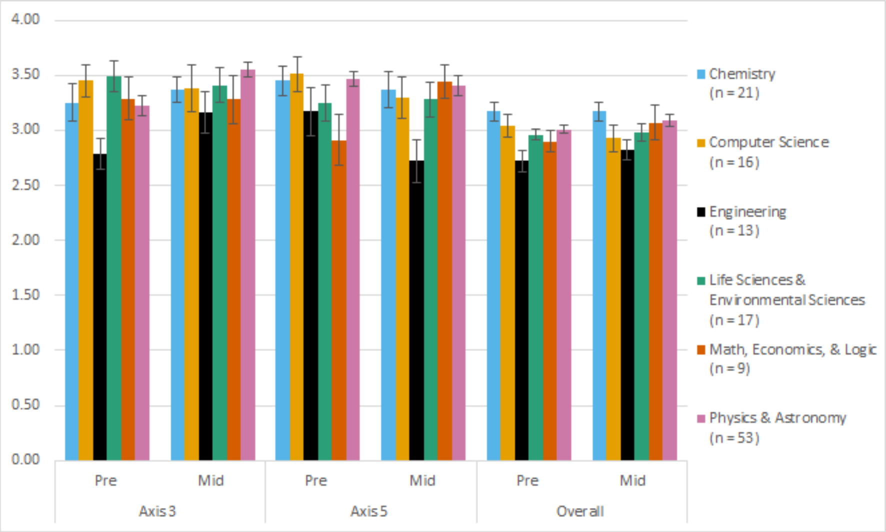
- For Axis 1 - *structure of scientific knowledge*, neither the main effect for discipline, F (5, 123) = 0.90, p = .48, η_p^2 = .04, nor the interaction between discipline and time, F (5, 123) = 1.42, p = .22, η_p^2 = .05, were significant.

- For Axis 2 - *nature of knowing and learning*, neither the main effect for discipline, F (5, 123) = 1.40, p = .23, η_p^2 = .05, nor the interaction between discipline and time, F (5, 123) = 0.31, p = .91, η_p^2 = .01, were significant.
- For Axis 3 - *real-life applicability*, there was a significant interaction between time and discipline, F (5, 123) = 2.48, p = .035, η_p^2 = .09, such that PLs assigned to engineering courses scored significantly lower at the pre-testing (m = 2.79) than did PLs assigned to chemistry (m = 3.25, p = .047), life science and environmental science (m = 3.50, p = .004), computer science (m = 3.45, p = .007), and physics and astronomy courses (m = 3.22, p = .031) at the pre-testing. PLs assigned to engineering courses also scored significantly lower at the mid-testing (m = 3.16) than did PLs assigned to physics and astronomy courses (m = 3.55, p = .391) at the mid-testing. The main effect for discipline was not significant, F (5, 123) = 1.44, p = .22, η_p^2 = .06.
- For Axis 4 - *evolving knowledge*, neither the main effect for discipline, F (5, 123) = 1.11, p = .36, η_p^2 = .04, nor the interaction between discipline and time, F (5, 123) = 1.11, p = .36, η_p^2 = .04, were significant.

- For Axis 5 - *Source of ability to learn*, there as a marginally significant main effect for discipline, F (5, 123) = 2.06, p = .075, η_p^2 = .08, such that PLs assigned to engineering (m = 2.95) courses scored significantly lower than PLs assigned to chemistry (m = 3.41, p = .017), computer science (m = 3.41, p = .025), and physics and astronomy (m = 3.44, p = .004) courses. There was a marginally significant interaction between discipline and time, F (5, 123) = 2.20, p = .059, η_p^2 = .08, such that PLs assigned to mathematics, economics, and logic courses scored significantly lower on the pre-testing (m = 2.91) than did PLs assigned to chemistry (m = 3.45, p = .029), computer science (m = 3.51, p = .020), and physics and astronomy courses (m = 3.47, p = .013) at the pre-testing. PLs assigned to engineering courses scored significantly lower on the mid-testing (m = 2.72) than did PLs assigned to chemistry (m = 3.37, p = .008), math, economics, and logic (m = 3.44, p = .017), life science and environmental science (m = 3.28, p = .029), computer science (m = 3.30, p = .026), and physics and astronomy courses (m = 3.41, p = .002) at the mid-testing.
- Overall - There was a significant main effect for discipline, F (5, 123) = 2.63, p = .027, η_p^2 = .10, such that PLs assigned to engineering courses (m = 2.77) scored significantly lower than PLs assigned to chemistry (m = 3.17, p < .001) and

physics and astronomy (*m* = 3.05, *p* = .006) courses. The interaction between discipline and time was not significant, F (5, 123) = 1.03, p = .40, η_p^2 = .04.

