ANNETTE CALDWELL SIMMONS
SCHOOL OF EDUCATION & HUMAN DEVELOPMENT

RESEARCH IN MATHEMATICS EDUCATION

STEM Academy Teacher Outcomes Program Evaluation: Cohorts 1 and 2

Technical Report 20-15

STEM Academy Teacher Outcomes Program Evaluation: Cohorts 1 and 2

Karen Pierce •Anthony Sparks • Elizabeth Adams • C. Taylor Cox • Cassandra Hatfield •Leanne Ketterlin Geller

Southern Methodist University

Summer 2020

Published by
Southern Methodist University
Department of Education Policy & Leadership
Simmons School of Education & Human Development
PO Box 750114
Dallas, TX 75275-0114
Contact information: rme@smu.edu

This research was supported by Texas Instruments Foundation and O'Donnell Foundation; GT00317. Opinions expressed herein do not necessarily reflect those of The Texas Instruments Foundation; O'Donnell Foundation or individuals within.

Abstract

Dallas Independent School District (ISD) is a large school system serving a majority percentage of students who identify as Black or Hispanic and from low socio-economic backgrounds. In an effort to increase Dallas ISD's student interest and achievement in STEM, district leadership partnered with the Texas Instruments Foundation, Southern Methodist University (SMU), and the O'Donnell Foundation to develop and implement the *STEM Academy for Science Teachers and Leaders*. As a part of the Academy, teachers and leaders engage in two primary components including: (a) intensive summer professional development on the SMU campus, and (b) one-on-one coaching with an SMU coach during the school year. Teachers participated in the STEM Academy for up to three years.

The purpose of this report is to summarize trends across time on the measured constructs for the participating teachers. These include the teachers' implementation of active learning, pedagogical content knowledge (PCK), science self-efficacy, and perceived STEM importance and confidence. In addition to summarizing the trends on these measured constructs, this report includes a brief description of the three summer academies, the coaching model, and information on the participating teachers.

The teacher implementation of active learning as measured by three different tools indicated that the frequency and quality of instruction may be more impacted by the timepoint within the year or by other year specific factors such as the perspective of the rater, than by the progression across years in the STEM Academy program. While the teachers self-reported a slight, but non-significant, increase in frequency of active learning strategy use, the UTeach Observation Protocol (UTOP) indicated that external raters found implementation to have increased across time. Although the external raters captured increased scores in all four measured domains of the UTOP across the two captured years of participation, the STEM Teacher Observation Protocol (STEM TOP) scale scores indicated that management and discipline may be cyclic within a given academic year, with an overall negative trend. STEM TOP integrated STEM instruction scale scores generally increased over time, and significantly improved for some teachers across certain timepoints. The STEM TOP was completed by program coaches, who had an established relationship with the teacher. The coaches tended to report less consistent gains and in some cases no change or decrease across time in teachers' implementation of active learning.

Regarding teacher science self-efficacy, and the teachers' perceived importance concerning the implementation of active learning in their classrooms, no consistent increases were measured across cohorts. Informed by Bandura's social learning theory, self-efficacy relates to teachers' beliefs that an action will have a favorable result (outcome expectation) and that they can perform the action successfully (self-efficacy expectation) (Bleicher, 2004). Teachers with higher self-efficacy for teaching are more likely to implement innovative, evidence-based instructional practices. We hypothesized that we would observe increases in teachers' science self-efficacy, which would contribute to increased implementation of STEM instructional strategies. During the first year of participation, cohort 2 teachers did show an improvement in their confidence level regarding the implementation of active learning strategies, but this growth was not maintained through the second year with a smaller sample. Relative to cohort 2 teachers, cohort 1 teachers consistently scored higher on measures of frequency, implementation,

and PCK, so it is possible that they have a more positive outlook on their own pedagogical abilities.

Content knowledge (CK) is defined as content-specific knowledge about a discipline; whereas, PCK is defined as the integration of CK and appropriate pedagogy for teaching that knowledge (Shulman, 1986; van Driel et al., 1998). This type of knowledge is necessary to make learning accessible to students. Teacher confidence in their ability to teach STEM correlates with teachers' CK and enacted science instruction (Munck, 2007; Nadelson et al., 2013). Teachers who were more confident in their ability to implement STEM instructional strategies were also more likely to incorporate inquiry-based instruction; whereas, other teachers expressed hesitation and doubt (Nadelson et al., 2013). The STEM Academy includes a strong emphasis on building teachers' confidence for implementing STEM instruction. As such, we anticipated that we would observe increases in teachers' confidence for teaching STEM across time. There were significant increases in science content as measured by the UTOP and PCK. In addition, teachers' confidence seemed to increase slightly, although only significantly for cohort 2 year 1. This may suggest that the STEM Academy influences teachers' science content implementation and PCK.

Finally, in a different study that utilized qualitative interviews, it was found that teachers implemented active learning strategies in differential patterns (Adams, Knox, Hatfield, & Ketterlin-Geller, 2020). These differential patterns stem from variation in the frequency and quality of the teachers' implementation, which could also be linked to the student motivation in those classes, and ultimately the teachers' overall scores captured on a protocol such as the STEM TOP or UTOP. Although this work is ongoing, it may provide an additional lens through which to consider the results presented in this report. One limitation of this study is that no comparison group of teachers was evaluated. Only the teachers participating in the STEM Academy and receiving treatment were given the surveys and observed.

Four recommendations for improving the STEM Academy in the future are suggested, based on the results and analysis within this report. First, future iterations or similar programs to the STEM Academy should include external raters during the observation process. The trends observed in this report suggest that coaches' scores may have been influenced by their relationship with the teacher, since the UTOP increased across time while the STEM TOP decreased. Second, in order to further alleviate the possibility of rater drift, future implementations of programs similar to the STEM Academy should consider specific coach training in order to reduce possible bias. Next, we recommend to the extent possible that teachers teach same grades across years to encourage developing expertise is specific content and process standards by grade and reduce possible obstacles in navigating new grade levels. Finally, it is important to examine and understand the profiles of schools and teachers who are successful in implementing active learning. This might include a needs assessment or measure of readiness for active learning. This report focuses on describing teacher trends in the aggregate, but is likely that extensive variability exists in teachers' experiences within the program depending on teacher and school factors. Efforts should be made to better understand individual teachers' experiences based on teacher interviews and other factors.

Table of Contents

STEM Academy Teacher Outcomes Program Evaluation: Cohorts 1 and 2

Background

The question of whether education is a private or a public good has been considered for decades. In 1997, Labaree outlined three major goals within education that, in his opinion, have shifted in importance over time. Democratic equality and social efficiency classify education as a public good, whereas social mobility emphasizes the advantages to the private individual (Labaree, 1997). Support and prioritization of STEM education in the United States is increasing, largely because industry, governmental actors, and individual citizens have all recognized the potential advancement across all sectors that improvements in this field could provide. Given the popularity of STEM advancement, it is surprising that, until recently, the STEM potential of many Americans remained relatively untapped.

The number of STEM related jobs grew at three times the rate of non-STEM jobs between the years 2000 and 2010 (Smithsonian, 2018), and both the American and global economies require more individuals with STEM related degrees to fill professional positions in an increasingly high-tech job market (DeJarnette, 2012). Although the United States has experienced growth in this field, it has not seen the same growth in qualified STEM workers as its global competitors in Europe and Asia (National Science Board, 2010).

While increasing the number of STEM qualified workers is relevant for improving American social efficiency, STEM education is also a powerful mechanism for social mobility. In 2012, individuals who held a bachelor's degree in a STEM major were more likely to be employed following graduation and had higher salaries compared to other students (The National Center for Education Statistics, 2014). In the same year, the full-time employment rate for STEM majors was seven percentage points higher than the rate for graduates overall (77% compared to 70%). Furthermore, the median salary for STEM majors was $14,000 higher than the median for students overall ($60,000 compared to $46,000). This evidence shows that the pursuit of a STEM career is both promising in terms of job attainment and salary opportunities.

Despite this evidence, individuals who belong to historically underrepresented subgroups continue to be less likely to pursue STEM careers. These underrepresented subgroups include women (Mau, 2016, Sassler, Glass, Levitte, & Michelmore, 2017), students with disabilities (Basham & Marino, 2013), students from low-income households (Chen, 2009; Dika & D'Amico, 2016), and individuals who identify as Black or Hispanic (Fealing, Lai, & Myers, 2015, Mau, 2016). From a democratic equality perspective, these findings are especially troubling. If one of the goals of education is to produce an informed citizenry, then every member of the population should have access to the STEM field. Significant disparities in STEM career attainment between subgroups is indicative of a deeper problem within the American education system, which only exacerbates the problems of social mobility and efficiency described above.

In 2013, Texas House Bill 5 (HB 5) required that Grade 8 students select an endorsement area, including STEM, Business and Industry, Public Services, Arts & Humanities, or Multidisciplinary Studies. During the 2014-2015 school year, just 16.9% of Dallas Independent School District (ISD) students selected the STEM pathway, despite the fact that a wide range of STEM industries are based in Dallas.

In response to these statistics, a partnership between the Texas Instruments Foundation, the O'Donnell Foundation, Southern Methodist University (SMU), and Dallas Independent School District (ISD) was established. A primary goal of this partnership was to determine how students' interest and perseverance in STEM could be significantly improved, ultimately affecting the STEM pipeline and equity in the technical fields. Four key areas were identified, including (a) active learning which includes inquiry-based STEM instructional strategies such as project based learning (PBL) and maker-based instruction (MBI), (b) scientific process standards, (c) teacher content knowledge, and (d) differentiated support for all learners, with an emphasis on social and emotional learning (Perry, Reeder, Brattain, Hatfield, & Ketterlin-Geller, 2017). Through these conversations, desired outcomes were determined that would help initiate and refine the goals of this 4-year project. The primary desired outcomes included (a) an increase in student science achievement and motivation, and (b) an increase in teacher implementation of active learning experiences.

Overview of the Project

There are two main components of the STEM Academy for Science Teachers and Leaders (STEM Academy hereafter). For teachers, the first component is an intensive 90-hour professional development academy each summer, and the second is on-site coaching and professional learning community (PLC) facilitation that occurs during the academic year at the campus level. For additional detail about the project, please reference previous evaluation reports (Adams, Hatfield, Cox, & Ketterlin-Geller, 2018a; Adams, Hatfield, Cox, Mota, Sparks, & Ketterlin-Geller, 2018b; Perry et al., 2017; Pierce, Adams, Rhone, Hatfield, & Ketterlin-Geller, 2019a; Pierce, Cox, Hatfield, Adams, & Ketterlin Geller, 2019b).

The program followed a cohort model, with the first cohort of teachers completing three years of program (cohort 1), and a second cohort of teachers completing two years (cohort 2). Cohort 1 teachers began participation in summer 2017 and cohort 2 teachers began participation in summer 2018.

The courses and other online supplemental components administered by the STEM Academy facilitators over the summer will hereafter be referred to as Academy 1, Academy 2, and Academy 3, for each year of program participation respectively. All Academy content was structured around four key areas that were identified during the development of the STEM Academy goals as being especially influential in fostering both student and teacher interest and success. These key areas are depicted in Figure 1.

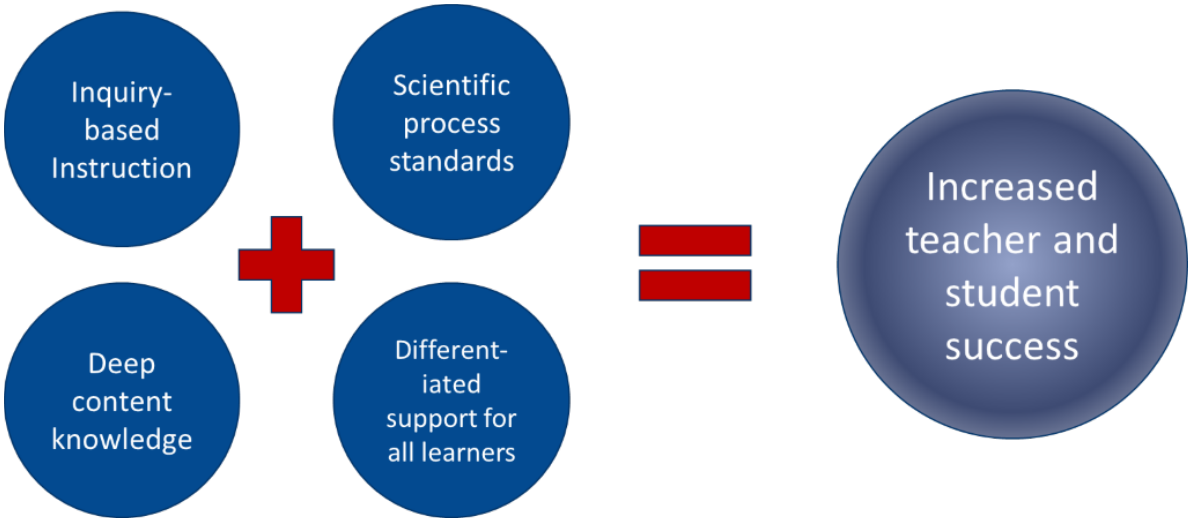


Figure 1. Key Areas of the STEM Academy

The main outcomes of the STEM Academy focus on teachers and students (i.e., increased teacher and student success). The STEM Academy focused on the development of participating teachers as leaders in their departments as a means of achieving this goal. Active learning and inquiry-based instruction in the science classroom lead to a better conceptual understanding by students according to studies conducted between 1984 and 2002 (Minner, Levy, & Century, 2010). Furthermore, a 2017 study identified sustained professional development in inquiry-based instructional strategies for teachers as having a positive trend on student growth in science mastery and a narrowed achievement gap within the scientific fields for students (Marshall, Smart, & Alston, 2017; Geier, Blumenfeld, Marx, Krajcik, Fishman, Soloway, & Clay-Chamber, 2008).

An understanding of the scientific process standards and deep content knowledge is required for effective implementation of inquiry-based and active learning instruction (NRC, 2000). Therefore, utilization of the Texas Essential Knowledge and Skills (TEKS) was embedded in several aspects of lesson design during all three years of the program. In addition, the program emphasized differentiated support for all learners by attending to students' social and emotional learning (SEL). It has been suggested that teachers who have higher social and emotional competence have better teacher-student relationships and better management of their classrooms through differentiated support structures, ultimately leading to better student learning (Jennings & Greenburg, 2009). As outlined in the next section, inquiry, content knowledge, process standards, and SEL components permeated all three years of the STEM Academy.

Academy 3 content built on the key areas, which were cultivated during Academies 1 and 2 (see Perry et al., 2017 and Pierce et al., 2019 for more detail). Each subsequent Academy continued to develop teachers' understanding and implementation of maker-based instruction (MBI) and project-based learning (PBL), as well as community-based resources and Professional Learning Community (PLC) meetings (see Figure 2 below).

Academy 1 was held in 2017 for 16 cohort 1 teachers from six Dallas ISD schools. The following year 12 cohort 1 teachers from seven schools continued with STEM Academy and participated in Academy 2. In 2019, ten cohort 1 teachers from six Dallas ISD schools participated in Academy 3. Cohort 2 teachers joined the STEM Academy during the summer of 2018, and 30 teachers from ten Dallas ISD schools participated in Academy 1. During the final year, 16 cohort 2 teachers from seven schools participated in Academy 2.

Summer Academies

The SMU project team designed the content and structure of the three summer Academies to meet the goals and objectives of the project and align with the needs and constraints of instructional leaders and science teachers. At the beginning of the project, in order to facilitate collaboration, the SMU team met monthly with representatives from Dallas ISD's STEM curriculum department and the research and evaluation department.

To meet the practical constraints of the science teachers, the SMU project team conducted some portions of the STEM Academies online and other portions face-to-face. Overall each summer, participating teachers received 90 hours of professional development: 70 hours of face-to-face coursework on the SMU campus during two weeks in June or July, and 20 hours of learning through the online modules.

The major themes and the different strategies that were included in each of the three summer Academies are shown in Figure 2 below. Some of the items, specifically the ones that align with the core pillars, were present across all three years. Other items were utilized during only one or two years as mechanisms for advancing the main goals of the STEM Academy.

<table>
<tr><th></th><th>Academy 1</th><th>Academy 2</th><th>Academy 3</th></tr>
<tr><td rowspan="5">Major Themes</td><td colspan="3">Inquiry</td></tr>
<tr><td colspan="3">Content Knowledge</td></tr>
<tr><td colspan="3">Process Standards</td></tr>
<tr><td></td><td colspan="2">Differentiation</td></tr>
<tr><td></td><td></td><td>Teacher Leadership</td></tr>
<tr><td rowspan="8">Emphasized Strategies</td><td colspan="3">Maker Based Instruction</td></tr>
<tr><td colspan="3">Project Based Learning</td></tr>
<tr><td colspan="3">Social and Emotional Learning</td></tr>
<tr><td colspan="3">Community-based Resources</td></tr>
<tr><td></td><td>The 5E Model of Instruction</td><td></td></tr>
<tr><td></td><td></td><td>The Science of Learning in STEM Education</td></tr>
<tr><td></td><td></td><td>Teacher Leadership Through Coaching</td></tr>
<tr><td></td><td></td><td>Professional Learning Communities</td></tr>
</table>

Figure 2. Focus of Each Summer Academy Organized by Major Theme and Emphasized Strategies

Academy 1

In keeping with the goals of the project, Academy 1 focused on inquiry-based instruction, scientific process standards, and teacher content knowledge. After considering current research in STEM education, MBI and PBL were selected as the inquiry-based pedagogical approaches for the STEM Academy.

MBI has two components: proficiency and purpose. MBI requires developing proficiency using a tool (e.g., vinyl cutter, 3D printer). Then that tool is used for a purpose (e.g. to make something that solves a problem or answers a question). MBI is hands-on, student-directed, iterative, and does not have a pre-determined outcome. PBL is organized around a central, driving question. Students work collaboratively to answer an authentic, STEM-related question. The driving question is typically broad enough to cover a variety of instructional standards. Students lead the project by asking more questions, testing hypotheses, conducting research, and designing and carrying out experiments. Usually, the assessment is embedded throughout the PBL unit, both formally and informally. Often, the final artifact is presented to peers, educators, and outside experts. During Academy 1, teachers generated instructional units that employed PBL and MBI approaches to teach high-priority TEKS. The instructional units were intended to be delivered during the following academic year.

In addition to MBI and PBL, instruction was also provided on SEL and community-based STEM educational resources by participating in field experiences such as the Dallas Zoo, the Trinity River Audubon Center, the Frontier of Flight Museum, or the Dallas Area Rapid Transit (DART) headquarters. Additional information about the two years of Academy 1 delivery can be found in two previous evaluation reports (Adams et al., 2018a; Perry et al., 2017).

Academy 2

This Academy was structured similarly to Academy 1, in that participating teachers attended 90 hours of professional development during the summer. During Academy 2, MBI and PBL strategies were revisited and additional techniques were offered by the STEM Academy instructors along with suggested modifications based on what the teachers reported had occurred in their classrooms during the previous year. Additionally, teaching strategies for using 5E, a questioning and engagement technique used by Dallas ISD and intended to interest and engage students in the learning process, were discussed and incorporated into lesson planning.

Academy 2 continued the focus on inquiry methods, process standards, and deep content knowledge, and also more heavily emphasized differentiation for all types of student learners compared to Academy 1. SEL and new community-based resources were included as well. For more details on Academy 2 see the two previous evaluation reports (Pierce et al., 2019a; Sparks, Adams, Mota, Simon, Burton, Hatfield, & Ketterlin Geller, 2019b).

