STEM Academy for Science Teachers and Leaders: UTOP Research Observation Procedures

Internal Report

RESEARCH IN MATHEMATICS EDUCATION

STEM Academy for Science Teachers and Leaders: UTOP Research Observation Procedures

Internal Report

Carlene Burton • Cassandra Hatfield • Elizabeth Adams • Leanne Ketterlin Geller

Southern Methodist University

This report was updated with an addendum in spring 2020.

Spring 2020

Abstract

This technical report describes the procedures used to collect observable data from 41 middle school science teachers ranging from grades 6 through 8. The teachers were observed using the UTeach Observational Tool. This observational tool was selected because it measures 22 aspects of classroom behavior and was designed based on theories and practices that emphasized teacher preparation and student engagement. The use of the tool seeks to answer the research question: How do teachers progress in designing and implementing active learning opportunities in the classroom during the STEM Academy for Teachers and Leaders? This report describes: a) the hiring of Research Assistants, b) the process of gaining necessary approvals to conduct observations c) the training procedures and content for the UTeach Observation Protocol for Math and Science, d) the scheduling of teacher observations and double observations for validity evidence, and e) the data collection processes.

Internal Report

Table of Contents

STEM Academy for Science Teachers and Leaders: UTOP Research Observation Procedures

Introduction

The STEM Academy for Middle School Science Teachers and Leaders (STEM Academy) is a systems-level program that focuses on increasing students' interest, motivation, and achievement in STEM and to increase middle school science teachers' success in teaching with STEM integrated practices. Teacher and leader professional development academies were provided during the summer months to develop their understanding of pedagogical approaches that emphasize student-led investigation through student engagement, exploration, explanation, elaboration, and evaluation (Bybee, 2015). The Academies were designed to emphasize four foundational pillars of STEM learning including (a) active learning or inquiry-based instruction, (b) the scientific process standards, (c) deep content knowledge, and (d) differentiated support for all learners. During the academic year, teachers and leaders were supported by instructional coaches for up to seven coaching cycles. The cycles were comprised of one-on-one coaching and professional learning community meetings.

The purpose of this technical report is to identify the processes and procedures used to collect observable data from the target classes[1] of each participating teacher, to be called 'participants,' in the STEM Academy. This report describes the processes of hiring research assistants to collect data, collecting data, and verifying data obtained using the UTeach Observation Protocol.

UTeach Observation Protocol

The UTeach program was founded at the University of Texas in 1997 to allow science and mathematics majors to obtain a degree in their discipline and teaching certification without adding extra time to degree completion. The UTeach Observation Protocol (UTOP) was designed to inform STEM teacher education by assessing the teaching practices of in-service teachers who had gone through the UTeach program, and who had been teaching for variable amounts of time. Thus, the instrument is connected to the theories that undergird the UTeach model and was developed based on a set of six fundamental foci that are heavily emphasized in STEM teaching. The UTOP was adopted by the STEM Academy to satisfactorily answer the questions:

- How do teachers progress in designing and implementing active learning opportunities in the classroom during the STEM Academy?

When compared with other observational tools, the UTOP was valued as the appropriate choice as it contains indicators that were specific in identifying pedagogical inquiry-based concepts and

---

[1] Target classroom is defined as a class that best represents typical general instruction.

would help to answer the research questions identified. The UTOP has four sections that gather information on the Classroom Environment, Lesson Structure, Implementation, and Mathematics/Science Content. Each section has an average of six indicators. Samples of the indicators for each section are shown in Table 1.

Table 1. Sample of Indicators by Section

UTOP Section	Sample Indicators
Classroom Environment	The teacher's classroom management strategies enhanced the classroom environment. The classroom is organized appropriately such that students can work in groups easily, get to lab materials as needed, teacher can move to each student of student group, etc.
Lesson Structure	The lesson was well organized and structured (e.g. the objectives of the lesson were clear to students, and the sequence of the lesson was structured to build understanding and maintain a sense of purpose). The lesson included an investigative or problem-based approach to important concepts in mathematics or science.
Implementation	The teacher involved all students in the lesson (calling on non-volunteers, facilitating student-student interaction, checking in with hesitant learners, etc.) The instructional strategies and activities used in this lesson clearly connected to students' prior knowledge and experience.
Mathematics/Science Content	The mathematics or science content chosen was significant, worthwhile, and developmentally appropriate for this course (includes content standards covered, as well as examples and activities chosen by teacher). During the lesson, it was made explicit to students why the content is important to learn.

Each indicator within the sections is rated on a scale of 1-5. Once a score is given, observers are required to provide written objective evidence of about three sentences, thus justifying each score. There is also a rubric that provides detailed explanations with examples for each score. The website *https://utop.uteach.utexas.edu/* provides a complete and updated version of the UTOP and its training guide/rubric.

