# Improving Integrated STEM Education: The Design and Development of a K-12 STEM Observation Protocol (STEM-OP) (RTP)

## Introduction

The collective education of science, technology, engineering, and mathematics has been commonly referred to as STEM education. For K-12 education in the United States, the use of the phrase "STEM education" has repeatedly been used in reference to a more integrated curricular and pedagogical approach, wherein the STEM disciplines are interconnected with one another and contextualized by real-world problems [1] – [3]. This shift from teaching the STEM disciplines in isolation to integrating across disciplines was signaled most prominently in the National Research Council's [NRC] [4] publication of *A Framework for K-12 Science Education*, which included engineering as part of science teaching and learning. The *Next Generation Science Standards* (*NGSS*) [5], created using the NRC's *Framework* [4], sparked a flurry of change in K-12 science education. To date, 20 states and the District of Columbia have adopted the *NGSS* in full and another 24 have adopted standards similar to the *NGSS* based on the original NRC framework [4]. The purposeful addition of engineering within these science standards, along with the inclusion of computational and mathematical thinking, has resulted in the continued focus of integrated STEM in K-12 education.

Policy calls for integrated STEM approaches to teaching and learning require significant changes to existing curricula and pedagogy in K-12 classrooms, which creates challenges for integrated STEM implementation. For one, professional development opportunities for in-service teachers to engage in learning about integrated STEM are still limited, and the opportunities that do exist vary across different contexts [6]. Changes to pre-service programs appear to be even slower, with few models or suggestions of how to teach integrated STEM education available to teacher educators [7]. To complicate this further, there are a multitude of definitions and frameworks that exist when it comes to integrated STEM education. Although educators and researchers are beginning to better articulate "what counts" as part of integrated STEM education [3], there are still challenges in clearly guiding how educators engage in or enact integrated STEM education in the classroom. This, in turn, creates challenges for teachers who need to attend to changes with respect to both content and pedagogy [8], [9].

Observational tools constructed to embrace the integrated nature of STEM education are one possible solution to this problem, as they have the potential to provide critical guidance to researchers, professional development facilitators, classroom coaches, and others working toward improving integrated STEM instruction. These tools may help push forward a common language to be used among stakeholders. However, prior to the work reported here, observational tools that focus on the implementation of integrated STEM in K-12 classrooms are not available. The work presented in this paper provides a detailed description of how one team of STEM educators and educational researchers developed an integrated STEM observational protocol for use in K-12 science and engineering classrooms – the STEM Observation Protocol (STEM-OP). In addition to providing an in-depth description of the instrument development process, which started with the development of a conceptual framework for integrated STEM education [9], we share the final items and the challenges faced in assuring the validity and usability of the instrument.

## Literature Review

### Defining STEM education.

The research literature provides several models as potential ways to conceptualize integrated STEM education. For instance, Bybee detailed nine commonly used models, ranging from STEM as a single discipline to STEM as a transdisciplinary course or program [10]. Breiner *et al.* defined the practice of STEM integration as the shift from traditional lecture-based classrooms to the implementation of pedagogy that involves more inquiry and problem-based learning approaches [11]. Still, others define integrated STEM as curricula that integrate science, technology, engineering, and mathematics concepts in ways that most authentically reflect the practice of professionals currently working in STEM fields in an effort to graduate more students who are prepared to work in STEM professions [12], [13]. While this approach to integrating STEM appears to have caught the most traction, it can also be interpreted in different ways, making it difficult for teachers to implement [14] – [17]. These differences in definitions and interpretations surrounding STEM education can perpetuate problems related to effective communication among different stakeholders, such as teachers and their administrators [9]. This creates further issues for educational researchers who seek to explore the implementation of integrated STEM in K-12 classrooms when multiple definitions and interpretations of STEM education are in play.

Some have voiced the need for clarity in a definition but warn against a single definition or model [10]. Others have noted the need for a global definition, cautioning that without one, progress with respect to education and research will be hindered [3], [16]. Without a clear, tangible framework (and observational tools associated with such a framework), additional research may continue to exacerbate the issues surrounding different definitions of STEM education. In other words, the existence of such a framework and observational tool may help the education community advance a definition of what integrated STEM education can look like in science and engineering classrooms.

### Challenges in observing integrated STEM instruction.

As the education community works toward refining our understanding of integrated STEM educational approaches, there is a parallel need to document these approaches when they occur in the K-12 classroom. Capturing exactly what STEM education looks like in the classroom has been enigmatic. In large part this is due to the fact that there are no observation instruments available that were specifically designed to document integrated STEM instruction in the classroom. While multiple instruments that focus on good teaching practice or teaching isolated STEM disciplines exist, none of the instruments consider frameworks for integrated STEM teaching to account for multiple STEM disciplines within a given unit of observation (e.g., day, lesson, curriculum unit).

Instruments such as the Reformed Teaching Observation Protocol (RTOP) [18], Uteach Observation Protocol (UTOP) [19], and the Classroom Observation Protocol for Undergraduate STEM (COPUS) [20] were designed to be used in science and mathematics classrooms where a single discipline is present. These instruments are practical and useful in their respective disciplinary spaces, but they were not designed for observing lessons that integrate across STEM disciplines or lessons that include engineering. This can then be prohibitive as the items in these instruments do not directly address scientific practices, engineering practices, and/or the interplay across different STEM content and practices. Other K-12 STEM classroom observation instruments include the Science Classroom Observation Protocol [21], the Science and Engineering Classroom Learning Observation Protocol [22], the Classroom Observation Protocol for Engineering Design (COPED) [23], and the Classroom Assessment Scoring System [24], which are science-specific, engineering-specific, engineering design process focused, and content agnostic, respectively. Although these instruments provide a variety of different designs to support the development of a new observational instrument, they are ill-suited for use in classrooms that feature STEM integration, as they do not directly address instruction or learning that occurs at the intersection of multiple disciplines. Further, these instruments were not designed with an integrated STEM education framework in mind, thus missing key characteristics of integrated STEM. The desire and need for such as instrument motivated the work reported here, which includes the development of a new observation protocol designed specifically for use in K-12 science and engineering classrooms implementing integrated STEM lessons and units.

## Methods

We developed the STEM Observation Protocol (STEM-OP) by drawing from knowledge of observation protocol development as described in published works [18] – [24]. Since the RTOP [18] is one of the most widely used observational instruments in K-12 science classrooms, it served as a guide for our instrument development and what the final product would look like. By using it as a model, our hope was to design an instrument with a similar user experience; in particular, we aimed to develop an instrument that included up to 25 items, each of which would be rated on an ordinal scale. However, one component we wished to improve upon from previous instruments such as the RTOP [18] was the inclusion of descriptive levels for the ordinal scoring scale. To share the development process, we first provide a description of the overarching project's context and available resources. What follows this overview is a detailed description of the steps we took to develop the final 10-item STEM-OP.

### Context.

The work described here is part of a four-year federally funded project that seeks to improve the quality of K-12 integrated STEM education in science and engineering classrooms through the development and dissemination of a classroom observation instrument for integrated STEM instruction. The STEM-OP was intentionally designed for research purposes and for use as a formative educational tool for improving integrated STEM education. As part of this development, we utilized a suite of over 2000 video-recorded classroom observations that were collected as part of a previous federally funded project. That project generated and refined 54 teacher-created integrated STEM curriculum units over the course of five years. These curriculum units were developed, piloted, and refined as part of an intensive professional development, and drew upon two frameworks for integrated STEM education that feature engineering as the integrator of content [25], [26]. These two frameworks centralized engineering by using an engineering design challenge as context to teach STEM content. Since engineering in K-12 education is a relatively new endeavor, the focus on engineering reflects the needs of teachers, who often do not have knowledge of engineering and its associated practices [27]. As teachers implemented these units in their classrooms, which took anywhere from one to several weeks of instruction, project staff observed and video-recorded each day of implementation.

As a result, we have access to over 2000 classroom observations that represent a variety of classroom settings, including different grade levels, teachers, student demographics, science content, and engineering design challenges. The data set includes observations from 106 unique teachers' classrooms from five school districts that include urban, inner-ring suburban, and outer-ring suburban K-12 settings in the Midwestern United States. The majority of observations focus on grades 4-8, although early elementary (K-3) and high school (grade 9 in particular) were included to a lesser extent. The science content covered in these units spanned several topics in physical, life, and earth sciences. Table 1 provides an overview of these curriculum units based on the grade-band and science content breakdown.

Table 1. Summary table of observed curriculum units available

| | **Disciplinary Topics** | **Curricula by Grade Band*** |
|---|---|---|
| Physical Science | Heat Transfer and States of Matter<br>Force and Motion<br>Waves and Electromagnetism | 3 Elem, 3 MS<br>4 Elem, 1 MS, 1 HS<br>5 Elem, 4 MS, 1 HS |
| Life Science | Ecosystems<br>Natural Selection and Evolution<br>Genetics | 4 Elem, 3 MS<br>2 Elem, 1 Elem/MS, 3 MS<br>1 Elem, 1 Elem/MS, 1 MS |
| Earth Science | Plate Tectonics and Landforms<br>Weather and Water Cycle<br>Rocks and Soil & Renewable Energy | 4 Elem, 3 MS<br>2 Elem, 2 MS<br>4 Elem (1 pre-K), 1 MS |

* Elem = grades K-5, MS = grades 6-8, HS = grades 9-12

**The design and development process.**

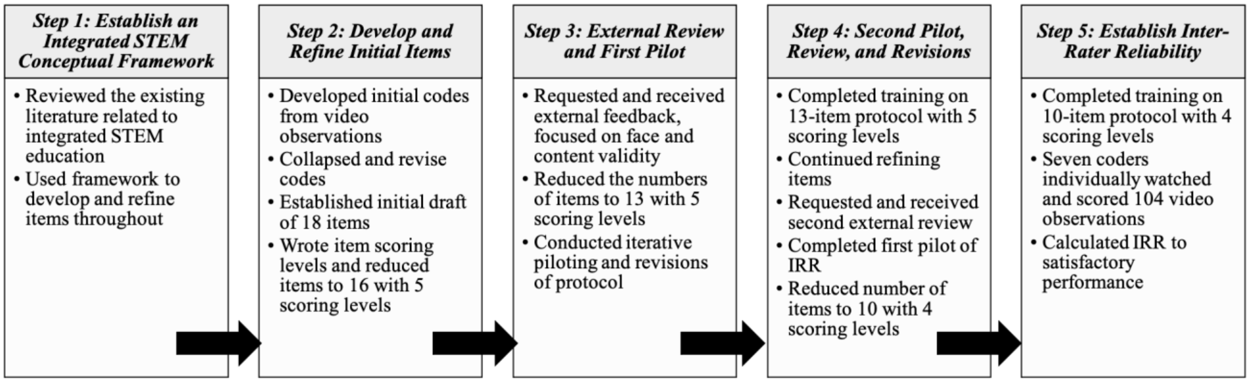
*Overview.*

Over the course of two years, we developed the STEM-OP and established interrater reliability among our coding team. Over this period of time, our team consisted of five principal investigators (four STEM education experts and one psychometrician), three post-doctoral researchers, and five doctoral students in STEM education. The process included various stages, which consisted of reviewing the existing literature on integrated STEM education, developing a conceptual framework to guide item development, creating initial lists of observable features of integrated STEM education, defining levels of each observable item, meeting with external advisors to discuss the items, piloting and refining items and item levels, and drafting a set of user guidelines. What follows is a description of the development process, broken down into five main steps; an overview of this process is provided in Figure 1.