Figure 6.
Scores for PLs by discipline comparing beginning (pre) and end (mid) of the fall semester. Error bars represent 1 SEM.



Discussion

The EBAPS survey appears reliable for STEM PLs based on the Cronbach alpha scores and the fact that the beginning-of-semester responses across three cohorts of PLs did not differ significantly. PLs scored highest on Axis 5 (*source of ability to learn*) and Axis 3 (*real world applicability*) and lowest on Axis 1 (*structure of scientific knowledge*) and Axis 4 (*evolving knowledge*). PLs appeared to score at the more sophisticated end of the scales, which might be expected, given that they have previously performed well in their assigned course.

As for the positive shifts that were observed during the fall semester, epistemological development could be the result of one or more factors:

- PLs were students enrolled in STEM courses where they would be learning more content and potentially be developing their epistemological beliefs.
- PLs were helping other students learn, and, therefore, they were learning more content themselves as well as developing their epistemological beliefs.
- PLs learned about teaching and learning through the pedagogy course, which contributed to the development of their epistemological beliefs.

For the last factor listed, the majority of the PLs in this study–118 out of 129 (91%) who completed the pre and mid surveys–were co-enrolled in the pedagogy course because they were required to take it as first-semester Learning Assistants or because they were participating in the PSL program. It is interesting that no additional gains were observed for the 20 students who completed the post-survey, which might suggest that the epistemological development was more affected by being a PL and taking a pedagogy course than simply taking more coursework as a student. Fortunately, it appears that epistemological gains made by the PLs in the fall semester did not regress during the spring semester based on the fact that their scores on the mid- and post- tests were not statistically significantly different across the various groups. If, in fact, the pedagogy course was a significant contributor to the epistemological development of PLs in the fall semester (and, as noted above, additional research is needed to determine the effect of factors contributing to epistemological development on any given Axis), it is encouraging that gains were not lost during the spring semester when PLs were no longer enrolled in the pedagogy course.

While there were some observed differences based on PL gender and discipline, most of the differences were observed on only a subset of Axes or were marginally statistically significant. When comparing PL disciplines, the statistical power of the analyses is likely limited by our sample sizes; for example, there were nearly 6 times as many physics PLs ($n = 53$) as math, logic, and economics PLs ($n = 9$). In most cases, the effect sizes were small, with partial eta squared values ranging from 0.03 to 0.10. Although the survey instrument was designed for physics and chemistry disciplines, we were able to administer the instrument to PLs in mathematics, life sciences, environmental science, computer science, and engineering disciplines. Anecdotally, during administration, PLs in computer science complained about taking the survey by saying it was not relevant to their field; however, these PLs did not score significantly higher or lower than PLs in other disciplines. The PLs assigned to engineering courses scored lower than other disciplines on some subscales of the survey instrument. Additional work is needed to understand the effects of peer leader development when they serve in courses within or outside of the discipline of their intended major.

Implications for Peer Leader Training and Professional Development

Within our context, some epistemological growth was observed during the fall semester, and this growth did not appear to regress at the end of the spring semester. Although 91% of the PLs in this study (1) were co-enrolled, or had completed, a pedagogy course and (2) would have been required to complete at least two training or professional development workshops each semester, this study was not designed to parse out the effects of the pedagogy course and/or training requirements from the effects of being a peer leader and taking additional coursework as a student. However, we are encouraged to observe development related to the *nature of knowing and learning*, *real-life applicability*, and *source of ability to learn*. It is possible that many of the topics included in the pedagogy course contributed directly to developing epistemological beliefs, such as Human Constructivism and Meaningful Learning Theory, Cognition, Metacognition, Effective Questioning, and Cooperative Learning. If a future study attempts to make claims about the specific impact of the pedagogy course, or specific training workshops, on epistemological development, care should be taken to ensure sufficient numbers of PLs in various sub-groups, such as discipline of PL assignment, nature of assigned PL course (e.g., inquiry-based, active learning, cooperative learning), intended major, concurrent coursework while serving as a PL, year in school, pedagogy course enrollment, and training workshops completed.