The manual/training guide provides "resources to instruct, develop, and sustain inter-rater reliability" (UTeach Observation Protocol, n.d.). Through training using videos to score recorded examples, users learn how to use the tool so that each rater can reliably score the tool using the rubric and indicators. Comparisons and adjustments among raters are made to observation ratings to calibrate scoring between raters. The goal is that a different rater can go into the same classroom and come out with the same score. Data collected from multiple observations when using the UTeach Protocol should "reduce the impact of a low score on any single indicator" (UTeach Observation Protocol, n.d.).

Hiring Research Assistants

External Research Assistants (RA) were identified and hired, to conduct a UTOP research observation in each participant's classroom for one month data collection window. The number of RAs hired was commensurate with the number of observations required. For a rationale on the number of RAs hired, see the section entitled *Research Observation Schedule Design*. Research Assistants were hired based on experience working in classroom settings. Most had worked as teachers or administrators and understood the dynamics and content of a science classroom. Most had formally observed classrooms before and had used a rubric to guide their scores. With this previous experience, the RAs could be trained to use the UTOP with minimal bias, understanding the importance of the standardized protocol. While they do have classroom experiences, they have no real working knowledge of the program or the growth/or lack of potential growth of the teachers they observe.

Job Posting

To find qualified candidates, a job description was disseminated via multiple channels, including the SMU student job board, email blasts to SMU STEM area colleagues and departments, and social media channels. For a complete description of the advertisement see Appendix A. The position remained open for two weeks.

Applicant Selection

A rubric was designed to rate applicants based on three (3) components: contingencies, experiences, and availability. These components were further broken down in to eight (8) specific criteria that had a rating of one (1) point each (8 in total). The criteria are listed in Table 2.

Table 2. Application Components with each Criteria

Component 1 Contingencies	Component 2 Experiences	Component 3 Availability
• Submitted a resume • Submitted a cover letter	• K-12 teaching • STEM teaching • Evaluation/Mentoring • Research	• Work authorization • Flexibility of schedule • CITI Training

3

Applicants with a six (6) or above were invited for an interview. It was not mandatory for an applicant to have already taken CITI training. If necessary, CITI training was provided to applicants who were hired. The purpose of the interview was to provide a summary of the project, make a personal connection with the interviewee, share the potential observation schedule so that applicants could ensure their availability prior to hiring, and to allow the interviewer to ask pertinent questions about the position. Interview questions included:

- Tell us about yourself and your experiences in schools.

- What experiences do you have with using observational protocols for evaluation or research purposes?

- What types of experiences have you had with research data collection?

- What types of experiences have you had with data confidentiality and privacy? Are you CITI trained?

- Should you be hired, do you foresee any challenges that will prevent you from being available between the hours of 8 am and 4 pm from Monday through Friday?

In the fall of 2018, 12 applicants applied for the position and four (4) were offered the job for fall 2018 observations. The goal was for these same RAs to return in the Spring of 2019 for the second set of observations. However, only one returned. Rationale for not continuing into the spring included scheduling conflicts and health concerns. After reposting the job advertisement for the spring data collection, 10 applications were received and offers were made to three (3) new RA in addition to the one returning RA. This group of RAs continued with data collection through the 2019-20 school year.

For biographical information about the RAs, see Appendix B.

Research Observation Approvals

Compliance with SMU and Dallas Independent School District (Dallas ISD) research compliance protocols were required for project staff and RAs to understand the responsibility in working with human subjects in research. Institutional Review Board (IRB) requirements included training and signatures on:

- Completing an online CITI Training course: The Collaborative Institutional Training Initiative (CITI Program) seeks to uphold "...the public's interest during the research process by providing web-based courses in research, ethics, regulatory oversight, responsible conduct of research..." for all personnel involved in the conduction of a research in the United States and internationally (https://about.citiprogram.org, 2019). The certification is valid for three years.

- Signing the Dallas Independent School District (Dallas ISD) Research Review Board (RRB) Research Proposal Confidential Data Form. Signing this form legally informs the signee that participants' data are protected under the federal Family Educational Rights

4

and Privacy Act of 1974 (FERPA) and failure to do ensure the rights of the participants will warrant immediate revocation of any contract or research project.

- Completing the Dallas ISD Volunteer Form: This form is accessible electronically and is a Dallas ISD requirement. The submission and confirmation of acceptance grants the approved volunteer access into a Dallas ISD classroom to offer support as a mentor, tutor, chaperone, etc. The district reserves the right to re-check the criminal history of any volunteer during the school year (http://www.volnow.org/media/Voly/VOLY-Volunteer-Quick-Start-Guide.pdf, 2019). All RAs involved in the project are expected to have a copy of the volunteer form with them at campus visits. The volunteer form is valid for one school year.

To ensure all RAs completed this process, the project specialist worked with the RAs, Research Coordinator and the Principal Investigator to facilitate training and necessary signatures and submissions to SMU IRB and Dallas ISD RRB.