Academy 3

Academy 3 was held for the cohort 1 teachers in summer 2019. It had a similar structure to Academies 1 and 2, however most of the face-to-face content was delivered by SMU professors and was affiliated with SMU courses. In keeping with the goals of the project, Academy 3 focused on inquiry-based instruction, scientific process standards, teacher content knowledge, and differentiated support for all learners. During Academy 3, MBI, PBL, and SEL were revisited, and additional components of teacher leadership and PLC structure and function were added. In the previous Academy 3 evaluation report, these instructional approaches are described in detail (Pierce et al., 2019b).

Coaching

In the first year of implementation, one coach provided instructional coaching and led PLC meetings with the participating 16 teachers. In the second year of implementation, four additional coaches served the increased number of schools and teachers with the original coach serving as the lead coach. During the third year, due to teacher attrition, only three coaches and one lead coach served the participating teachers and campuses. The lead coach was responsible for coach training and supporting problem solving with scheduling or challenges observed in classrooms.

The structure of the STEM Academy coaching includes a one-on-one pre-conference, observation, and post-conference, which is defined as a full cycle of coaching. Each year teachers engaged in up to seven coaching cycles and PLC meetings with their instructional coach from SMU. The coaching model offered teachers support in understanding and implementing aspects that were learned during the summer professional development academy. Extant research shows that this type of individualized support (i.e., coaching and PLC meetings) facilitates the effective implementation of instructional strategies and practices (Kraft and Blazar, 2017). In a meta-analysis examining existing experimental and quasi-experimental research, Kraft and Blazar (2017) found that coaching affects long-term, sustained change in teachers' instructional practices and student outcomes. Whereas the summer professional development occurred over one specific period of time, coaching and PLC meetings were sustained throughout the school year. In addition, the one-on-one nature of the coaching allowed dialogue to be highly individualized, context specific, and targeted toward specific pedagogical skills (Kraft et al., 2017).

For more information about the coaching and PLC implementation across all years, see Adams et al. (2018b) and Sparks, Adams, Cox, Hatfield, and Ketterlin Geller (2019a).

Purpose of this Report

The purpose of this report is to summarize trends across time on the measured constructs for participating teachers. These include the teachers' implementation of active learning, pedagogical content knowledge, science self-efficacy, and perceived STEM importance and confidence. The STEM Academy project also included leader coaching and support, but that is outside the scope of this report. For information on the leader components of the project, see Pierce, Adams, Sparks, Cox, Knox, Hatfield, and Ketterlin Geller (2020).

Evaluation Questions

This report is guided by the following evaluations questions:

1. To what extent did teachers implement active learning strategies in their classrooms? Did teachers increase in their implementation across time?

2. The STEM Academy emphasized teacher pedagogical content knowledge (PCK), which integrates teachers' content knowledge with how to teach that content knowledge. To what extent did teachers demonstrate science PCK? Did teachers' PCK increase across time?

3. To what extent did teachers feel efficacious in teaching science? Did teachers' science self-efficacy increase across time?

4. To what extent did teachers report importance and confidence in using active learning in their classrooms? Did teachers' perceived importance and confidence increase across time?

For teacher results, we report results for only teachers who completed data at all timepoints, which facilitates comparisons across time. Trends across time for all participating teachers are available upon request.

Participating Teachers

Who are the teachers who participated in the STEM Academy?

The participating teachers completed the Teacher Information Form in order to provide background and demographic information to the STEM Academy research team. This questionnaire included items about the participants' gender, race, and ethnicity, as well as gathering the number of years the participants had been teaching, teaching science, working in education generally, working in other career fields, and working at their specific campus. Finally, information on the teacher's education and certifications was collected, along with the number of hours and content of the in-service training they had attended over the past year. The information obtained from this form is summarized below.

For cohort 1 teachers, we summarize teacher characteristics across three years of implementation. For cohort 2 teachers, we summarize teacher characteristics across two years of implementation. For cohort 1, the years of participation are 2017-18 (Y1), 2018-19 (Y2), and 2019-20 (Y3), and for cohort 2 the years of participation are 2018-19 (Y1) and 2019-20 (Y2).

Teacher Recruitment

During the first year of implementation, STEM Academy was able to successfully recruit 16 teachers at six middle schools. Over the course of the first year, one teacher withdrew from the program. During 2018-19, the STEM Academy enrolled the 15 persisting cohort 1 teachers, 6 additional Grades 6, 7, and 8 science teachers at the six previously participating middle schools and 24 science teachers at seven new middle schools. Two cohort 1 teachers moved to new schools, but persisted in the program, bringing STEM Academy to 15 different schools during year two.

To incentivize participation starting in 2018-19, funds were allocated for teachers to receive a $1,000 stipend in addition to course credit. This stipend was designed to compensate teachers for the 70 hours of face-to-face coursework on the SMU campus in the summer. This decision was designed to encourage participation for teachers who already held a master's degree or were currently enrolled in a master's program. In addition, based on teacher feedback, SMU and Dallas ISD tailored the communication plan to meet the needs of teachers and leaders in the district. For more details on the recruitment efforts and the teacher participation interest survey, see the previous report from Adams, Hatfield, and Ketterlin Geller (2018).

Cohort 1

Table 1 shows that 16 cohort 1 teachers participated in 2017-18. Of those teachers, 12 continued in 2018-19 and 9 continued in 2019-20. The majority of teachers were Female, Black, and not Hispanic or Latino. In 2019-20, cohort 1 teachers reported an average of 6 years of teaching experience. On average, cohort 1 teachers taught science for five years.

Table 1

Cohort 1 Teacher Demographic Information

Characteristic		2017-18 # of Teachers	2017-18 % of Teachers	2018-19 # of Teachers	2018-19 % of Teachers	2019-20 # of Teachers	2019-20 % of Teachers
Gender	Male	4	25%	3	25%	2	22%
	Female	12	75%	9	75%	7	78%
Race	Alaska Native	0	0%	0	0%	0	0%
	Asian	0	0%	0	0%	0	0%
	Black	9	56%	7	58%	5	56%
	Native Hawaiian	0	0%	0	0%	0	0%
	Other Pacific Islander	0	0%	0	0%	0	0%
	White	7	44%	5	42%	4	44%
Ethnicity	Hispanic or Latino	4	25%	3	25%	2	22%
	Not Hispanic or Latino	12	75%	9	75%	7	78%
Total		16	100%	12	100%	9	100%

Table 2 shows the grade levels taught by participating cohort 1 teachers. In 2017-18, 2018-19, and 2019-20, most cohort 1 teachers taught Grades 7 and 8.

Table 2

Grade-levels Taught by Cohort 1 Teachers

Current Grade Level	2017-18 Number of Teachers	2018-19 Number of Teachers	2019-20 Number of Teachers
6th	1	0	1
7th	9	10	6
8th	10	9	6

Note: Four teachers taught more than one grade level in 2017-18, seven teachers taught more than one grade level in 2018-19, and four teachers taught more than one grade level in 2019-20.

For additional descriptive information about teachers in cohort 1, please reference published academy evaluations (Adams et al., 2018a; Perry et al., 2017; Pierce et al., 2019a; Pierce et al., 2019b; Sparks et al., 2019a).

Cohort 2

Table 3 shows that 29 cohort 2 teachers participated in 2018-19. Of those teachers, 17 continued in 2019-20. Similar to cohort 1, the majority of teachers were Female, Black, and not Hispanic or Latino. In 2019-20, cohort 2 teachers reported an average of 9 years of teaching experience. On average, cohort 1 teachers taught science for 8 years.

Table 3

Cohort 2 Teacher Demographic Information

	Characteristic	**2018-19** # of Teachers	**2018-19** % of Teachers	**2019-20** # of Teachers	**2019-20** % of Teachers
Gender	Male	9	31%	5	29%
	Female	20	69%	12	71%
Race	Alaska Native	0	0%	1	6%
	Asian	2	7%	0	0%
	Black	15	52%	9	53%
	Native Hawaiian	0	0%	0	0%
	Other/Pacific Islander	0	0%	1	6%
	White	9	31%	4	24%
	Two or More Races	3	10%	2	12%
Ethnicity	Hispanic or Latino	2	7%	2	12%
	Not Hispanic or Latino	27	93%	15	88%
Total		29	100%	17	100%

Table 4 shows the grade levels taught by participating teachers. In 2018-19, 16 teachers taught Grade 6, 11 teachers taught Grade 7, and 7 teachers taught Grade 8. In 2019-20, 13 teachers taught Grade 6, 11 teachers taught Grade 7, and 8 teachers taught Grade 8.

Table 4

Grade-level Taught

Current Grade Level	**2018-19** Number of Teachers	**2019-20** Number of Teachers
6th	16	13
7th	12	7
8th	7	8

Note: In 2018-19 four teachers taught two grade levels, and one taught all three grade levels. In 2019-20 five teachers taught two grade levels and three taught all three grade levels.

For additional descriptive information about teachers in cohort 2, please reference published academy evaluations (Adams et al., 2018a; Sparks et al., 2019b). For information about the characteristics of exiting teachers and their reasons for exit, please refer to Cox, Adams, & Ketterlin-Geller (2020).

Measures

Throughout the STEM Academy, the participating teachers completed several surveys aimed at assessing various aspects of their content knowledge, attitudes, and enacted practice. Both SMU coaches and external observers gathered data during in-class observations on two different observational tools. This section describes the five different measurement tools used in detail and indicates where each measure can be viewed. Additionally, to maintain continuity, the results in the next section are organized by each measure in the order that they are listed here.

Teacher STEM Perceptions, Practice, and Culture (STEM PPC) survey

Our team developed the STEM PPC survey to measure teachers' perceived importance, confidence, and frequency in using active learning strategies. This survey includes 24 items describing active learning strategies (e.g., learning experiences encourage student ownership). Three strategies were excluded from analysis because they do not describe active learning strategies (e.g., students learn from teacher-led lecture or activities). To measure frequency of implementation, teachers reported their use of each practice on a six-point scale ranging from "less often than 1 time per month" to "everyday." Additionally, the teachers indicated on a four-point scale the importance they placed on each practice, ranging from "not important at all" to "very important." Similarly, the teachers indicated their confidence in implementing each practice on a four-point scale, ranging from "not confident at all" to "very confident."

Teachers completed the STEM PPC at three timepoints: before the face-to-face summer academy (i.e., approximately July), near the beginning of the school year (i.e., approximately October), and near the end of the school year (i.e., approximately February). Because SMU was awaiting IRB approval from Dallas ISD, our team was not able to collect fall 2017 teacher survey responses. The STEM PPC is included in Appendix A. A confirmatory factor analysis supported the internal structure of the STEM PPC (Sparks, Adams, Perry, & Ketterlin-Geller, 2020).

STEM Teacher Observation Protocol (STEM TOP)

Our team developed the STEM TOP based on existing observational measures to capture a snapshot of teachers' implementation of active learning strategies. SMU coaches used the STEM TOP during each coaching observation as a tool to support their feedback to teachers. In order to encourage consistency in the type and quality of feedback teachers received and facilitate the use of STEM TOP scores in reports, SMU coaches were trained and calibrated on the STEM TOP (see Pierce, Adams, Sparks, Burton, Hatfield and Ketterlin Geller, 2020 for information on the STEM TOP development and rater consistency).

SMU coaches completed the STEM TOP approximately seven times across the course of the school year for each participating teacher. These observations, referred to as cycles, occurred between October and April for both the 2018-19 and 2019-20 school years. The STEM TOP includes 22 items, which are organized into two scales including STEM Instruction and Management and Discipline (see Adams, Sparks, and Ketterlin Geller, 2019c for information on the scale development). Each item is rated on a four-point scale ranging from "not observed" to "exemplary." The STEM TOP was developed and piloted during 2017-18; thus, our team was not able to collect STEM TOP data during the 2017-18 academic year. The STEM TOP is included in Appendix B. Exploratory factor analyses and evidence of concurrent validity support the use of the STEM TOP for measuring enacted STEM instruction (Adams, Sparks, & Ketterlin-Geller, 2019c; 2020).

UTeach Observation Protocol (UTOP)

The UTOP is an observational protocol "designed to allow individuals to evaluate teaching effectiveness while valuing different modes of instruction" (UTeach, 2019, par. 1). The UTOP was developed by the UTeach College of Natural Sciences program at the University of Texas at Austin and has 27 items divided into four sections, including classroom culture, lesson structure, implementation, and science content, which are intended to evaluate instruction and provide meaningful feedback. The version of the UTOP used in this study is available at https://utop.uteach.utexas.edu/ (UTeach, 2019). Relative to the STEM TOP, the UTOP is more comprehensive, including lesson purpose, lesson preparation, teaching method/learning activity, and student grouping structure. Each of 27 items scored on a 1 "Not observed at all" to 5 "Observed to a great extent" Likert scale, with some items including an "NA" option if not observed. As a result of the comprehensive nature of the tool, the UTOP requires considerably more time to score. Our team hired and trained external raters to complete the UTOP in the fall and spring in participating teachers' classrooms (see Burton, Hatfield, Adams, and Ketterlin Geller, 2020 for information on the UTOP training and calibration). Because SMU was awaiting RRB approval from Dallas ISD, our team was not able to collect fall 2017 teacher UTOP data.

Pedagogy of Science Teaching Test (POSTT)

Teachers completed an existing teacher assessment of science PCK called the Pedagogy of Science Teaching Test (POSTT) (Cobern, Schuster, Adams, Skjold, Muğaloğlu, Bentz, & Sparks, 2014) at two timepoints each year; before and after each face-to-face summer academy. An item bank for the POSTT is available at https://wmich.edu/science/inquiry-items (Cobern, Schuster, Adams, Skjold, Muğaloğlu, Bentz, & Sparks, 2014). The pre and post assessments included 10 items, which were selected based on relevancy to middle school science content. Each item included a vignette, which described a lesson, and teachers made a recommendation for the lesson from four possible responses. Each response ranged in its alignment to inquiry-oriented instruction on a four-point scale with the categories of response classified as a) open inquiry, b) guided inquiry, c) active direct, and d) didactic direct. The versions of the POSTT used in this study are included in Appendix C.

Science Teacher Efficacy Beliefs Inventory (STEBI)

Teachers completed an existing measure of teacher self-efficacy for teaching science called the Science Teacher Efficacy Beliefs Inventory (STEBI) (Riggs & Enochs, 1989; Enochs & Riggs, 1990). This survey includes two subscales:

- Personal science teaching efficacy beliefs (PSTE) defined as a teacher's belief in their own ability to perform a behavior, and

- Science teaching outcome expectancy (STOE) defined as a teacher's expectation that "certain behaviors [will] produce desirable outcomes" (Riggs & Enochs, 1989, p. 4).

Teachers reported agreement with 25 items on a five-point scale ranging from "strongly disagree" to "strongly agree". The version of the STEBI used in this study is available at https://cpb-us-w2.wpmucdn.com/u.osu.edu/dist/2/5604/files/2014/09/Science-TE-2fbsc7e.pdf (Enochs & Riggs, 1990). Teachers completed the STEBI at three timepoints: before the face-to-face summer academy (i.e., approximately July), near the beginning of the school year (i.e., approximately October), and near the end of the school year (i.e., approximately February). Because SMU was awaiting IRB approval from Dallas ISD, our team was not able to collect fall 2017 teacher survey responses.

Results

The following section includes the results from the five measures described above. This section is organized into four subsections:

- teacher implementation of active learning, which includes teachers' self-reported enacted instruction based on the STEM PPC and observed enacted instruction based on observations conducted by coaches using the STEM TOP and external raters using the UTOP;

- teacher pedagogical content knowledge as measured by the POSTT;

- teacher science self-efficacy as measured by the STEBI; and

- teachers' reported importance and confidence of STEM practices as measured by the STEM PPC.

Additionally, all data is disaggregated by cohort, and for cohort 2 by year. This is because the cohort 1 teachers persisted at a higher rate, and their data can be considered across multiple years. Cohort 2 experienced high levels of attrition, so the data collected from participating cohort 2 teachers is separated into teachers who participated in one and two years of the program. We include significance testing based on a repeated measures analysis of variance (ANOVA) or *t*-tests.

Teacher Implementation of Active Learning

Our team measured implementation of active learning strategies using three tools including a teacher survey called the STEM PPC and two observational measures including the STEM TOP and the UTOP.

These three measures capture information from different perspectives (i.e., teachers, coaches, external raters) about teachers' implementation of active learning. More specifically, teachers report on the lessons that they themselves taught on the survey; teachers' coaches rate teachers' lessons using the STEM TOP; and external raters, who have no existing relationship with the teacher, rate teachers' lessons using the UTOP. Each of these perspectives uniquely contributes to understanding teachers' implementation of active learning at each time point and across time.

STEM PPC

*Based on **teacher self-report** on the STEM PPC, to what extent did teachers implement active learning strategies in their classrooms? Did teachers increase in their implementation across time?*

Cohort 1

In this analysis, we investigated the extent of cohort 1 teachers' self-reported frequency of active learning. Figure 3 includes only teachers who completed the survey at all timepoints.

Cohort 1 teachers' self-reported frequency increased from summer 2017 to spring 2020. Although the teachers' perceived frequency fluctuated across all time points, it increased from the summer to the spring of each of the three years included in this data, and each subsequent year showed increases from the previous year.

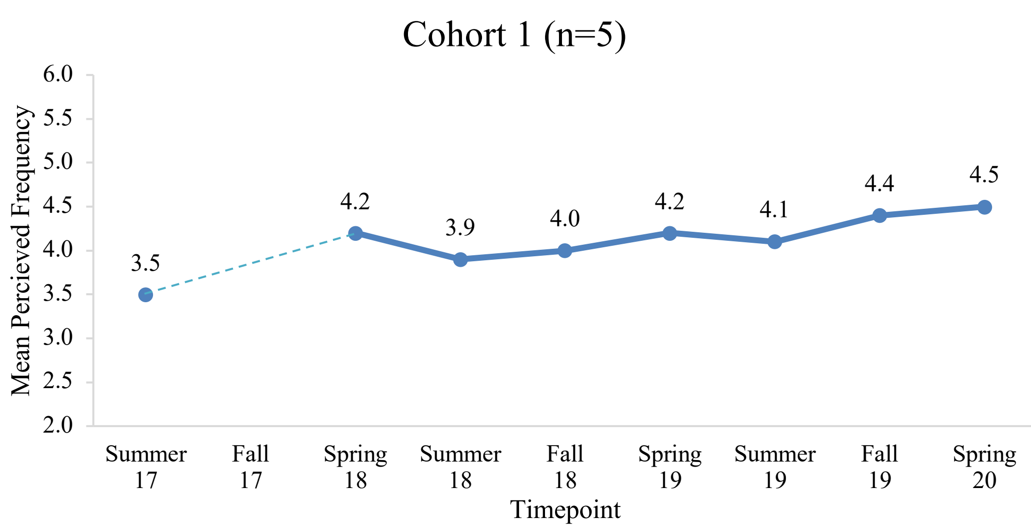


Figure 3. Cohort 1 Teachers' Mean Self-Reported Frequency in using Active Learning Strategies Across Time based on the Teacher STEM PPC Survey

Note: Cohort 1 Fall 2017 survey data were not collected because SMU was awaiting approval from Dallas ISD. Dotted lines connect summer 2018 and spring 2018 to indicate that this timepoint was not collected for cohort 1 teachers.

As shown in Table 5, these increases were not significant.

Table 5

Repeated Measures ANOVA Cohort 1 for Teacher-Reported Frequency

Source	Sums of Squares	df	MS	F	*P*-Value
Time	3.504	7	0.501	2.185	0.067
Error	6.414	28	0.229		

Cohort 2 Year 1

In this analysis, we investigated the extent of cohort 2 teachers' self-reported frequency of their use of classroom active learning for cohort 2 teachers who participated in one year of the STEM Academy. Figure 4 shows that the extent of cohort 2 teachers' self-reported frequency increased slightly across the year.