*Step 1: Establishing an integrated STEM conceptual framework.*

To begin this work, we reviewed manuscripts related to the development of other observation protocols (including those previously described) for content, language, and development procedures. This part of the literature review helped us better understand the nature of protocol design. The second step in this process was a thorough review of the literature related to integrated STEM education. This included the examination of a variety of relevant K-12

Figure 1. Overview of the observation protocol development process




standards documents and frameworks [4], [5], [26], [28], [29], as well as the broader literature on curriculum integration and integrated STEM education [2], [8], [30], [31]. The primary purpose of this review was to identify frequent, research-based practices for teaching the STEM disciplines and common themes surrounding the different definitions of STEM education. In this review, we determined common features of integrated STEM education that are pervasive throughout the literature and for which agreement exists. While a full description of our conceptual framework can be found in Authors [9], the following sections highlight key areas.

First, researchers agree that integrated STEM education is driven by engaging students in developing solutions to real-world problems [2], [11], [13], [25], [32], [33]. Real-world problems used to motivate and contextualize learning should be complex enough to foster multiple solutions [34] and require students to apply and expand their knowledge of the STEM disciplines [35]. Given the need to increase diversity in the STEM fields [36], [37], these real-world problems should be personally motivating and additionally connect STEM content to students' lives; this has been shown to enhance student engagement by making learning more meaningful and relevant [38] – [40].

Second, given the prominence of engineering within science and STEM standards [4], [5], [41], the real-world problems used to contextualize learning should include an engineering context, most often presented as an engineering design problem or challenge through the specific needs of a client [25], [42] – [45]. In this engineering design-based approach to integrated STEM education, students are expected to engage in an engineering design process to develop and justify design solutions [2], [4], [5], [25], [46]. The iterative engineering design process requires students to develop and fine-tune problem-solving skills as "testing the most promising solutions and modifying what is proposed on the basis of the test results leads to greater refinement and ultimately to an optimal solution" [2, p. 210]. Further, it is important that students have opportunities to engage in redesign [25], [42] to understand how engineering uses evidence and STEM content when evaluating proposed design solutions [46] – [48].

Third, in addition to explicitly connecting STEM content to a real-world problem, it is important that connections between the STEM disciplines are made explicit to students [2], [15], [16], [25], [33], [49]. Although it is not necessary that all four STEM disciplines are integrated within a STEM unit [25], [33], it is important to note that the emphasis of policies such as the *NGSS* on science and engineering often relegate mathematics and technology to the background [50] – [53]. Most often mathematics is presented as a tool for data analysis and measurement [14], [54] instead of more meaningful engagement in mathematical thinking relevant to STEM such as mathematical modeling [55], [56]. Technology integration is even less well defined [16], [57], often limited to a pedagogical tool or a product of engineering [2], [57]. Given the emphasis on engaging *students* in authentic scientific and engineering practices [4],[5] and modelling STEM fields as viable career options [15], [16], [33], representing technology within integrated STEM education as the technology tools and practices used by STEM professionals is particularly fruitful [57] and consistent with the definition we subscribe to in our work. As such, in this work technologies are conceptualized as "tools for knowledge construction rather than media of conveyance and knowledge acquisition" [58, p. 2] that are analogous to those used by practitioners of science, engineering, and mathematics [2], [57].

Fourth, teamwork and communication are key components to integrated STEM education [17], [25], [33], [59]. Students are expected to work collaboratively and cooperatively within small groups to co-construct knowledge of STEM content and design solutions to real-world problems. Given the ill-structured nature of real-world problems, students need to negotiate their understanding and decision-making, all of which require a high level of pedagogical facilitation and support to make sure that all students' voices are equitably heard [60], [61].

Finally, the implementation of integrated STEM education requires the use of student-centered pedagogies to support the development of 21st century skills [11] – [13], [17]. These types of skills include the four C's of creativity, communication, critical thinking, and collaboration [62]. Minner *et al*. reported "that having students actively think about and participate in the investigation process increases their science conceptual learning" [63, p. 493]. Further, engaging in the same practices as STEM professionals is important to understand the nature of the field [64]. Integrated STEM is grounded in social constructivist theories of learning [65] and students are expected to apply STEM content and practices to the design of solutions to real-world problems [2], [3], [11], [13], [25], [32], [33]. Central to integrated STEM is learning from failure [25], [66], [67], using iterative test data to improve and refine their design solutions [25], [42] and justifying scientific claims and design decisions with evidence and reasoning [32], [46].

Supporting all of these features is the desire to provide students with more authentic learning experiences in the classroom that allow students opportunities to develop science and STEM identities [68]. Given how identities are "produced through practices, relationships and interactions within specific sites and spaces" [68, p. 619], there is a need to increase the diversity of those persisting in STEM fields through college and beyond to include historically underrepresented students. *A Framework for K-12 Science Education* notes that not all students will choose to pursue STEM careers, but that, "a science education based on the framework will motivate and inspire a greater number of people - and a better representation of the broad diversity of the American population - to follow these paths than is the case today" [4, p. 9-10]. With this important goal in mind, integrated STEM education should allow students to not only engage in work that resembles that of STEM professionals, but also allow students to develop STEM identities and be exposed to STEM careers in concrete ways. When the perception of STEM careers does not match the image students have of someone they are familiar with, it is likely that they abandon their prior goal and change their career trajectory [69]. Because of this, integrated STEM lessons can empower teachers to intervene and help students develop accurate perceptions of STEM professions and foster the development of their students' STEM identities.

In addition to the authentic experiences described above, explicit exposure to STEM careers must also be included.

In summary, our working definition of integrated STEM is consistent with that of Kelley and Knowles, who define STEM education in general as, "an approach to teaching the STEM content of two or more STEM domains, bound by STEM practices within an authentic context for the purpose of connecting these subjects to enhance student learning" [3, p. 3]. Beyond this, however, our conceptual framework provides the detail necessary to operationalize this broad definition. Given that the proposed use of the STEM-OP is in science classrooms, many of which in the United States are now guided by the NRC's *Framework* [4] and the *NGSS* [5], our protocol requires the presence of science and/or engineering content.

*Step 2: Developing and refining initial items.*

Using this conceptual framework, we designed the STEM-OP to indicate the presence of integrated STEM education within an observed classroom period, on average 50 minutes. The items were developed to reflect the degree to which characteristics of STEM integration (as described in our conceptual framework) are observed; the intention was not to assess the quality of the teachers' pedagogy. With this in mind, we focused the development of the STEM-OP on observable characteristics of integrated STEM education.

*Drafting preliminary codes*. Drafting a list of potential items to include in the instrument occurred in several stages. We first used a deductive process to identify key concepts from the literature that were used to develop the aforementioned conceptual framework. We then followed an inductive approach through which items were generated from classroom videos that featured exemplary instruction; this overall process reflects methods often found in qualitative research related to coding [70], [71]. This second step allowed us to identify and grasp the more observable features of integrated STEM education. In practice, we individually watched a purposefully selected video from the suite of videos available; this video was selected based on the project PIs' prior knowledge of the observation, which was deemed as theoretically representative of "good STEM teaching". Each member of our team took field notes, paying particular attention to features that characterize integrated STEM education; afterward, we used individual field notes to list as many observed features as possible. These features served as initial codes and were collected into a single list where duplicate items were removed. This task was repeated with a second purposefully selected video, using the first list as a provisional set of deductive codes, adding more codes as needed.

*Revision and collapsing of preliminary codes*. The process above resulted in a total of 79 codes. These codes varied in focus as some attended to actions of the teacher or students, while others focused on features of the lesson design. We sorted these codes into a visual display composed of a three-column table with the headers: *teacher, student,* and *lesson*. Three small teams were created to collate similar codes, categorize them according to pedagogy (teacher), student practices (student), and integration of STEM (lesson) by using the conceptual framework as a guide. Constantly referring to the conceptual framework allowed us to identify and remove "just good teaching" codes in the Teaching column, as the goal of this instrument was to go beyond good pedagogy and to examine aspects of integrated STEM education. In addition, each small team added new codes that were not initially captured but were found in the literature and conceptual framework and further collapsed codes that overlapped within their given column. For each code that remained, the small teams explicitly linked it to the conceptual framework and supporting literature, focusing on the nuances of integrated STEM education.

After our small group work on the individual columns, 49 codes remained (19 Teacher, 20 Student, 10 Lesson). As a project team, we collectively looked across the three columns to examine the codes for common themes related to characteristics of STEM education (i.e., overarching ideas or goals), which we then grouped together. For instance, codes related to addressing real-world problems existed across both the student and lesson columns, but centered on providing a real-world problem context (Table 2); other codes focused on multiple solutions cut across all three dimensions. Once grouped together, the codes were reduced and rewritten as initial items with roughly worded item levels. This process resulted in 18 conceptualized items with a varying degree of tentative scoring levels and descriptions.

