2021 ASEE ANNUAL CONFERENCE
Virtual Meeting | July 26–29, 2021 | Pacific Daylight Time
Paper ID #33366

Review of In-class Active Learning Observation Protocols

Ms. Allison Van Beek, University of Toronto

Allison is a third year PhD candidate in the Department of Mechanical and Industrial Engineering. She's part of the collaborative Engineering Education program. Her supervisor is Prof. Susan McCahan.

Dr. Susan McCahan, University of Toronto

Susan McCahan is a Professor in the Department of Mechanical and Industrial Engineering at the University of Toronto. She currently holds the positions of Vice-Provost, Innovations in Undergraduate Education and Vice-Provost, Academic Programs. She received her B.S. (Mechanical Engineering) from Cornell University, and M.S. and Ph.D. (Mechanical Engineering) from Rensselaer Polytechnic Institute. She is a Fellow of the American Association for the Advancement of Science in recognition of contributions to engineering education has been the recipient of several major teaching and teaching leadership awards including the 3M National Teaching Fellowship and the Medal of Distinction in Engineering Education from Engineers Canada.

Review of In-Class Active Learning Observation Protocols

Abstract

This review paper examines the literature on classroom observation methods with an emphasis on observation protocols that are appropriate for active learning classrooms in engineering. Classroom observations have been used for professional development (e.g., formative feedback, evaluation of teaching) and administrative assessment (e.g., program evaluation). The focus of this work is to identify a protocol for collecting observation data to provide insight into active learning activity in STEM education and to inform design decisions for future active learning classroom space and technology design. With the emergence of purpose-built active learning classrooms, observations can capture active learning pedagogies and characterize the fit between teaching strategies and space affordances. This paper provides an overview of classroom observation protocols, and particularly those that were designed for active learning pedagogies. The review of these protocols identifies the advantages of each, and the aspects of the protocols that are suited to providing information on space design. Active classrooms typically include a physical layout that supports collaborative learning, and technology that supports interaction. To produce feedback on space design, a protocol should provide insight on the way STEM instructors makes use of both the physical layout and the technology to realize their teaching goals. We found that the existing protocols meet many, but not all of the requirements. We propose a hybrid protocol, that combines two existing frameworks, specifically aimed at providing information for active classroom design.

Introduction

Active learning, and spaces purpose-built for active learning pedagogies, are becoming increasingly prevalent in STEM fields, and engineering in particular. One method to learn more about the utilization of active learning pedagogies is to observe teaching and learning in real time using a protocol that describes the interaction between pedagogy, space, and technology [1]. The use of observation to provide information on teaching practices is a well documented concept [2], [3], [4]. Teaching and Learning Centers have a long history of offering classroom observations and these are frequently used to provide formative and summative feedback to instructors, teaching teams, and academic leaders [4]. Classroom observation can also play a role in program evaluation [1]. Observation also helps determine to what degree a pedagogy or intervention has been implemented [1]. Classroom observation can reveal subtleties of classroom contexts that are difficult to measure using other methodologies [3]. Observations of teaching allow us to learn more about the degree to which instructors are, or are not, fully utilizing the affordances of their classroom facilities as well as why incorporating specific teaching strategies, such as active learning, are not happening as hoped and expected [3].

Further, documenting which active learning strategies are being used in STEM learning environments could provide valuable insights for teaching centers, administrators, and classroom space designers. Studies typically compare traditional to active strategies but a comparison between implementations of active learning strategies will help researchers understand how the change in instructional techniques and space affordances is changing STEM education and what types of classroom space work best [5]. Turpen and Finkelstein [6] note that while teaching observation protocols are being adapted to new forms of pedagogy, including active learning techniques, not enough has been done to document education practices and what work has been done focuses too narrowly on certain parameters (e.g., reformed or not reformed, student centered or not student centered).

As more classrooms are renovated, or built, to support active learning, more needs to be known about what type of pedagogies are employed by instructors and teaching teams in these spaces and how teaching philosophies are manifested in the learning activities that are happening in these classrooms [5]. There are few existing protocols that speak to the relationship of space and technology with teaching approaches [4]. Hora [7] references Henderson & Dancy [8] and PCAST [9] to suggest that instructors in active learning classrooms (ALCs) are not incorporating active learning instructional practices as widely as expected [7]. Additionally, Birdwell et al. [4] report an increased frequency of faculty requests to support the transition of teaching from traditional classrooms to active learning classrooms. This literature, taken in aggregate, suggests that information gathered through classroom observation would contribute substantially to our understanding of space and technology needs.

To provide effective feedback for individual instructors, teaching teams, and administrators, we are undertaking a research project that explores the use of technology in active learning in purpose-designed active learning spaces through classroom observation. To achieve this goal, we are seeking a protocol that effectively captures the nuances of the interactions between the students, instructors, space, and technology in active STEM learning settings. An effective teaching observation protocol should capture holistically the complex teaching moves that are inherent in active learning activities. In our work, we are also interested in gathering information to inform the design of new active learning classrooms. In this paper, we review existing protocols and identify their advantages and disadvantages with this goal in mind. Although this review was undertaken for our study, the resulting review provides a roadmap for others who are considering the use of classroom observation in their research on active learning.

Activity Theory

We are seeking an observation protocol that captures patterns of behavior and connections holistically. Activity Theory provides a framework and will help us to systematically describe who is acting in the room, what their motivations and goals are, and how (or if) they achieve their goals. This theory suggests that the use, design, and evaluation of technology is socially co-constructed and is mediated by human interaction and communication [10]. Each class context is a unique socially and culturally complex environment. As instructors, teaching assistants, students, and staff – as a community of people - interact, they change as a group, and the system evolves over time [4]. Understanding this context begins when we see how different people, with their unique histories, informed by social norms and rules, externalize their goals and motives through tools (i.e., technology, language, objects, etc.).