UTOP Training

The Research Coordinator trained the RAs on using the UTOP in September 2018 and February 2019. The training included up to eight hours of face-to-face training. Four RAs attended the fall 2018 training and four RAs, one of whom was a returning RA, attended the spring 2019 training. In 2019-20, three returning RAs participated in an abbreviated training in fall 2019 and spring 2020. The Research Coordinator and a graduate research assistant adapted the 2018-19 UTOP trainings based on a UTOP training, which was led by Dr. Candace Walkington, who was part of the team who developed the UTOP. Videos of Dr. Walkington's training are available: https://smu.box.com/s/mguj5mvh7wmv28mntg5zedum0zl0kd07

In fall 2018, prior to the training, RAs read a book chapter focused on the validity evidence of the UTOP. The training included a description of the overall protocol, a description of each indicator, and a group activity during which RAs summarized the essential observable behaviors for each section of the UTOP. Following the activity, RAs scored two practice videos. Dr. Walkington provided access to the master scores of these practice videos. After scoring each of the videos, RAs discussed scores and compared their justification with the expert raters' justification. After two practice videos, RAs watched and scored a calibration video. The Research Coordinator subsequently compared their scores with the master scores to determine inter-rater agreement. The fall 2018 training PowerPoint is available: https://smu.box.com/s/o573fgpip9gpnrnlgf5gpqqp5z45zhjh

In spring 2019, minor changes were made to the training. Prior to the training, RAs reviewed the UTOP measure and the UTOP User's Guide, rather than reading a book chapter. In Spring 2019, the three new RAs attended the eight-hour training; the returning RA attended the calibration session only. In Spring 2019, the training was revised to emphasize the descriptions and definitions of each of the indicators. The group activity was removed and the trainees spent less time on the validity evidence for the tool. The spring 2019 training PowerPoint is available: https://smu.box.com/s/l7tovcv2hgtvik09r9zreonvk70ahbs8

In fall 2019 and spring 2020, UTOP data collection continued with three returning raters. Returning raters engaged in a two day four-hour re-certification training each semester. During the re-certification training, RAs reviewed the UTOP indicators and scored one calibration video. The fall 2019 re-certification PowerPoint is available: https://smu.app.box.com/file/507968060993. The spring 2020 re-certification PowerPoint is available: https://smu.app.box.com/file/607546477155.

RA inter-rater agreement is depicted in Table 3. The goal was for each RA to score each indicator and section the same or within one point (+/-) of the master score 80% of the time or more, which was achieved in spring 2019. The goal for inter-rater agreement was not met in fall 2018 or spring 2020. Following the fall 2018 training, RAs attended two additional hours of UTOP training focusing on the expert raters' score justifications. In addition, fall 2018 RAs re-read the UTOP User's Guide, specifically focusing on eight indictors that were scored with the least consistency. Similarly, following the spring 2010 training, RAs attended one additional hour of UTOP training focused on the expert raters' score justifications. Raters 4 and 6 watched a second video and matched within one above 90%.

Internal Report

6

Table 3. RA Inter-Rater Reliability

	Fall 2018					Spring 2019					Fall 2019					Spring 2020			
Rater	Exact		Within one		Rater	Exact		Within one		Rater	Exact		Within one		Rater	Exact		Within one	
Rater 1	8/26	31%	15/26	58%	Rater 4	16/26	62%	26/26	100%	Rater 4	9/26	42%	24/26	92%	Rater 4	9/26	46%	20/26	77%
Rater 2	1/26	4%	7/26	27%	Rater 5	8/26	31%	23/26	88%	Rater 5	13/26	50%	25/26	96%	Rater 6	5/26	19%	16/26	61%
Rater 3	9/26	35%	23/26	88%	Rater 6	16/26	62%	26/26	100%	Rater 7	13/26	50%	23/26	88%	Rater 7	11/26	46%	23/26	88%
Rater 4	7/26	27%	19/26	73%	Rater 7	13/26	50%	23/26	88%	--	--	--	--	--	--	--	--	--	--
Overall	6/26	23%	16/26	62%	Overall	13/26	50%	25/26	94%	Overall	12/26	46%	24/26	92%	Overall	9/26	35%	20/26	77%

Note: Fall 2018 calibration scores: https://smu.box.com/s/mrc0qcs1tdff52qoqokoa3i1nfip7kfq; spring 2019 calibration scores: https://smu.box.com/s/zzzg3bxsbnf8g6z283gioh5eslazyxa3; fall 2019: https://smu.box.com/s/kvrae44ifgc32i6pnaekbj1pwzbiv4c2; spring 2020 calibration scores: https://smu.app.box.com/file/608147603358.