Table 2. Example overlaps in codes in a preliminary draft

| Characteristic | Teacher | Student | Lesson |
|---|---|---|---|
| **Real-World Context** | | Students develop a broader understanding of real-world problems through the context of the lesson. | The lesson is situated in a relevant real-world science or engineering problem. |
| **Multiple Pathways** | Teacher emphasizes multiple solution pathways to solving a problem. | Students communicate their procedural decisions used to complete the task. | Lesson activities have multiple viable solutions. |
| **STEM Content** | | Students develop procedural fluency in the application of science, technology, engineering, and/or mathematics.<br><br>Students use content from multiple disciplines to complete a task.<br><br>Students explore relationships between science, technology, engineering, and/or mathematics concepts. | Lesson activities require the use of multiple disciplinary practices to complete a task [e.g., data analysis and/or scientific investigation are used to support student decisions on a design]. |

*Creating item scoring levels*. With this shorter list of 18 conceptualized items and some notes about potential item levels, we next determined how many levels to include for each item. After several discussions about the feasibility of creating discrete, measurable levels that could be observed without additional classroom materials (e.g., lesson plans, student work, interviews with teachers), we decided on five levels with a 0-4 scale, similar to the RTOP [18]. Items were assigned to small teams to continue refining the wording of each item and its observable levels. Part of this process included a critical examination of the items, which sometimes resulted in removing an item completely or recognizing that we could further collapse items. This continued work occasionally included the splitting of a single item into two items, as some ideas were "too big" or unwieldy to address in a single item. In other cases, items were removed due to the perceived challenges in observing an item. For example, "Students' positive STEM identities are promoted," was first extracted from an item related to creating awareness of STEM beyond the classroom, but then was removed due to the difficulty in directly observing this from a whole classroom observation. After roughly a month of working on this task and organizing the items, the first usable draft protocol was available. It consisted of 16 items with a root item, description of the item and its purpose, 5 levels (scored 0-4), and an area for written comments/feedback.

*Step 3: External review and first pilot.*

As part of checking the face and content validity, the working draft of the 16-item STEM-OP underwent external review. An advisory board consisting of three STEM education experts provided feedback on the item content, the wording of the items and item levels, and our general process for the instrument's development up to that point. As shown in Tables 3 and 4, this feedback resulted in removing three items relating to *Connections to Big Ideas*, *Learning from Failure*, and *Creativity* and splitting one item to differentiate between drawing from students' prior knowledge and acknowledging students' own experiences outside of the classroom. We then piloted the 14-item STEM-OP using more classroom video observations. Through an iterative process of using the instrument and adjusting wording, we removed an additional item related to *Multiple Representations* (Table 4), as it required too much interpretation by the observer to score it consistently; this reduced the number of items to 13.

*Step 4: Second pilot, review, and revisions.*

After a year of developing the protocol items and item levels, the coding team (composed of four graduate students and two post-doctoral researchers) participated in intensive training with the PIs on using the STEM-OP. Part of this was done as two new post-doctoral researchers and one new graduate student replaced graduating members of our project team. This training was also intended to help all of us further refine our own understanding of the 13 items to improve our ability to score reliably. The training included various iterations of (1) collectively watching a video and then going through the entire instrument together to score it, (2) individually watching and scoring a video, and (3) meeting as a whole group to come to consensus. Throughout this process, we continued to refine the wording of the items and item levels for clarity and usability. One of the most significant changes was in adjusting the wording of the items to focus on the actions and words of the teacher. This decision was made primarily to ensure the reliability and consistency of observations, as the teacher is the vehicle by which instruction is made explicit. The only item for which this did not happen was with an item related to technology, as our conceptual framework dictated its focus on students' use of technology.

After two months of continued scoring and discussions towards consensus, we met again with the project's external advisory board for further feedback on the face and content validity of the items. This resulted in making several minor revisions to the wording of the 13 items and levels as outlined in Table 3. This modified version was tested again as before, and although discussions to come to consensus were still conducted, no further revisions were made to the items or item levels. After several more weeks, we tested the protocol for reliability among coders to identify further problematic items.

Table 3. Major changes made between the 16-item draft and the final 10-item protocol

| 16 Item Draft | Changes Made & 13-Item Result | Changes Made & 10-Item Result | 10-Item Final Name |
|---|---|---|---|
| **Prior Knowledge**<br>*Students' prior knowledge is addressed.* | This item was split to delineate the difference between prior knowledge from previous lessons and students' lived experiences.<br><br>**Prior Knowledge**<br>*The teacher incorporates concepts from previous STEM lessons.* | Prior Knowledge was removed after the second Pilot due to low IRR (α = 0.381) and the recognition that accurately observing this would require more information outside of the observation (e.g., lesson plans from prior days). Iterative changes were made to the levels of Relating to Students' Lives. | |
| | **Relating to Students' Lives**<br>*The teacher elicits and incorporates students' lived experiences from outside the classroom related to STEM.* | *The teacher elicits and incorporates students' experiences from outside the classroom related to STEM.* | **Item 1**<br>**Relating Content to Students' Lives** |
| **Real-World Context**<br>*The lesson is situated in a real-world context that is relevant to students.* | The language of this item was adjusted for readability.<br><br>*The lesson is contextualized in a real-world problem that is connected to student learning.* | Modifications were made to focus on how the teacher presents the lesson and student activities. The explicit connection to engineering design challenges was added.<br><br>*The teacher contextualizes student learning within a real-world problem or design challenge.* | **Item 2**<br>**Contextualizing Student Learning** |
| **Multiple Pathways**<br>*Students are encouraged to explore multiple pathways to solve a problem.* | The language of this item was adjusted for readability.<br><br>*The teacher encourages students to develop and evaluate multiple solutions.* | Iterative changes were made to the levels of this item.<br><br>*The teacher encourages students to develop and evaluate multiple solutions.* | **Item 3**<br>**Developing Multiple Solutions** |
| **Integrating STEM Content and Practices**<br>*Students integrate content and engage in practices from multiple STEM disciplines to complete an activity.* | The focus was changed to just content as practices were already covered in a separate item.<br><br>*The teacher requires students to integrate content from multiple disciplines.* | This item was split into two items after our Second Pilot due to low IRR (α = 0.278), but foundational need to include an item about content integration.<br><br>*5a. The teacher provides opportunities for students to learn S/T/E/M concepts.*<br><br>*5b. The teacher integrates content from multiple disciplines in STEM activities.* | **Item 4**<br>**Cognitive Engagement in STEM**<br>**Item 5**<br>**Integrating STEM Content** |

| | | | |
|---|---|---|---|
| **Use of STEM Practices**<br>*Students develop an understanding of how science, technology, engineering, and/or mathematics knowledge is developed through their use of STEM practices.* | Overall, this item stayed the same, but was modified slightly to better reflect the student-centered pedagogies employed as students engage in STEM practices.<br><br>*The teacher provides opportunities for students to engage in STEM practices and develop an understanding of how they are used.* | Iterative changes were made to the levels of this item.<br><br>*The teacher provides opportunities for students to engage in STEM practices and develop an understanding of how they are used.* | **Item 6**<br>**Student Agency** |
| **Collaboration and Teamwork**<br>*Students collaborate with one another to complete learning activities, understand STEM content, and develop teamwork skills.* | The wording of this item was modified to reflect the actions of the teacher and the expectations of collaborative group work.<br><br>*The teacher requires students to collaborate with one another to co-construct knowledge of a phenomenon, real-world problem, and/or design solutions to a real-world problem.* | Iterative changes were made to the levels of this item.<br><br>*The teacher requires students to collaborate with one another to co-construct knowledge of a phenomenon, real-world problem, and/or design solutions to a real-world problem.* | **Item 7**<br>**Student Collaboration** |
| **Evidence-Based Reasoning**<br>*Students use evidence-based reasoning to develop their understanding of a real-world phenomenon and to justify claims and decisions.* | The wording of this item was modified to reflect the actions of the teacher and the expectations of collaborative group work.<br><br>*The teacher requires students to use evidence-based reasoning.* | Iterative changes were made to the levels of this item.<br><br>*The teacher requires students to use evidence-based reasoning.* | **Item 8**<br>**Evidence-Based Reasoning** |
| **STEM-Specific Technology**<br>*STEM-specific technology and tools are used to promote student learning of STEM content and/or practices.* | This item was modified slightly to focus on facilitating STEM practices.<br><br>*Technology is used to model phenomena and/or design solutions to a real-world problem.* | Unlike the other items, this item was modified to specify that students in particular are the users of technology.<br><br>*Students use STEM-specific technologies.* | **Item 9**<br>**Technology Practices in STEM** |
| **STEM Awareness Beyond the Classroom**<br>*Classroom activities promote student awareness of STEM opportunities beyond the classroom.* | This item was modified to better reflect the conceptual framework and focus on the awareness of STEM careers in particular.<br><br>*The teacher promotes awareness of STEM careers/pathways in connection to the lesson.* | Slight modifications were made to improve readability.<br><br>*The teacher promotes awareness of STEM careers.* | **Item 10**<br>**STEM Career Awareness** |

Table 4. Items removed between the 16-item draft and the final 10-item protocol.

| 16 Item Draft | Changes Made & 13-Item Result | Justification for Removing from 13-Item Draft |
|---|---|---|
| **Effective Questioning**<br>*Teacher uses effective questioning strategies to reveal, support, and challenge students' understanding of STEM concepts and practices.* | Slight modifications were made to improve readability and specify different foci in classrooms.<br>*The teacher uses questioning strategies to reveal, support, and challenge students' understanding of the phenomenon, the real-world problem, and/or design challenge.* | This item was eventually removed after our Second Pilot due to low IRR (α = 0.374) and the acknowledgement that this was representative of good teaching in general and not STEM-specific. |
| **Communicating Understanding**<br>*Students communicate their understanding of a real-world phenomenon.* | Slight modifications were made to improve readability and focus on the teacher's actions.<br>*The teacher provides opportunities for students to communicate their understanding of a phenomenon, a real-world problem, and/or a solution to that problem.* | This item was removed after our Second Pilot due to low IRR (α = 0.264). |
| **Constructive Feedback**<br>*Students receive constructive feedback from others.* | Slight modifications were made to improve readability and focus on the teacher's actions.<br>*The teacher provides opportunities for students to receive and apply constructive feedback.* | This item was removed after our Second Pilot due to low IRR (α = 0.289) and the acknowledgement that this was representative of good teaching in general and not STEM-specific. |
| **Multiple Representations**<br>*Students communicate their understanding of a real-world phenomenon or problem through multiple representations.* | This item was removed shortly after the external advisory board meeting as this item seemed reduced to counting the number of representations rather than focusing on what students were doing with those representations. Some wording changes were made before it was removed, as noted below.<br>*The lesson activities provide opportunities for students to translate between multiple representations of a phenomenon and/or real-world problem.* | |
| **Connections to Big Ideas**<br>*STEM lesson content and activities are connected to the overarching goals/Big Idea of the unit.* | This item was removed at the recommendation of the external advisory board as it required knowledge outside of the single observation and this was too hard to observe if not stated explicitly by the teacher. | |
| **Learning from Failure**<br>*Students learn through failure and iteration.* | This item was removed at the recommendation of the external advisory board as the item was too specific to engineering. However, the iterative process is reflected in the levels of the item related to multiple solutions. | |
| **Creativity**<br>*The lesson promotes student creativity and creative expression.* | This item was removed at the recommendation of the external advisory board as observing creativity was too subjective due to many definitions of creativity. Creativity is also reflected in the *Developing Multiple Solutions* item. | |

Each week for a total of six weeks, five videos were scored independently by the coding team and one of the PIs. These videos were previewed by another PI and purposefully sampled to ensure a variety of grade levels, content areas, teachers, day within the unit of instruction (i.e., not all the first day), and engineering design challenges. Coders completed this task independently, submitting scores, field notes, and evidence/justification for their scores using a Google form.