These actions, and divisions of labor, can cause tensions and contradictions as people learn new norms and rules. Instructors and students redefine what it means to be an “instructor” or a “student” in an active learning classroom [11]. Instructors and students use the affordances of the space and technology to support their teaching and learning goals. If the major elements of the active learning classroom are considered only individually, the global view of the interactions is lost. The transformation of the roles and goals of the people participating in the activity are only observable when viewed as tool-mediated interactions, in context, as a whole [10]. Further, Kaptelinin and Nardi indicate that Activity Theory can be used descriptively to explore how people use technology [12]. Therefore, for an observation protocol to support the gathering of information on this interaction, it must be both detailed and identify the relationships between people, space and technology (i.e., holistic).

Introduction to Active Learning Ecosystem

Radcliffe et al. [13] developed the Pedagogy-Space-Technology (PST) Framework to explain how active learning classrooms (ALCs) embody a new learning ecosystem where these relationships are paramount. The PST Framework evaluates how the three elements (pedagogy, space, and technology) influence each other to ensure that each is considered strategically (i.e., during the design and evaluation phases of projects), (see Fig. 1). Elements of pedagogy, technology and space not only intersect but are deeply intertwined in these classrooms. For example, the affordances of a learning environment will shape what pedagogies can be employed by teaching teams. The ecosystem is presented as a nexus, not as a hierarchy with one element taking precedence over another. Radcliffe et al. [13] designed the PST Framework to evaluate existing active learning spaces at their home institution; we use their articulation of the three major elements of an active learning space to operationalize the theoretical framework (i.e., Activity Theory), that underpins this research project.

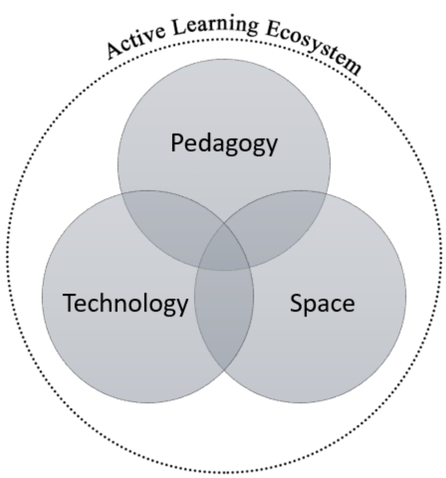

Figure 1. Active Learning Ecosystem diagram, adapted from [14]

Review of existing active learning classroom observation protocols and tools

A review of existing classroom observation protocols was undertaken to ascertain if there is an existing protocol designed for gathering information on the three elements of the active learning ecosystem and their interaction. Alternatively, is there any existing classroom observation protocol could be extended to use within an active learning space to holistically capture these interactions? In 2013, Hora and Ferrare [3] published a review of classroom observation techniques in postsecondary settings. Using their work as a baseline, we started by identifying the characteristics of a protocol that would meet the requirements for our project. We identified four primary attributes required for a protocol suitable for collecting information regarding space design for active engineering classrooms.

1. The protocol should be descriptive, not evaluative. This distinction is the major determining factor of whether a protocol would be effective for this purpose. Many observation protocols are meant to evaluate the quality of teaching, rather than simply provide a description of teaching moves [15], sometimes referred to as Teacher Discourse Moves (TDMs). Evaluative protocols tend to require subjectivity and inference and work well in situations where observations are completed by peers, versus external observers [4]. Evaluative protocols are often unstructured and reflective, which does not provide a standardized base for comparison or aggregation of data between class sessions or courses that we are seeking [3].
2. The protocol should be pedagogically agnostic, not specific. We are interested in capturing the full spectrum of teaching activity. Some protocols are designed to observe a specific pedagogy in practice. The goal of this project is not to research the effectiveness of one particular pedagogy, it is to research the effectiveness of the space and tools designed to enable the instructor's teaching plan. If a protocol is too heavily based on observing an expected pedagogy, important observations might go untracked. Active learning classrooms should support a variety of teaching perspectives and pedagogies and be designed in a way that encourages the use of a broad spectrum of learning activities, so the instructor has the flexibility to pick the best teaching approach for their intended goals. Also, if the protocol is not holistic, we risk capturing specific elements of the teaching practice, but not the full context of the relationship between elements. We are seeking a protocol that has this breadth.
3. The protocol captures quantitative data. While no observation protocol is truly objective, our focus is on observable behaviors that can be quantified (e.g., duration, frequency, etc.). The researchers decided to pursue observable behavior (e.g., physical movement around the room, lecturing with demonstration equipment, dividing a large class into smaller break out activities, etc.) versus focusing on observations like "student engagement", which is difficult to quantify and, for context, typically relies on a complementary survey or interview with the learners to map the observer's perceptions with the learner's perceptions. Research shows that an observer's pre-existing beliefs about high-quality teaching can lead to high variability in ratings [16]. Protocols that focus on observable behavior also increase inter-rater reliability [16] while decreasing rater bias [17].
4. The protocol is designed for use within higher education STEM disciplines. While it is possible to use protocols that were designed for other domains, it would be advantageous to use an existing protocol validated within STEM education, and engineering specifically, given that this is the context of interest. While much teaching activity is

universal, a protocol that specifically captures behaviors typically used by STEM instructors (e.g., working out problems, modeling, design activities, etc.) would advantageous. We also consider the possibility of modifying a protocol designed for K-12 education. Wainwright, Flick, & Morrell [18] argue that adopting classroom observation tools developed for K–12 settings is not recommended because the types of student-teacher interactions observed in K–12 classrooms differ in nature to those observed in higher education settings. Nonetheless, we have included K-12 observation protocols in this review because they are widely used and often have been rigorously validated.