Ultimately, the results from this study with the EBAPS survey were used to develop training and professional development workshops that address epistemological development across all the EBAPS survey constructs and specifically the constructs of the structure of scientific knowledge and evolving knowledge. The results were also being used by the pedagogy course coordinators during their normal practice of updating and modifying the course curriculum each semester. Although there is always more to include in the pedagogy course than time permits, there are ways to incorporate ideas around epistemological development within existing topics, such as mental models, cognition, and metacognition. For PL program directors and pedagogy course instructors, it is important to identify the types of epistemological development we value in peer leader positions, such as awareness of one's own knowledge and learning process, sources of knowledge, and how people learn new knowledge in general. Once these values are identified, and assuming we have instruments, like the EBAPS, that are valid and reliable enough to enable us to assess these constructs, program directors and course instructors can focus on developing class activities, training workshops, and professional development programming that exposes PLs to these concepts and supports their continued development towards more sophisticated beliefs.

Implications for Future Research

While data from this study demonstrated that the EBAPS can be used reliably with STEM academic peer leaders, additional research is needed to explore the sensitivity of the instrument. Although some gains were observed on some Axes overall or for a subset of the PL population, this dataset cannot be used to investigate potential ceiling effects on some of the Axes or the sensitivity of the instrument to various degrees of epistemological development. Because epistemological development was not included as an explicit topic in the pedagogy course or for any training or professional development workshops, any observed increase to more sophisticated beliefs would have resulted from indirect instruction, exposure to, or interaction with, the associated concepts or constructs. It would be interesting to design a study that includes a sub-group of PLs who had explicit training, professional development workshops, or pedagogy course readings on epistemology and compare that to PLs who do not participate in any of these sessions explicitly related to epistemology, which could then provide information related to the potential sensitivity of the EBAPS instrument.

As mentioned above, this current study was not designed to investigate the specific contributions or influences of various training and professional development requirements, including the pedagogy course, on epistemological development. An interesting area for future research would be to design a mixed-methods study to attempt to better understand the factors that contribute to epistemological development for STEM peer leaders. An extension of this work could include investigating the impact, if any, that the nature of the assigned course plays in a PL's epistemological development. In other words, do STEM PLs assigned to courses that include more active, collaborative, or inquiry-based learning demonstrate more epistemological development overall, or on certain Axes, than PLs assigned to more traditional courses? Ultimately, further research is needed to explore whether the effect of epistemological development is based on being a PL, the course disciplines of the PL position, the course design and curriculum of the PL position, the pedagogy course and/or specific professional development workshops and having additional coursework and growth as a student. While PLs' participated in surveys and interviews as part of the larger PSL study, that data is outside the scope of the research questions posed in this article.

Conclusions

The analysis of EBAPS survey responses from 135 PLs, across six STEM disciplines, over a three-year period suggests that the survey is valid and reliable with this population of undergraduate students: the internal consistency of the items administered at the beginning and end of the fall semester was 0.61 and 0.79, respectively; responses were not statistically significantly different between cohorts of PLs across three years. PLs' EBPAS scores increased on some axes over the course of the fall semester (namely: *nature of knowing and learning; real-life applicability, source of ability to learn*), but little to no change was observed by the end of the spring semester. Although one might expect additional epistemological growth from students continuing to both take coursework in their major and serve as peer leaders, it was encouraging that the responses do not suggest a regression of development after the first semester. While there were some observed differences between male and female students and among PL disciplines, additional research is needed to parse out the effects of any differences and to explore the effects of being a PL, completing additional coursework, or completing a pedagogy course. Ultimately, these findings inform the development of PL training and professional development workshops to address epistemological development within the PLs themselves and associated with their role supporting student development.

Acknowledgements

The authors gratefully acknowledge funding support from the National Science Foundation (NSF), Directorate of Undergraduate

Education (DUE)/Education and Human Resources (EHR), Improving Undergraduate STEM Education (IUSE; award 1432394) for this research. The opinions, findings, conclusions or recommendations expressed in this report are solely those of the authors and do not necessarily reflect those of the NSF. The authors thank Sari Katzen and Nipa Onulak for their work as program coordinators on the grant-funded project.