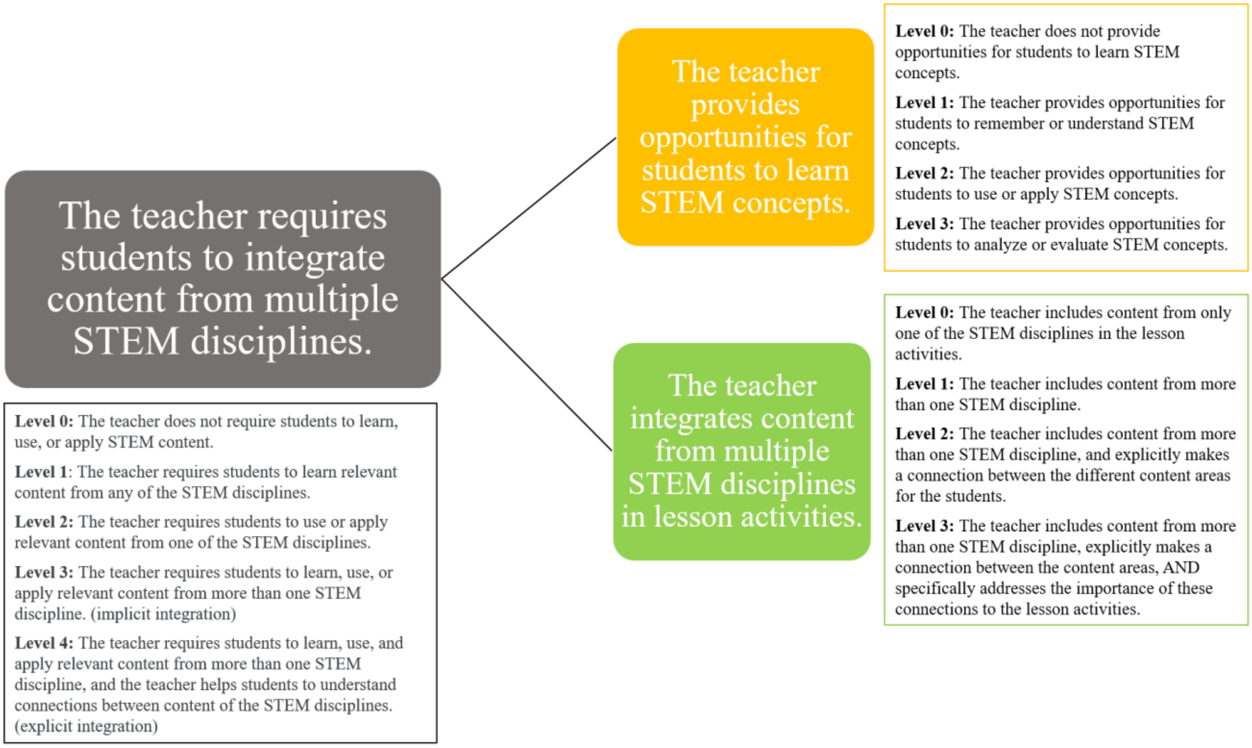
After 30 videos were scored, we used Krippendorff's alpha to calculate Inter-Rater Reliability (IRR) for each individual item across all the videos scored by all the individual coders. A Krippendorff's alpha value of 0.0 indicates rater agreement is no better than the agreement one would expect from random assignment of codes whereas a Krippendorff's alpha value of 1.0 indicates complete agreement across all raters. Krippendorff's alpha was chosen over other IRR measures, such as Cohen's kappa, because it is a more versatile statistic that is not as sensitive to missing data and it can better accommodate multiple raters and non-ratio (categorical or ordinal) levels of measurement [72]. Although a threshold of 0.67 is commonly considered best practice when performing reliability testing, considering the early-stage and exploratory nature of this work, we deemed a threshold of 0.60 to be appropriate.

While the reliability of certain items (*Multiple Pathways*, *Collaboration and Teamwork*, *STEM-Specific Technology*, and *STEM Awareness Beyond the Classroom*) approached our selected acceptable threshold of $\alpha \geq 0.6$, four more items (*Relating to Students' Lives*, *Real-World Context*, *Use of STEM Practices*, and *Evidence-Based Reasoning*) showed progress with respect to IRR ($\alpha > .4$). These preliminary findings suggested that at minimum, these eight items were approaching reliability. However, the remaining five items showed low agreement among coders ($\alpha < .4$) (Tables 3 and 4), suggesting inherent problems with the items. In addition to unsatisfactory IRR, items related to *Prior Knowledge*, *Effective Questioning*, *Communicating Understanding*, and *Constructive Feedback* had all attempted to measure pedagogical issues that, while important for effective instruction, are not specific to integrated STEM education.

Despite the unsatisfactory IRR result of the item related to *Integrating STEM Content and Practices*, this item represents a core component of integrated STEM within our conceptual framework and therefore could not be removed from the protocol. The low agreement among coders reflects the challenges of observing integrated STEM education consistently across multiple coders in different contexts; in other words, identifying the integration of STEM content and practices is difficult. To decrease the difficulty and reduce the number of variables in the 13-item version of this item, we split it into two items to reflect (1) the depth of learning and cognitive engagement and (2) the degree of integration between disciplines (Figure 2).

To further address issues related to IRR, as shown in Figure 2, we adjusted the items from a 5-point scale to a 4-point scale. For most items, this was done by removing the highest level, which was rarely observed during the piloting of the STEM-OP; thus, our inability to observe these items negatively impacted our ability to establish reliability for 5-point items. An example of this, taken from Item 8 related to *Evidence-Based Reasoning* is: "The teacher requires students to make claims and/or design decisions based on evidence, justify them using reasoning, and evaluate the quality of the evidence." Rarely did we observe opportunities for students to assess evidence quality. In three cases, middle levels were removed due to issues of clarity and

Figure 2. Evolution of STEM concept integration item




distinction between levels. For example, for Item 3 related to *Developing Multiple Solutions*, level 0 ("The teacher does not promote the development of any solution") and level 1 (The teacher promotes only one solution) were simplified to the new level 0 "The teacher does not encourage the development of multiple solutions". As part of this revision, some item descriptions and accompanying levels were reworded to simplify the language and reduce the potential for inconsistent scoring so that improved IRR could be achieved.

To pilot these new revisions and provide feedback related to the new draft's usability, coders rewatched and rescored five of the video observations from the previously watched set of 30. Feedback concerning usability revealed that the revised 10-item STEM-OP with only four distinct levels was easier to manage during observations due to a decrease in the total number of levels to account for (40 instead of 65 with the 5-level, 13-item version). Iterative modification of the instrument's wording continued for approximately another two months to prepare the coders for the next step of formally establishing IRR on the final 10-item protocol.

*Step 5: Establishing Inter-Rater Reliability (IRR).*

The formal IRR phase was initiated with a two-day training on the finalized STEM-OP. The training included independent scoring of a classroom video, followed by discussion to clarify the items and levels as needed. In the succeeding months, six coders watched and scored six videos per week using a new Qualtrics version of the protocol. In addition to the six coders, one of the PIs also scored the videos to increase the number of perennial coders from six to seven. This seven-coder team then scored 104 videos with the 10-item STEM-OP before calculating IRR. Our selected acceptable Krippendorff's alpha levels ($\alpha$ >.6) were achieved for all but one item – Item 5: *Integrating STEM Content* (Table 5). Given the centrality of integration across the disciplines in the integrated STEM conceptual framework, and after another consultation with the external advisory board, we decided to keep this crucial item in the protocol; implications of this decision are included in the discussion section. The following sections share the details of the final protocol, including the levels and connections to supporting literature.

Table 5. Inter-Rater Reliability of final protocol items

| Item | Item Name | Krippendorff's Alpha (α) |
|---|---|---|
| 1 | Relating Content to Students' Lives | 0.654 |
| 2 | Contextualizing Student Learning | 0.736 |
| 3 | Developing Multiple Solutions | 0.805 |
| 4 | Cognitive Engagement in STEM | 0.634 |
| 5 | Integrating STEM Content | 0.580 |
| 6 | Student Agency | 0.725 |
| 7 | Student Collaboration | 0.724 |
| 8 | Evidence-Based Reasoning | 0.699 |
| 9 | Technology Practices in STEM | 0.725 |
| 10 | STEM Career Awareness | 0.870 |

**Description of the Final Protocol**

The final observation protocol includes ten items that reflect key components of integrated STEM education, aligned to our conceptual framework. Similar to other observation instruments, such as the RTOP [18], each item is rated on a Likert-scale from 0 to 3. What is unique in the structure of our instrument is that each level additionally includes criteria-specific descriptions. Further, each item has a short, descriptive title, and a longer description that describes the intent of the item. The sections that follow share the ten items and their scoring levels, plus a description of their alignment to the literature and conceptual framework.

*Item 1 - Relating content to students' lives.*

Item 1 (Figure 3) focuses on the extent to which the lesson content is connected to students' lives. The literature suggests that student engagement increases when the content is related to their lives and prior experiences outside the classroom [3], [38] – [40]. While noting the importance of culturally responsive teaching in STEM, the STEM-OP is not intended to measure the nature of cultural relevance in instruction, which can be assessed in other instruments such as the CRIOP [82], [83]. As a result, this item focuses on the ways in which the teacher relates to or incorporates students' everyday and personal experiences from outside the classroom. During the development of the STEM-OP, the level descriptions for Item 1 did not change significantly except for the addition of concrete examples in level 1 and moving away from the phrase "lived experiences" to personal and/or everyday experiences. At the lowest level, the item attends to mentions of concrete examples or personal experiences that students may have had, but explicit connections to the lesson are not necessarily made. The levelling progresses to consider whether students' experiences were activated and elicited in some form or another and then whether these experiences were explicitly connected to the lesson.