Although the inter-rater agreement was less than desirable in fall 2018 and spring 2020, these rates of agreement are consistent with what others in the field have observed (Hill, Charalambos, & Kraft, 2012). In order to prevent rater drift (Casabianca, Lockwood, & McCaffrey, 2015), co-calibration sessions were conducted at each timepoint, resulting in 20% of UTOP observations scored by two RAs. During the co-calibration, two RAs observed the same lesson, scored the lesson using the UTOP without interacting with one another, subsequently discussed their scores, and determined final scores for the lesson.

Due to scheduling issues in fall 2018, the Research Coordinator conducted two of the co-calibrations. The Research Coordinator examined inter-rater agreement mid-way through scoring based on the available completed co-calibration data by looking at RAs individual scores prior to their discussions. To the extent possible, it is desirable to conduct co-calibrations during the first few weeks of data collection rather than near the end of data collection.

Table 4 shows that in fall 2018, the inter-rater agreement statistics improved from the training calibration, with RAs agreeing on scores exactly or within one of each other for 88% of the scores on average. Given concerns about the agreement rates in fall 2018, the Research Coordinator prepared a detailed description of the inter-rater agreement, which raters then read prior to conducting additional observations. This detailed description is available: https://smu.box.com/s/vuscvf67nchc4vtio5nelobqxjhprjm5

At other timepoints, the co-calibration sessions indicated acceptable inter-rater agreement, as evidenced by RAs agreeing exactly or within one for an average of over 90% of indicators.

Internal Report

Table 4. Inter-Rater Agreement During Data Collection

	Fall 2018					Spring 2019					Fall 2019					Spring 2020				
Rater Match	Exact		Within one			Rater Match	Exact		Within one		Rater Match	Exact		Within one		Rater match	Exact		Within one	
R2/R1	16/26	62%	25/26	96%		R4/R7	11/26	42%	21/26	85%	R4/R7	22/26	85%	26/26	96%	R6/R4	23/26	88%	26/26	100%
R4/R3	13/26	50%	21/26	80%		R4/R6	19/26	73%	26/26	100%	R5/R7	17/26	65%	26/26	100%	R6/R4	19/26	73%	26/26	100%
R3/R1	5/26	19%	18/26	69%		R5/R6	21/26	81%	26/26	100%	R7/R4	16/26	62%	26/26	100%	R4/R7	19/26	73%	26/26	100%
R4/R2	16/26	62%	25/26	96%		R5/R6	21/26	81%	25/26	100%	R7/R4	24/26	92%	26/26	100%	R7/R6	17/26	65%	26/26	100%
R4/RC	18/26	69%	24/26	92%		R7/R4	24/26	92%	26/26	100%	R5/R4	19/26	73%	25/26	96%	R7/R6	22/26	85%	26/26	100%
R3/RC	20/26	77%	26/26	100%		R7/R6	21/26	81%	26/26	100%	R7/R4	18/26	69%	26/26	100%	R7/R4	24/26	92%	26/26	100%
R4/R2	19/26	73%	25/26	96%		R5/R4	20/26	77%	25/26	98%	--	--	--	--	--	--	--	--	--	--
Overall	15/26	58%	23/26	88%		Overall	20/26	77%	25/26	98%	Overall	19/26	74%	26/26	100%	Overall	21/26	79%	26/26	100%

Note: Rater (R1); Rater 2 (R2); Rater 3(R3); Rater 4 (R4); Research Coordinator (RC); Rater 5 (R5); Rater 6 (R6); Rater 7 (R7). Fall 2018 calibration results are available: https://smu.box.com/s/fnea87u3xsbiyiojkdgwwbulpwsqocg8; spring 2019 calibration results are available:

https://smu.box.com/s/h52bwnlwz7rz60px4we2zww1pg2d15ln; fall 2019 calibration results are available: https://smu.app.box.com/file/636526619233; spring 2020 calibration results are available: https://smu.box.com/s/kq35c1pqj2gk54o44eysmoiyejd7l6o3

Internal Report

The calibration and co-calibration results indicate that the training resulted in overall acceptable inter-rater agreement. Evidence suggests that the follow-up training resulted in increased levels of inter-rater agreement as evidenced by the co-calibration lessons. The calibration and co-calibration results suggest that re-training was effective. The following section of this report outlines the UTOP data collection procedures.

Research Observation Scheduling

The research observations were scheduled at the same time as a coaching observation. The purpose of conducting the coaching observation and the research observation at the same time was to reduce the number of communications sent to the participants and decrease the amount of time spent by the project specialist in trying to schedule an additional time point in which our team would interact with the participants. Getting responses from participants in a timely manner via email or text was challenging at times because participants were extremely busy with their teaching responsibilities. This section provides details of the process of creating schedules and revising schedules throughout the observation window.