References

Allenbaugh, R. J., & Herrera, K. M. (2014). Pre-assessment and peer tutoring as measures to improve performance in gateway general chemistry classes. *Chemical Education Research and Practice, 15*(4), 620-627. https://doi.org/10.1039/c4rp00094c

Alzen, J. L., Langdon, L., & Otero, V. K. (2017). The Learning Assistant model and DFW rates in introductory physics courses. *2017 Physics Education Research Conference (PERC) Proceedings*, 36-39. https://doi.org/10.1119/perc.2017.pr.004

Alzen, J. L., Langdon, L. S., & Otero, V. K. (2018). A logistic regression investigation of the relationship between the Learning Assistant model and failure rates in introductory STEM courses. *International Journal of STEM Education, 5*(1), 1-12. https://doi.org/10.1186/s40594-018-0152-1

Arco-Tirado, J. L., Fernández-Martín, F. D., & Fernández-Balboa, J.-M. (2011). The impact of a peer-tutoring program on quality standards in higher education. *Higher Education*, *62*(6), 773-788. https://doi.org/10.1007/s10734-011-9419-x

Atieh, E. L., & York, D. M. (2020). Through the Looking CLASS: When Peer Leader Learning Attitudes Are Not What They Seem. *Journal* of Chemical Education, 97(8), 2078–2090. https://doi.org/10.1021/acs.jchemed.0c00129

Barbera, J., Adams, W. K., Wieman, C. E., & Perkins, K. K. (2008). Modifying and Validating the Colorado Learning Attitudes about Science Survey for Use in Chemistry. *Journal of Chemical Education, 85*(10), 1435-1439. http://www.jce.divched.org/Journal/Issues/2008/Oct/abs1435.html

Barrasso, A. P., & Spilios, K. E. (2021). A scoping review of literature assessing the impact of the learning assistant model. *International Journal of STEM Education, 8*(1), 1-12. https://doi.org/10.1186/s40594-020-00267-8

Bauer, C. (2005). Beyond "Student Attitudes": Chemistry Self-Concept Inventory for Assessment of the Affective Component of Student Learning. *Journal of Chemical Education, 82*, 1864-1870. https://doi.org/10.1021/ed082p1864

Bauer, C. F. (2008). Attitude toward Chemistry: A Semantic Differential Instrument for Assessing Curriculum Impacts. *Journal of Chemical Education, 85*(10), 1440-1445. https://doi.org/10.1021/ed085p1440

Becker, A. P., Goldberg, B., & Jariwala, M. (2016). Self-Perception of Teaching Fellows and Learning Assistants in Introductory Physics Classes. *2016 Physics Education Research Conference (PERC) Proceedings*, 48-51. https://www.per-central.org/items/detail.cfm?ID=14190

Blackwell, S., Katzen, S., Patel, N., Sun, Y., & Emenike, M. (2017). Developing the Preparation in STEM Leadership Programs for Undergraduate Academic Peer Leaders. *The Learning Assistant Review*, 22(1), 49-84.

Blanc, R. A., & et al. (1983). Breaking the Attrition Cycle: The Effects of Supplemental Instruction on Undergraduate Performance and Attrition. *Journal of Higher Education, 54*(1), 80-90. https://doi.org/10.1080/00221546.1983.11778153

Bourne, S. M., Limfat, S., Dilip, S., Han, S., Robak, M. T., Marsden, P., & Baranger, A. M. (2021). The Undergraduate Teacher-Scholar Program: Comparing Near-Peer and Non Near-Peer Instructors in Laboratory Courses. *Journal of College Science Teaching, 50*(5), 10-17. https://www.nsta.org/journal-college-science-teaching/journal-college-science-teaching-mayjune-2021/undergraduate

Cano, F. (2005). Epistemological beliefs and approaches to learning: their change through secondary school and their influence on academic performance. *British Journal of Educational Psychology, 75*(Pt 2), 203-221. https://doi.org/10.1348/000709904x22683

Cao, Y., Smith, C., Lutz, B., & Koretsky, M. (2018). Cultivating the Next Generation: Outcomes from a Learning Assistant Program in Engineering. American Society for Engineering Education, Salt Lake City, UT.

Carberry, A., Ohland, M., & Swan, C. (2010). A Pilot Validation Study Of The Epistemological Beliefs Assessment For Engineering (Ebae): First Year Engineering Student Beliefs. American Society for Engineering Education 2010 Annual Conference & Exposition Louisville, Kentucky.

Close, E. W., Close, H. G., & Donnelly, D. (2013). Understanding the learning assistant experience with physics identity [Article]. *AIP Conference Proceedings, 1513*(1), 106-109. https://doi.org/10.1063/1.4789663

Close, E. W., Conn, J., & Close, H. G. (2016). Becoming Physics People: Development of Integrated Physics Identity through the Learning Assistant Experience. *Physical Review Physics Education Research, 12*(1), 010109-010101-010109-010118. http://dx.doi.org/10.1103/PhysRevPhysEducRes.12.010109

Colver, M., & Fry, T. (2016). Evidence to Support Peer Tutoring Programs at the Undergraduate Level. *Journal of College Reading and Learning, 46*(1), 16-41. https://doi.org/10.1080/10790195.2015.1075446

Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology, 29*(2), 186-204. https://doi.org/https://doi.org/10.1016/j.cedpsych.2004.01.004

DeBacker, T. K., Crowson, H. M., Beesley, A. D., Thoma, S. J., & Hestevold, N. L. (2008). The Challenge of Measuring Epistemic Beliefs: An Analysis of Three Self-Report Instruments. *The Journal of Experimental Education, 76*(3), 281-312. https://doi.org/10.3200/JEXE.76.3.281-314

Duell, O. K., & Schommer-Aikins, M. (2001). Measures of People's Beliefs About Knowledge and Learning. *Educational Psychology Review, 13*(4), 419-449. www.jstor.org/stable/23363499

Elby, A. (2001). Helping physics students learn how to learn. American Journal of Physics, 69(S1), S54-S64. https://doi.org/10.1119/1.1377283

Elby, A. (2006a). *EBAPS home*. Retrieved 7/15/2020 from http://www2.physics.umd.edu/~elby/EBAPS/home.htm

Elby, A. (2006b). *Idea Behind EBAPS*. Retrieved 7/15/2020 from http://www2.physics.umd.edu/~elby/EBAPS/idea.htm

Eren-Sisman, E. N., Cigdemoglu, C., & Geban, O. (2018). The effect of peer-led team learning on undergraduate engineering students' conceptual understanding, state anxiety, and social anxiety [10.1039/C7RP00201G]. *Chemistry Education Research and Practice, 19*(3), 694-710. https://doi.org/10.1039/C7RP00201G

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. http://dx.doi.org/10.1002/he.20008

Fayowski, V., & MacMillan, P. D. (2008). An Evaluation of the Supplemental Instruction Programme in a First Year Calculus Course. *International Journal of Mathematical Education in Science and Technology, 39*(7), 843-855. http://www.informaworld.com/openurl?genre=article&id=doi:10.1080/00207390802054433

Ganser, S. R., & Kennedy, T. L. (2012). Where It All Began: Peer Education and Leadership in Student Services. *New Directions for Higher Education*, 17-29. http://dx.doi.org/10.1002/he.20003

Gok, T. (2012). The Impact of Peer Instruction on College Students' Beliefs about Physics and Conceptual Understanding of Electricity and Magnetism. *International Journal of Science and Mathematics Education, 10*(2), 417-436. http://dx.doi.org/10.1007/s10763-011-9316-x

Gray, K. E., & Otero, V. K. (2009). Analysis of Former Learning Assistants' Views on Cooperative Learning [Article]. *AIP Conference Proceedings, 1179*(1), 149-152. https://doi.org/10.1063/1.3266700

Grove, N., & Bretz, S. L. (2007). CHEMX: An Instrument to Assess Students' Cognitive Expectations for Learning Chemistry. *Journal of Chemical Education, 84*(9), 1524-1529. http://jchemed.chem.wisc.edu/Journal/Issues/2007/Sep/abs1524.html

Halloun, I., & Hestenes, D. (1998). Interpreting VASS Dimensions and Profiles for Physics Students. *Science and Education, 7*(6), 553-577. https://doi.org/10.1023/A:1008645410992

Hite, R. L., Childers, G., Gottlieb, J., Velasco, R., Johnson, L., Williams, G. B., Griffith, K., & Dwyer, J. (2021). Shifts in learning assistants' self-determination due to COVID-19 disruptions in Calculus II course delivery. *International Journal of STEM Education, 8*(1), 55. https://doi.org/10.1186/s40594-021-00312-0

Hockings, S. C., DeAngelis, K. J., & Frey, R. F. (2008). Peer-Led Team Learning in General Chemistry: Implementation and Evaluation. *Journal of Chemical Education, 85*(7), 990-996. https://doi.org/10.1021/ed085p990

IBM SPSS Statistics for Windows. In. (2021). (Version 28) IBM Corp. https://www.ibm.com/products/spss-statistics

Jena, A. K., & Chakraborty, S. (2018). Epistemological Beliefs: Its Relationship with Learning Styles, Learning Approaches, and Achievement. *Asia Pacific Journal of Education, Arts and Sciences, 5*(1), 60-70. http://apjeas.apjmr.com/wp-content/uploads/2018/02/APJEAS-2018.5.1.07.pdf

Johnson, K., & Willoughby, S. (2018a). Changing epistemological beliefs with nature of science implementations. *Physical Review Physics Education Research, 14*(1), 010110. https://doi.org/10.1103/PhysRevPhysEducRes.14.010110