Figure 3. Item 1 - Relating content to students' lives

| 1 | **Relating Content to Students' Lives** |
|---|---|
| *Students' everyday and personal experiences from outside the classroom are activated, meaningfully incorporated into the lesson, and related to the development of STEM knowledge.* | |
| 0. The teacher does not acknowledge students' everyday and/or personal experiences related to STEM.<br>1. The teacher mentions personal experiences or provides concrete examples to illustrate the STEM content in the lesson.<br>2. The teacher elicits students' everyday and/or personal experiences related to STEM during the lesson.<br>3. The teacher elicits students' everyday and/or personal experiences related to STEM and explicitly connects these to the lesson. | |

*Item 2 - Contextualizing student learning.*

Item 2 (Figure 4) was constructed to focus on motivating student learning through contextualizing the lesson with a real-world problem [2], [33] that makes learning more relevant for students [35], [73]. The presence of a real-world problem or engineering design problem, however, only makes a surface-level connection to our conceptual framework. As noted above, it is important that students explicitly connect and apply science and mathematics to the real world-problem or engineering design challenge [25], [26]. The increasing levels of this item focus on the teacher's efforts not only to contextualize students' learning, but to explicitly emphasize the connections between the real-world problem or the design challenge with what students are learning. This goes beyond just having the connection present, but ensures that students understand the connection between the lesson content and the context.

Figure 4. Item 2 - Contextualizing student learning

| 2 | **Contextualizing Student Learning** |
|---|---|
| *Learning is contextualized within an appropriate (e.g., age, gender, race, etc.) real-world problem or design challenge that connects to the content of the lesson. Connections between students' learning and the context are explicit so that students understand the importance of their learning.* | |
| 0. The teacher does not contextualize the lesson within a real-world problem or design challenge.<br>1. The teacher contextualizes the lesson by alluding to a real-world problem or design challenge, but does not connect to what the students are learning.<br>2. The teacher contextualizes the lesson by briefly connecting a real-world problem or design challenge with what the students are learning.<br>3. The teacher contextualizes the lesson by emphasizing the connections between the real-world problem or design challenge and what students are learning and helps them make explicit connections between the content and the context. | |

*Item 3 - Developing multiple solutions.*

Item 3 (Figure 5) highlights the importance of divergent thinking and multiple solutions, concepts particularly central to engineering design. This item acknowledges the importance of engineering design within our conceptual framework, which enables students to come up with multiple solutions [34]. This type of divergent problem-solving requires the development and use of critical thinking skills and creativity [62]. Further, design-based integrated STEM education features learning from failure [42] wherein students have opportunities to iterate their solutions. The conceptual framework highlights that not only should the real-world problem or engineering design challenge allow for the development of multiple solutions, but that students should also engage in iterative testing, learning from failure and using evidence, to refine their solutions. As such, the levels of this item, built from initial draft codes and items, reflect these components with the highest level not only promoting multiple solutions to a given problem, but also providing opportunities for students to evaluate these solutions and redesign them.

Figure 5. Item 3 - Developing multiple solutions

| 3 | **Developing Multiple Solutions** |
|---|---|

*The teacher promotes students' development of multiple solutions during the STEM lesson. Students are encouraged to develop multiple design alternatives and evaluate them, identifying the relative advantages and disadvantages of each possible solution.*

0. The teacher does not encourage the development of multiple solutions.
1. The teacher encourages students to develop multiple solutions, but does not provide opportunities for students to evaluate these solutions.
2. The teacher encourages multiple solutions and provides opportunities for students to evaluate the viability of different solutions.
3. The teacher encourages multiple solutions and provides opportunities for students to not only evaluate the viability of different solutions, but also use this information to redesign their solution.

*Item 4 – Cognitive engagement in STEM.*

Item 4 (Figure 6) reflects STEM learning as a dynamic process that requires student engagement at a variety of cognitive levels. The revised Bloom's taxonomy [74] was used as a framework for developing Item 4's levels, which reflect the need for students to develop conceptual knowledge of STEM content through a variety of means. At the lowest level, students engage in tasks that require lower-order skills such as remembering facts and demonstrating their understanding of concepts. Lessons that provide students the opportunity to apply what they have learned, analyze concepts, and evaluate ideas are given higher scores. In design-centric settings such as an engineering lesson, this kind of progression of cognitive engagement is mirrored as well [75]. Previous research gives high regard to activities wherein students develop design solutions by applying what they have learned and improve these ideas by analyzing their solutions and evaluating them [26], [75] – [77]. As a result, cognitive engagement in design-based tasks gradually progresses from lower to higher levels of thinking along the design process.

Figure 6. Item 4 – Cognitive Engagement in STEM

| 4 | **Cognitive Engagement in STEM** |
|---|---|

*Students engage in learning within a STEM lesson at different cognitive levels. While it is appropriate for students to be expected to learn facts and definitions, it is important that students have opportunities to work at higher levels of cognitive engagement such as applying concepts in new situations, and evaluating and analyzing concepts. In other words, students should experience all levels of Bloom's taxonomy when in a STEM classroom.*

0. The teacher does not provide opportunities for students to learn S/T/E/M concepts.
1. The teacher provides opportunities for students to remember or understand S/T/E/M concepts and/or a design problem.
2. The teacher provides opportunities for students to use or apply S/T/E/M concepts and/or a design plan.
3. The teacher provides opportunities for students to analyze or evaluate S/T/E/M concepts and/or design solutions.

*Item 5 - Integrating STEM content.*

Item 5 (Figure 7) represents the heart of the STEM-OP. As noted above, this item arose out of a more complex item that had attempted to attend to: how students use content (learn vs. use and apply), the number of disciplines present (one vs. multiple), and the degree of emphasis on the connection among the disciplines. The final version of Item 5 reflects the degree to which the teacher helps students recognize and appreciate the relationship of the STEM disciplines by making the connections between the disciplines explicit [2], [15], [16], [25], [33]. This reflects the need to make these connections visible as noted in the conceptual framework; otherwise, students struggle to understand how the various disciplines present in classroom activities relate to one another. As a result, this item focuses on the degree to which the teacher makes connections among the STEM disciplines explicit to the students, regardless of how many STEM disciplines are present in the lesson or how STEM content is used. The levels of this item instead build on the explicitness and specificity of the connections.

Figure 7. Item 5 - Integrating STEM content

**5 Integrating STEM Content**

*Within the lesson, multiple content areas are represented that cut across two or more STEM disciplines. The tasks assigned to students should make it clear that students need to draw from these multiple areas and recognize that they are drawing upon multiple disciplines.*

0. The teacher does not include STEM content or includes content from only one of the STEM disciplines in the lesson activities.
1. The teacher includes content from more than one STEM discipline.
2. The teacher includes content from more than one STEM discipline and explicitly makes a connection between the different content areas for the students.
3. The teacher includes content from more than one STEM discipline and includes specific and/or sustained connections between these content areas within the lesson.

*Item 6 - Student agency.*

Item 6 (Figure 8) focuses on students' engagement in and use of STEM practices. Although initial versions of this item focused on students developing an understanding of what STEM practices *are*, the final version emphasized *how* students engage in those practices to better reflect student-centered pedagogies. While the heart of this item is centered on the use of STEM practices in the classroom, the item assesses the degree to which students guide their own use of those practices, which include, but are not limited to the definition of science and engineering practices provided by the *NGSS* [5] and the NRC's *Framework* [4]. As students engage in these disciplinary practices, they add to the knowledge building process that is related to the disciplinary practices, becoming an epistemic agent [78]. The levels of this item reflect the gradual increase of student autonomy and agency, focusing on the student-centered nature of integrated STEM education while moving away from lessons directed by the teacher.

Figure 8. Item 6 - Student agency

| 6 | **Student Agency** |
|---|---|

*Epistemic agency refers to students' ability to shape and evaluate knowledge and knowledge building practices in the classroom. Within STEM, these knowledge building practices call for students to engage in STEM practices (behaviors that STEM professionals engage in - e.g., problem scoping, developing and using models, planning and carrying out investigations) as they develop their knowledge of STEM concepts. In addition to using STEM practices, students should also reflect on the use of these practices to better understand how STEM knowledge is developed.*

0. STEM practices are not evident in the lesson.
1. The teacher presents STEM practices as directions for the students to follow.
2. The teacher provides opportunities for students to exercise agency when engaging in STEM practices.
3. The teacher provides opportunities for students to reflect upon their use of STEM practices within the activity.

*Item 7 - Student collaboration.*

Item 7 (Figure 9), an item relatively unchanged throughout the process, highlights the importance of collaboration and teamwork emphasized in the STEM education literature base [26], [42], [79]. This item is not intended to examine student engagement and discourse in small groups, which would require focused observations of student groups [60], [61], but rather the nature of the tasks and the teachers' expectations of students when they engage in small group work. As students work together, the extent of students' collaboration is ranked from simply completing procedural tasks (such as following the steps of a "cookbook" lab) as a group to sharing ideas to co-constructing knowledge of STEM. The highest level of group work includes collaborative processes where students must negotiate to come to a group consensus [80]. The levels of this item build in complexity and reflect the greater cognitive effort needed to complete these different tasks as well as a stronger connection to student-centered pedagogies.

Figure 9. Item 7 - Student collaboration

| 7 | **Student Collaboration** |
|---|---|

*Students have opportunities to collaborate with one another as they complete learning activities and develop a deeper understanding of STEM content. Students are encouraged to consider ideas from multiple individuals, critiquing these ideas and integrating new ideas into their existing understanding to co-construct a deeper understanding of STEM content. Students' voices and ideas are represented, and students are empowered to participate and contribute to the collective learning taking place.*

0. The teacher does not provide opportunities for students to collaborate with one another in a group setting.
1. The teacher places students in groups and requires them to complete a procedural task related to STEM content.
2. The teacher places students in groups and requires them to collaborate with one another by sharing ideas related to a phenomenon, real-world problem, design solution (e.g., brainstorming to generate ideas), and/or STEM content.
3. The teacher places students in groups and requires them to collaborate with one another to co-construct knowledge of a phenomenon, real-world problem, design solution, and/or STEM content.

*Item 8 - Evidence-based reasoning.*

Item 8 (Figure 10), another item that was relatively unchanged throughout the process, requires students to engage in evidence-based reasoning (EBR). EBR is important for both scientists [32] and engineers [46] and requires students to develop and exercise critical thinking skills and justify their claims and design decisions with evidence; this can look slightly different in science and engineering spaces. Notably, predictions are excluded from our definition of claim since stand-alone predictions may not be supported by substantiated evidence. In engineering, making design decisions is analogous to making a claim about a design and supporting it with evidence

[4], [46]. Item 8 mirrors the argumentation and evidence-based reasoning literature where a claim with evidence is considered higher in hierarchy than a claim without evidence [81].