After creating this informal checklist, we completed a search for observation protocols in two stages. First, we performed a search in relevant academic databases, including ERIC and Compendex. Then, to provide some breadth, the search terms were expanded, and a broader search of the literature was conducted:

- Classroom Observation Protocols
- Classroom Observation Tools
- Classroom Observation Instruments
- Classroom Observation Methods
- Active Learning in STEM
- Instrument Development

During the search process, it was challenging to separate the publications detailing the development of the instrument versus articles describing the use of the instrument using a simple algorithm, so this sorting process was manual. Protocols were not included if they were not well-grounded and tested. The strength of classroom observation data relies on the psychometric soundness of the protocol. In addition to this search process, the usual approach to a literature review was undertaken (i.e., identifying sources that reference key papers in the field).

In conducting the search of the literature, we considered 15 protocols designed and tested in both K-12 and higher education contexts. A full compilation of the protocols is shown in Appendix A, which documents the alignment between each protocol and the identified attributes.

Review of existing Classroom Observation Protocols

In conducting the search of the literature, we considered protocols designed and tested in both K-12 and higher education contexts. A full compilation of the protocols is shown in Appendix A, which documents the alignment between each protocol and the identified attributes.

Observation protocols that do not specifically mention active learning or STEM observation and are evaluative were not considered in depth. There are some widely used protocols that did not meet these requirements: The Teaching Behaviors Inventory (TBI) [19], the Classroom Assessment Scoring System (CLASS) [20] and the Framework for Teaching [21]. Complementary techniques, such as teaching inventories, are also not reviewed in this paper, as their process typically involves self-assessment by the instructor (pre and/or post observation).

Evaluative Protocols

Several prominent protocols are evaluative, not descriptive. While we are not seeking an evaluative protocol, a review of classroom observation tools would not be complete without reviewing these established protocols. One example is the well-known Reformed Teaching Observation Protocol (RTOP), developed in 2000 [2]. This summative evaluation tool focuses on lesson design and implementation, content, and classroom culture and captures modern teaching practice, specifically the current move towards teaching practices grounded in constructivism. The RTOP scores how well instructors implement and adhere to reformed teaching practice. It was developed by the Arizona Collaborative for Excellence in the Preparation of Teachers (ACEPT) to improve the preparation of math and science teachers in K-12 education (specifically pre-service teachers). While designed for K-12, it has also been employed in higher education. The protocol itself is composed of 25 questions, divided into 3 sections: Lesson Design and Implementation, Content (which is further sub-divided into propositional knowledge and procedural knowledge), and Classroom Culture (sub-divided into communicative interactions and student-instructor relationships). Each item is scored on a scale (0 to 4) from “Never Occurred” to “Very Descriptive.”

The culmination of the observation process is an RTOP score that can be used for comparison, identifying firmly that the protocol is evaluative versus descriptive [22]. Research is showing that an observer’s pre-existing beliefs about high-quality teaching can lead to high variability in ratings [15]. In RTOP there is no tracking for frequency or duration of each activity nor of the tools used to enable that activity. Summative observations do not allow for analyzing the transformation of activity over time, which is key to observing the classroom from an Activity Theory perspective. Other researchers have reported that the RTOP includes observational judgements that can be awkward to share as feedback (e.g., “There was a climate of respect for what others had to say.”)

Although RTOP is widely used, we found that it did not meet two major attributes; it is evaluative and focuses on a specific pedagogy (reformed teaching). The Likert scale rating collapses complex activity into generalized feedback [5], removing the ability to analyze how elements (people, tools, goals, etc.) in an active learning classroom interact and evolve.

The RTOP has spawned several other protocols, such as the OCEPT Classroom Observation Protocol (O-TOP) [18], which also evaluates reformed-teaching practice. Like the RTOP, the O-TOP uses a Likert scale to evaluate the effectiveness of professional development training for new teachers and instructors. In 2012, Dringenberg et al. [23] developed the Science and Engineering Classroom Learning Observation Protocol (SEcLO), modeled on the RTOP. This protocol links STEM learning outcomes to teaching practice. Observers using SEcLO mark exemplary behaviors and vocabulary on an observation sheet, followed by a summary to provide scoring and additional details. The SEcLO, like the RTOP, is evaluative, not descriptive. Classroom observation tools are often based on existing protocols and then modified for specific research purposes [23]. We noted this practice frequently in the course of conducting our review.

The Global Real-time Assessment Tool for Teaching Enhancement (G-RATE; formerly VaNTH Observation System (VOS), released in 1999) [24] was designed for use within STEM higher education. This protocol is aligned with the How People Learn (HPL) framework [25], observing specific learner-centered, assessment-centered, knowledge-centered, and community-centered pedagogical dimensions (as well as classroom organization) [24]. G-RATE was developed to provide multidimensional feedback to graduate teaching assistants. This protocol is the only protocol in this review that focuses on providing feedback to graduate teaching assistants, especially those working within active learning environments where their role is expanded to technical support, liaisons, and collaborators. A second notable and forward-thinking characteristic is that the original observation protocol, VOS, was built as an application and not a paper-based instrument. VOS was updated to G-RATE to reduce observer training time. G-RATE includes a built-in list of codes (with built-in support) and allows for customization [24].