Johnson, K., & Willoughby, S. (2018b). Epistemic belief structures within introductory astronomy. *Physical Review Physics Education Research, 14*(1), 010135. https://doi.org/10.1103/PhysRevPhysEducRes.14.010135

Kampa, N., Neumann, I., Heitmann, P., & Kremer, K. (2016, 2016/07/01/). Epistemological beliefs in science—a person-centered approach to investigate high school students' profiles. *Contemporary Educational Psychology, 46*, 81-93. https://doi.org/https://doi.org/10.1016/j.cedpsych.2016.04.007

Keen-Rocha, L. S. (2008). *Personal epistemological growth in a college chemistry laboratory environment* University of South Florida]. https://scholarcommons.usf.edu/etd/327

Knight, J. K., Wise, S. B., Rentsch, J., & Furtak, E. M. (2015). Cues Matter: Learning Assistants Influence Introductory Biology Student Interactions during Clicker-Question Discussions. *CBE—Life Sciences Education, 14*(4), ar41,41-14. https://doi.org/10.1187/cbe.15-04-0093

Koenig, J. A. (2011). *Assessing 21st Century Skills: Summary of a Workshop*. National Academies Press. http://www.nap.edu/catalog.php?record_id=13215

Lekhi, P. A. (2018). *International students' epistemologies and learning experiences in specially designed first-year chemistry courses* University of British Columbia]. University of British Columbia. http://hdl.handle.net/2429/68086

Lewis, S. E. (2011). Retention and Reform: An Evaluation of Peer-Led Team Learning. *Journal of Chemical Education, 88*(6), 703-707. https://doi.org/10.1021/ed100689m

Lockie, N. M., & Van Lanen, R. J. (2008). Impact of the Supplemental Instruction Experience on Science SI Leaders. *Journal of Developmental Education, 31*(3), 2-4.

Martin, D. C., & Arendale, D. R. (1992). *Supplemental Instruction: Improving First-Year Student Success in High-Risk Courses. The Freshman Year Experience: Monograph Series Number 7.* https://eric.ed.gov/?id=ED354839

May, D. B., & Etkina, E. (2002). College physics students' epistemological self-reflection and its relationship to conceptual learning. *American Journal of Physics, 70*(12), 1249-1258. https://doi.org/10.1119/1.1503377

Moore, R., & LeDee, O. (2006). Supplemental Instruction and the Performance of Developmental Education Students in an Introductory Biology Course. *Journal of College Reading and Learning, 36*(2), 9-20. https://doi.org/10.1080/10790195.2006.10850184

Mutanyatta-Comar, J., & Mooring, S. R. (2019). Evaluation of a Peer-Led Team Learning-Flipped Classroom Reform in Large Enrollment Organic Chemistry Courses. In *From General to Organic Chemistry: Courses and Curricula to Enhance Student Retention* (Vol. 1341, pp. 145-157). American Chemical Society. https://doi.org/10.1021/bk-2019-1341.ch011

Nadelson, L. S., & Fannigan, J. (2014). A Path Less Traveled: Fostering STEM Majors' Professional Identity Development through Engagement as STEM Learning Assistants. *Journal of Higher Education Theory and Practice, 14*(5), 29-41.

Ordoñez, X. G., Ponsoda, V., Abad, F. J., & Romero, S. J. (2008). Measurement of Epistemological Beliefs: Psychometric Properties of the EQEBI Test Scores. *Educational and Psychological Measurement, 69*(2), 287-302. https://doi.org/10.1177/0013164408323226

Otero, V. (2015). Nationally scaled model for leveraging Course Transformation with Physics Teacher Preparation: The Colorado Learning Assistant Model. In C. Sandifer & E. Brewe (Eds.), Recruiting and Educating Future Physics Teachers: Case Studies and Effective Practices (pp. 107-116). American Physical Society.

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*(2), 020104-020101--020104-020107. http://dx.doi.org/10.1103/PhysRevSTPER.4.020104

Parkinson, M. (2009). The Effect of Peer Assisted Learning Support (PALS) on Performance in Mathematics and Chemistry. *Innovations in Education and Teaching International, 46*(4), 381-392. https://doi.org/10.1080/14703290903301784

Peffer, M. E., & Ramezani, N. (2019). Assessing epistemological beliefs of experts and novices via practices in authentic science inquiry. *International Journal of STEM Education, 6*(1), 3. https://doi.org/10.1186/s40594-018-0157-9

Pellegrino, J. W. (2016). 21st Century Science Assessment: The Future Is Now (SRI Education White Paper, Issue. S. International.