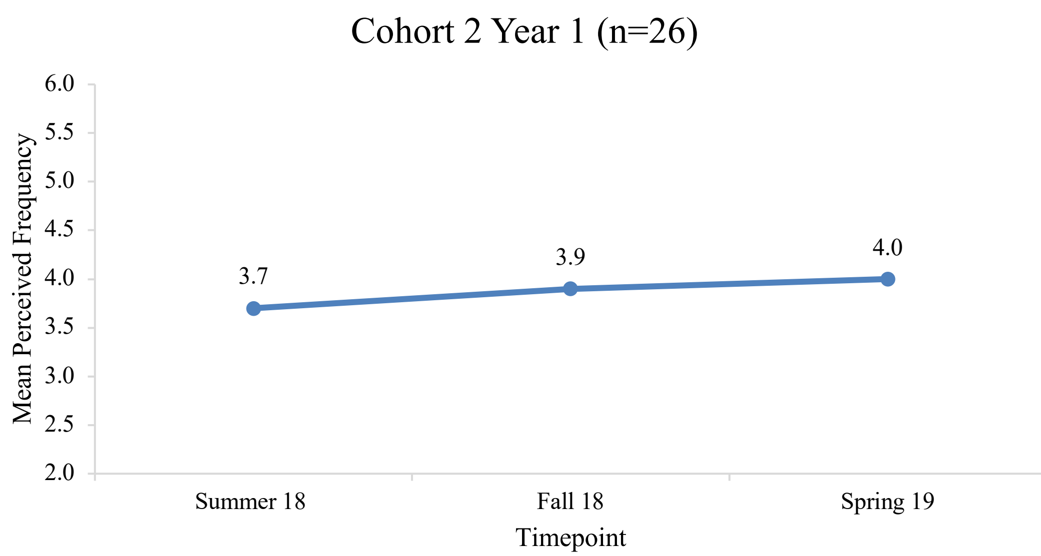


Figure 4. Cohort 2 Year 1 Teachers' Mean Self-Reported Frequency in using Active Learning Strategies Across Time based on the Teacher STEM PPC Survey

A repeated measures ANOVA analysis, in Table 6, showed that this increase was non-significant.

Table 6

Repeated Measures ANOVA Across STEM PPC Frequency Timepoints

Source	Sums of Squares	df	MS	F	*P*-Value
Time	1.084	2	0.542	1.699	0.194
Error	15.32	48	0.319		

Cohort 2 Year 2

The next analysis shows cohort 2 teachers' self-reported frequency of their use of classroom active learning for cohort 2 teachers who participated for two years of the STEM Academy. Figure 5 shows that the extent of cohort 2 teachers' perceived frequency increased slightly across the two years.

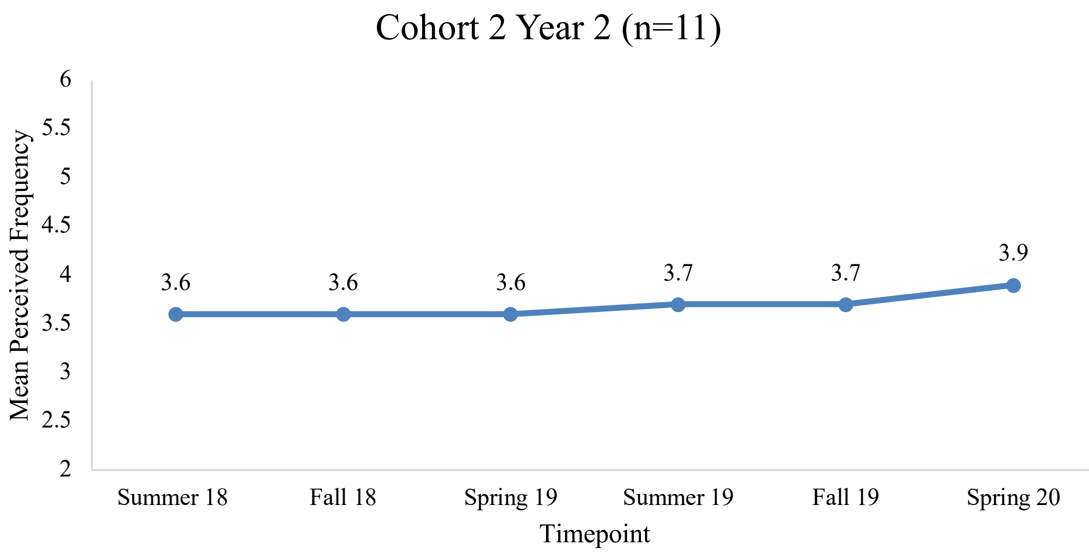


Figure 5. Cohort 2 Year 2 Teachers' Mean Self-Reported Frequency in using Active Learning Strategies Across Time based on the Teacher STEM PPC Survey

A repeated measures ANOVA analysis, in Table 7, showed that this increase was not significant.

Table 7

Repeated Measures ANOVA Across STEM PPC Frequency Timepoints

Source	Sums of Squares	df	MS	F	*P*-Value
Time	0.064	5	0.0129	0.0129	0.999
Error	19.495	45	0.4332		

For the mean, standard deviation, minimum, and maximum of the scores indicting self-reported frequency for each cohort, see the data tables in the Supplemental Appendix document (available upon request).

<u>STEM TOP</u>

Based on ***coach observational scores*** *on the UTOP, to what extent did teachers implement active learning strategies in their classrooms? Did teachers increase in their implementation across time?*

Cohort 1

For cohort 1 teachers who completed all STEM TOP cycles in years 2 and 3, both the Management and Discipline and the STEM Instruction scale fluctuated. Figure 6 shows that the items comprising the Management and Discipline scale were consistently scored higher on average than the items comprising the STEM instruction scale. Additionally, the Management and Discipline score was at its highest during cycle 1 of each year and decreased as each year progressed. When considered by year, the Management and Discipline scale score decreased, while the STEM Instruction scale score showed no directional trend in change.

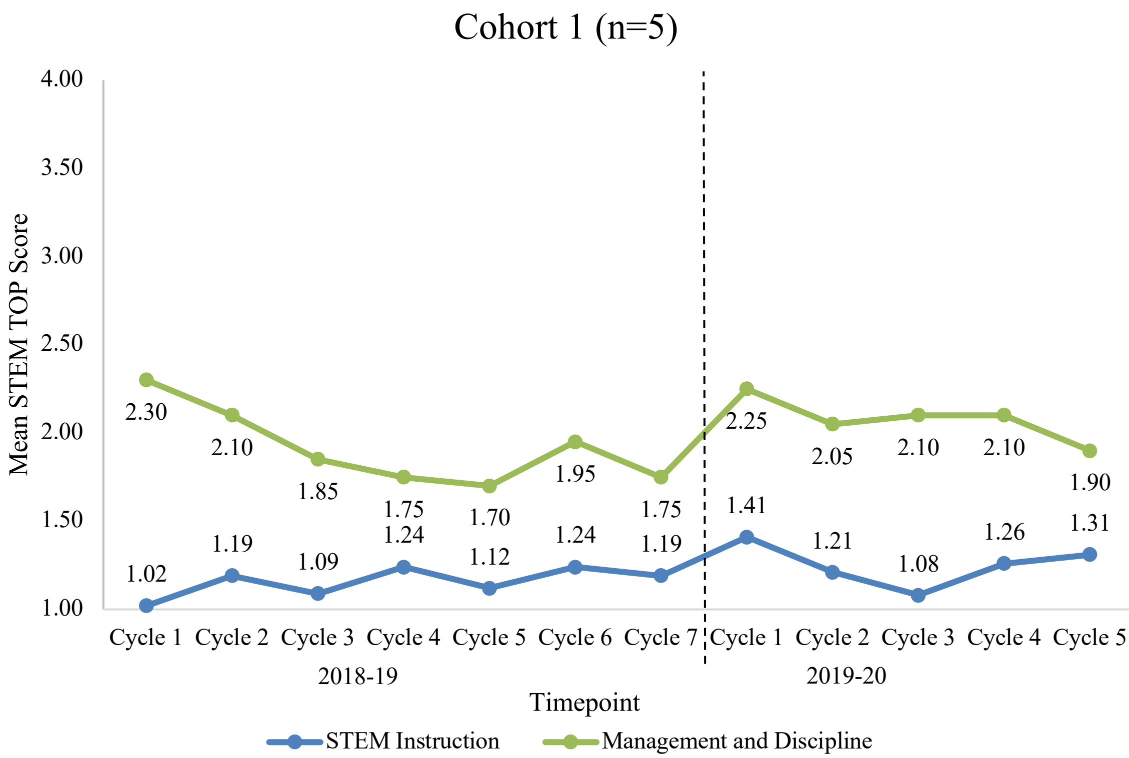


Figure 6. Cohort 1 Teachers' Mean STEM TOP Scores Across Time based on SMU Coach Observations by Scale and Observation Cycle

Over the course of cohort 1 teachers' participation in the STEM Academy, no significant changes were captured on the STEM TOP measure, as shown in Table 8. Post-hoc analyses revealed an adequately powered analysis with five teachers across twelve time points, assuming a high correlation between measures (r = 0.85), and an intended effect size of 0.80.

Table 8

Repeated Measures ANOVA for Cohort 1 STEM TOP Overall

Source	Sums of Squares	df	MS	F	*P*-Value
Time	0.441	11	0.0401	0.426	0.936
Error	4.147	44	0.0942		

Cohort 2 Year 1

Figure 7 shows that for cohort 2 teachers in their first year of the STEM Academy, the Management and Discipline scale scores were consistently higher than the STEM instruction scale scores. This is similar to cohort 1, although, unlike cohort 1 the overall Management and Discipline scales score decrease across the year is more moderate.

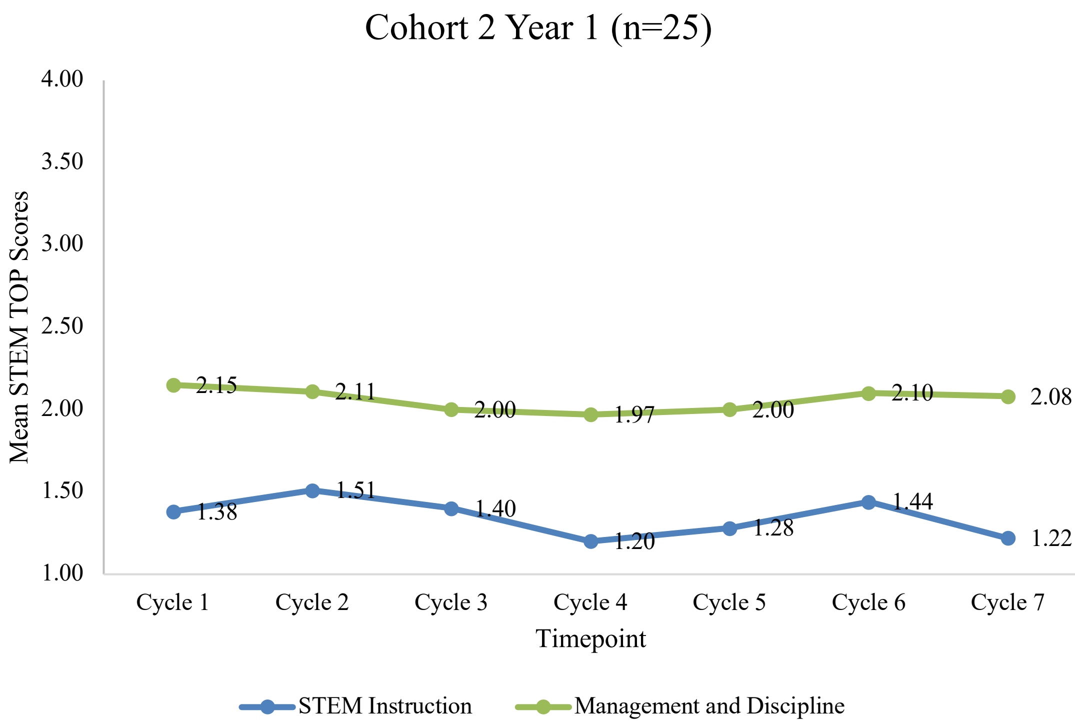


Figure 7. Cohort 2 Year 1 Teachers' Mean STEM TOP Scores Across Time based on SMU Coach Observations by scale and observation cycle

Table 9 provides the results of the repeated measure ANOVA test across the seven timepoints for cohort 2 during year 1. Table 10 provides the post-hoc pairwise analyses using a Tukey-adjusted *p*-value. We observed significant differences between cycles two and four, two and seven, and four and six.

Table 9

Repeated Measures ANOVA Across STEM TOP Timepoints Overall

Source	Sums of Squares	df	MS	F	*P*-Value
Time	1.241	6	0.2069	3.587	0.002
Error	9.113	158	0.0577		

Table 10

Significant Tukey-Adjusted Pairwise Comparisons

Cycle Pairs	Mean Difference	*P*-Value

2-4	-0.233	0.008**
2-7	-0.209	0.031*
4-6	0.196	0.046*

Note: *<0.05, **<0.01

Cohort 2 Year 2

Figure 8 shows the scale scores for each timepoint for cohort 2 teachers who completed both years of the STEM Academy. While the Management and Discipline scale scores decreased across all timepoints, the change is not as dramatic as that seen in cohort 1, and the intra-year drops were not as pronounced. Furthermore, while cohort 1 had peaks in their Management and Discipline scale score at cycle 1 of each year followed by decreasing scores, this trend is not apparent in cohort 2. STEM Instruction shows slight decreases over the two year period.

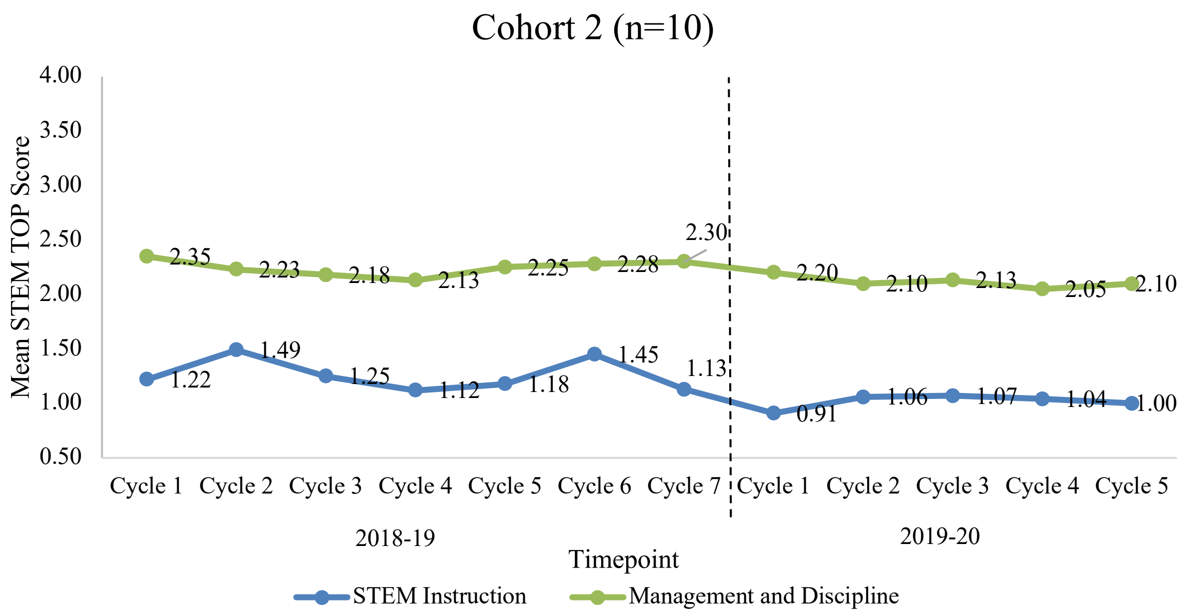


Figure 8. Cohort 2 Year 2 Teachers' Mean STEM TOP Scores Across Time based on SMU Coach Observations by scale and observation cycle

Overall, we observed significant differences between timepoints on the STEM TOP as evidence by the repeated measures ANOVA in Table 11. Table 12 provides the Tukey-adjusted pairwise comparisons for cohort 2 in year 2.

Table 11

Repeated Measures ANOVA Across STEM TOP Timepoints

Source	Sums of Squares	df	MS	F	*P*-Value
Time	2.419	11	0.2200	3.407	<.001
Error	6.392	99	0.0646		

Table 12

Significant Tukey-adjusted pairwise comparisons

Cycle Pairs	Mean Difference	*P*-Value
Y1 Cycle 2-Y2 Cycle 4	-0.395	0.034*
Y1 Cycle 2-Y2 Cycle 5	-0.419	0.018*
Y1 Cycle 6-Y2 Cycle 5	-0.395	0.034*

Note: *$p < 0.05$

For each cohort, STEM TOP means, intra-class correlations, standard deviations, minimum scores, and maximum scores by STEM TOP item across time are available in the Supplemental Appendix document (available upon request).

UTOP

*Based on **external rater observational scores** on the UTOP, to what extent did teachers implement active learning strategies in their classrooms? Did teachers increase in their implementation across time?*

Cohort 1

The external raters evaluated the participating teachers on the four sections on the UTOP. This builds on the previous enacted instruction data that we collected because it represents a different perspective. Whereas, teachers self-reported their enacted instruction and coaches, who had a relationship with teachers, observed their instruction. The data presented in this section represents scores observed by external raters who had little knowledge of the program and did not have existing relationships with the teachers. Figure 9 shows that the average score of all UTOP sections increased across all timepoints.

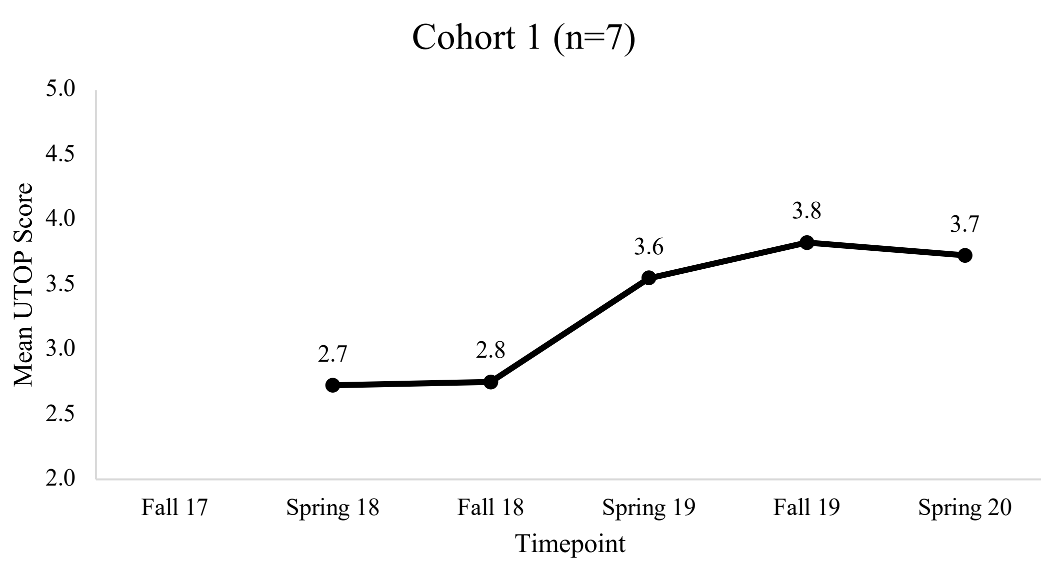


Figure 9. Cohort 1 Teachers' Mean UTOP Scores Across Time based on External Rater Observations

Note: Cohort 1 fall 2017 observational data were not collected because SMU was awaiting approval from Dallas ISD.