As the use of technology in classrooms has become more pervasive, classroom observation protocols have been developed to capture this aspect of the learning environment. In 2014, The International Society for Technology in Education (ISTE) published the Classroom Observation Tool (ICOT) [1] which is a protocol that focuses on student interaction with technology during learning activities. The ICOT captures seven attributes of the learning environment, including technology integration, student groupings, teacher roles, learning activities and technologies used. This protocol was of particular interest as it represents a different focus versus the other protocols in this review because it emphasizes tool adoption as the central focus of the observation activity. However, while the focus on technology use aligns with the goals of our work, the protocol is evaluative. Observers are asked to rate the use of the technology based on how essential it was to the learning activity. Unfortunately, the protocol does not holistically describe the full scope of teaching moves. While it captures the use of classroom technology, it does not capture all, nor does it consider the context of the other two elements of the active learning ecosystem: space and pedagogy.

Birdwell et al. [4] mention the ICOT in their review and conclude that there is “an absence of an observation protocol explicitly designed to address instructional approaches within active learning classrooms.” Birdwell et al. designed the Active Learning Classroom Observation Protocol (ALCOT) to fill this gap. The ALCOT provides developmental feedback to instructors on their use of active learning spaces, specifically including space and technology utilization. The authors use a prescriptive approach, designing observation categories based on what should be happening in an ideal active learning classroom.

The protocol is reflective and was designed in response to the increase in faculty members approaching the Indiana University, Bloomington, Centre for Teaching and Learning for assistance to re-design and re-think their courses [4]. It is described as a holistic protocol, considering the entirety of the learning experience. The authors of the protocol identify three areas for reflection elicited during the observation: awareness of the interconnected relationships between space, technology, and pedagogy, classroom management, and the instructor role. The protocol also has an aspirational element, incorporating categories that should be observed in effective use of an active learning classroom [4]. The result is a protocol composed of four categories: support of active learning, creation and implementation of student collaborative learning activities, formative assessment in the classroom, and classroom management [4].

In addition to the observation tool, which guides reflection at the end of an observation, there is also a chronological note-taking instrument for real-time tracking. This is largely informal and is intended to provide qualitative feedback. The open-ended nature of the protocol, while likely invaluable as individual feedback, does not easily afford quantitative data collection. As such, the ALCOT meets some of the attributes we are seeking, in that it accounts for technology usage and it is aligned with active learning.

The evaluative protocols reviewed here are designed to provide formative feedback to the instructor; not for formative feedback for space design or other purposes [4]. However, we see value in reviewing the types of questions the protocols included as this could identify behaviours missing from the descriptive protocols. This is specifically true for the ALCOT. For example, aspects of the ALCOT contribute valuable dimensions that are absent in many of the descriptive protocols.

Descriptive Protocols

Descriptive classroom observation protocols are designed to document activity in a learning environment objectively. This aligns with our goal of collecting data that can be analyzed through an Activity Theory lens to inform instructional space and technology design [26]. We are interested in analyzing what is happening in active learning classrooms, versus evaluating what is happening. As mentioned previously, an advantage of a descriptive protocol is the ability to compare and contrast information about the frequency and duration of teaching moves across class sessions and courses. This is an important feature for the design of classrooms, where the space and technology must work effectively for a range of instructional styles and teaching teams.

The Teaching Dimensions Observation Protocol (TDOP) [7], captures the multi-dimensional nature of classroom activity, reflecting a complex environment consisting of multiple actors and artifacts. This descriptive observation system stresses the importance of studying educational settings at a granular level and encourages researchers to customize the protocol for their specific research goals, modifying as necessary the 46 starting codes used to document behavior (in two-minute intervals) [7]. The TDOP is divided into five aspects of classroom dynamics: teaching methods, pedagogical strategies, types of student-instructor interactions, types of cognitive engagement, and use of instructional technology. Each category has specific codes to track observable behaviors. After observation, these codes can be combined to represent pedagogical constructs, if and as needed. By combining TDOP codes and analyzing them through the Differentiated Overt Learning Activities (DOLA) [27] framework, researchers have explored the teaching practices employed in active learning classrooms and the prevalence and nature of active learning [7]. The focus on task performance, classroom norms, and the level of granularity of data collection align with the attributes we have identified. The TDOP does not include value judgements on the effectiveness of instructional practice nor does it seek to comment on alignment with specific pedagogies [22].

The Classroom Observation Protocol for Undergraduate STEM (COPUS) [4] is a simplified version of the Teaching Dimensions Observation Protocol (TDOP) and it is widely used in STEM fields. COPUS is designed to allow faculty members to observe how active learning classrooms are used by teaching teams and students. Knowing that instructors are unlikely to dedicate days to internalizing a protocol, the COPUS authors' developed the protocol to decrease the training needed without sacrificing the integrity of the instrument. More intensive training is necessary for more complex protocols such as the RTOP (to ensure inter-rate reliability) and the TDOP (which suggests three days to use the TDOP effectively). COPUS reduced the training requirement to 1.5 hours, allowing STEM faculty members to act as peer observers [22].

As a descriptive protocol, COPUS does not measure the efficacy of teaching practice nor does it provide feedback to improve teaching per se [4]. COPUS, like TDOP, measures activity in two-minute increments, and is limited to 25 codes divided into two categories: "What the students are doing" and "what the instructor is doing" [22]. After the observation, instructors are given a pie chart with a breakdown of their class orchestration, allowing them to review their patterns in a non-judgemental way. While COPUS reduces the overall dimensions for coding, it expands information on student behaviors. COPUS does not have some of the key dimensions of observation that are of interest in space and educational technology design, such as coding for instructional technology and other classroom tool use [22].

Many observation protocols, such as COPUS, concentrate on instructor teaching behavior, rather than tool use per se, which the authors posit is not necessary for faculty observation programs [22]. Excluding the use of instructional artifacts misses tool mediation, a key tenet of Activity Theory and an essential point in our work. It is critical to capture when (and what) technologies are used, if they successfully support the intent of the activity, and what the ramifications of this use are because tools shape the way people interact with their reality; knowing how to use a tool demonstrates transmission of social and cultural knowledge [12].