Figure 10. Item 8 - Evidence-based reasoning

| 8 | **Evidence-Based Reasoning** |
|---|---|
| | *As students develop their understanding of a STEM phenomenon, real-world problem, or design challenge, they use and evaluate evidence generated by themselves and others. This evidence is used to support their claims about phenomena and/or justify design decisions; a claim is different from a hypothesis, as a claim is supported by collected evidence and a hypothesis is a prediction.* |
| | 0. The teacher does not provide students with opportunities to make claims and/or design choices.<br>1. The teacher provides opportunities for students to make claims and/or design choices, but these claims/choices are unsupported by evidence.<br>2. The teacher requires students to make claims and/or design choices based on evidence, but does not require them to justify their reasoning.<br>3. The teacher requires students to make claims and/or design choices based on evidence and justify them using reasoning. |

*Item 9 - Technology practices in STEM.*

Item 9 (Figure 11) reflects our conceptual framework and our previous work that suggests productive ways to define the role of technology in STEM education be based on reflecting the work of STEM professionals [57]. Although there are many ways to conceptualize technology, it has been particularly challenging to define the role of technology in STEM education [16], [57]. The use of educational technology by teachers, while important, is not specific to STEM instruction. Therefore, rather than focusing on *what* technology is being used, Item 9 emphasizes on *how* the technology (whether analog or digital) is being used in ways parallel to that of STEM professionals, namely, “to record, manage, and analyze data; and to model ever more complex systems with greater precision” [4, p. 203]. All of these are reflected in the levels of Item 9, which represents tools for knowledge construction [58] through problem-solving and decision-making [28]. Furthermore, the levels become more complex the higher they go as they require students to be more cognitively engaged in the knowledge construction process.

Figure 11. Item 9 - Technology practices in STEM

| 9 | **Technology Practices in STEM** |
|---|---|
| | *Students engage in technology practices that are analogous to those used by practitioners of science, mathematics, and engineering. Students should use a variety of technological tools and techniques to identify and solve problems by creating new, useful, or imaginative solutions. Students should also develop and employ strategies for understanding the natural world in ways that leverage the power of technological methods to represent complex phenomena.* |
| | 0. Students do not use technology to collect, analyze or represent data, or create or modify scientific models and/or design solutions.<br>1. Students use technology to collect data.<br>2. Students use technology to analyze and/or represent data.<br>3. Students use digital technology to create or modify a scientific model or design solution (e.g., CAD software). |

*Item 10 - STEM career awareness.*

Item 10 (Figure 12) emphasizes the importance of raising STEM career awareness among students to promote STEM career interests with the intention to help students develop STEM identities [4], [68]. Although this item originally focused more heavily on developing STEM identities through exposure to STEM "beyond school", this was not clearly observable. By shifting the focus to explicitly sharing STEM career trajectories and their job descriptions, students can learn about such careers, which may positively influence the development of their STEM identities. Integrated STEM lessons allow teachers to intervene and help students develop more realistic perceptions of STEM as students engage in authentic STEM practices (see Item 6). Item 10 was constructed to address reform efforts and help students develop more accurate understandings of STEM careers. The levels build on one another and reflect details about STEM careers within a lesson. Such details range from simply mentioning an example of STEM careers to also sharing information about the work of STEM professionals.

Figure 12. Item 10 - STEM career awareness

| 10 | **STEM Career Awareness** |
| --- | --- |
| | *Students are made aware of STEM careers at age-appropriate levels. These opportunities may be promoted in different ways, ranging from brief mentions of types of STEM careers to explicitly relating what students are doing in class to specific STEM careers. This can be done directly by the teacher or through the teachers' active use of other resources (e.g., videos) in the room.* |
| | 0. The teacher does not promote awareness of STEM careers.<br>1. The teacher promotes awareness of STEM careers by simply naming a STEM career.<br>2. The teacher promotes awareness of STEM careers by broadly describing the types of things that STEM professionals do.<br>3. The teacher promotes awareness of STEM careers by sharing specific examples and details about one or more STEM careers. |

## Discussion

The work presented here describes the design and development of the STEM observation protocol for use in K-12 science and engineering classrooms that leverage integrated STEM instruction. Driven by our conceptual framework, our process used a combination of drawing from the literature base and available classroom observational video data to develop the protocol items. Over the course of approximately two years and various rounds of revisions and external review, the STEM-OP was reduced from an initial list of 79 codes down to 10 observable items. Considering the novelty of this protocol, we are the first to attempt any sort of instrument for use in observing integrated STEM education and in defining distinct, observable levels of these ten items. While the STEM-OP does not attempt to measure the quality of the teacher's pedagogy while implementing integrated STEM instruction as described in our ten items, it lays the groundwork to develop an instrument to do so. The following sections address our reflections on the process as well as potential uses of the final protocol, including limitations.

### Validity of the design and development process.

This iterative and rigorous process included multiple avenues to consider the validity of the final STEM-OP. For one, our initial work in reviewing the literature to develop a conceptual framework that highlights common characteristics of integrated STEM education helps to establish credibility and trustworthiness for our overall process. Each of the final ten items were reviewed and revised with respect to how well they reflected the conceptual framework. This process helped us better understand the instrument from a user experience by paying attention to only observable actions in the classroom and removing the subjectivity that may be found in other observation instruments. External to the project team, our external advisory experts provided further feedback on the overall process, but more importantly on the items themselves, helping to establish face and content validity. Writing the items was an iterative process that required separating items and reconfiguring them in meaningful ways. This overall process highlights the complexity of integrated STEM education as it is not easy to capture, and we anticipate that our instrument may not fit the needs of all those who want to teach integrated STEM in their classrooms (see Intended instrument use and Limitations sections below).

Although statistical work related to validity and reliability are underway (see Conclusions and Future Directions), our instrument shows to be internally consistent and reliable among our coders for all but one item (Table 5); however, it should be noted that Item 5 was close to reaching our agreed upon threshold for acceptability. Item 5 relates to the integration of STEM content, focusing on the extent to which the connections between different content areas are made apparent to students. This particular aspect of integrated STEM education has been seen as central to various definitions but appears to be one of the hardest to actualize [2], [15], [16], [28], [33], [49]. This item was shaped and revised numerous times throughout the process, reflecting the difficulty in capturing and defining the essence and nature of integrated STEM education. Although the item may be statistically less sound than the other nine items, its inclusion in our protocol is important for the improvement of integrated STEM education.

**Intended instrument use.**

Although the primary use of this instrument is for educational research purposes to better understand what integrated STEM education looks like in K-12 science and engineering classrooms, there are several considerations to be made. First, although each item (except for Item 9) focuses on the actions of the teacher, the observed lessons themselves should not be teacher-centered, but rather focus on student-centered pedagogies. The choice to focus on the actions of the teacher reflects the need for each item to be clearly observable with little room for subjectivity; as such, the teacher becomes the vehicle by which the curriculum is enacted.

Second, although the STEM-OP will allow an observer to note when "good STEM" occurs in the classroom through high scores on the protocol, it does not necessarily allow observers to measure the pedagogical quality. Although the quality of instruction matters, the STEM-OP was not designed to assess this aspect. Further, because no other tool yet attempts to measure integrated STEM instruction, our protocol could set the bar for specifying what exactly integrated STEM education entails. This instrument was not intended to reflect "good teaching", but rather the "STEM-iness" (the degree of STEM integration) within an observed lesson. In other words**,** our instrument is not intended to capture the pedagogical quality of STEM implementation, but the extent to which certain aspects of integrated STEM are present in a lesson (or the degree of "STEM-iness" of a lesson). Therefore, this instrument is intended to be used in conjunction with other measures (e.g., other observation instruments, debriefing or coaching discussions, examination of curricular materials, assessment of student work). For instance, although we include an item related to student collaboration, we would need a separate instrument to effectively capture effective and productive student collaboration.

Third, outside of research purposes, this protocol may be used in more formative rather than summative spaces. While the instrument generates numerical data, that is not the only intended use. The conceptual framework and protocol could be used to design pre-service courses or professional development for in-service teachers who are looking to embed integrated STEM instruction into their practice. Additionally, this instrument could be used by teacher educators and classroom coaches to provide formative feedback to pre- and in-service teachers. In addition to course development, the protocol could also be used in a coaching or mentoring conversation to identify what aspects of integrated STEM education might be missing in a lesson or repeatedly missing from a complete unit of instruction. This has the potential to alleviate tensions surrounding communication among various stakeholders by using a common language and understanding [1]. As noted above, what should be emphasized is that this instrument does not assess pedagogical quality; this is where additional instruments or qualitative comments could be utilized to complement the scores on the ten items.

**Limitations.**
The STEM-OP is not without limitations. Although our definition and conceptual framework of integrated STEM education reflect much of what is commonly addressed in the literature due to the focus on solving real-world contextualized problems, we prioritize science and engineering. This reflects the educational climate in the United States, where engineering and integrated STEM education are most often present in science classrooms due to current reforms and standards [4], [5]. Not all STEM educators adopt our focus on engineering and engineering design. For example, mathematics and computer science educators may find the STEM-OP to be flawed for their particular environments. Although this instrument was not designed for K-12 mathematics and computer science classrooms, nor was it piloted within such classrooms, it is possible that the instrument could be modified for use in those spaces. Similarly, based on our conceptual framework, we narrowed our vision of technology to focus on students' use of technology to support their learning of science, engineering, and mathematics.

Another limitation relates to the suite of videos collected through a prior project. The prior project used specific frameworks [25], [26] to develop lessons and curriculum units, which included specific engineering-centric design features, such as the use of client letters. Although our own conceptual framework incorporated these characteristics of integrated STEM education, other items within our conceptual framework were not explicit in the prior project. As a result, items related to the use of technology and raising STEM career awareness were not predominantly featured in these lessons. Further, the suite of videos did not include many examples of lower elementary (i.e., K-2) or high school (i.e., 9-12) classrooms. While the STEM-OP may be used for these grade bands, modifications may be necessary to make it more appropriate for teaching younger students, whose cognitive abilities are still developing, or for older students, whose cognitive abilities allow them to engage in abstract thinking. The video recorded observations themselves are also limited to the camera's position in the room, and the observer is limited by what they can see and hear; this can be overcome with live observations, which may offer additional visual and audio cues.

**Conclusions and Future Directions**
The STEM-OP is intended to allow its users to observe and characterize integrated STEM instruction. We acknowledge that it cannot identify or describe all aspects of an integrated STEM lesson, a fact that reflects the complex nature of integrated STEM instruction. However, the STEM-OP does attend to many of the features of integrated STEM repeated in the literature.

It allows users to reflect on these features and continue the dialogue surrounding integrated STEM education. Specifically, our instrument focuses on the observable dimensions of integrated STEM education.