Table 13 provides the results for the repeated measures ANOVA over the five time points. Significant results prompted post-hoc analyses. Table 14 describes the Tukey-adjusted pairwise analyses. We detected significant difference in mean UTOP scores between Spring 2018 and Fall 2019, Spring 2018 and Spring 2020, Fall 2018 and Fall 2019, and Fall 2018 and Spring 2020.

Table 13

Repeated Measures ANOVA for UTOP Overall Cohort 1

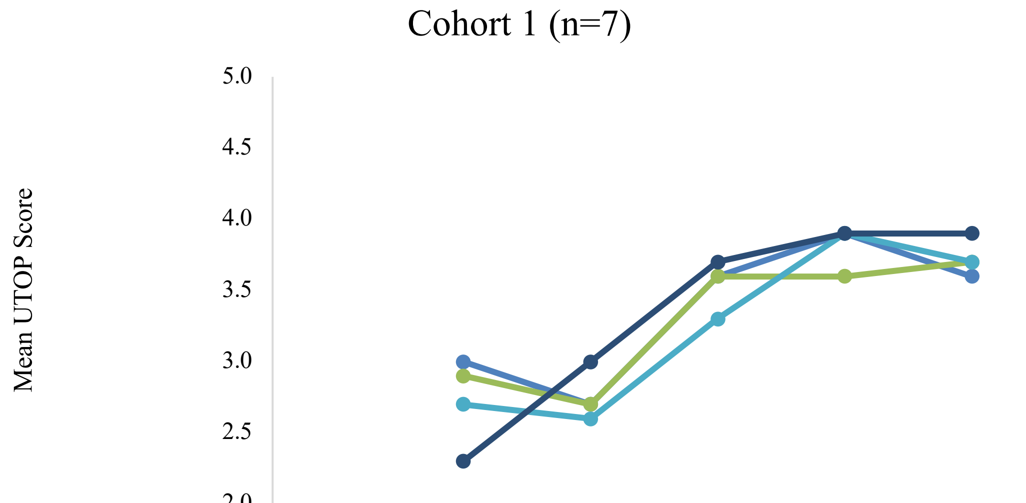
Source	Sums of Squares	df	MS	F	*P*-Value
Time	7.761	4	1.940	5.405	0.003
Error	8.614	24	0.359		

Table 14

Significant Tukey-adjusted pairwise comparisons for UTOP Cohort 1

Pairs	Mean Difference	*P*-Value
Spring 18 – Fall 19	1.071	0.021*
Spring 18 – Spring 20	1.000	0.034*
Fall 18 – Fall 19	1.036	0.027*
Fall 18 – Spring 20	0.964	0.043*

Figure 10 shows the average score for each domain at each timepoint, and only includes scores for teachers who completed all years. Science content increased during the first two years and plateaued during year three. The other three sections all increased across year two, but both classroom culture and implementation average scores decreased during year three.



	Fall 17	Spring 18	Fall 18	Spring 19	Fall 19	Spring 20
Classroom Culture		3.0	2.7	3.6	3.9	3.6
Lesson Structure		2.9	2.7	3.6	3.6	3.7
Implementation		2.7	2.6	3.3	3.9	3.7
Science Content		2.3	3.0	3.7	3.9	3.9

Figure 10. Cohort 1 Teachers' Mean UTOP Scores Across Time based on External Rater Observations
Note: Cohort 1 fall 2017 observational data were not collected because SMU was awaiting approval from Dallas ISD.

A repeated measures ANOVA (Table 15) revealed significant mean differences by time for implementation and science content scales of the UTOP. The post-hoc analyses are described in Table 16.

Table 15

Repeated Measures ANOVA by UTOP Scales for Cohort 1

Scale	Source	SS	df	MS	F	*P*-Value
Classroom Culture						
	Time	6.171	4	1.543	2.932	0.042*
	Error	12.63	24	0.526		
Lesson Structure						
	Time	6.000	4	1.500	2.250	0.094
	Error	16.00	24	0.667		
Implementation						
	Time	9.314	4	2.329	5.653	0.002**
	Error	9.886	24	0.412		
Science Content						
	Time	13.31	4	3.329	6.955	<.001***
	Error	11.49	24	0.479		

Table 16

Significant Tukey-adjusted pairwise comparisons for UTOP Components: Cohort 1

Component	Pairs	Mean Difference	*P*-Value
Classroom Culture			
	Fall 18 – Fall 19	1.143	0.050
Implementation			
	Spring 18 – Fall 19	1.143	0.021*
	Fall 18 – Fall 19	1.286	0.008**
	Fall 18 – Spring 20	1.143	0.021*
Science Content			
	Spring 18 – Spring 19	1.429	0.006**
	Spring 18 – Fall 19	1.571	0.002**
	Spring 18 – Spring 20	1.571	0.002**

Cohort 2 Year 1

Figure 11 shows that cohort 2 teachers who completed the first year showed an increase in average UTOP score across the four sections.

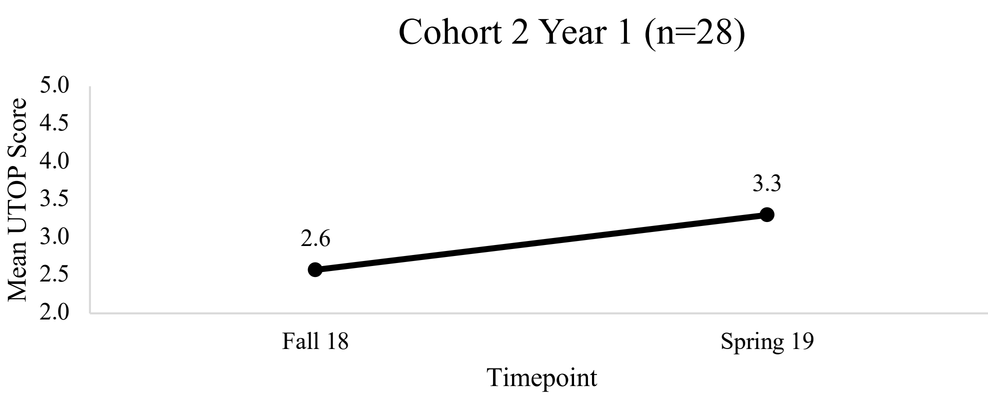


Figure 11. Cohort 2 Year 1 Teachers' Mean UTOP Scores Across Time based on External Rater Observations

Table 17 shows that this increase for cohort 2 teachers in their first year of participation was significant.

Table 17

Paired t-test for cohort 2 teachers year 1

Fall 18 M(SD)	Spring 19 M(SD)	Difference	*t*-statistic	*P*-Value	95% CI
2.6 (0.55)	3.29 (0.82)	0.73 (.77)	5.03	<.001	(0.434,1.031)

This pattern of increase was present in all four domains, as shown in Figure 12, and was most pronounced in implementation scores and least pronounced in science content scores.

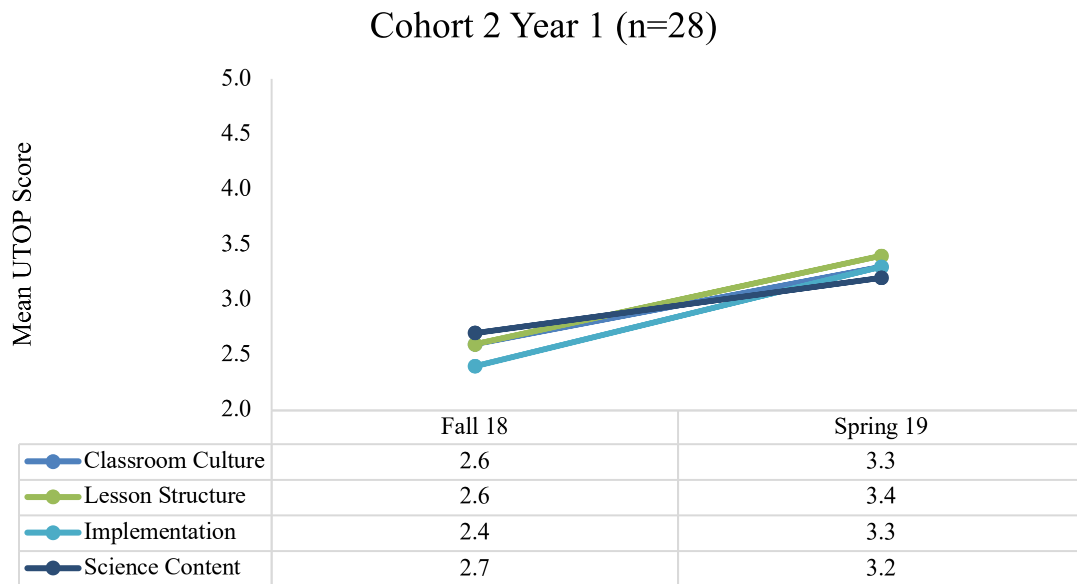


Figure 12. Cohort 2 Year 1 Teachers' Mean UTOP Scores Across Time based on External Rater Observations

The *p*-values in Table 18 indicate that all of these increases were significant.

Table 18

Dependent t-test for UTOP scales

Scale	Mean Difference	*t*-statistic	*P*-Value	95% CI
Classroom Culture	0.679	3.52	0.002	(0.283,1.074)
Lesson Structure	0.821	4.80	<.001	(0.471,1.172)
Implementation	0.929	5.02	<.001	(0.549,1.308)
Science Content	0.500	3.81	<.001	(0.231,0.769)

Cohort 2 Year 2

For cohort 2 teachers who completed both years, the overall UTOP score increased during year 1, but decreased during year 2. Figure 13 shows that the teachers received their lowest score at the first timepoint in the fall of 2018. Then they achieved their high score during the spring of 2019, and the scores decreased during the 2019-20 academic year.

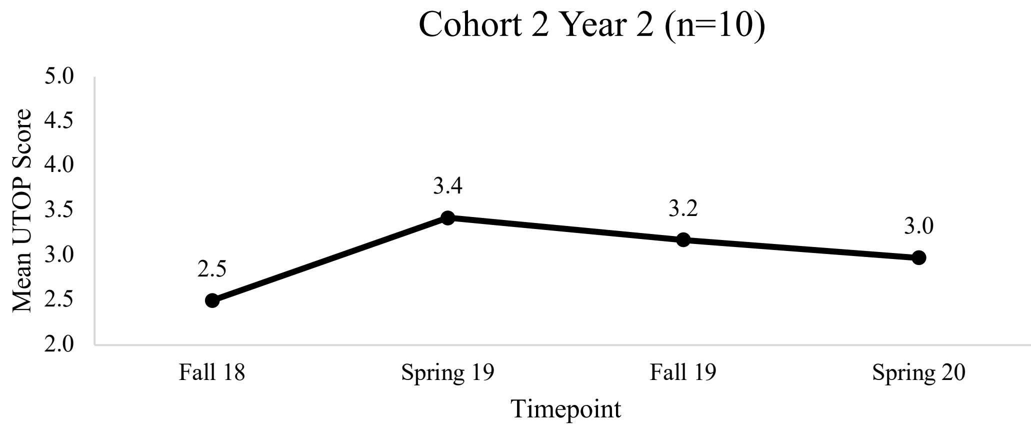


Figure 13. Cohort 2 Year 2 Teachers' Mean UTOP Scores Across Time based on External Rater Observations

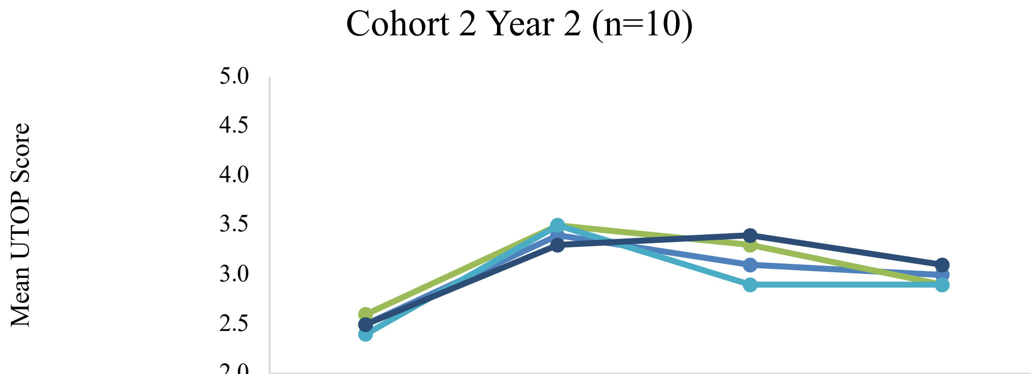
We observed significant mean differences by time point, shown in Table 19. Post-hoc analyses with a Tukey-adjusted *p*-value revealed a significant difference between the first and second time points ($t(27) = 3.29$, $p = 0.01$). Furthermore, the trend is similar to that seen across the second and third year for cohort 1 teachers, with the exception being that cohort 2 teachers achieved their highest score during the spring of 2019 while cohort 1 teacher peaked during the fall of 2019.

Table 19

Repeated Measures ANOVA

Source	SS	df	MS	F	*P*-Value
Time	4.605	3	1.535	3.876	0.02
Error	10.69	27	0.396		

Figure 14 shows the average scores for each of the domains, which all follow similar trends as the overall average scores, except science content, which decreased only from fall 2019 to spring 2020.



	Fall 18	Spring 19	Fall 19	Spring 20
Classroom Culture	2.5	3.4	3.1	3.0
Lesson Structure	2.6	3.5	3.3	2.9
Implementation	2.4	3.5	2.9	2.9
Science Content	2.5	3.3	3.4	3.1

Figure 14. Cohort 2 Year 2 Teachers' Mean UTOP Scores Across Time based on External Rater Observations

Table 20 describes the repeated measures ANOVA by UTOP scale. We observed no significant mean differences across four time points on the classroom culture scale. For the lesson structure scale, we observed a significant mean difference between the first two time points using a Tukey-adjusted p-value for multiple comparisons ($t(27) = 2.97$, $p = 0.03$). Similarly, we detected a significant mean difference between the first two time points on the implementation scale of the UTOP ($t(27) = 3.10$, $p = 0.02$). Lastly, we detected a significant mean difference between time points one and three for the science content scale of the UTOP ($t(27) = 2.86$, $p = 0.04$).

Table 20

Repeated Measures ANOVA by UTOP Scale

Scale	Source	SS	df	MS	F	*P*-Value
Classroom Culture						
	Time	4.20	3	1.40	2.39	0.091
	Error	15.8	27	0.59		
Lesson Structure						
	Time	4.88	3	1.63	3.55	0.028*
	Error	12.4	27	0.46		
Implementation						
	Time	6.08	3	2.03	3.01	0.048*
	Error	18.2	27	0.67		
Science Content						
	Time	4.88	3	1.63	3.28	0.036*
	Error	13.4	27	0.50		

For the mean, standard deviation, minimum, and maximum scores by item and by each cohort, see the data tables in the Supplemental Appendix document here: https://smu.box.com/s/ombb4wlkkrnfzj8y3jup3dm972gyk2ig. In addition, these descriptive statistics are available by item for all participating teachers across time in the supplemental appendix document.

The teacher implementation of active learning as measured by three different tools indicated that the frequency and quality of instruction may be more impacted by the timepoint within the year or by other year specific factors such as the perspective of the rater, than by the progression across years in the STEM Academy program. While the teachers self-reported a slight, but non-significant, increase in frequency of active learning strategy use, the UTeach Observation Protocol (UTOP) indicated that external raters found implementation to have increased across time. Although the external raters captured increased scores in all four measured domains of the UTOP across the two captured years of participation, the STEM Teacher Observation Protocol (STEM TOP) scale scores indicated that Management and Discipline may be cyclic within a given academic year, with an overall negative trend. STEM TOP integrated STEM Instruction scale scores generally increased over time, and significantly improved for some teachers across certain timepoints. The STEM TOP was completed by program coaches, who had an established relationship with the teacher. The coaches tended to report less consistent gains and in some cases no change or decrease across time in teachers' implementation of active learning.

Overall, teachers tended to report slight changes over time, coaches' scores tended to show little change which fluctuated over the school year, and external raters tended to score teachers higher across time, with the exception of cohort 2 during year 2. The difference between coaches and external raters may warrant further drift checks to see if the coaching relationship is influencing scores in a way that biased scores and may have resulted in coaches scoring teachers more harshly across time due to increased expectations.

Overall, the UTOP indicated increased scores in all four measured sections across all time points, although slight decreases were seen in the second year for both cohorts. The STEM TOP scale scores indicated that Management and Discipline may be cyclic within a given academic year, with an overall negative trend, and STEM Instruction scale scores were variable with some increases, some decreases, and some stagnation observed. On the STEM PPC teachers generally did not change their perceptions about the frequency of their active learning strategy use in the classroom.

Finally, there are differences between the mean scores of the two cohorts of teachers. Cohort 1 scored higher on both the STEM PPC and the UTOP than cohort 2. Even when the year is adjusted for duration of participation, the cohort 1 teachers self-reported higher frequencies of active learning strategy use and the external UTOP observers captured higher levels of implementation, lesson structure, classroom culture, and science content.

Teacher Pedagogical Content Knowledge

*Based on the **POSTT**, to what extent did teachers demonstrate science pedagogical content knowledge (PCK)? Did teachers' PCK increase across time?*

The results of the POSTT are summarized in this section.

Cohort 1

Figure 15 shows the mean scores for the POSTT pre- and post- academies in summers 2017, 2018, and 2019. We see an increase in scores between summer 2017 and summer 2018, but a slight decrease in summer 2019.

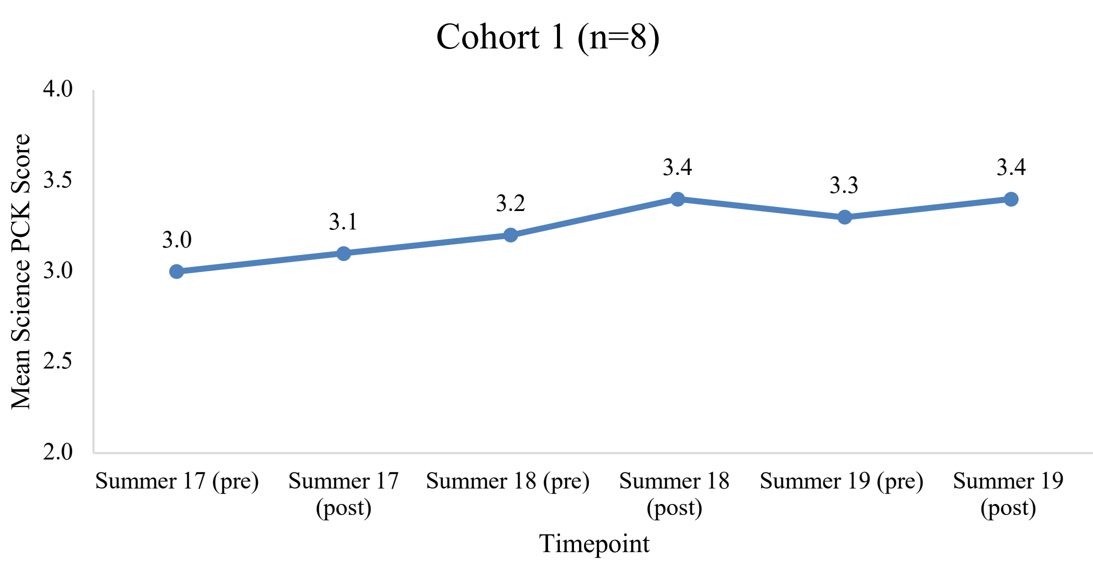


Figure 15. Cohort 1 Teachers' Mean Science Pedagogical Content Knowledge (PCK) Across Time

To determine the existence of a significant change across time points, a repeated measures ANOVA was conducted. The repeated measures ANOVA accounts for the dependency involved with surveying the same individuals at multiple time points. Table 21 illustrates that the repeated measures ANOVA with sphericity assumed across the six time points was significant. Post-hoc analyses, in Table 22, revealed significant mean differences between Summer 2017 Pre and Summer 2018 Post, Summer 2017 Pre and Summer 2019 Pre, and Summer 2017 Pre and Summer 2019 Post.