Research Observation Schedule Design

At the beginning of the academic year, a coaching observation calendar was created. Since coaching observations and research observations occurred simultaneously, the project specialist worked with the SMU coach to confirm the dates and times of the coaching observations and scheduled observers to visit classrooms during those times. Table 5 shows a weekly calendar that allowed the project specialist to confirm each RAs schedule met these required specifications.

- Every teacher in the program had to be observed by one of the four RAs.

- Each RA could only conduct two observations per day, with at least one hour between each observation for score assignment.

- 20% of the total observations needed to have a second RA observer to allow for co calibration between the RAs.

- RAs travel time between schools was a minimum of 30 minutes.

At least one other member of the research team verified that the scheduled draft matched the dates of the coaches' visit. The Project Specialist shared the observation schedule with the RAs during a logistics meeting. In the logistics meeting it was made clear that the schedule would likely have revisions based on confirmation communications with teachers closer to the observation date.

During the logistics meeting of spring 2020, the Project Specialist informed RAs on the use of Outlook calendar invites to mitigate scheduling challenges. Accepting of the invites signals availability. Additionally, weekly emails provided reinforcement and opportunities for coaches and RAs to verify their availability. It was made clear that the schedule would likely have revisions based on confirmation communications with the teachers closer to the observation date.

Table 5. A Week's View of the Observation Schedule

**MARCH				
MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY
25	26	27	28	29
School 5: **RA 1** 10:25-11:56-Teacher 1 12:30-2:04-Teacher 2 *School 2*: **RA 2** 9:44-10:59-Teacher 3 1:07-2:22-Teacher 4 *School 7*: **RA 3** 10:15-11-Teacher 5	*School 10*: **RA 1** 8:25-9:14-Teacher 6 10:08-10:54-Teacher 7 *School 10*: **RA 4** 10:58-11:44-Teacher 8 1:10-1:56-Teacher 9 ***RA 1/RA 4** *School 10*: **RA 2** 11:48-12:34-Teacher 10 2:00-2:46-Teacher 11 ***RA 2/RA 3** *School 5*: **RA 3** 10:04-11:56-Teacher 12	*School 3*: ***RA 2/RA 3** 11:35-12:20-Teacher 13 *School 3*: **RA 1** 12:24-1:09-Teacher 14 *School 3*: **RA 4** 1:13-1:58-Teacher 15	*School 7*: **RA 4** 10:08-10:53-Teacher 16 *School 7*: **RA 3** 10:57-11:42-Teacher 17 *School 8*: **RA 1** 9:58-10:38-Teacher 18 *School 8*: **RA 2**- 9:14-9:54-Teacher 19 10:42-12:39-Teacher 20	RAs and RME Staff in office

Note. *RA X/RA Y represents double calibration of a teacher. During this calibration, each RA observes the same teacher, collects separate notes, and meet to discuss their observations. Together, they arrive at conclusions for each indicator represented on the UTeach Protocol.

**See link to complete March calendar of observations https://smu.app.box.com/file/455749227078

Reminder Communications

Through the use of emails, the AppToTo messaging system and the support of the coaches, teachers, and RAs were reminded of the observations and the submission of a lesson plan for the lesson observed. The lesson plan was a requirement for the practicum course in which teachers were enrolled in as a part of the treatment.

The first email reminded the teacher of the observation by the coach and informed them that the RA would also be present. Additionally, the email requested that the teacher upload the lesson plan associated with the observation to Canvas, the learning management system utilized for the STEM Academy teachers at SMU. Teachers were asked to reply to confirm the date or provide an alternate date. See Appendix C for example.

Most teachers confirmed by responding to the email reminder. Rescheduling was typically due to teacher absence or campus testing schedule changes.

Schedule Revisions Procedures

When schedule changes occurred from the initial schedule, it was crucial for the coaches and Project Specialist to communicate and revise the schedule. This was followed by communication with teachers and RAs. Once changes were finalized, the Project Specialist shared the updated schedule with the RAs and coaches adjusted calendar invites accordingly.