Future work with the STEM-OP is still needed and we are continuing work on statistical validity and reliability through exploratory and confirmatory factor analysis. Other work related to the usability of the protocol is also needed. Specifically, although the expected use of the STEM-OP will be primarily for a single, one-day classroom observation (reflecting current observation practices such that observations typically occur on a day-by-day basis), we expect that collective observations over the course of a conceptually coherent lesson or unit of instruction will be illuminating. This line of work is influenced by our previous work in which we noted that over the course of a curriculum unit, teachers attend to various aspects of STEM education differently throughout implementation [14], [49]. Additionally, continued research with the STEM-OP also includes better understanding of differences across grade levels and content areas (e.g., physical versus life science). This research will help the education community better comprehend the complexity of integrated STEM education with the eventual goal of exploring student outcomes, whether achievement or affect based. The STEM-OP provides the foundation for identifying observable components of integrated STEM education, making this goal achievable.

In addition to on-going research, we are developing STEM-OP training procedures for prospective users of the instrument. As part of this, we have begun to develop a set of user guidelines that would assist future users in understanding each items' intention, including non-exhaustive examples of how to score or what to look for with respect to each scoring level for any given item. This set of user guidelines will support new users in better understanding each item. Furthermore, this on-going work will be made available to educators and educational researchers wishing to better understand the instrument provided here and learn how to use it in classroom settings as part of a formal training program.

**Acknowledgements**
This work was supported by the National Science Foundation under Award Numbers 1854801, 1812794, and 1813342. The findings, conclusions, and opinions herein represent the views of the authors and do not necessarily represent the view of personnel affiliated with the National Science Foundation.

**References**

[1] E. A. Dare, E. A. Ring-Whalen, and G. H. Roehrig, "Creating a continuum of STEM models: Exploring how K-12 science teachers conceptualize STEM education," *International Journal of Science Education*, vol. 41, no. 12, pp. 1701-1720, Jul. 2019.

[2] National Academy of Engineering and National Research Council. *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*. Washington, DC, USA: National Academies Press, 2014.

[3] T. J. Moore, A. C. Johnston, & A. W. Glancy, "STEM integration: A synthesis of conceptual frameworks and definitions," in *Handbook of Research in STEM Education,* C. C. Johnson, M. J. Mohr-Schroeder, T. J. Moore, & L. D. English, Eds. New York, USA: Routledge, 2020, pp. 3-16.

[4] National Research Council, *A Framework for K-12 Science Education: Practice, Crosscutting Concepts, and Core Ideas*. Washington, DC, USA: National Academies Press, 2012.

[5] NGSS Lead States, *Next Generation Science Standards: For States, by States.* Washington, DC, USA: National Academies Press, 2013.

[6] J. A. Luft, J. M., Diamond, C. Zhang, and D. Y. White, "Research on K-12 STEM professional development programs," in *Handbook of Research in STEM Education,* C. C. Johnson, M. J. Mohr-Schroeder, T. J. Moore, & L. D. English, Eds. New York, USA: Routledge, 2020, pp. 337-348.

[7] S. L. Bartels, K. M. Rupe, and J. S. Lederman, "Shaping preservice teachers' understandings of STEM: A collaborative math and science methods approach," *Journal of Science Teacher Education*, vol. 30, no. 6, pp. 666-680, Apr. 2019.

[8] E. A. Ring, E. A., Dare, E. A., Crotty, and G. H. Roehrig, "The evolution of teacher conceptions of STEM education throughout an intensive professional development experience," *Journal of Science Teacher Education,* vol. 28, no. 5, pp. 444-467, Aug. 2017.

[9] G. H. Roehrig, E. A. Dare, J. A. Ellis, E. A. Ring-Whalen, and M. D. Rouleau, in preparation.

[10] Bybee, R. W. *A Case for STEM Education*. Arlington, VA, USA: National Science Teachers' Association Press, 2013.

[11] J. M. Breiner, S. S. Harkness, C. C. Johnson, and C. M. Koehler, "What is STEM? A discussion about conceptions of STEM in education and partnerships," *School Science and Mathematics,* vol. 112, no. 1, pp. 3-11, Jan. 2012.

[12] J. B. Labov, A. H. Reid, and K. R. Yamamoto, "Integrated biology and undergraduate science education: A new biology education for the twenty-first century?", *CBE Life Science Education,* vol. 9, pp. 10–16, Oct. 2010.

[13] M. Sanders, "STEM, STEM education, STEM mania," *Technology Teacher,* vol. 68, no. 4, pp. 20–26, Jan. 2009.

[14] E. A. Ring-Whalen, E. A. Dare, G. H. Roehrig, P. Titu, and E. A. Crotty, "From conception to curricula: The role of science, technology, engineering, and mathematics in integrated STEM units," *International Journal of Education in Mathematics, Science and Technology,* vol. 6, no. 4, pp. 343-362, 2018.

[15] L. D. English, "STEM education in K-12: Perspectives on integration," *International Journal of STEM Education,* vol. 3, no. 1, pp. 1-8, Mar. 2016.

[16] D. R. Herschbach, "The STEM initiative: Constraints and challenges," *Journal of STEM Teacher Education,* vol. 48, no. 1, pp. 96-112, Mar. 2011.

[17] C. R., Rinke, W. Gladstone-Brown, C. R. Kinlaw, and J. Cappiello, J, "Characterizing STEM teacher education: Affordances and constraints of explicit STEM preparation for elementary teachers," *School Science and Mathematics,* vol. 116, no. 6, pp. 300–309, Oct. 2016.
[18] D. Sawada, M. D. Piburn, E. Judson, J. Turley, K. Falconer, R. Benford, and I. Bloom, "Measuring reform practices in science and mathematics classrooms: The reformed teaching observation protocol," *School Science and Mathematics*, vol. 102, no. 6, pp. 245-253, Oct. 2002.
[19] C. Walkington, P. Arora, S. Ihorn, J. Gordon, M. Walker, L. Abraham, L., and M. Marder, "Development of the UTeach observation protocol: A classroom observation instrument to evaluate mathematics and science teachers from the UTeach preparation program," Unpublished manuscript, 2012.
[20] M. K. Smith, F. H. Jones, S. L. Gilbert, and C. E. Wieman, "The classroom observation protocol for Undergraduate STEM (COPUS): A new instrument to characterize university STEM classroom practices," *CBE-Life Sciences Education*, vol. 12, no. 4, pp. 618-627, Dec. 2013.
[21] RMC Research Corporation, *Science Classroom Observation Protocol: Washington State's Vision of Effective Science Learning Experiences for Students*, 2010.
[22] E. Dringenberg, R. Wertz, S. Purzer, and J. Strobel, "Development of the science and engineering classroom learning observation protocol," In *Proc. 2012 ASEE Annual Conf. & Expo.,* (San Antonio, TX, USA), June 2012, 25.4651-25.465.8
[23] L. B. Wheeler, S. L. Navy, J. L. Maeng, and B. A Whitworth, "Development and validation of the classroom observation protocol for engineering design (COPED)," *Journal of Research in Science Teaching,* vol. 56, no. 9, pp. 1285-1305, May 2019.
[24] J. Allen, A. Gregory, A. Mikami, J. Lun, B. Hamre, and R. Pianta, "Observations of effective teacher-student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system-secondary," *School Psychology Review*, vol. 42, no. 1, p. 76-98, Mar. 2013.
[25] T. J. Moore , M. S. Stohlman, H.-H. Wang, K. M. Tank, A. W. Glancy, and G. H. Roehrig, "Implementation and integration of engineering in K–12 STEM education ," In *Engineering in Pre-College Settings: Synthesizing Research, Policy, and Practices*, S. Purzer, J. Strobel, and M. Cardella, Eds. West Lafayette, IN, USA: Purdue University Press, 2014.
[26] T. J. Moore, A. W. Glancy, K. M. Tank, J. A. Kersten, and K. A. Smith, "A framework for quality K-12 engineering education: Research and development," *Journal of Pre-College Engineering Education Research*, vol. 4, no. 1, pp. 1-13, 2014.
[27] C. M. Cunningham and W. S. Carlsen, "Teaching engineering practices," *Journal of Science Teacher Education,* vol. 25, no. 2, pp. 197-210, Feb. 2014.
[28] International Society for Technology in Education, "ISTE Standards," ISTE Standards. https://www.iste.org/standards
[29] National Council of Teachers of Mathematics, *Principles to Actions: Ensuring Mathematical Success for All*, NCTM, 2014.
[30] R. Fogarty, "Ten ways to integrate the curriculum," *Educational Leadership*, vol. 49, no. 2, pp. 61-65, Oct. 1991.
[31] R. W. Bybee, "Advancing STEM education: A 2020 vision," *Technology and Engineering Teacher*, vol. 70, no. 1, pp. 30–35, Sept. 2010.