Perkins, K. K., Adams, W. K., Pollock, S. J., Finkelstein, N. D., & Wieman, C. E. (2005). Correlating Student Beliefs With Student Learning Using The Colorado Learning Attitudes about Science Survey [Article]. *AIP Conference Proceedings, 790*(1), 61-64. https://doi.org/10.1063/1.2084701

Peterfreund, A. R., Rath, K. A., Xenos, S. P., & Bayliss, F. (2008). The Impact of Supplemental Instruction on Students in STEM Courses: Results from San Francisco State University. *Journal of College Student Retention: Research, Theory & Practice, 9*(4), 487-503. http://baywood.metapress.com/link.asp?target=contribution&id=A51P5824Q7868624

Qualtrics. In. (2020). Qualtrics. https://www.qualtrics.com/

Rath, K. A., Peterfreund, A., Bayliss, F., Runquist, E., & Simonis, U. (2012). Impact of Supplemental Instruction in Entry-Level Chemistry Courses at a Midsized Public University. *Journal of Chemical Education, 89*(4), 449-455. http://dx.doi.org/10.1021/ed100337a

Rath, K. A., Peterfreund, A. R., Xenos, S. P., Bayliss, F., & Carnal, N. (2007). Supplemental Instruction in Introductory Biology I: Enhancing the Performance and Retention of Underrepresented Minority Students. *CBE - Life Sciences Education, 6*(3), 203-216. https://www.lifescied.org/doi/10.1187/cbe.06-10-0198

Redish, E. F., Saul, J. M., & Steinberg, R. N. (2000). Student Expectations in Introductory Physics. *American Journal of Physics, 66*(112). https://doi.org/10.1119/1.18847

Sabella, M. S., Van Duzor, A. G., & Davenport, F. (2016). Leveraging the expertise of the urban STEM student in developing an effective LA Program: LA and Instructor Partnerships. *2016 Physics Education Research Conference (PERC) Proceedings*, 288-291. https://www.per-central.org/items/detail.cfm?ID=14250

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*(3), 498-504. https://doi.org/10.1037/0022-0663.82.3.498

Schommer, M. (1993). Epistemological development and academic performance among secondary students. *Journal of Educational Psychology, 85*(3), 406-411. https://doi.org/https://doi.org/10.1037/0022-0663.85.3.406

Schraw, G., Dunkle, M. E., & Bendixen, L. D. (1995). Cognitive processes in well-defined and ill-defined problem solving. *Applied Cognitive Psychology, 9*(6), 523-538. https://doi.org/10.1002/acp.2350090605

Semsar, K., Knight, J. K., Birol, G., & Smith, M. K. (2011). The Colorado Learning Attitudes about Science Survey (CLASS) for Use in Biology. *CBE - Life Sciences Education, 10*(3), 268-278. https://www.lifescied.org/doi/full/10.1187/cbe.10-10-0133

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. https://doi.org/10.1002/he.20002

Talbot, R. M., Hartley, L. M., Marzetta, K., & Wee, B. S. (2015). Transforming Undergraduate Science Education With Learning Assistants: Student Satisfaction in Large-Enrollment Courses [Article]. *Journal of College Science Teaching, 44*(5), 24-30. http://digital.nsta.org/publication/?i=254738&article_id=1986947

Tingson-Gatuz, C. R. (2009). *Mentoring the Leader: The Role of Peer Mentoring in the Leadership Development of Students-of-Color in Higher Education* Michigan State University]. https://www.proquest.com/docview/304952859

Top, L. M., Schoonraad, S. A., & Otero, V. K. (2018). Development of pedagogical knowledge among learning assistants. *International Journal of STEM Education, 5*(1), 1. https://doi.org/10.1186/s40594-017-0097-9

Van Dusen, B., & Nissen, J. (2020). Associations between learning assistants, passing introductory physics, and equity: A quantitative critical race theory investigation. *Physical Review Physics Education Research, 16*(1), 010117. https://doi.org/10.1103/PhysRevPhysEducRes.16.010117

Van Dusen, B., & Nissen, J. M. (2017). Systemic inequities in introductory physics courses: the impacts of learning assistants. *2017 Physics Education Research Conference (PERC) Proceedings*, 400-403. https://doi.org/10.1119/perc.2017.pr.095

Warren, A. R. (2018). Quantitative critical thinking: Student activities using Bayesian updating. *American Journal of Physics, 86*(5), 368-380. https://doi.org/10.1119/1.5012750

Warren, A. R. (2020). Impact of Bayesian updating activities on student epistemologies. *Physical Review Physics Education Research, 16*(1), 010101. https://doi.org/10.1103/PhysRevPhysEducRes.16.010101

Weidler-Lewis, J., Pierce, D., & Walter, C. (2013). Conflicting Discourses: An Examination of "Time" in Online Learning Courses Facilitated by Learning Assistants. Proceedings of EdMedia 2013--World Conference on Educational Media and Technology, Victoria, Canada.