The flow of communication among the Project Specialist, teachers, coach, and RA was always consistent. The RAs communicate only with the Project Specialist, while coaches and teachers communicate with each other throughout the month of observation. This level of communication ensures that the schedule remains current and updated. An example of a situation that occurred and the steps to reschedule are shown Figure 1.
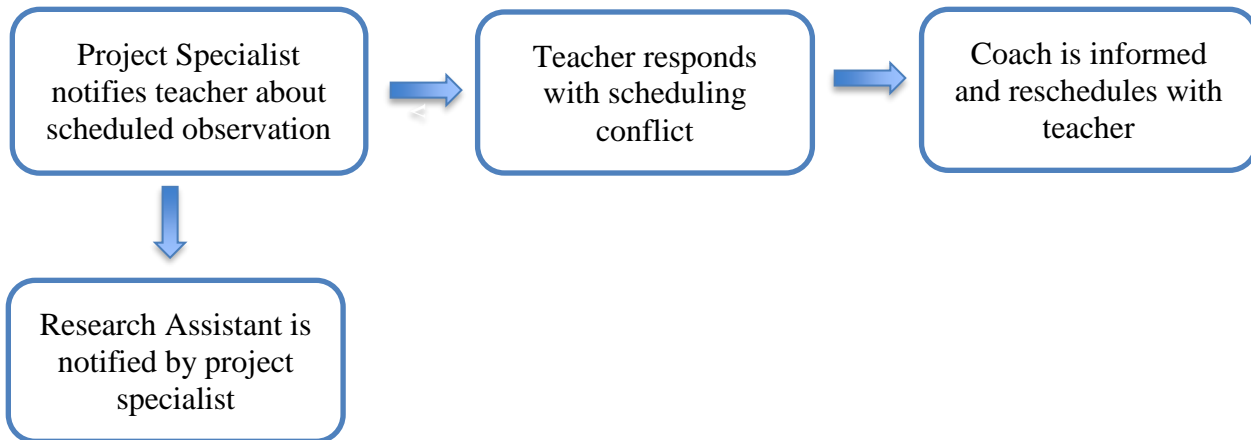

Figure 1. Example Communication Steps for Rescheduling an Observation

Research Observation Data Collection Procedures

Data collection is a collective effort of the RAs and the RME research team. Clearly defined roles, responsibilities, and procedures ensure the data collection validity and the privacy of the participants.

Research Assistants (RA)

Two specific roles of an RA are to be a professional representative of the RME team upon entering the school and subsequently the classroom, and also to be objective in his or her observation protocol. These expectations were shared with the RAs during the logistics meeting and are included in the field guide. The RA responsibilities ensure that their job as an observer to collect data also include:

- monitoring emails, group texts for updates to schedule, and accept Outlook invites,

- communicating issues to RME that may have prevented the expected flow of a lesson,

- arriving to campus at least 15 minutes prior to start of lesson,

- ensuring that no identifiable data is left on the laptop,

- completing the UTOP forms objectively,

- uploading documents according to security specifications,

- uploading data in a timely manner, and

- clarifying discrepancies if needed.

The RME Research Team

The RAs are all dependent on the effective execution of all relevant actions needed to conduct observations and complete the data entry process. It is imperative that the research team take into consideration that while the RAs are in the field, they have all the relevant materials and information needed to complete the assignment. As such, the team's responsibilities included:

- sharing schedules with RAs a week in advance,

- support coaches with Outlook invites to RAs,

- communicating with RAs when there were changes to the schedule,

- providing computers and observation forms needed to effectively collect data,

- granting Box access to RAs for secure data entry and time keeping,

- providing tracking documents for RAs to log data entry, hours, mileage, and

- verifying data in a timely manner and seeking clarification from RAs to ensure accuracy of data.

UTOP Observation Form Data Collection and Verification

The RAs complete a paper or electronic Word version of the UTOP form during their observations and then finalize it after the observation. Then the RAs transfer the data into the online version in Qualtrics. It is the responsibility of each RA to ensure that both the UTOP form and Qualtrics have the same data recorded. The RAs then complete their portion of the UTOP Observation Tracking Checklist (Figure 2) to track the completion of the observation and data entry.

Following the RAs responsibilities, the RME team verifies the data entry as a two-step process on a weekly basis during the observation window. The first RME verifier corroborates that the RA has accurately uploaded the UTOP forms and completes the Qualtrics version. The verifier initials the document with the date. The first verifier downloads the Qualtrics data set for the week and shares with the second RME verifier. The second verifier's responsibility is to confirm that all information on the UTOP form matches the Qualtrics version. If there are discrepancies, the Project Specialist communicates with the RA who made the error. The RA responds with the correction(s). The Project Specialist notes all corrections in an Excel document. All verification is completed within a week of observation. At the conclusion of data collection for each timepoint, the Research Coordinator cleans the data in Stata using the corrections listed in the Excel document.

Resources and Tracking

The research team designed resources and tracking documents that were shared with the RAs during the UTOP training and logistics meeting. The resources and tracking documents included a Research Observation Field Guide, research observation tracking checklist, hours and mileage tracker and contact list.

Research Observation Field Guide

The Research Observation Field Guide is a resource that outlines the expectations of each member of the UTOP data collection team. This team comprises of the coaches, the RAs, and the RME team. This document includes essential steps to be taken before, during, and after observations. During a logistics meeting, the document was reviewed and discussed with the RAs before observations began. Each RA was equipped with electronic and hard copy versions.