A more recent protocol, the Classroom Discourse Observation Protocol (CDOP) [5], developed in 2019, references the RTOP, TDOP, and COPUS, and distinguishes itself from those protocols by focusing specifically on quantifying the nature of dialogue techniques. This protocol seeks to quantify teacher discourse moves (TDMs) in undergraduate STEM learning environments by using COPUS coding to quantify teaching activity and to further extrapolate the nature of TDMs [5]. TDMs are defined by the authors as techniques that mediate classroom discussions, and the authors acknowledge that one of the reasons this area is under studied is the difficulty of validating the data in a reliable manner. There are 15 codes in this protocol. Five are instructor-centric: sharing, real-worlding, evaluating, linking, and forecasting. The other ten codes are student-centric: generative, checking in, clarifying, connecting, contextualizing, representing, constructing, requesting, explaining, and challenging [5]. The underlying principles of this protocol align with the theoretical framework of our project but focusing specifically of teacher discourse moves narrows the observation space and therefore limits the collection of data on instructor-space-technology interactions. The authors acknowledge that their protocol is designed to be used in conjunction with, and to extend, other protocols (e.g., RTOP, TDOP, COPUS) and they used COPUS to analyze data collected during their observations. Their stated future work includes using the DOLA framework to categorize TDMs for the purpose of exploring different levels of cognitive engagement.

The authors of the TDOP, COPUS, and CDOP use (or plan to use) the DOLA framework to analyze their data. This is possible for protocols that emphasize tracking observed behaviors and then analyzing the underlying processes associated with these behaviors (e.g., asking questions could be linked to active student engagement; open ended problem solving with peers could be linked to constructive student engagement, etc.) [7]. DOLA was first described by Chi in 2009. In more recent publications, Chi and Wylie [28] refer to the taxonomy as the Interactive, Constructivist, Active Passive (ICAP) Framework. This framework can be used for classroom observations directly or to interpret information collected using other protocols. The ICAP Framework is based on the ICAP hypothesis, which states that students learn more from interactive learning activities than passive activities. It is designed to describe modes of cognitive engagement, based on categorizing student interactions during class sessions. Used extensively in STEM fields, ICAP is a taxonomy that describes and evaluates the effectiveness of active learning activities used in Engineering Education [29]. As a protocol, ICAP is descriptive. However, it does not meet our needs as it focuses on student activity only. However, we see potential in the use of the DOLA taxonomy as an analytic lens for our work in the future.

Recent STEM-focused, descriptive classroom observation protocols have built on the RTOP, TDOP, and COPUS. Typically, these protocols focus on one specific element of pedagogical design. The three protocols we examine next were all designed for use in the K-12 environment and have not been validated in higher education. The Inquiring into Science Instruction Observation Protocol (ISIOP) [30] aligns with some of our identified needs. The authors attempt to capture instruction as a complex, dynamic, and creative interaction between teacher and student. However, the teacher-student interaction is only one dynamic interaction that happens within an active learning classroom and the lengthy (70-75 items) protocol does not capture the interaction of tools and space [30]. The Design-based Science Teaching Observation Protocol (EDSTOP) [31], released in 2012, is an adapted version of the ISIOP and was designed to specifically measure engineering design vocabulary and behaviors in K-12 science teaching using discourse analysis. This protocol focuses on the instructor, including tracking questions and responses, transitions from one activity to another, physical movement and the set up of the classroom space, which is appealing. However, there is no provision for the coding of tool use in this protocol [31]. Subsequently, the Classroom Observation Protocol for Engineering Design (COPED) was designed to evaluate engineering design curriculum integration in K-12 classrooms [32]. The authors focus their protocol on emphasizing engineering design processes and habits of mind. The COPED is an incremental protocol designed to observe one aspect of engineering education in K-12 classes. Wheeler [32] states that other protocols (e.g., ALCOT and COPUS) are not specific enough to capture characteristics of engineering education. Both the SEcLO and EDSTOP protocols are descriptive, but they focus on observing one specific pedagogy in practice.

In addition to these protocols which we reviewed in depth, we also examined 3 protocols that were excluded due to a major misalignment between the goals of the protocol and the goals of the research project. These include the Classroom Observation Protocol (COP), which was referenced by SEcLO but we could not find a published article regarding the protocol; the Classroom Observation Rubric [6], which describes solely the interaction during peer instruction (but is one of the only protocols that captures tool use, e.g., the use of clicker); the UTEACH Observation Protocol (UTOP) [33], which evaluates new K-12 teachers on the quality of instruction (measured by criterion drawn from the UTEACH program).

Discussion

We identified four attributes necessary for an observation protocol to be used to inform active classroom space and technology design. To align with these attributes, the protocol should be descriptive, pedagogically agnostic, and capture the active learning ecosystem holistically. Ideally, this protocol would be designed for STEM higher education contexts, allowing quantitative comparisons and aggregation of data across teaching teams, courses, and terms.

Protocols that are evaluative do not align well with these attributes. This removes several protocols (e.g., RTOP, O-TOP, ICOT, ALCOT) from consideration. However, there are aspects of these protocols that are intriguing. For example, the ICOT was of particular interest because its purpose is to record student use of technology within the learning environment, focusing specifically on why technology is used so differently across different people and classrooms [1]. In comparing the ICOT to the TDOP, a descriptive protocol, we found that the aspects of technology use captured in the ICOT were equally or better captured by the TDOP.