Table 21

Repeated Measures ANOVA Across POSTT Timepoints

Source	Sums of Squares	df	MS	F	*P*-Value
Time	1.083	5	0.217	3.91	.006
Error	1.942	35	0.055		

Table 22

Significant Tukey-adjusted comparisons for POSTT: Cohort 1

Pairs	Mean Difference	*P*-Value
Summer 17 Pre – Summer 18 Post	0.388	0.026*
Summer 17 Pre – Summer 19 Pre	0.363	0.043*
Summer 17 Pre – Summer 19 Post	0.400	0.020*

Cohort 2 Year 1

Figure 16 shows the mean scores for the POSTT in summer 2018 for all teachers who completed the academy. The post-POSTT score was slightly higher than the pre-POSTT score.

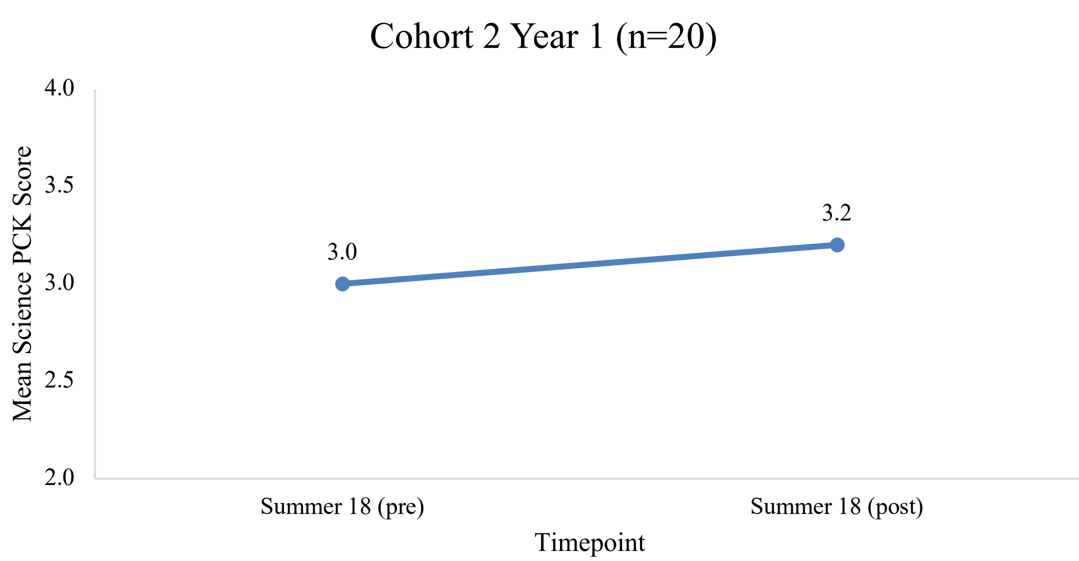


Figure 16. Cohort 2 Year 1 Teachers' Mean Science Pedagogical Content Knowledge (PCK) Across Time

When a *t*-test was conducted, this increase was found to be significant, as shown in Table 23.

Table 23

Paired t-test for POSTT Cohort 2 Year 1

Pre M (SD)	Post M (SD)	Difference (SD)	*t*-statistic	*P*-Value	95% CI
3.0 (0.44)	3.2 (0.32)	0.24 (0.44)	2.41	0.03*	(0.03,0.44)

Cohort 2 Year 2

Figure 17 shows the mean scores for the POSTT across time for all teachers who completed both years of the academy. In both years, the post-POSTT score was slightly higher than the pre-POSTT score.

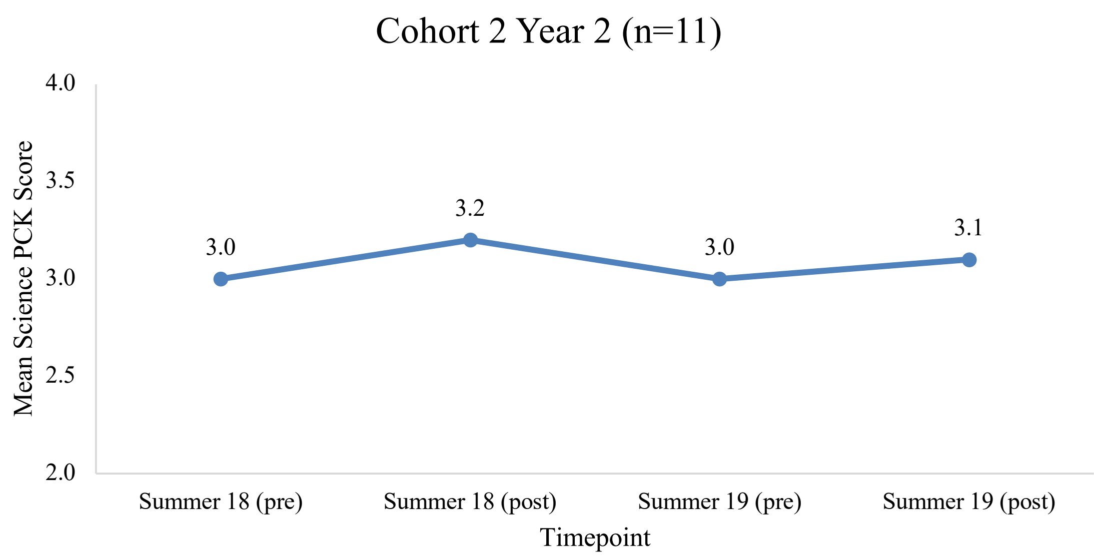


Figure 17. Cohort 2 Year 2 Teachers' Mean Science Pedagogical Content Knowledge (PCK) Across Time

A repeated measures ANOVA was conducted to investigate significant changes in mean content knowledge over time. Table 24 illustrates the results of the ANOVA test. The insignificant *p*-value suggests no change in teachers' content knowledge across the four time points.

Table 24

Repeated Measures ANOVA POSTT C2

Source	Sum Squares	df	Mean Squares	F	*P*-Value
Time	0.30	3	.098	1.19	0.33
Error	2.48	30	.083		

For the mean, standard deviation, minimum, and maximum of the scores indicting self-reported frequency for each cohort, see the data tables in the Supplemental Appendix document here https://smu.box.com/s/ombb4wlkkrnfzj8y3jup3dm972gyk2ig. Means and standard deviations on individual items by cohort are also included.

Cohort 1 teachers improved in their science PCK across the three years, which was also captured by the science content domain on the UTOP discussed in the previous section. Cohort 2 teachers did not exhibit the same level of growth as cohort 1 teachers, as measured on the POSTT, but cohort 2 teachers' PCK increased significantly during the first summer of the program.

Teacher Science Efficacy

*Based on the **STEBI**, to what extent did teachers feel efficacious in teaching science? Did teachers' science self-efficacy increase across time?*

Cohort 1

Figure 18 shows the mean scores for cohort 1 teachers' personal science teaching efficacy beliefs, their science teaching outcome expectancy, and items on the STEBI overall. At all timepoints, the personal science teaching efficacy beliefs were higher than the outcome expectancy. From the first year to the third year, the teachers' personal science teaching efficacy beliefs increased, while their science teaching outcome expectancy decreased.

Of note, one teacher in fall 2018 rated more than half of the items for personal science teaching efficacy on the survey with a value of 2 (disagree). This same teacher at all other timepoints rated all of these items as 4 (agree) or 5 (strongly agree). Due to the small sample size, this disagreement in the fall of 2018 led to a lower personal science teaching efficacy score, which in turn contributed to a lower overall score. We decided to retain this teacher given that we believe that all teachers responded in ways that reflected their beliefs at the time of the survey.

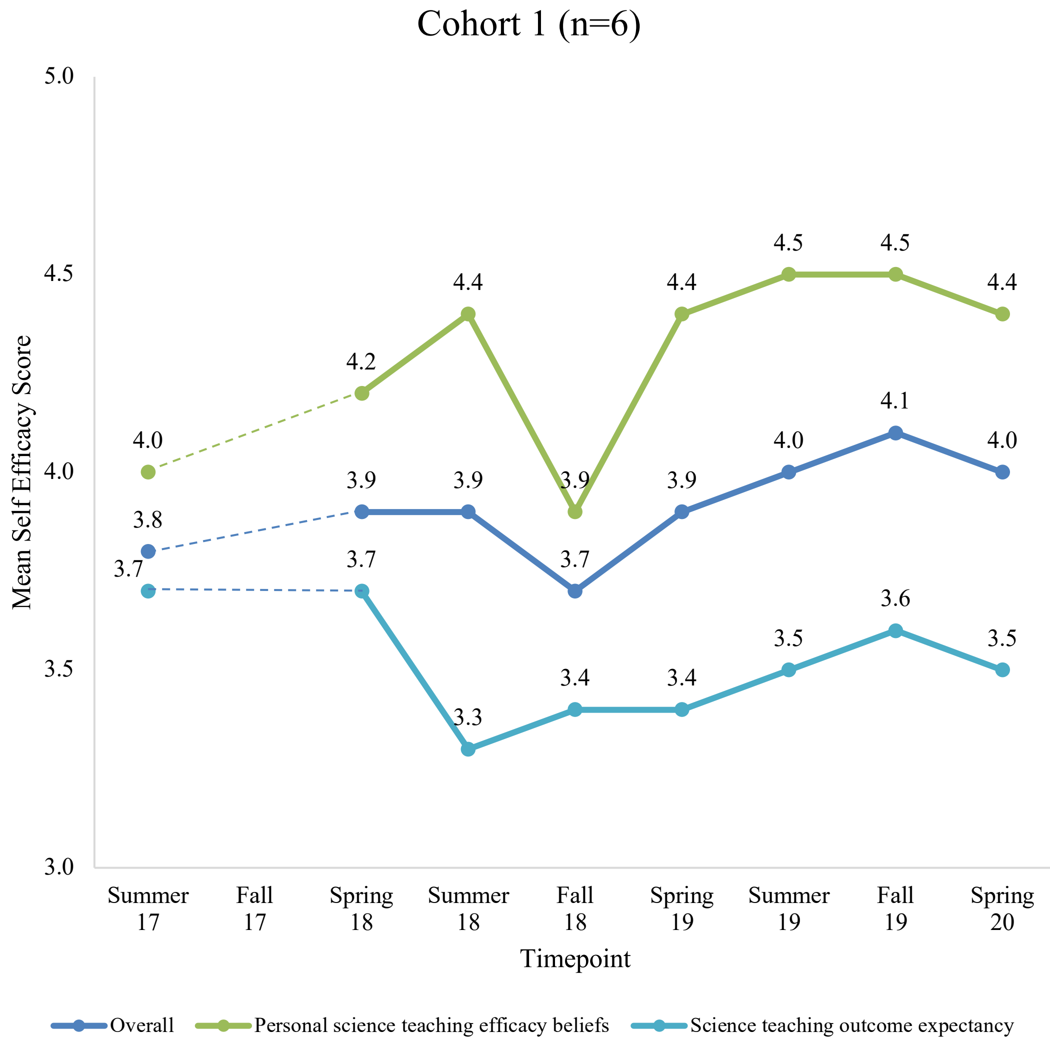


Figure 18. Cohort 1 Teachers' Mean Self-Efficacy for Teaching Science Across Time
Note: Cohort 1 fall 2017 survey data were not collected because SMU was awaiting approval from Dallas ISD. Dotted lines connect summer 2018 and spring 2018 to indicate that this timepoint was not collected for cohort 1 teachers.

Table 25 details the results of a repeated measures ANOVA analysis, which revealed no mean differences between time points.

Table 25

Repeated Measures ANOVA for STEBI Overall: Cohort 1

Source	Sum Squares	df	Mean Squares	F	*P*-Value
Time	0.743	7	0.106	2.037	0.078
Error	1.824	35	0.052		

Cohort 2 Year 1

Figure 19 shows that for cohort 2 teachers who completed the first year, the personal science teaching efficacy beliefs decreased from summer 2018 to fall 2018 but returned to baseline in spring 2019. For science teaching outcome expectancy, the means remained stable during the three timepoints measured during this year.

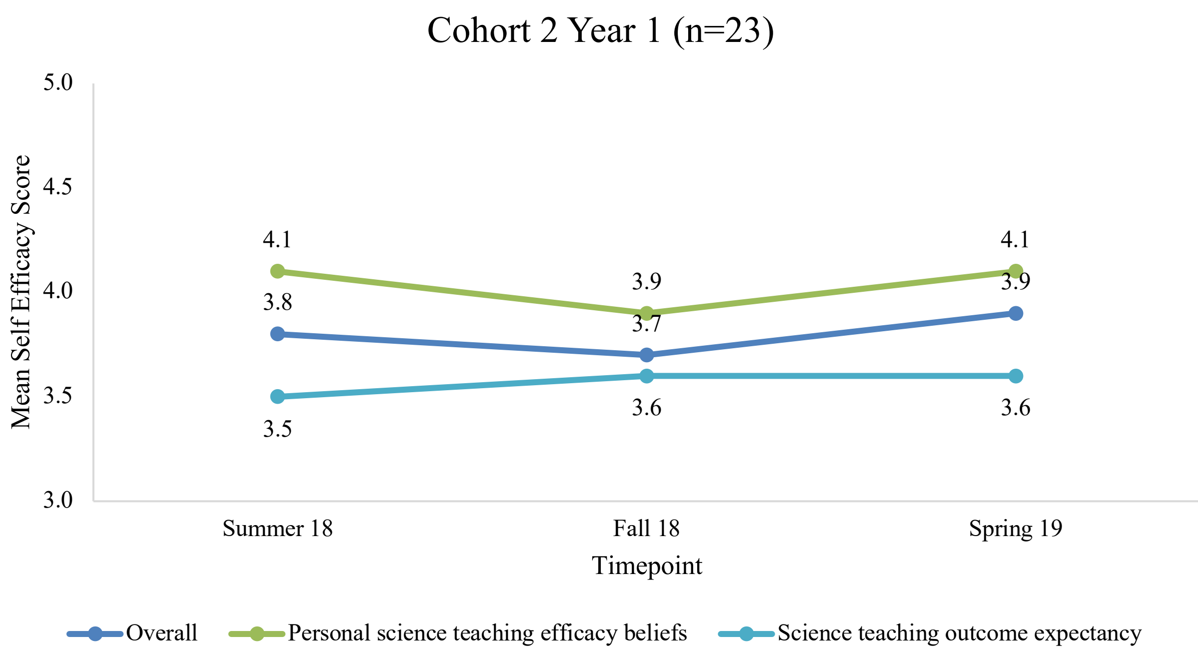


Figure 19. Cohort 2 Year 1 Teachers' Mean Self-Efficacy for Teaching Science Across Time

A repeated measures ANOVA was conducted to investigate significant changes in mean overall teacher beliefs over time. Table 26 illustrates the results of the ANOVA test. The insignificant *p*-value suggests no change in teachers' overall self-efficacy.

Table 26

Repeated Measures ANOVA Across STEBI Timepoints: Overall

Source	Sums of Squares	df	MS	F	*P*-Value
Time	0.190	2	0.095	1.365	0.266
Error	3.066	44	0.070		

Cohort 2 Year 2

Figure 20 shows that for the cohort 2 teachers who completed both years, the personal science teaching efficacy beliefs remained relatively constant across all six timepoints. The decrease during the fall of 2018 noted in the cohort 1 year 1 only data is also present here, and the second year shows no change in mean personal science teaching efficacy beliefs. For science teaching outcome expectancy, the means increased slightly during the three timepoints measured during year 1, and then decreased slightly during year 2.

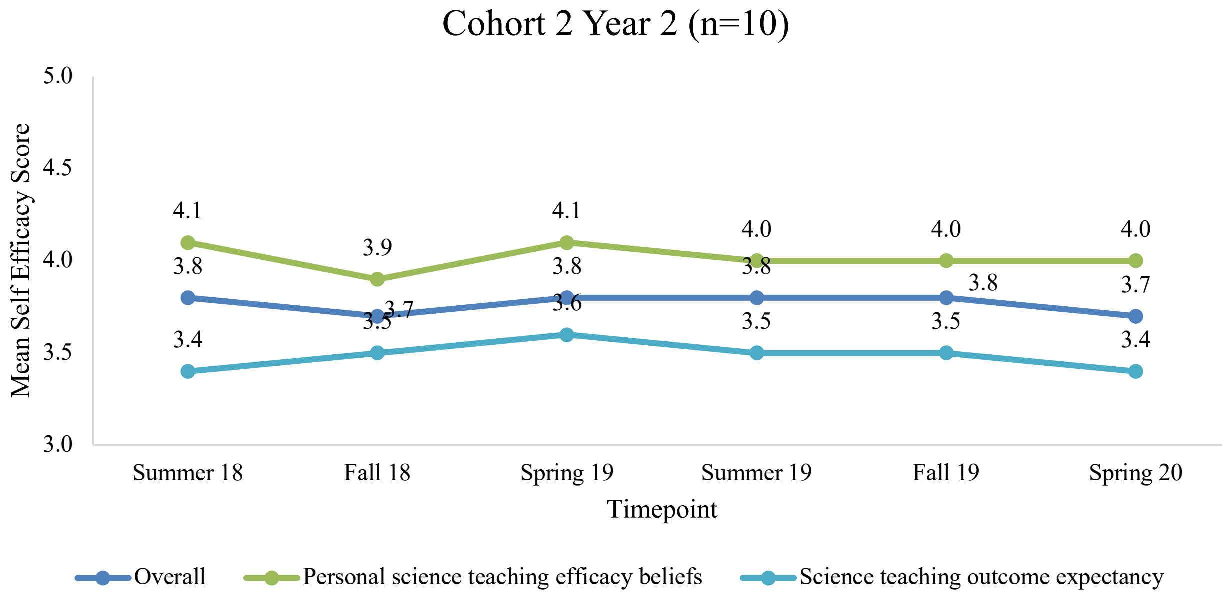


Figure 20. Cohort 2 Year 2 Teachers' Mean Self-Efficacy for Teaching Science Across Time

A repeated measures ANOVA was conducted to investigate significant changes in mean teacher science efficacy beliefs over time. Table 27 illustrates the results of the ANOVA test. The insignificant *p*-value suggests no change in teachers' personal science teaching efficacy beliefs or science teaching outcome expectancy.

Table 27

Repeated Measures ANOVA Across STEBI Timepoints: Overall

Source	Sums of Squares	df	MS	F	*P*-Value
Time	0.152	5	0.030	0.717	0.614
Error	1.904	45	0.042		

For the mean, standard deviation, minimum, and maximum of STEBI scores for each cohort, see the data tables in the Supplemental Appendix document here (available upon request). In addition, descriptive statistics on individual items are included.