Research Observation Tracking Checklist & Contact List

A spreadsheet was created with three worksheets. One sheet was used to track the data collection process for each observation. There is a total of eleven columns, 9 completed by the RA and 2 completed by two RME staff members. As shown in Figure 2, the spreadsheet was used to track when a teacher was observed, the date the UTOP forms were saved on Box and also data entered to Qualtrics. Additionally, a column was provided for RAs to share any unplanned events that may have interrupted the lesson. Once RAs completed their columns for an observation, RME staff members assumed the role of verifier and completed the other two columns once those steps were complete

A second sheet included a contact of RME team telephone and email contact information of all the coaches, other members of the RA team, and RME staff who can assist if there are issues in the field. A third sheet was also used by RME staff to track the laptop distribution to the RAs.

Figure 2. UTOP Observation Form for Data Verification

See link to document https://smu.app.box.com/file/37702131415

Research Observation Tracking Hours and Mileage Lists

The Project Specialist created spreadsheets to track and hours and mileage (Table 6). Each RA had individually secured access to monitor mileage and hours used. Logged information was used by the RAs to submit their hours and mileage to Kelly Services[2] for payment, while the Project Specialist used the data to ensure that hours and mileage are accurately logged. RAs made weekly submissions to Kelly Services. The Project Specialist is then notified by Kelly Services to confirm the weekly hours and mileage submitted. If there are no issues, then the notification is confirmed. The Project Specialist used the hours logged and submitted to update the observation budget.

Table 6. An Actual Representation of Mileage Logged by RA 5

UTOP Observation Mileage Log						
Name of RA						
Date	Point of Dept	School Name	School Address	One way/Return Trip	# of miles (must match Google Maps)	Comments
3/5/19	*Office*	*School 1*	*School 1*	One way	14.20	
3/6/19	*School 2*	*School 3*		One way	13.30	
3/6/19	*School 3*	*Office*		One way	6.10	
3/19/19	*Office*	School 4	*School 4*	Return Trip	9.80	Pick up & return laptop
3/26/19	*Office*	*School 5*		One way	13.90	Pick up laptop on the way to *School 5*
3/28/19	*School 6*	*Office*		One way	19.70	Return laptop & data entry

Note: RAs were provided a copy of the Google Maps distances for each school.

Following a closing out meeting, RAs shared that duplicating their hours and mileage both RME tracking sheets and Kelly timesheets seemed unnecessary. The RME team reviewed the purpose of the forms and then then determined that eliminating the mileage and hours tracking sheets was acceptable for future observation time points.

---

[2] Kelly Services are the providers of internal (within the university) and external temporary staff. They also facilitate the hiring process for temporary employees. RAs receive payments through Kelly Services.

Conclusions

This technical report describes the collection of observational data from a total of 41 middle school science teachers ranging from grades 6 through 8 in the Dallas Independent School District as part of the STEM Academy for Teachers and Leaders. Teachers participated in summer academies that were designed to emphasize four foundational pillars of STEM learning including (a) active learning or inquiry-based instruction, (b) the scientific process standards, (c) in-depth content knowledge, and (d) differentiated support for all learners. RAs collected data at two points in the school year, approximately October/November and February/March using the UTeach Observational Protocol (UTOP). The purpose of data collection was to provide additional data to support the research question: How do teachers progress in designing and implementing active learning opportunities in the classroom during the STEM Academy? The communication plan adjusted to participant needs as teachers had emergencies or school operational adjustments that warranted rescheduling. These challenges did not adversely affect the RAs' ability to collect data. Based on the acceptable inter-rater agreement statistics, the data collected can be analyzed to provide evidence of teachers' instructional practice as measured by the UTOP in future project reports and manuscripts.

Internal Report

References

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

Bybee, R. W. (2015). *The BSCS 5E instructional model: Creating teachable moments.* Thousand Oaks, CA: Corwin.

Kilpatrick, J., Swafford, J. & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics.* Washington, DC: National Academy Press.

Texas Education Agency (2013). *Texas Response to the Curriculum Focal Points for Kindergarten through Grade 8 Mathematics: Revised 2013.* Austin, TX: Author. Casabianca, J. M., Lockwood, J. R., & McCaffrey, D. F. (2014). Trends in classroom observation scores. *SAGE Journal, 75*(2), 311-337. doi:https://doi.org/10.1177/0013164414539163

Hill, H. C., & Charalambos, C. Y. (2012). When reliability is not enough: teacher observation systems and a case for generalizability study. *SAGE Journals, 41*(2), 56-64. doi:https://doi.org/10.3102/0013189X12437203

Research ethics and compliance training. (n.d.). Retrieved from https://citiprogram.org: https://about.citiprogram.org/en/mission-and-history

UTeach observation protocol. (n.d.). Retrieved from https://utop.uteach.utexas.edu: https://utop.uteach.utexas.edu/content/utop-mathematics-and-science-training-guide

Volunteer with Dallas ISD. (n.d.). Retrieved from https://www.dallasid.org: https://www.dallasisd.org/volunteer#calendar78778/20190528/month