Descriptive protocols that focus on a specific pedagogy (i.e., CDOP, ISIOP, EDSTOP, COPED) focus only on one actor group (i.e., SEcLO, G-RATE, ICAP), or otherwise too narrowly focus on one aspect of the active learning ecosystem (i.e., COR) do not fully align with the identified attributes. In addition, most protocols (with the exception of the ICOT, ALCOT, TDOP and technically the COR, which tracks clicker use) exclude entirely, or describe minimally, the use of technology. These protocols are extremely successful in documenting observed teaching behaviours, but they do not capture the active learning ecosystem holistically.

To collect information on space and technology design the selected protocol should allow us to capture the interaction of actors in the space, with a focus on instructors and how their teaching is enabled or disabled by the active learning ecosystem. Protocols that focus specifically on the students (e.g., SEcLO, ICAP) or teaching assistants (e.g., G-RATE) do not meet the needs for this project. Neither do protocols that only evaluate or describe one specific type of teaching strategy (e.g., CDOP, ISIOP, EDSTOP, COPED). However, in other projects, such as research focused on student-centered experiences in active learning, one of these observation protocols may be a good fit.

There are challenges and limitations of using classroom observation protocols to describe teaching and learning activity. In addition to determining the fit of the protocol with the identified attributes, there are several other considerations that are important when selecting an observation framework. These considerations include: the intended audience for the resulting information (e.g., the instructor, or administrators, etc.), training requirements, and validation and testing (i.e., are the psychometric constructs of the protocol valid?). These considerations will guide the selection of a particular protocol for a given research project. In our specific project, we are looking for a protocol that allows for comparison of our data with other published research on active learning and allows for compilation of data across courses. We are less concerned with the training requirements, but in other projects this may be an overriding concern.

Capturing classroom dynamics in a comprehensive, granular, and rigorous manner necessitates a detailed classroom observation tool. The two tools that provide this level of granularity are the TDOP and the ISIOP. These protocols require days of training and are not designed for an evaluative peer observation so they may not be suitable for a project that relies on volunteers. However, for projects where the researcher is the primary observer, these two protocols offer the potential for deeper information collection. The ISIOP is designed for K-12 and while it tracks teacher moves, unfortunately, it does not incorporate a component to technology use.

The protocols in this review are all validated (except SEcLO, which was not explicit about this information) and have had their psychometric constructs evaluated by external experts to ensure construct validity (e.g., COPED). Demonstrating validity (i.e., how well the instrument captures or measures the domain of activity) and reliability (i.e., how consistently the activity is measured, regardless of who is collecting the data) for classroom observation protocols is an area that requires more external investigation [3].

A useful add on to any protocol is the DOLA framework for interpreting the resulting data. This framework has been used with observation protocols (e.g., TDOP, ICAP) after classroom observation is complete. DOLA supports exploration of patterns in teaching behavior and references them to other metrics related to instructor (e.g., teaching evaluation scores) or student (e.g., grades) success. As the use of classroom observation tools becomes more common, careful consideration will have to be given as to how the data are analyzed and benchmarked. If these results are used to support institutional change, it could influence the entire teaching and learning activity system [3].

After reviewing these 15 observation protocols, we found that one protocol, the Teaching Dimensions Observation Protocol (TDOP) most closely aligned to our research goals. The TDOP, which uses a uses a socio-cultural theoretical framework, tracks all the dimensions of a learning ecosystem. However, while this protocol met most of the attributes, we identified gaps in TDOP because it was not designed explicitly for active learning environments. To fill this gap, we suggest borrowing codes from a protocol that has a strong foundation in active learning such as the ALCOT. The TDOP was designed to be modified to fit the context of interest. Therefore, TDOP offers a more flexible protocol that can be tailored to the needs of the researcher [3].

The ALCOT can be used to extend the TDOP to encompass active learning strategies by broadening the scope of observed activity to include three additional areas for observation: awareness of the interconnected relationships between space, technology, and pedagogy; classroom management; and the role of the instructor [4]. Some of these reflective items prescribe “best practice” activity in that they have been identified through other research as being behaviors that exist in successful active learning classrooms [4]. To incorporate them into TDOP would require reconceptualizing them as codes that describe observable behaviors. Merging the Teaching Dimensions Observation Protocol (TDOP) and the Active Learning Classroom Observation Tool (ALCOT) creates one protocol that meets all the attributes that we defined at the outset of this review. It has the potential to provide feedback on each element of the active learning ecosystem, and interactions between these elements, for use by active learning space designers.

Conclusion

Classroom observation offers an approach to collect information that could be highly valuable to STEM active learning classroom and technology designers. Many protocols exist but collecting this data is a challenge and institutions wrestle with how to operationalize this information to foster pedagogical transformation [22]. Protocols are typically evaluative or descriptive in nature. They tend to focus on either pedagogy broadly or one specific strategy. The data they yield is often both quantitative and qualitative. Evaluative protocols are typically used as mechanisms for formative feedback or to measure program implementation. In contrast, descriptive protocols do not provide judgement regarding teaching quality or effectiveness, but rather capture teaching behaviour as data for analysis [3].

Given the increased prevalence of active learning classrooms, there is a need for a descriptive classroom observation tool that can be used to capture the elements of teaching practice as well as providing insight into the intersection of pedagogy, space, and technology (i.e., the active learning ecosystem) holistically. To offer evidence-based recommendations for design, more needs to be known about what specific type of pedagogies are employed, how they manifested in practice, and how classroom spaces and technologies enable or hinder these practices. As we learn more about what type of activity happens in practice through increased understanding about different contexts of active learning classrooms, we can improve the effectiveness and usability of new active learning classrooms, as well as other formal and informal learning environments.

After completing a comprehensive review of existing protocols, we concluded that the TDOP, modified using dimensions from the ALCOT, will potentially fit the identified attributes. The TDOP is an established descriptive classroom observation protocol and enhancing it using the ALCOT adds missing dimensions relating to active learning best practices. However, other engineering education researchers, for example those interested in the student experience of active learning, may find one of the other protocols described in this paper better aligns with their research goals.