In general, the STEBI results do not indicate significant changes in cohort 1 or 2 teachers' science self-efficacy. The results from the STEBI appear to be unaffected by the cohort 2 teachers' participation in the STEM Academy, because no significant changes were observed. For the cohort 1 teachers, the personal science teaching efficacy beliefs increased over the course of the academy, but their science teaching outcome expectancy decreased slightly. However, these changes were not statistically significant.

Teachers' Perceived Importance and Confidence in Active Learning

Based on the ***STEM PPC****, to what extent did teachers report importance and confidence in using active learning in their classrooms? Did teachers' perceived importance and confidence increase across time?*

Cohort 1

In this analysis, we investigated the extent of cohort 1 teachers' perceived importance and confidence. Figure 21 shows that the extent of cohort 1 teachers' perceived confidence increased from summer 2017 to spring 2020 with a slight decrease from spring 2018 to summer 2018. The cohort 1 teachers' perceived importance remained constant across this timeframe and was higher at each timepoint than their perceived confidence. The STEM PPC was measured on a four-point scale, so a ceiling effect may be responsible for the static responses of the cohort 1 teachers, especially since they consistently rated the importance as near the top available score.

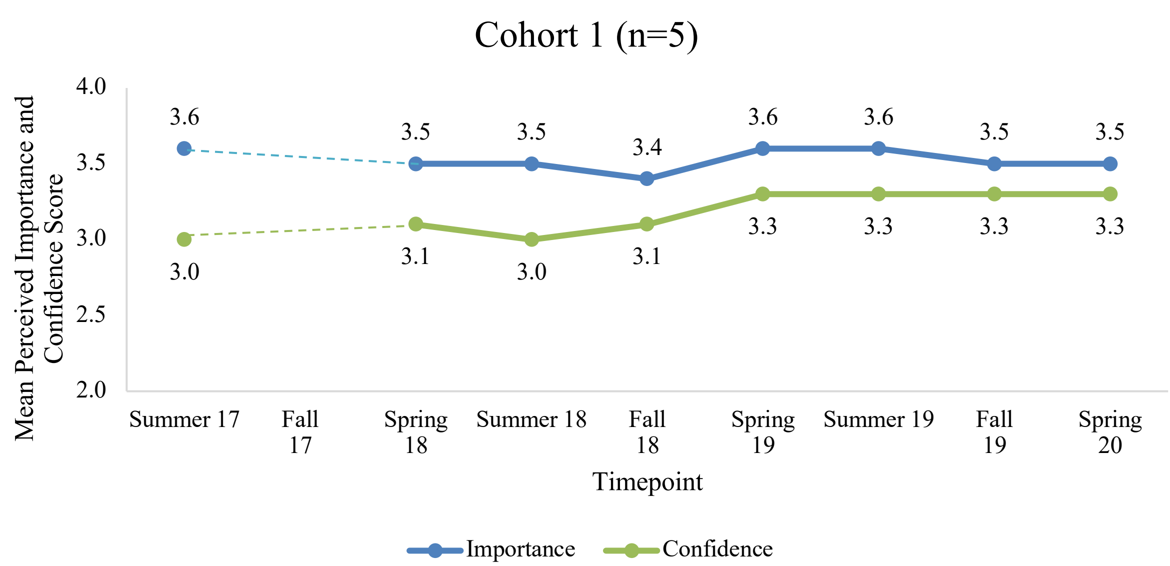


Figure 21. Cohort 1 Teachers' Mean Perceived Importance and Confidence in using Active Learning Strategies Across Time
Note: Cohort 1 Fall 2017 survey data were not collected because SMU was awaiting approval from Dallas ISD. Dotted lines connect summer 2018 and spring 2018 to indicate that this timepoint was not collected for cohort 1 teachers.

Table 28 describes the insignificant mean differences by time on both the importance and confidence scales.

Table 28

Repeated Measures ANOVA for Importance and Confidence: Cohort 1

Source	Sums of Squares	df	MS	F	*P*-Value
Importance					
Time	0.283	7	0.040	0.800	0.594

Error	1.415	28	0.501		
Confidence					
Time	0.746	7	0.107	1.359	0.261
Error	2.196	28	0.078		

Cohort 2 Year 1

In this analysis, we investigated the extent of cohort 2 teachers' perceived importance and confidence in using classroom active learning. Figure 22 shows that cohort 2 teachers' perceived importance decreased over the year, while the extent of confidence increased between these seasons.

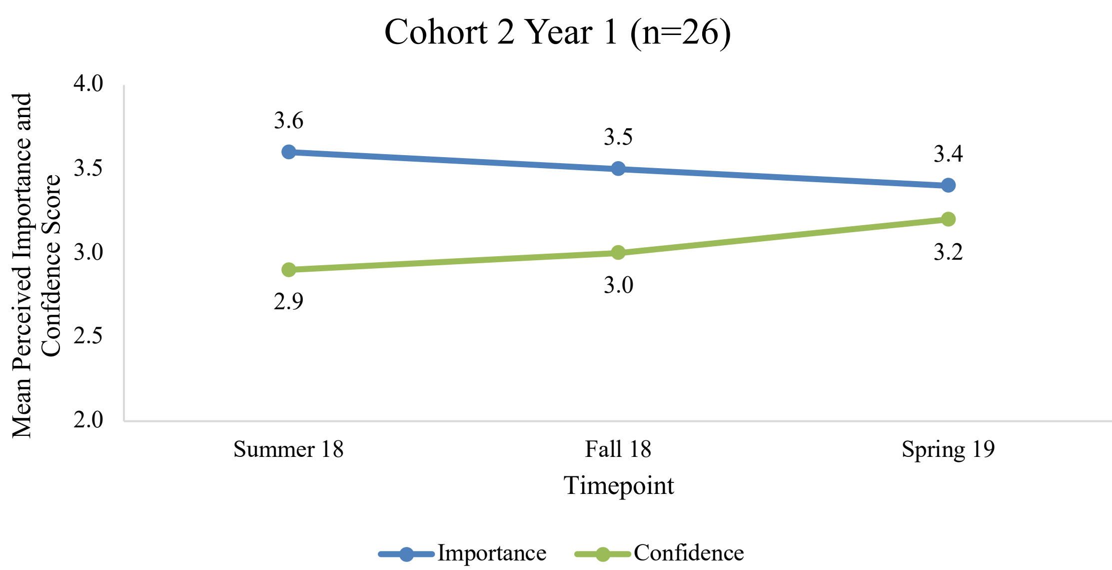


Figure 22. Cohort 2 Year 1 Teachers' Average Perceived Importance and Confidence in using Active Learning Strategies Across Time

A repeated measures ANOVA was conducted to investigate significant changes in mean importance and confidence over time. Table 29 illustrates the results of the ANOVA test. The significant *p*-value suggests a significant change in both teachers' importance and confidence in using active learning strategies. Similar to cohort 1, there may be a ceiling effect due to the use of a four-point scale on STEM PPC.

Table 29

Repeated Measures ANOVA Across Importance and Confidence Timepoints

Source	Sums of Squares	df	MS	F	*P*-Value
Importance					
Time	0.443	2	0.217	4.788	0.013
Error	2.218	49	0.045		
Confidence					
Time	0.884	2	0.442	4.839	0.012
Error	4.476	49	0.091		

Cohort 2 Year 2

In this analysis, we investigated the extent of cohort 2 teachers' perceived importance and confidence in using classroom active learning. Figure 23 shows that cohort 2 teachers' perceived importance decreased slightly over the two years while their perceived confidence increased slightly over the timeframe.

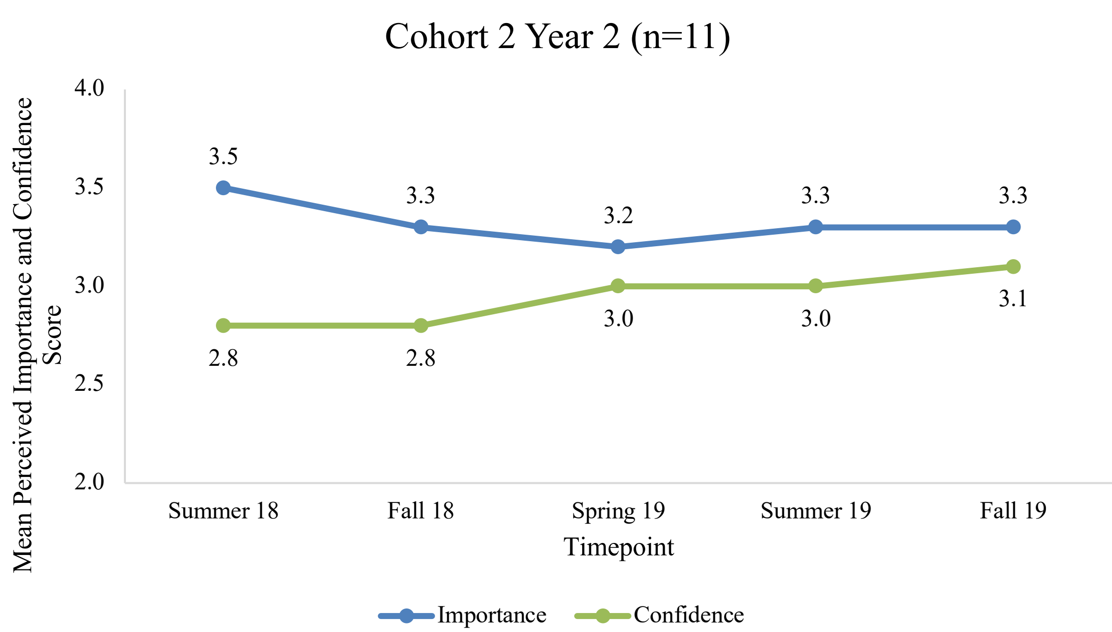


Figure 23. Cohort 2 Year 2 Teachers' Average Perceived Importance and Confidence in using Active Learning Strategies Across Time

A repeated measures ANOVA was conducted to investigate significant changes in perceived importance and confidence, and Table 30 illustrates the results of the ANOVA test. The insignificant *p*-value suggests no change in teachers' perceived importance and confidence across the five time points.

Table 30

Repeated Measures ANOVA Across STEBI Timepoints

Source	Sums of Squares	df	MS	F	*P*-Value
Importance					
Time	0.295	5	0.059	1.138	0.354
Error	2.331	45	0.052		
Confidence					
Time	0.519	5	0.104	1.335	0.267
Error	3.501	45	0.078		

There is some evidence that teachers' confidence increased slightly over time. Cohort 2 teachers who participated in the first year of the program increased significantly in their confidence, but this trend was not observed during the second year of the program with the smaller sample of cohort 2 teachers. Teachers' perceived importance either did not change or decreased across the course of the program. Cohort 2 teachers who participated in the first year of the program decreased significantly in perceived importance, but this trend was not observed during the second year.

Perceived importance was consistently higher than the perceived confidence. This indicates that across all timepoints and within each cohort, the teachers involved in the STEM Academy thought active learning strategies were important, regardless of how well they felt they were able to implement them.

Conclusions

The STEM Academy supported the participating teachers in several ways including summer professional development and mid-year coaching. It was expected that by offering training and support in active learning and inquiry-based instruction, teachers would increasingly implement these pedagogical strategies in their classrooms. In order to assess the success of the STEM Academy, the teacher outcomes were evaluated based on teacher implementation of active learning, teacher PCK, teacher science self-efficacy, and the teachers' perceived importance and confidence regarding the implementation of active learning in their classrooms.

Teacher implementation of active learning was determined through the use of three different measures. These measures included a teacher self-report, an evaluation by the coach, and an evaluation by external observers with no relationship to the teacher. Self-reports and external raters tended to indicate that teachers'enacted active learning practices increased across time with significant increases in lesson structure, implementation, and science content observed by external raters depending on cohort. However, coaches' observations tended to indicate fluctuation or slight decreases in scores across time. This suggests that the relationship between the rater and teacher may influence scores. In general, increases across time were more consistent for cohort 1 teachers, suggesting that these teachers differ somehow from cohort 2 teachers, but it is not possible make comparisons in this report, because these two groups are comprised of different individuals.

Two of the measures used resulted in different trends of teacher implementation. The UTOP indicated increased scores in all four measured domains across the full range of captured time points. Conversely, the STEM TOP evaluated by coaches did not show this same trend of improvement. The difference in trends between the coaches' observations with the STEM TOP and the external evaluators' observations with the UTOP suggest that rater drift phenomenon may need to be examined to better understand the results of this study, and in future implementations of programs like the STEM Academy.

Regarding teacher science self-efficacy, and the teachers' perceived importance concerning the implementation of active learning in their classrooms, no consistent increases were measured across cohorts. Informed by Bandura's social learning theory, self-efficacy relates to teachers' beliefs that an action will have a favorable result (outcome expectation) and that they can perform the action successfully (self-efficacy expectation) (Bleicher, 2004). Teachers with higher self-efficacy for teaching are more likely to implement innovative, evidence-based instructional practices. For example, Tschannen-Moran and McMaster (2009) observed increases in teachers' implementation of evidence-based reading strategies in classrooms of teachers with higher self-efficacy. Existing research shows that teacher self-efficacy is responsive to professional development (PD) in reading (Nadelson et al., 2013; Tschannen-Moran & McMaster, 2009) and mathematics (Ross & Bruce, 2007). Authors of both studies identified statistically significant increases in self-efficacy attributable to teacher participation in PD. We hypothesized that we would observe increases in teachers' science self-efficacy, which would contribute to increased implementation of STEM instructional strategies.

During the first year of participation, cohort 2 teachers did show an improvement in their confidence level regarding the implementation of active learning strategies, but this growth was not maintained through the second year with a smaller sample. Relative to cohort 2 teachers, cohort 1 teachers consistently scored higher on measures of frequency, implementation, and PCK, so it is possible that they have a more positive outlook on their own pedagogical abilities.

Following STEM-focused PD, teachers often experience unanticipated challenges translating STEM practices to their specific contexts (Kelly, Gningue, & Qian, 2015). Teachers who believe in the importance of STEM tend to overcome those challenges more efficiently compared to other teachers (Allen, Webb, & Matthews, 2016). As such, viewing STEM as important is especially critical when teachers encounter colleagues, parents, or school or district leadership who are less knowledgeable about STEM.   Cohort 1 teachers taught predominantly Grade 8 students in both 2017-18 (71%) and 2018-19 (59%) and then majority Grade 7 students in 2019-20 (73%). Cohort 2 teachers taught majority Grade 8 students in 2018-19 (40%), but then majority Grade 6 students in 2019-20 (51%). It is possible that differences exist in the content standards of the grade levels, that there could be challenges associated with changing grade levels, or that there are inherent differences in the training and ability levels of teachers who are assigned to Grade 8 versus Grade 7 or Grade 6 classes.

Content knowledge (CK) is defined as content-specific knowledge about a discipline; whereas, PCK is defined as the integration of CK and appropriate pedagogy for teaching that knowledge (Shulman, 1986; van Driel et al., 1998). This type of knowledge is necessary to make learning accessible to students. Existing research demonstrates a strong connection between teachers' science and mathematics PCK and student academic achievement (Gess-Newsome, 2013; Keller, Neumann, & Fischer, 2017). Teachers with stronger mathematics PCK were more likely to engage students cognitively in higher-level tasks during instruction, rather than focus on lower-level procedural knowledge (Baumert et al., 2010). A recent meta-analysis found that STEM-focused teacher PD focused on both CK and PCK increased students' academic achievement (Lynch et al., 2019).

Teacher confidence in their ability to teach STEM correlates with teachers' CK and enacted science instruction (Munck, 2007; Nadelson et al., 2013). Teachers who were more confident in their ability to implement STEM instructional strategies were also more likely to incorporate inquiry-based instruction; whereas, other teachers expressed hesitation and doubt (Nadelson et al., 2013). In addition, teachers' confidence for teaching STEM is responsive to PD. Nadelson et al. (2013) observed a statistically significant increase in teachers' confidence with an effect size of 0.48 following a three-day PD experience designed to increase teachers' STEM confidence. The STEM Academy includes a strong emphasis on building teachers' confidence for implementing STEM instruction. As such, we anticipated that we would observe increases in teachers' confidence for teaching STEM across time. There were significant increases in science content as measured by the UTOP and PCK. In addition, teachers' confidence seemed to increase slightly, although only significantly for cohort 2 year 1. This may suggest that the STEM Academy influences teachers' science content implementation and PCK.

Finally, in a different study that utilized qualitative interviews, it was found that teachers implemented active learning strategies in differential patterns (Adams, Knox, Hatfield, & Ketterlin-Geller, 2020). These differential patterns stem from variation in the frequency and quality of the teachers' implementation, which could also be linked to the student motivation in those classes, and ultimately the teachers' overall scores captured on a protocol such as the STEM TOP or UTOP. Although this work is ongoing, it may provide an additional lens through which to consider the results presented in this report

One limitation of this study is that no comparison group of teachers was evaluated. Only the teachers participating in the STEM Academy and receiving treatment were given the surveys and observed. Therefore, we are unable to determine whether the change observed in certain indicators is attributable to teachers' participation in the STEM Academy.

Recommendations

Four recommendations for improving the STEM Academy in the future are suggested, based on the results and analysis within this report.

1. For future iterations or similar programs to the STEM Academy, it is important to include external raters during the observation process. The trends observed in this report suggest that coaches' scores may have been influenced by their relationship with the teacher, since the UTOP increased across time while the STEM TOP decreased. External raters are essential for identifying if this discrepancy exists in the observational data and allowing researchers to quantifying this gap.

2. In order to further alleviate the possibility of rater drift, future implementations of programs similar to the STEM Academy should consider specific coach training in order to reduce possible bias. Although both the STEM TOP and the UTOP observational data produced interesting trends that this study could examine, the overall messages were at times contradictory.

3. In general there was not consistent evidence of change across years in the outcomes included in this study. One reason for this might be change in grade levels or other school factors. We recommend to the extent possible that teachers teach same grades across years to encourage developing expertise is specific content and process standards by grade and reduce possible obstacles in navigating new grade levels.

4. Finally, it is important to examine and understand the profiles of schools and teachers who are successful in implementing active learning. This might include a needs assessment or measure of readiness for active learning. This report focuses on describing teacher trends in the aggregate, but is likely extensive variability exists in teachers' experiences within the program depending on teacher and school factors. Efforts should be made to better understand individual teachers' experiences based on teacher interviews and other factors.

References

Adams, E. L., Hatfield, C., & Ketterlin Geller, L. R. (2018). *STEM Academy for Science Teachers and Leaders: Participation Interest Survey February 2018* (Tech. Rep. No. 18-01). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Adams, E. L., Hatfield, C., Cox, C. T., & Ketterlin-Geller, L. R. (2018a). *STEM Academy for Science Teachers and Leaders: 2018 Teacher Academy I Evaluation* (Tech. Rep. No. 18-02). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Adams, E. L., Hatfield, C., Cox, C. T., Mota, A., Sparks, A., & Ketterlin Geller, L. R. (2018b). *STEM Academy for Teachers and Leaders: 2017-18 Coaching and PLC Evaluation* (Tech. Rep. No. 18-03). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Adams, E. L., Sparks, A., & Ketterlin-Geller, L. R. (2020, April). *Introducing the STEM teacher observation protocol (STEM TOP): Evidence of internal structure.* Paper presented at the National Council on Measurement in Education (NCME), San Francisco, CA. (Conference canceled)

Adams, E. L., Sparks, A., & Ketterlin-Geller, L. (2019c). *Multi-level exploratory factor analysis: Evidence of internal structure for a measure of STEM instructional practice.* Paper presented at the Annual Texas Universities' Educational Statistics and Psychometrics (TUESAP) Alliance Conference, Dallas, TX.