Appendix A – Research Assistant Job Advertisement

Research in Mathematics Education - Southern Methodist University: Research in Mathematics Education (RME) is searching for motivated graduate students and/or former educational professionals within the Simmons School of Education or other schools on the SMU campus with specific interest in mathematics, mathematics education and/or STEM education. We are a research unit focused on supporting teachers and students in K-8 mathematics and STEM education. Read more at:
www.smu.edu/rme

About the project: Dr. Leanne Ketterlin-Geller, Professor in SMU's Simmons School of Education and Human Development, Department of Education Policy & Leadership, is the Director of the Research in Mathematics Education unit and is leading a research project: *STEM Academy for Middle Science Teachers and Leaders*. The purpose of the project is to determine impact of providing middle-school science teachers with intensive, professional development academies during the summer and on-going coaching throughout the school year. The project is a four-year partnership between Southern Methodist University (SMU) and Dallas ISD.

Position: Research Assistant

Location: Middle Schools within Dallas Independent School District

Job Summary: The Research Assistant will engage in activities such as participating in trainings to become a calibrated rater using a research-based science observational protocol, conduct observations using the protocol, and/or administer surveys in classrooms. The Research Assistant will assist with the preparation of physical materials, data entry procedures related to data collection, safeguarding the confidentiality of subjects, and other duties as necessary. The Research Assistant should expect to work between 10 to 20 hours a week during the month March 2019. Availability should be at least 4 days a week during school hours (8 am - 4 pm) Monday – Friday. Days and times will vary during these time periods. Flexibility is required.

Required Education and Experience (**Knowledge, skill, and ability**)**:**
- At least one full year of experience as a teacher (three or more years preferred)
- Access to a secure laptop computer with MS Word and email
- Access to reliable transportation to travel to Dallas ISD schools
- Flexible schedule with availability Monday through Thursday during the work/school day (full day availability preferred)
- Completion of previous CITI training preferred, but not required

Travel: Mileage will be reimbursed at the 2019 federal rate from SMU to data collection sites and back.

To Apply: Please send your cover letter and an updated resume or CV by January 31st, 2019 to Dr. Elizabeth Adams (stemresearch@smu.edu). **Priority consideration for applications on or before January 4th, 2019.**

Salary: Commensurate with experience
Contact: Elizabeth Adams (stemresearch@smu.edu) for additional information regarding this position (Benefits are not included)

Appendix B – Research Assistant Biographies

The qualifications of the RAs are as follows:

- One RA held a Master's degree in Education and a Bachelor's degree in Science. She has a Texas Educator Certificate in Science for grades 8-12. She is fluent in Spanish and Portuguese and has prior research experience as a Graduate Research Assistant. She is currently an Independent Education Consultant with Teach for America.
- Another RA held a Master's degree in Elementary Education with a Bachelor's degree in Elementary Education, and a Bachelor's degree in Biology. He has over 16 years as an educator within Grades first through 12th grade and a Technology Instructional Specialist. He is currently retired and volunteers as a chaperone in the Dallas Independent School District.
- Another RA held a Bachelor' degree in Science and Chemistry and is a certified Bilingual teacher. She was a Dyslexia Specialist with experience in planning instructional/learning strategies. She is currently is a Bilingual Data Collector.
- Another RA is currently completing his Master's degree in Bilingual Education at the Southern Methodist University. He also tutors Middle School Math and Science in the Garland ISD and is a certified Red Cross Lifeguard. He has prior research experience as a Data Collector.
- Another RA held a Master's degree in Education and a Bachelor's degree in Science. She has ten years of classroom experience as a math teacher and was a calibrated rater for TELPAS. She created the first computer science course for middle school in her district and has prior experience training teachers. She is currently an Edtech and Mathematics Education Consultant.
- Another RA held a Master's degree in Education with an emphasis on Curriculum and Instruction, a Bachelor's degree in Science, and possesses an instructional coach certificate She is TEI and T-TESS trained, has trained other trainers. She has over 35 years of classroom experience. Currently, she is an adjunct professor in the Deaf Education Department at the tertiary level.
- Another RA held a Bachelor's degree in Science Psychology. She has experience as a Data Collector and over 4 years' experience in the classroom. She is currently a substitute bilingual teacher.

Appendix C – Initial Observation/Lesson Plan Email to Teachers

Re: SMU STEM Academy Spring Observation

Hello _____ ,

An RME observer will be visiting your class on (date) during your (class period). The lesson plan should include:

- lesson objectives

- targeted TEKS

- materials

- instructional plan

- methods for monitoring student learning

Please upload your weekly/monthly lesson plan to Canvas prior to the day of observation.

This is 18% of your *Practicum for STEM Teachers* course grade. See Canvas link for submission. https://smu.instructure.com/courses/53385/assignments/189963

Please let us know if you have any questions,

The STEM Academy and Research Teams