White, J.-S. S., Van Dusen, B., & Roualdes, E. A. (2016). The impacts of learning assistants on student learning of physics. *2016 Physics Education Research Conference (PERC) Proceedings*, 384-387. https://doi.org/10.1119/perc.2016.pr.091

Wilson, S. B., & Varma-Nelson, P. (2016). Small Groups, Significant Impact: A Review of Peer-Led Team Learning Research with Implications for STEM Education Researchers and Faculty. *Journal of Chemical Education, 93*(10), 1686-1702. https://doi.org/10.1021/acs.jchemed.5b00862

Wood, P., & Kardash, C. (2002). Critical elements in the design and analysis of studies of epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 231–260). Lawrence Erlbaum Associates Publishers.

Appendix A

Table A1
Instruments and Techniques Used to Measure Attitudes and Epistemological Beliefs

Instrument	Source
Attitudes, beliefs, and expectations	
Views about Science Survey (VASS)	Halloun & Hestenes (1998)
Maryland Physics Expectations survey (MPEX)	Redish et al. (2000)
Chemistry Expectations about learning survey (CHEMX)	Grove & Bretz (2007)
CLASS-Phys	Perkins et al. (2005)
CLASS-Chem	Barbera et al. (2008)
CLASS-Bio	Semsar et al. (2011)
Attitude towards the Subject of Chemistry Inventory (ASCI)	Bauer (2008)
Chemistry Self-Concept Inventory (CSCI)	Bauer (2005)
General Epistemological Surveys and Instruments	
Epistemological Beliefs Inventory (EBI)	Schraw et al. (1995)
Epistemological Beliefs Questionnaire (EBQ)	Jena & Chakraborty (2018); Schommer (1993)
Epistemological Questionnaire (EQ)	Cano (2005)
Epistemic Beliefs Survey (EBs)	Wood & Kardash (2002)
combined Epistemological Questionnaire and Epistemic Beliefs Inventory (EQEBI)	Ordoñez et al. (2008)
Schommer's Beliefs about Knowledge and Learning Questionnaire	Schommer (1990)
Written reflections	May & Etkina (2002)
Practices in authentic science inquiry	Peffer & Ramezani (2019)
STEM-Specific Epistemological Instruments	
Epistemological Beliefs about Physical Science Survey (EBAPS)	Elby (2006a)
Epistemological Beliefs Assessment for Engineering (EBAE)	Carberry et al. (2010)
Questionnaires	Conley et al., 2004; Kampa et al. (2016)

Appendix B

Table B1

Description of EBAPS Survey Components (Elby, 2006a)

Axis	Label	Description	Items*
1	**Structure of scientific knowledge**	Students' view physics and chemistry knowledge as "a bunch of weakly connected pieces without much structure and consisting mainly of facts and formulas" or as "a coherent, conceptual, highly-structured, unified whole"	10
2	**Nature of knowing and learning**	Students' view learning science as "consist[ing] mainly of absorbing information" or "constructing one's own understanding" through active engagement, experiences, and reflection	8
3	**Real-life applicability**	Students' view scientific knowledge and ways of thinking as limited only to specific academic/scientific settings or more broadly to real life settings. "These items tease out students' views of the applicability of scientific knowledge *as distinct from* the student's own desire to apply science to real life, which depends on the student's interests, goals, and other non-epistemological factors."	4†
4	**Evolving knowledge**	Students' view scientific knowledge along a continuum from absolutism (e.g., "all scientific knowledge is set in stone") to extreme relativism (e.g., "no distinctions between evidence-based reasoning and mere opinion").	3‡
5	**Source of ability to learn**	Students' views of being "good" at science exist along a continuum from fixed natural ability to the result of hard work and effective study strategies. Note: these views are intended to be "*distinct from* [students'] self-confidence and other beliefs about themselves."	5†‡

* Two items do not map onto any Axis.

† One item is included in both Axis 3 and Axis 5.

‡ One item is included in both Axis 4 and Axis 5.