Adams, E. L. Knox, T., Hatfield, C. & Ketterlin-Geller, L. R. (2020, March). *Teacher beliefs and practice within the context of an intensive teacher STEM professional development.* Roundtable to be presented at the National Association for Research in Science Teaching (NARST) Research Conference, Portland, OR. (Conference canceled)

Allen, M., Webb, A. W., & Matthews, C. E. (2016). Adaptive teaching in STEM: Characteristics for effectiveness. *Theory Into Practice*, *55*(3), 217-224.

Basham, J.D. & Marino, M.T. (2013). Understanding STEM education and supporting students through universal design for learning. *TEACHING Exceptional Children, 45*(4), 8-15.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., ... & Tsai, Y. M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American educational research journal*, *47*(1), 133-180.

Bleicher, R. E. (2004). Revisiting the STEBI-B: Measuring self-efficacy in preservice elementary teachers. *School Science and Mathematics*, *104*(8), 383-391.

Burton, C., Hatfield, C., Adams, E. L., & Ketterlin Geller, L. R. (2020). *STEM Academy for Science Teachers and Leaders: Research Observation Administration Procedures*

Technical Report (Tech. Rep. No. 19-05). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Chen X., & Weko, T. (2009). Students who study science, technology, engineering, and mathematics (STEM) in postsecondary education. *National Center for Education Statistics, 161,* 1-24.

Cobern, W., Schuster, D. G., Adams, B., Skjold, B. A., Muğaloğlu, E. Z., Bentz, A., & Sparks, K. (2014). Pedagogy of Science Teaching Test. *International Journal of Science Education. Volume 36*(13), 2265-2288.

Cox, C. T., Adams, E. L., & Ketterlin Geller, L. R. (2019). *STEM Academy for Teachers and Leaders: Teacher Exit Survey* (Tech. Rep. No. 19-09). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Dika, S.L., & D'Amico, M.M. (2016). Early experiences and integration in the persistence of first-generation college students in STEM and non-STEM majors. Journal of Research in Science Teaching, 53(3), 368-383.

DeJarnette, N. (2012). America's children: Providing early exposure to STEM (science, technology, engineering and math) initiatives. Education, 133(1), 77-84.

Enochs, L. G., & Riggs, I. M. (1990). Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *School science and mathematics, 90*(8), 694-706.

Fealing, K.H., Lai, Y., & Myers, S.L. (2015). Pathways vs. pipelines to broadening participation in the stem workforce. Journal of Women and Minorities in Science and Engineering, 21(4), 271-293. doi: 10.1615/JWomenMinorScienEng.2015004760

Geier, R., Blumenfeld, P. C., Marx, R. W., Krajcik, J. S., Fishman, B., Soloway, E. & Clay-Chamber, J. (2008). Standardized test outcomes for students engaged in inquiry-based science curricula in the context of urban reform. Journal of Research in Science Teaching, 45(8), 922–939.

Gess-Newsome, J. (2013). Pedagogical content knowledge. *International guide to student achievement*, 257-259.

Houser, K. (2017). *How to wrap-up a coaching cycle.* Retrieved from https://www.mshouser.com/organization/how-to-wrap-up-a-coaching-cycle

Keller, M. M., Neumann, K., & Fischer, H. E. (2017). The impact of physics teachers' pedagogical content knowledge and motivation on students' achievement and interest. *Journal of Research in Science Teaching*, *54*(5), 586-614.

Kelly, A. M., Gningue, S. M., & Qian, G. (2015). First-year mathematics and science middle school teachers: Classroom challenges and reflective solutions. *Education and Urban Society, 47*(20), 132-159.

Kraft, M., A. & Blazar, D. (2017). Individualized coaching to improve teacher practice across grades and subjects: New experimental evidence. *Educational Policy, 31*(7), 1033-1068.

Labaree, D. F. (1997). Public goods, private goods: The American struggle over educational goals. *American educational research journal*, *34*(1), 39-81.

Landivar, L. C. (2013). *Disparities in STEM employment by sex, face, and Hispanic origin.* Washington, DC: U.S. Department of Commerce.

Lynch, S. J., Burton, E. P., Behrend, T., House, A., Ford, M., Spillane, N., Matray, S., Han, E., & Means, B. (2018). Understanding inclusive STEM high schools as opportunity structures for underrepresented students: Critical components. *Journal of Research in Science Teaching*, *55*(5), 712-748.

Marshall, J. C. & Alston, D. M. (2014). Effective, sustained inquiry-based instruction promotes higher science proficiency among all groups: A five-year analysis. *Journal of Science Teacher Education, 25*(7), 807– 821. doi:10.1007/s10972-014-9401-4.

Marshall, J., Smart, J., & Alston, D. (2017). Inquiry-Based Instruction: A Possible Solution to Improving Student Learning of Both Science Concepts and Scientific Practices. *International Journal of Science and Mathematics Education*, *15*(5), 777–796. https://doi.org/10.1007/s10763-016-9718-x

Mau, W.J. (2016). Characteristics of US students that pursued a STEM major and factors that predicted their persistence in degree completion. *Universal Journal of Educational Research, 4*(6), 1495-1500. doi: 10.13189/ujer.2016.040630

Minner, D., Levy, A., & Century, J. (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, *47*(4), 474–496. https://doi.org/10.1002/tea.20347

Miyake, A., Kost-Smith, L. E., Finklestein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science, 330*(6008)*,* 1234-1237.

Munck, M. (2007). Science pedagogy, teacher attitudes, and student success. *Journal of Elementary Science Education*, *19*(2), 13-24.

Nadelson, L. S., Callahan, J., Pyke, P., Hay, A., Dance, M., & Pfiester, J. (2013). Teacher STEM perception and preparation: Inquiry-based STEM professional development for elementary teachers. *The Journal of Educational Research*, *106*(2), 157-168.

National Science Board. (2010). *Science and engineering indicators 2010*. Arlington, VA: National Science Foundation.

Palmer, R. T., Maramba, D. C., & Dancy, T. E. (2010). A qualitative investigation of factors promoting the retention and persistence of students of color in STEM. *The Journal of Negro Education, 80*(4), 491-504.

Perry, L., Reeder, M.J., Brattain, K., Hatfield, C., & Ketterlin Geller, L. (2017). *STEM Academy for Teachers and Leaders: 2017 Academy Evaluation.* (Tech. Rep. No. 17-01). Dallas, TX: Research in Mathematics Education, Southern Methodist University.

Pierce, K., Adams, E.L., Rhone, A.M., Hatfield, C., & Ketterlin Geller, L. (2019a). *STEM Academy for Science Teachers and Leaders: 2018 Teacher Academy 2 Evaluation* (Tech. Rep. No. 18-07). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Pierce, K., Adams, E. L., Sparks, A., Cox, C. T., Knox, T., Hatfield, C., and Ketterlin Geller, L. (2020). *STEM Academy for Science Teachers and Leaders: Leader Academy and Coaching Evaluation* (Tech. Rep. No. 20-13). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Pierce, K., Cox, C. T., Hatfield, C., Adams, E. L., & Ketterlin Geller, L. R. (2019b). *STEM Academy for Science Teachers and Leaders: 2019 Teacher Academy 3 Evaluation* (Tech. Rep. No. 19-08). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Pierce, K., Adams, E. L., Sparks, A., Burton, C., Hatfield, C., & Ketterlin Geller, L. R. (2020). *STEM Teacher Observation Protocol Instrument Development and Data Collection.* (Tech. Rep. No. 19-11). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Riggs, I. M., & Enochs, L. G. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education*, *74*(6), 625-637.

Ross, J., & Bruce, C. (2007). Professional development effects on teacher efficacy: Results of randomized field trial. *The journal of educational research*, *101*(1), 50-60.

Sassler, S., Glass, J., Levitte, Y., & Michelmore, K. (2017). The missing women in STEM? Assessing gender differentials in the factors associated with transition to first jobs. *Social Science Research, 63,* 192-208.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher*, *15*(2), 4-14.

Smithsonian (2018). *The STEM Imperative.* Retrieved December 12, 2018 from https://ssec.si.edu/stem-imperative

Sparks, A., Adams, E.L., Cox, C.T., C., Hatfield, C., & Ketterlin Geller, L. (2019a). *STEM Academy for Teachers and Leaders: 2018-19 Coaching and PLC Evaluation* (Tech. Rep. No. 19-07). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Sparks, A., Adams, E.L., Mota, A., Simon, E., Burton, C., Hatfield, C., & Ketterlin Geller, L. (2019b). *STEM Academy for Science Teachers and Leaders: 2019 Teacher Academy 2*

Evaluation (Tech. Rep. No. 19-10). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Sparks, A., Adams, E. L., Perry, L., & Ketterlin-Geller, L. R. (2020). *The development of an instrument to measure teachers' perceptions of STEM practice.* Paper to be presented at the National Association for Research in Science Teaching (NARST) Research Conference, Portland, OR. (Conference canceled)

Tschannen-Moran, M., & McMaster, P. (2009). Sources of self-efficacy: Four professional development formats and their relationship to self-efficacy and implementation of a new teaching strategy. *The elementary school journal*, *110*(2), 228-245.

Tytler, R. & Osborne, J. (2012). Student attitudes and aspirations toward science. In B. Fraser, K. Tobin, C. & McRobbie (Eds.) *Second International Handbook of Science Education.* Dordrecht: Springer.

UTeach. (2019). *The UTeach Observation Protocol.* UTeach College of Natural Sciences program at the University of Texas Austin. Retrieved from https://utop.uteach.utexas.edu/

van Driel, J. H., Verloop, N., & de Vos, W. (1998). Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching, 35*(6), 673-695.

Appendix A – STEM PPC

STEM Academy for Teachers: Perceptions, Practice, and Culture Survey

Listed in the left column are descriptions of practices that teachers may or may not engage in when providing instruction or collaborating with others. Using the scales described below, please circle the number that represents how important you think that practice is, how confident you feel implementing such a practice, and the average frequency with which you implement that practice.

Practice	Importance of practice 0 = Not important at all 1 = Not very important 2 = Important 3 = Very important					Confidence in implementing practice 0 = Not confident at all 1 = Not very confident 2 = Confident 3 = Very confident					Frequency A: Less often than 1 time per month B: 1 time per month C: 2-3 times per month D: 1 time per week E: 2-3 times per week F: Everyday					
Students engage with one another in small groups to accomplish assigned tasks.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Instruction includes open-ended questions that encourage different approaches.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Students learn from teacher-led lecture or activities.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Instruction allows students to connect science concepts to real-life situations.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Students engage with hands-on materials.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Lessons encourage students to use a common method or procedure to solve problems.	0	1	2	3		0	1	2	3		A	B	C	D	E	F

Practice	Importance of practice 0 = Not important at all 1 = Not very important 2 = Important 3 = Very important					Confidence in implementing practice 0 = Not confident at all 1 = Not very confident 2 = Confident 3 = Very confident					Frequency A: Less often than 1 time per month B: 1 time per month C: 2-3 times per month D: 1 time per week E: 2-3 times per week F: Everyday					
Students engage with technology during lesson.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Students engage with science textbook.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Learning experiences foster students' personal interest in science.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Activities allow for multiple solutions or representations.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Formative feedback is provided to students.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Lessons connect to students' prior knowledge and experiences.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Lessons explicitly connect to mathematics ideas and terminology.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Lessons explicitly connect to engineering ideas and terminology.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Lessons explicitly connect to technology ideas and terminology.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Lessons expose students to information about STEM careers.	0	1	2	3		0	1	2	3		A	B	C	D	E	F

Practice	Importance of practice 0 = Not important at all 1 = Not very important 2 = Important 3 = Very important					Confidence in implementing practice 0 = Not confident at all 1 = Not very confident 2 = Confident 3 = Very confident					Frequency A: Less often than 1 time per month B: 1 time per month C: 2-3 times per month D: 1 time per week E: 2-3 times per week F: Everyday					
Learning experiences encourage student ownership.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Learning experiences foster a maker mindset (e.g., curiosity, resilience, and an eagerness to collaborate and share).	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Students have opportunities to showcase their achievements.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Instruction calls students' attention to the mathematics concepts that they may have learned in their math class.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Lessons utilize informal learning spaces.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Lessons utilize curricular materials developed by industry (e.g., TI, NASA).	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Students' cultural and ethnic backgrounds inform instructional planning.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Lessons incorporate community building activities.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Teachers collaborate during professional learning community (PLC) meetings.	0	1	2	3		0	1	2	3		A	B	C	D	E	F

Practice	Importance of practice 0 = Not important at all 1 = Not very important 2 = Important 3 = Very important					Confidence in implementing practice 0 = Not confident at all 1 = Not very confident 2 = Confident 3 = Very confident					Frequency A: Less often than 1 time per month B: 1 time per month C: 2-3 times per month D: 1 time per week E: 2-3 times per week F: Everyday					
Learning experiences foster students' motivation to learn science.	0	1	2	3		0	1	2	3		A	B	C	D	E	F
Teachers review assessment data with other teachers.	0	1	2	3		0	1	2	3		A	B	C	D	E	F

	Strongly Disagree	Disagree	Agree	Strongly Agree
I am familiar with the science TEKS for my grade level.				
I use project-based learning when teaching science lessons.				
I am familiar with the mathematics TEKS for my grade level.				
I feel comfortable integrating mathematical concepts into the science lessons I teach.				
My classroom culture encourages all students to share ideas without fear of being ridiculed or judged.				
Teachers should be aware of and respectful of students' emotions.				
My school emphasizes to students the importance of STEM.				
My school emphasizes to parents the importance of STEM.				
My school encourages interest in STEM.				

Describe characteristics of project-based learning.
Describe characteristics of maker-based education.
Describe an example of how you have integrated mathematics into a science lesson.
Describe a typical professional learning community meeting on your campus.
What informal learning settings have you used to enhance your instruction and engage your students?

Appendix B – STEM TOP

SMU | Annette Caldwell Simmons School of Education & Human Development

STEM Teaching Observation Protocol

Teacher: ______________________________ **Coach:** _______________ **Observation Date:** ________ **Double Observation**

Campus: ______________________________ **Start Time:** _________ **End Time:** ___________ Y N

Instructional format (circle all that apply): **Number of Students:** ________ **Grade Level:** _______ **Cycle:** ______

Whole class Individual Small group **Day of Week:** ________ **Class Type (Regular/PreAP):** __________

Rate the extent to which each indicator was:

0= Not observed: Not demonstrated at all

1= Emerging: This is an opportunity for growth; demonstrated at a low level

2= Proficient: Demonstrated at an expected level

3=Exemplary: Demonstrated at a high level

Domain	Indicator	Score (0-3)	Notes
1. Lesson Structure	a. The lesson objectives are clear to students		
	b. The lesson is structured to build understanding and maintain a sense of purpose		
	c. The lesson includes an investigative or problem-based approach (e.g., students investigate or discover scientific ideas)		
	d. The lesson is clearly connected to students' prior knowledge and experiences		
2. Learner Centered Instruction	a. Students explain and justify their thinking		
	b. Students engage in behaviors reflective of the process standards		
	c. Students direct their own learning (e.g., are provided with flexibility or choices during the lesson)		
	d. Teacher engages students in appropriately challenging content (e.g., critical thinking, problem-solving strategies)		
	e. Teacher openly welcomes discussion about mistakes or misconceptions		
	f. Teacher poses cognitively demanding, open-ended questions		
	g. Teacher explicitly connects learning to the real world (e.g., careers, current events)		
	h. Teacher explicitly connects learning to other disciplines (e.g., social studies, mathematics)		
	i. Teacher involves all students (e.g., calling on non-volunteers, facilitating student-student interaction, checking in with hesitant learners, etc.)		
	j. Teacher is attentive to students' academic and social/emotional needs (e.g., use of cooperative learning, language-appropriate strategies and materials, awareness of student comfort)		
3. Evaluation and Feedback	a. Teacher uses a variety of assessment strategies (e.g., large group questions, one on one discussion, small group feedback, exit tickets, quiz or test, informal progress check)		
	b. Teacher provides feedback focused on expanding learning and understanding (formal and informal formative), not correctness or the end product (summative)		
	c. Students evaluate their own or other's work		
4. Management and Discipline	a. Students are on task throughout the class		
	b. Students demonstrate an understanding of expectations for behavior		
	c. Students demonstrate an understanding of classroom procedures/routines		
	d. Teacher efficiently manages time (e.g., transitions, wait time, pacing)		
	e. Teacher redirects off task or disruptive behavior *(NA if no disruptive behavior)*		

STEM Teacher Observation Report

Summary of Lesson (optional):
Praise/Reinforcement:
Recommendations/Questions:
Polish/Refinement:

Coach Initials: _________

Appendix C – POSTT

Pre-Test:

Pedagogy of Science Teaching Tests (POSTT):
Grade 6-8 Items
(Cobern et al., 2014)

Instructions: This assessment is composed of classroom science teaching vignettes similar to teaching practices one can find in any classroom today. Practicing teachers contributed ideas for many of the vignettes; others are based on teacher observations, or on science curriculum standards.

As you read each vignette, think about how you might teach science in a similar situation. Respond accordingly.

General wrap-up of unit

Mr. Nelson's 6th grade students have just completed a unit in their earth science class. As a "wrap-up," Mr. Nelson would like students to re-examine the three learning objectives that served as the focus for this entire unit.
Of the following, which is most similar to how you would like to conduct the wrap-up?

A. I would ask the students what the main things are that they have learned in the unit, according to their own ideas of what is important or interesting, and have them list these as the unit wrap-up.
B. I would restate the three learning objectives for the students, and then relate each of them to the specific concepts that arose in the unit.
C. I would ask the students to reflect back on their work, and identify for themselves what the important central ideas of the unit were, then have them relate these to the original learning objectives.
D. I would restate the three learning objectives, then ask the students to say how the various concepts that arose in the lesson related to each of these.

Predator and prey

Mr. Peoples is conducting a unit on food chains and is about to introduce his 7th grade students to predator/prey relationships. He has a good computer simulation game for this subject that he can use with his class.
Thinking about how you would teach this lesson, of the following, which is the best advice for conducting this lesson?

A. Mr. Peoples should explain to his students that balance typically exists in nature such that the numbers of predators and their prey are related. For example, he can tell them that a rabbit population will increase if disease reduces the coyote population in the same region. He should then project the simulation game to demonstrate relationships between rabbit and coyote populations.
B. Mr. Peoples should explain to his students that balance typically exists in nature such that the numbers of predators and their prey are related. For example, he can tell them that a rabbit population will increase if disease reduces the coyote population of the same region. Using the computer simulation game, he should have the students monitor and record the rabbit levels over a simulated ten year period during which the population of coyotes rises and falls, so that they can confirm the predator/prey concept he explained.
C. Mr. Peoples should ask what would happen with rabbits if many coyotes died suddenly of disease. After some discussion, Mr. Peoples should suggest that the students explore their ideas using the computer simulation game he has for this subject, by recording yearly counts over a simulated period of ten years. The students will then have data to be used in a class discussion on predator/prey relationships.
D. Mr. Peoples should begin by asking the students what they know about predators and prey. Without responding other than to encourage their ideas, Mr. Peoples should then show them the computer simulation game he has for this subject and invite them to use the simulation in any way they wish to explore their ideas. The lesson would end with students writing up their findings.

Soil porosity

Ms. Cubbage's 7th grade science class has been learning about soil types by observing soil color and texture (particle size). While making observations of soil samples, the students notice that some soil types seem more "fluffy" than others. Ms. Cubbage realizes that her students are referring to *porosity* (how densely the materials are packed together, ability to allow water to move through) which is one of the key concepts later in her unit.

Thinking about how you would teach this lesson, of the following, which is most similar to how you would respond to the students' observation?