References

[1] T. Bielefeldt, "Guidance for Technology Decisions from Classroom Observation," *Journal of Research on Technology in Education,* vol. 44, *(3),* pp. 205-223, 2012. Available: https://files.eric.ed.gov/fulltext/EJ976466.pdf.

[2] M. Piburn et al, "Reformed teaching observation protocol (RTOP) reference manual. ACEPT technical reportno. IN00-3." Arizona Board of Regents, Tempe, AZ, 2000.

[3] M. Hora and J. Ferrare, "A Review of Classroom Observation Techniques in Postsecondary Settings," WCER Working Paper no. 2013-1, 2013.

[4] T. Birdwell et al, "Active Learning Classroom Observation Tool: A Practical Tool for Classroom Observation and Instructor Reflection in Active Learning Classrooms," Journal on Centers for Teaching and Learning, vol. 8, pp. 28-50, 2016.

[5] P. Kranzfelder et al, "The Classroom Discourse Observation Protocol (CDOP): A quantitative method for characterizing teacher discourse moves in undergraduate STEM learning environments," PloS One, vol. 14, (7), 2019. Available: https://doi.org/10.1371/journal.pone.0219019. DOI: 10.1371/journal.pone.0219019.

[6] C. Turpen and N. D. Finkelstein, "Not all interactive engagement is the same: Variations in physics professors' implementation of Peer Instruction," Physical Review Special Topics. Physics Education Research, vol. 5, (2), 2009. . DOI: 10.1103/PhysRevSTPER.5.020101.

[7] M. T. Hora, "Toward a Descriptive Science of Teaching: How the TDOP Illuminates the Multidimensional Nature of Active Learning in Postsecondary Classrooms," Science Education, vol. 99, (5), pp. 783-818, 2015.

[8] C. Henderson and M. H. Dancy, "Physics faculty and educational researchers: Divergent expectations as barriers to the diffusion of innovations," American Journal of Physic, vol. 76, (1), pp. 79-91, 2008.

[9] President's Council of Advisors on Science and Technology, "Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering and mathematics," Executive Office of the President, Washington, D.C., 2012.

[10] G. Gay and H. Hembrooke, Activity-Centered Design: An Ecological Approach to Designing Smart Tools and Usable Systems. Cambridge, MA: MIT Press, 2004.

[11] Y. Engeström, "From Learning Environments and Implementation to Activity Systems and Expansive Learning," Actio: An International Journal of Human Activity Theory, vol. 2, pp. 17-33, 2009.

[12] V. Kaptelinin, B. Nardi and C. Macauley, "Methods and Tools: The Activity Checklist," Interactions, vol. 6, (4), pp. 27-39, 1999. Available: http://interactions.acm.org/archive/view/july-aug.-1999/methods-tools-the-activity-checklist1.

[13] D. Radcliffe et al, "Designing next generation places of learning: Collaboration

at the pedagogy-space-technology nexus: ALTC priority project #627 ," Australian Learning & Teaching Council, Australia, 2008.

[14] Steelcase Education, "Active learning center: Grant proposal guide," Steelcase Education, 2015.

[15] L. B. Nilson, Teaching at its Best: A Research-Based Resource for College Educators. Third Edition. San Francisco, CA: Jossey-Bass, 2010.

[16] C. Guarino and B. Stacy, "Review of gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains," National Educational Policy Center, Boulder, CO, 2012.

[17] A. H. Cash et al, "Rater calibration when observational assessment occurs at large scale: Degree of calibration and characteristics of raters associated with calibration," Early Childhood Research Quarterly, vol. 27, (3), pp. 529-542, 2012. . DOI: 10.1016/j.ecresq.2011.12.006.

[18] C. L. Wainwright, L. Flick and P. Morrell, "The Development of Instruments for Assessment of Instructional Practices in Standards-Based Teaching," 2003. Available: https://scholarscompass.vcu.edu/jmsce_vamsc/vol6/iss1/6.

[19] H. G. Murray, "Effective teaching behavior in the college classroom," in Effective Teaching in Higher Education: Research and Practice, R. P. Perry and J. C. Smart, Eds. New York, N.Y.: Agathon Press, 1997, pp. 171-204.

[20] R. C. Pianta, K. M. La Paro and B. K. Hamre, Classroom Assessment Scoring System™: Manual K-3. 2008.

[21] C. Danielson, "Enhancing professional practice: A framework for teaching," 2007.

[22] M. K. Smith et al, "The Classroom Observation Protocol for Undergraduate STEM (COPUS): A New Instrument to Characterize University STEM Classroom Practices," CBE Life Sci Educ., vol. 12, (4), pp. 618–627, 2013. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3846513/.

[23] E. Dringenberg and R. E. H. Wertz, AC 2012-3324: Development of the Science and Engineering Classroom Learning Observation Protocol (SEcLO). 2012.

[24] M. Cox et al, "Developing A global real time assessment tool for the teaching enhancement of engineering graduate teaching assistants," in ASEE Annual Conference and Exposition, June 2010.

[25] National Research Council and Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds), How People Learn: Brain, Mind, Experience, and School: Expanded Edition. Washington, D.C.: The National Academies Press, 2000.

[26] B. Nardi, Context and Consciousness: Activity Theory and Human-Computer Interaction. Cambridge, Mass.: MIT Press, 1996.

[27] M. T. H. Chi, "Active-Constructive-Interactive: A Conceptual Framework for Differentiating Learning Activities," Topics in Cognitive Science, vol. 1, (1), pp. 73-105, 2009. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1756-8765.2008.01005.x. DOI: 10.1111/j.1756-8765.2008.01005.x.