A. I would congratulate the students on such a good observation, then explain to them *porosity* is a description of how densely packed soils are. I would then tell students how to test soils for it, and follow up by doing tests on our soil samples for porosity.
B. I would congratulate the students on such a good observation, and ask them what they thought they were looking at. Through discussion I would try to get them to think about packing and how one might test for packing. We would do tests and based on their findings, I would introduce the concept of *porosity*.
C. I would recognize that what is most important here is that the students were being independent investigators, not necessarily that they were stumbling upon the idea of *porosity*. I would simply encourage their scientific attitudes and have them continue their investigations.
D. I would congratulate the students on such a good observation, then explain to them that what they observed was called *porosity*. Using a demonstration, I would show the students that more porous soils are less packed and that water moves more easily through porous soils.

Light reflection

Ms. Baker is teaching her 8th grade students the law of reflection: when a ray of light strikes a mirrored surface, it leaves at the same angle as when it arrived. Ms. Baker has to decide how she will teach the lesson.

Thinking about your own teaching, of the following, which is most similar to how you would teach the lesson?

A. I would write the law of reflection on the board and illustrate with a diagram. Next I'd show them a real example, using a light ray source, mirror, and protractor. Then we would discuss any questions the students might have.

B. I would ask students to find out what they can about light behavior around mirrors by exploring on their own with an assortment of available items, including light ray sources, mirrors, and protractors. Then the students would report back on what they did and what they found out.

C. I would first pose a question about reflection for the students to explore. The students could investigate using light ray sources, mirrors, and protractors, and then discuss their findings. I would close the lesson by giving them a summary of the law of reflection.

D. I would write the law of reflection on the board and illustrate with a diagram. Then I'd have the students verify the law using light ray sources, mirrors, and protractors. We would then discuss their findings.

Inheritance

Mr. Montgomery was teaching his 7th graders about inheritance. After introducing the topic and demonstrating how to use a Punnett square to determine genotypes and phenotypes of possible offspring, he asked students to solve a variety of application problems in small groups.

Thinking about how you would teach, how would you end this lesson?

A. Since students would have already discussed the problems in their small groups and developed their own understanding of the topic, I would end the lesson here.

B. I would give the students the right answers to the problems.

C. I would ask students to explain their answers to the class. Drawing on their explanations, I would guide them to the correct answers.

D. I would review the correct answers to the problems with the students as a class discussion.

Sundial

Ms. Navetta is planning a 7th grade lesson on the changing position of the sun in the sky during the day and how this is the basis of a simple 'sundial' to tell time of day. The basic sundial is a simply a vertical stick on a piece of board, and in sunlight the angle of the stick's shadow can be marked on the board. Ms. Navetta also has a larger demonstration model with lines marked at various angles and labeled with hour of day. Ms. Navetta considers various ways to conduct the lesson.

Thinking about how you would teach, of the following, which is most similar to how you would conduct the lesson?

A. Explain how a sundial works related to sun position in the sky. Have each group assemble a basic sundial, using a prepared handout sheet with lines and hour markings. Then take the students outside to try out their sundials and see that they indicate the correct time of day.

B. Do not explain sundials but take the students outside and have each group set up a stick and board. Ask them to brainstorm what this might be useful for, and to expand on their ideas. Have them come back every hour, anticipating that they will mark a series of shadow lines to make a sundial.

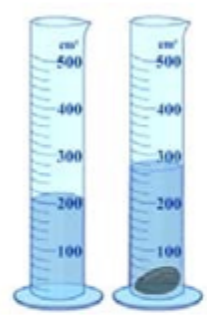
C. Explain how a sundial works, in relation to sun position in the sky. Then gather the class outside around the demonstration model, so they can see how the sundial indicates the correct time of day. Come back an hour later to see that the shadow has moved to the next marking.

D. Instead of explaining sundials take the students outside and note the location of the sun in the sky. Have each group set up a stick and board and mark the position of the shadow. Ask them to suggest how this might be used as a 'shadow clock' to tell time of day. Have them come back every hour and mark a new shadow angle, labeling it with the hour, to make a sundial.

Volume and displacement 1

Ms. Katinka is doing a lesson on volume in her 6th grade classroom. Part of the lesson will involve using a graduated cylinder partially filled with water for determining the volume of small irregular objects.

Thinking about how you would teach this lesson, of the following choices, how would you advise Ms. Katinka to structure her lesson?



A. Ms. Katinka should open the lesson by clearly stating the learning objective: the use of displacement as a measure of volume. The teacher then asks the students what happens to the water level in the bathtub when they climb in. She tells them that this is an example of displacement and then assigns an activity using graduated cylinders where the students measure the displacement caused by various objects.

B. Ms. Katinka should open the lesson by asking the students what happens to the water level in the bathtub when they climb in. She uses their ideas to introduce an activity using graduated cylinders where the students measure the displacement caused by various objects. Following further discussion of their observations, the teacher clarifies that the students have been measuring volume.

C. Ms. Katinka should open the lesson by having the students freely explore what happens when various objects are placed in the graduated cylinder. The students should first record their observations and then discuss their findings amongst themselves and with the teacher.

D. Ms. Katinka should assign an appropriate reading in a science textbook on volume and displacement. The students read in class, then the teacher shows the students how to determine the volume of an irregularly shaped object by water displacement in a graduated cylinder. The teacher then has the students find several objects around the room to test on their own using the displacement method.

Sediments and water

Ms. Downey would like her 8th grade students to understand the erosive effect of water on various types of sediment, and that running water erodes some types of sediments more easily than others.

Thinking about how you would teach this lesson, of the following, which one is most similar to how you would teach the lesson?

A. I'd ask the students if they think water erodes some sediments more easily than others, and allow them to complete an activity using sand and silt to help them determine the potential for each to erode.

B. I'd explain that the more water a sediment can hold, the more erosion will occur, and use samples of sand and silt to demonstrate.

C. I'd explain that the more water a sediment can hold, the more erosion will occur, and have the students do an activity themselves to verify this.

D. I'd ask the students to explore the effect of water on various sediment samples and come to their conclusions on each sediment's potential to erode.

Photosynthesis

Ms. Hamid has been teaching her 8th grade students about photosynthesis, and in particular that chlorophyll production in plant leaves is light-induced. She sets up an example to illustrate this. She has placed fast-growing seedlings where they are exposed to different levels of light intensity. The students observe the growing plants over several days and estimate the amount of chlorophyll using a color chart to record leaf color. They record their data in their science notebooks and on a classroom data table. On the last day, Ms. Hamid reviews the role of light in chlorophyll production as illustrated by the activity.

Thinking about how you would teach this topic, of the following, which is the best evaluation of her lesson?

A. This is a good lesson design overall because Ms. Hamid begins with an explanation of the concepts she wants the students to learn followed by an activity for students to confirm that chlorophyll production is light-induced.

B. Ms. Hamid begins appropriately with an explanation of the concepts she wants the students to learn. This being so, it is not clear that the activity is needed, especially since it requires so much class time.

C. Ms. Hamid's approach is too pre-organized and prescriptive. It would be better for students themselves to decide how to set up plants and lights, see what happens, and figure out a way to compare chlorophyll production in the leaves.

D. The instructional sequence would be better if the students do the plant observations first, showing that chlorophyll is light-induced, after which Ms. Hamid can explain the process more fully.

Planets and life

Having already learned the requirements for life on Earth, Ms. Taylor wants her 6th grade class to consider whether or not life might be possible on *other* planets. To do this, she has small groups of students research each of the other planets, paying special attention to the characteristics that would and would not make them suitable for life, and report back to the class on their findings.

Thinking about how you would teach this lesson, of the following, which best describes how you might *change* the lesson?

A. Instead of having students research the other planets, I would describe their characteristics and after that give them this group assignment.

B. I would instead simply pose the question of whether life might be possible on other planets, and allow the students to discuss and research this question in small groups, reporting back to the class.

C. I would instead describe the characteristics of the other planets for my class, highlighting whether the ones which are necessary for life are present or absent.

D. I would teach this lesson in a similar manner to Ms. Taylor, because she gives the students a focus but allows them to do their own research.

Post-Test:

Pedagogy of Science Teaching Tests (POSTT): Grade 6-8 Items

(Cobern et al., 2014)

Instructions: This assessment is composed of classroom science teaching vignettes similar to teaching practices one can find in any classroom today. Practicing teachers contributed ideas for many of the vignettes; others are based on teacher observations, or on science curriculum standards.
As you read each vignette, think about how you might teach science in a similar situation. Respond accordingly.

Fossils provide evidence for change

Ms. Jefferson's 6th grade is learning about how fossils can provide evidence of how life has changed over time. She poses a question: "How do fossils help show us changes on Earth over time?" Ms. Jefferson continues by asking the students to examine several different rock samples, all containing different types of fossils.
Thinking of how you would teach this lesson, of the following, how would you evaluate this lesson so far?

A. The students are making asked to make observations before being instructed on what to look for. Instead, Ms. Jefferson should have described how fossils provide evidence of change over time, using the fossil samples as examples to demonstrate her point.

B. The students are asked to make observations before being instructed on what to look for. Instead, Ms. Jefferson should have described how fossils provide evidence of change over time, and then have the students examine the different rock samples to verify Ms. Jefferson's explanation.

C. This lesson is fine the way it is. Ms. Jefferson states a question for the students to think about and then provides materials the students can utilize to explore this question.

D. Ms. Jefferson should have not posed such a detailed question prior to student investigations. The students should have been allowed to examine the rock samples and, as a class, discussed their ideas about the fossils.

Hand sanitizers: a teachable moment

Ms. Simmons' 8th grade class is learning about where bacteria are found and their influence on humans, when one of the students asks, "Why do we use hand sanitizers?" Ms. Simmons is not sure of the best way to respond to this student's question with respect to the lesson objectives.
Thinking about how you would teach, of the following, which is most similar to how you would deal with this teachable moment?

A. I would encourage the students to think of ways to answer this question, and give them time and materials to pursue their own investigations.

B. I would give my students a brief explanation of how hand sanitizers work and tie this information to what we had been studying about bacteria. Then we would get back to the lesson objectives.

C. I would give my students a brief explanation of how hand sanitizers work and then ask my students if they could think of ways to test their effectiveness. I would conclude the lesson by tying their ideas back to the objectives.

D. I would elicit the students' ideas about the question, including how to test those ideas. We would try some of those tests and I would conclude the lesson by tying their findings back to the objectives.

Stream table

Mr. Hamid would like his 8th grade class to explore the concept of sand erosion as a function of water volume. He has available stream tables, which can be used to illustrate this by taking measurements of sand height at various points before and after adding water. Mr. Hamid also has a film clip that explicitly explains and illustrates the erosion process. He is unsure which of these to use, and when.
Thinking about how you would teach this lesson, of the following, which is most like how you would structure this lesson?

A. I would begin the lesson with the film clip and explain to the students how sand erosion is a function of water volume. Then I would demonstrate this using the stream table, calling the students attention to what they saw in the film clip and are now observing in the stream table demonstration.

B. I would begin the lesson by having students explore with the stream tables, and then have them discuss what they discovered. The lesson would revolve around the students' explorations and conclusions, and not include the film.

C. I would begin the lesson by having the students investigate sand erosion by doing the stream table activity, followed by a class discussion of their observations. Then I would use the film clip to bring closure to the lesson.

D. I would begin the lesson with the film clip. Then I would have had the students do the water table activity so that they could see water erosion and the effect of water volume firsthand.

Bacteria

Ms. Simon's 8th grade class is studying bacteria. She tells her students that it is important to wash our hands before eating, even though we cannot see bacteria on them. Ms. Simon explains that, although we cannot see individual organisms, given nutrients, thousands more bacteria can grow in one place until the colony becomes visible. She demonstrates this by showing students the bacterial colonies that grew on some agar plates she had inoculated several days previously. She finishes the lesson with a reminder to wash the bacteria off your hands before eating.
Thinking about how you would teach this lesson, of the following, which one is most similar to what you would likely do?

A. I would begin by simply giving students fresh agar plates to inoculate (in any safe manner), and having them observe the plates over several days. I would end with a class discussion about what they learned from their observations and how it relates to being healthy.
B. I would begin and end the lesson the same as Ms. Simon. However, rather than showing them plates with bacteria already on them, I would allow students to touch their fingers onto their own fresh agar plates and then make observations over several days.
C. I would begin by having a class discussion on whether it is important to wash our hands before eating. Then I would have students touch their fingers onto their own fresh agar plates and make observations over several days. We would finish the lesson by relating student observations to the importance of hand-washing.
D. I would conduct this lesson in a similar way to Ms. Simon.

Frog dissection 1

Mr. Goodchild is doing a frog dissection with his 8th graders to help teach them about anatomy.
Thinking about how you would teach a lesson, of the following, which is most similar to what you believe is the best way to incorporate a dissection into a lesson?

A. It should be used as a step-by-step student activity while answering probing questions, followed up by teacher-led discussion and clarifications.
B. It should be used as a follow-up step-by-step student activity after Mr. Goodchild explains exactly what students will need to notice about the frog anatomy.
C. It should be used as a stand-alone step-by-step activity for students to explore the frog's anatomy and raise discussion questions on their own.
D. It should be used as a step-by-step demonstration by Mr. Goodchild while he explicitly points out what students need to know about frog anatomy.

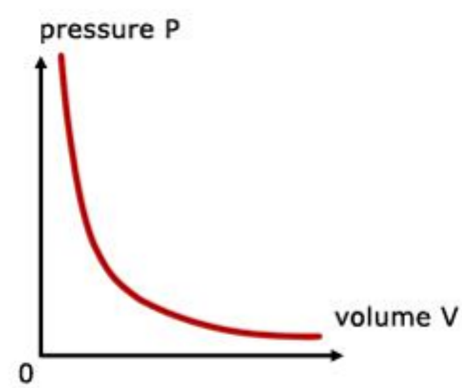
Comets

Ms. Thole's 7th grade class is doing an astronomy lesson on comets. She creates a dry ice model of a comet and has the students make observations. Then, using resources in the classroom, Ms. Thole has the students search for the characteristics of real comets, and create a list of properties found in the dry ice comet model that are also properties of real comets. Students will continue the list with additional properties they find of real comets. The lesson ends with a classroom discussion of their findings.
Thinking about how you would teach, of the following, which best describes your evaluation of Ms. Thole's lesson?

A. This is a good lesson because it allows the students to do their own investigating. However, the students would be more engaged if the teacher allowed them to design their own comet models.
B. Ms. Thole should have begun by describing the characteristics of real comets using photos and diagrams. Then she should have referred to the dry ice comet model to highlight some of the characteristics of real comets.
C. This is a good lesson because the students are able to explore a concrete model as well as using other resources to learn about the characteristics of comets.
D. Ms. Thole should have begun by describing the characteristics of real comets using photos and diagrams. Then, referring to the dry ice comet model, she should ask students to record their observations of how the model is similar to a real comet.

Boyle's law

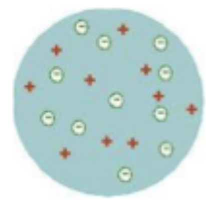
Mr. Lawton's 8th grade class is learning about Boyle's gas law. The volume of an enclosed gas is inversely proportional to pressure if temperature is kept constant. There is a gas law apparatus that can be used safely in the classroom. Mr. Lawton, however, is not sure in what order to do things: start by formally stating Boyle's law first or start by having the students do a gas experiment.
Thinking about how you would teach, of the following, which sequence would you advise Mr. Lawton to use?



A. He should first state Boyle's law so students encounter the correct science from the start, and then demonstrate it using the gas apparatus. This approach is efficient with no real need for students to repeat the experiment themselves.
B. He should first state Boyle's law so students encounter the correct science from the start. The students can then use the gas apparatus to confirm Boyle's Law for themselves.
C. He should first allow the students to experiment on their own to collect data on gas volume and pressure. Building on their data, he should then guide students toward Boyle's Law as a model that accounts for their findings.
D. He should first allow the students to explore gas behavior on their own and collect data. He should ask them to create models that account for their findings and be prepared to report back to the class.

What is an atom like?

Ms. Dalton taught her 8th grade students that matter is made of atoms. She then asked what an atom itself might be like. She wants to introduce the idea of models in science: though we can't see atoms we can envisage models to account for their properties. She said one clue about atoms was that scientists had discovered there are negative particles (electrons) in atoms, though atoms are neutral overall. She has available large posters of the older 'plum pudding' model and the subsequent 'nuclear atom' model.



"Plum pudding"

Nuclear Atom

Thinking about how you might teach, of the following, which is closest to how you would conduct the lesson? I would...

A. Begin by asking the students to suggest their own models for what an atom might be like, given that it has electrons but is neutral overall. I would then have them compare their ideas with the 'plum pudding' and 'nuclear atom' models. Drawing on their comments, I would explain how scientists arrived at the first model and why they later changed to the second model. Then I would then have the students sketch different example atoms drawing information from a Periodic Table.
B. Display the posters of both models and explain how scientists arrived at the first model and why they later changed to the second model. Referring to information from the Periodic Table, I would show how the 'nuclear atom' model can represent the atoms of different elements.
C. Display the posters of both models and explain how scientists arrived at the first model and why they later changed to the second model. I would then have the students sketch different example atoms drawing information from a Periodic Table.
D. Ask the students to think of as many possible models as they can of how an atom might look. I would have them report back on their ideas with sketches. Most of the lesson time would be given to students proposing and supporting their models. I would end the lesson by having the students compare and contrast their ideas with the 'plum pudding' and 'nuclear atom' models, noting that scientists now embraced the latter.

Temperature and solubility

Ms. Clark's 7th graders have learned that sugar becomes more soluble in water as the water temperature increases. Now she wants her students to learn that, unlike sugar, the solubility of salt does not increase with temperature. Graduated cylinders of hot and cold water, salt, sugar, and stir sticks are available.

Thinking about how you would teach, of the following, which one is most similar to how you would conduct this lesson?

A. I would explain that while we found that sugar is more soluble in hot water, not all solids behave the same way. I would demonstrate using salt instead of sugar in the graduated cylinders of hot and cold water.
B. I would pose the question of whether all solids might dissolve better in hot water like sugar did. I would ask them to design and do an experiment to test whether salt dissolves better in hot or cold water.
C. I would give my class sets of graduated cylinders, salt, sugar, and hot and cold water, and ask them if they could find out anything about salt versus sugar dissolving in water. I would not prescribe what they should do. Later, we would discuss what they did and what they found out.
D. I would explain that while we found that sugar is more soluble in hot water, not all solids behave the same way. I would then have them verify this in the lab using the same amount of salt in each cylinder of hot and cold water.

Chlorophyll

"Chlorophyll is a natural pigment found in green plants. It is the primary pigment that absorbs light energy from the sun for photosynthesis. This energy is then used by the plant to synthesize carbohydrates from carbon dioxide and water." This is what Ms. Pozner's 8th grade life science textbook says, and she is wondering how best to teach a lesson on chlorophyll and the process of photosynthesis. She has several ideas but each one has limitations.

Thinking about how you would teach, of the following ideas that Ms. Pozner is considering, which is most similar to how you would teach this lesson? I would...

A. Have the students read the textbook section on photosynthesis and chlorophyll. We would then have a classroom discussion where we draw out and clarify the important points. I would then summarize the concepts for the students.
B. Have the students read the textbook section on photosynthesis and chlorophyll. I would then explain these concepts while showing relevant slides.
C. Read aloud the textbook section on photosynthesis and chlorophyll to the class. Then I would explain these concepts while showing relevant slides.
D. Have the students use the textbook and online resources to investigate photosynthesis and the role of chlorophyll. I would then have the students summarize their findings for the class.