[28] M. T. Chi and R. Wylie, "The ICAP framework: Linking cognitive engagement to active learning outcomes," Educational Psychologist, vol. 49, (4), pp. 219-243, 2014.

[29] M. Menekse et al, "Differentiated Overt Learning Activities for Effective Instruction in Engineering Classrooms," Journal of Engineering Education (Washington, D.C.), vol. 102, (3), pp. 346-374, 2013. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/jee.20021. DOI: 10.1002/jee.20021.

[30] D. Minner and J. DeLisi, "Inquiring into Science Instruction Observation Protocol (ISIOP). User's Manual." 2012.

[31] B. M. Capobianco, J. DeLisi and J. Radloff, "Characterizing elementary teachers' enactment of high-leverage practices through engineering design-based science instruction," Science Education (Salem, Mass.), vol. 102, (2), pp. 342-376, 2018. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/sce.21325. DOI: 10.1002/sce.21325.

[32] L. B. Wheeler et al, "Development and validation of the Classroom Observation Protocol for Engineering Design (COPED)," Journal of Research in Science Teaching, vol. 56, (9), pp. 1285-1305, 2019. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/tea.21557. DOI: 10.1002/tea.21557.

[33] C. Walkington et al, Development of the UTeach Observation Protocol: A Classroom Observation Instrument to Evaluate Mathematics and Science Teachers from the UTeach Preparation Program. 2011.

Appendix A: Protocol Comparison Table

This comparison table lists the protocols reviewed. The table illustrates the alignment between each protocol and the identified attributes of interest.

Column Legend

Protocol Name: Full name and acronym
Type of Protocol: Protocols tend to be Descriptive (D) or Evaluative (E)
Type of Pedagogy: Protocols were identified as being pedagogically agnostic (A) or specific (S).
Type of Data: Protocols can collect quantitative data, qualitative data, or a combination of both.
STEM: Was the protocol designed for use within STEM?
Education Level: Protocols were either developed for K-12 or Higher Education.
Type of Actor: Who does the observation focus on? Students (S), instructors (I), teaching assistants (TA), or a combination (I/S/…)?
Type of Observer: Who will be doing the observations? Peer (P) or Researcher (R)

Protocol	Overview	Type of Protocol	Type of Peda.	Type of Data	STEM ?	Ed Level	Type of Actor	Type of Observer
Reformed Teaching Observation Protocol (RTOP) [2]	Implementation of Reformed (constructivist) Teaching Practice principle s	E	S	QN	Y	K-12	I&S	P or R
OCEPT Classroom Observation Protocol (O-TOP) [18]	Document the impact of reform-based professional development and its impact of teachers	E	S	QN	Y	K-12 or HE	I&S	R
Science and Engineering Classroom Learning Observation Protocol (SEcLO) [23]	Tracks Engineering design process vocabulary and behaviors (Scientific vocabulary and behaviors, Degree of frustration and understanding, gender difference)	E	A	QN	Y	K-12	S	R

Global Real-time Assessment Tool for Teaching Enhancement (G-RATE) [24]	Pedagogical Practice; Classroom Interactions; feedback to GTAs about the extent to which they are fulfilling their teaching roles is needed.	E	S	QN	Y	HE	TA	R
ISTE Classroom Observation Tool (ICOT) [1]	Determine whether information and communications technologies (ICT) were being integrated into instruction; emphasis was 21st century skills	E	A	QN	N	K-12	I&S	R
Active Learning Classroom Observation Tool (ALCOT) [4]	Holistic Review of Learning Experience; used in classrooms with dissimilar design, the ALCOT had to be inclusive of variations of space, furniture, and technologies	E	A	QL	N	HE	I	P or R
Teaching Dimensions Observation Protocol (TDOP) [7]	Key Aspects of Classroom Dynamics; an approach to thinking about teaching in science classrooms in more multidimensional terms	D	A	QN	Y	HE	I&S	R

Classroom Observation Protocol for Under-graduate STEM (COPUS) [4]	Range and frequency of teaching practice; there is increasing interest in collecting information on the range and frequency of teaching practices at department-wide and institution-wide scales.	D	A	QN	Y	HE	I&S	P
Classroom Discourse Observation Protocol (CDOP) [5]	Teacher Discourse Moves from observational data; focuses on dialogical techniques	Descriptive	S	QN	Y	HE	I&S	R
Interactive, Construct-ivist, Active, Passive (ICAP) Framework [28]	Defines cognitive engagement activities on the basis of students' overt behaviors and proposes that engagement behaviors can be categorized and differentiated into one of four modes: Interactive, Constructive, Active, and Passive.	D	A	QN	Y	HE	S	P or R
Inquiry Science Instruction Observation Protocol (ISIOP) [30]	developed to characterize the nature of inquiry-based instruction in science classrooms; ISIOP is designed to	D&E	A	QN	Y	K-12	I	R

	capture "teacher moves"							
Engineering Design-based Science Teaching Observation Protocol (EDSTOP) [31]	Design-based Science Teaching Observation Protocol; measure implementation of engineering design-based science teaching by elementary teachers who participated in a 2-week PD	D	S	QN	Y	K-12	I	R
Classroom Observation Protocol for Engineering Design (COPED) [32]	Characterize ED integration in secondary science classrooms; focus on ED instruction	D	S	QN	Y	K-12	I&S	R
Classroom observation rubric (COR) [06]	Examine interactive teaching methods (peer instruction) in undergraduate physics classrooms	D	S	QN	Y	HE	I&S	R
Uteach Observation Protocol (UTOP) [33]	Assess the overall quality of instruction but without preference or bias toward any one way of teaching	E	S	QN	Y	K-12	I&S	R