[32] R. Brown, J. Brown, K. Reardon, and C. Merrill, “Understanding STEM: Current perceptions,” *Technology and Engineering Teacher,* vol. 20, no. 6, pp. 5-9, Mar. 2011.
[33] T. R. Kelley, and J. G. Knowles, “A conceptual framework for integrated STEM education,” *International Journal of STEM Education,* vol. 3, no. 11, pp. 1-11, July 2016.
[34] C. Lachapelle, and C. Cunningham, C. “Engineering in elementary schools,” In *Engineering in Pre-College Settings: Synthesizing Research, Policy, and Practices*, S. Purzer, J. Strobel, & M. Cardella, Eds. West Lafayette, IN, USA: Purdue University Press, 2014, pp. 61-88.
[35] D. Monson, D., and D. Besser, “Smashing milk cartons: Third-grade students solve a real-world problem using the engineering design process, collaborative group work, and integrated STEM education,” *Science and Children*, vol. 52, no. 9, pp. 38-43, Summer 2015.
[36] C. Hill, C. Corbett, and. A. St. Rose, A. *Why So Few? Women in Science, Technology, Engineering, and Mathematics,* AAUW, 2010.
[37] S. Vakil, S., and R. Ayers, R., “The racial politics of STEM education in the USA: Interrogations and explorations,” *Race Ethnicity and Education*, vol. 22, no. 4, pp. 449–458, Apr. 2019.
[38] C. Djonko-Moore, J. Leonard, Q. Holifield, E. Bailey, E., and S. Almughyirah, “Using culturally relevant experiential education to enhance urban children’s knowledge and engagement in science,” *Journal of Experiential Education*, vol. 41, no. 2, pp. 137–153, Nov. 2017.
[39] L. Moll, C. Luis, C. Amanti, D. Neff, and N. Gonzalez, “Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms,” *Theory into Practice*, vol. 31, no. 2, pp. 132–141, Spring 1992.
[40] B. R. Upadhyay, “Using students’ lived experiences in an urban science classroom: An elementary school teacher’s thinking,” *Science Education*, vol. 90, no. 1, pp. 94–110, Dec. 2005.
[41] National Research Council, *Successful K-12 STEM education: Identifying Effective Approaches in Science, Technology, Engineering, and Mathematics*. Washington, D.C., USA: National Academies Press, 2011.
[42] L. A. Bryan, T. J. Moore, C. C. Johnson, and G. H. Roehrig, “Integrated STEM education,” In *STEM Roadmap: A Framework for Integrated STEM Education*, C. C. Johnson, E. E. Peters-Burton, and T. J. Moore, Eds., New York, NY, USA: Routledge, 2015, pp. 23-37.
[43] C. E. Hmelo, D. Holton, and J. L. Kolodner, “Designing to learn about complex systems,” *The Journal of the Learning Sciences,* vol. 9, no. 3, pp. 247–298, 2000.
[44] M. Mehalik, Y. Doppelt, and C. D. Schunn, “Middle school science through design-based learning versus scripted inquiry: Better overall science concept learning and equity gap reduction,” *Journal of Engineering Education*, vol. 97, no. 1, pp. 71-85, Jan. 2008.
[45] P. M. Sadler, H. P. Coyle, and M. Schwartz, “Engineering competitions in the middle school classroom: Key elements in developing effective design challenges,” *The Journal of the Learning Sciences,* vol 9, pp. 299–327, 2000.
[46] E. A. Siverling, A. Suazo-Flores, C. A. Mathis, and T. J. Moore, “Students' use of STEM content in design justifications during engineering design-based STEM integration,” *School Science and Mathematics*, vol. 119, pp. 457–474, Dec. 2019.
[47] C. A. Mathis, E. A. Siverling, T. J. Moore, K. A. Douglas, and S. S. Guzey, “Supporting engineering design ideas with science and mathematics: A Case Study of Middle School

Life Science Students," *International Journal of Education in Mathematics, Science and Technology*, vol. 6, pp. 424–442, 2018.
[48] C. A. Mathis, E. A. Siverling, A. W. Glancy, S. S. Guzey, and T. J. Moore, "Students' use of evidence-based reasoning in K-12 engineering: A case study (Fundamental)," in *Proc. 2016 ASEE Annual Conference & Expo.* (New Orleans, LA, USA), Jun. 2016.
[49] E. A. Dare, J. A., Ellis, G. H. Roehrig, "Understanding science teachers' implementations of integrated STEM curricular units through a phenomenological multiple case study," *International Journal of STEM Education*, vol. 5, no. 4, pp. 1-19, Feb. 2018.
[50] E. D. Baldinger, S. Staats, L. M. Covington-Clarkson, E. Gullickson, F. Norman, and B. Akoto, "Returning voice to the silent M: A review of conceptions of mathematics in Integrated STEM Education," in *Integrated Approaches to STEM Education: An International Perspective*, J. Anderson and Y. Li, Eds., in press.
[51] L. Coad, "The M in STEM: What is it really?," *Australian Mathematics Teacher,* vol. 72, no. 2, pp. 3-6, Summer 2016.
[52] J. Frykholm, and G. Glasson, "Connecting science and mathematics instruction: Pedagogical context knowledge for teachers," *School Science and Mathematics,* vol. 105, no. 3, pp. 127-141, Mar. 2005.
[53] W. S. Walker, III, "Integrated STEm or integrated STEM?," *School Science and Mathematics,* vol. 117, no. 6, pp. 225-227, Oct. 2017.
[54] G. H. Roehrig, E. A. Dare, E. A., Ring-Whalen, and J. R. Wieselmann, "Understanding coherence and integration in integrated STEM curriculum," *International Journal of STEM Education*, vol. 8, no. 2, pp. 1-21, Jan. 2021.
[55] W. A. Aminger, S. Hough, S. A. Roberts, V. Meier, A. D. Spina, H. Pajela, M. McLean, and J. A. Bianchini, "Preservice secondary science teachers' implementation of an NGSS practice: Using mathematics and computational thinking," *Journal of Science Teacher Education,* vol. 32, no. 2, pp. 188-209, Sep. 2020.
[56] M. Hoda Wilkerson, and M. Fenwick, "Using mathematics and computational thinking." In *Helping students make sense of the world: Using next generation science and engineering practices*, C. V. Schwarz, C. Passmore, and B. J. Reiser, Eds. Arlington, VA: National Science Teachers Association, 2017, pp. 181-204.
[57] J. A. Ellis, Wieselmann, J. R., R. Sivaraj, G. H. Roehrig, E. A. Dare, and E. A. Ring-Whalen, "Toward a productive definition of technology in science and STEM education," *Contemporary Issues in Technology and Teacher Education – Science*, vol. 20, no. 3, pp. 472-496, 2020.
[58] D. H. Jonassen, "Technology as cognitive tools: Learners as designers," Unpublished paper from *IT Forum*, ID 110390460, 1994.
[59] T. J. Kennedy, and M R. L. Odell, "Engaging Students in STEM Education," *Science Education International,* vol. 25, no. 3, pp. 246–258, 2014.
[60] J. R. Wieselmann, E. A. Dare, E. A. Ring-Whalen, and G. H. Roehrig, "'I just do what the boys tell me': Exploring small group student interactions in an integrated STEM unit," *Journal of Research in Science Teaching*, vol. 57, no. 1, pp. 112-144, Jan. 2020.
[61] J. R. Wieselmann, E. A. Dare, E. A. Ring-Whalen, and G. H. Roehrig, "'There are other ways to help besides using the stuff': Using activity theory to understand dynamic student participation in small group STEM activities," *Journal of Research in Science Teaching*, 2021.

[62] National Education Association, *Preparing 21st Century Students for a Global Society: An Educator's Guide to the "Four Cs"*, Washington, D.C., USA: National Education Association, 2012.
[63] D. D. Minner, A. J. Levy, and J. Century, "Inquiry-based science instruction—What is it and does it matter? Results from a Research Synthesis Years 1984 to 2002," *Journal of Research in Science Teaching,* vol. 47, no. 4, pp. 474–496, Apr. 2010.
[64] B. M. Reynante, M. F. Selbach-Allen, and D. R. Pimentel, "Exploring the promises and perils of integrated STEM through disciplinary practices and epistemologies," *Science & Education*, vol. 29, pp. 785–803, Apr. 2020.
[65] P. A. Asunda, "A conceptual framework for STEM integration into the curriculum through career and technical education," *Journal of STEM Teacher Education*, vol. 49, no. 1, pp. 3-16, Spring 2014.
[66] H. Petroski, "To engineer is human: The role of failure in successful design," *Defense AR Journal,* vol. 23, no. 1, pp. 106-109, 2016.
[67] E. Simpson, D. Bradley, and J. O'Keeffe, "Failure is an option: An innovative engineering curriculum," *International Journal of Building Pathology and Adaptation*, vol. 36, no. 3, pp. 268-282, July 2018.
[68] L. Archer, J. Dewitt, J. Osborne, J. Dillon, B. Willis, and B. Wong, "'Doing' science versus 'being' a scientist: Examining 10/11-year-old schoolchildren's construction of science through the lens of identity," *Science Education*, vol. 94, no. 4, pp. 617-639, Apr. 2010.
[69] J. D. Lee, "Which kids can 'become' scientists? Effects of gender, self-concepts, and perceptions of scientists," *Social Psychology Quarterly*, vol. 61, no. 3, pp. 199-219, Sept. 1998.
[70] J. Corbin, and A. Strauss, *Basics of Qualitative Research*, Thousand Oaks, CA, USA: SAGE, 2015.
[71] J. Saldana, *The Coding Manual for Qualitative Researchers*, Thousand Oaks, CA, USA: SAGE, 2013.
[72] K. Krippendorff, "Computing Krippendorff's alpha-reliability," University of Pennsylvania Departmental Papers Repository, Jan. 2011.
[73] C. A. Buxton, "Social problem solving through science: An approach to critical, place-based, science teaching and learning," *Equity & Excellence in Education*, vol. 43, no. 1, pp. 120-135, Feb. 2010.
[74] L. W. Anderson and D. R. Krathwohl, Eds. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*, New York, NY, USA: Longman, 2001.
[75] A. Sharunova, Y. Wang, M. Kowalski, and A. J. Qureshi, "Applying Bloom's taxonomy in transdisciplinary engineering design education," *International Journal of Technology and Design Education*, Sept. 2020.
[76] P. A., Asunda, and S. Ware, "Applying the congruence principle of Bloom's taxonomy to develop an integrated STEM experience through engineering design," *The Journal of Technology Studies,* vol. 41, no. 2, pp. 88-100, Nov. 2015
[77] P. Dolog, L. Thomsen, and B. Thomsen, "Assessing problem-based learning in a software engineering curriculum using Bloom's taxonomy and the IEEE software engineering body of knowledge," *ACM Transactions on Computing Education*, vol. 16, no. 3, pp. 1-41, May 2016.

[78] E. Miller, E. Manz, R. Russ, D. Stroupe, and L. Berland, "Addressing the epistemic elephant in the room: Epistemic agency and the Next Generation Science Standards," *Journal of Research in Science Teaching,* vol. 55, no. 7, pp. 1053-1075, Apr. 2018.

[79] L. Thibaut, S. Ceuppens, H. De Loof, J. De Meester, L. Goovaerts, A. Struyf, J. Boeve-de Pauw, W. Dehaene, J. Deprez, M. De Cock, L. Hellinckx, H. Knipprath, G. Langie, K. Struyven, D. Van de Velda, P. Van Petegem, and F. Depaepe, "Integrated STEM education: A systematic review of instructional practices in secondary education," *European Journal of STEM Education*, vol. 3, no. 1, pp. 1-12, May 2018.

[80] K. B. Wendell, C. G. Wright, and P. Paugh, "Reflective decision-making in elementary students' engineering design," *Journal of Engineering Education,* vol. 106, no. 3, pp. 356-397, July 2017.

[81] E. M. Furtak, I. Hardy, C. Beinbrech, R. J. Ravelson, and J. T. Shermwell, "A framework for analyzing evidence-based reasoning in science classroom discourse," *Educational Assessment*, vol. 15, no. 3-4, pp. 175-196, Dec. 2010.

[82] V. Malo-Juvera, R. Powell, and S. Cantrell, "Development, validation, and factor analysis of the culturally responsive instruction observation protocol," AERA Online Paper Repository, 2013.

[83] R. Powell, S. C. Cantrell, V. Malo-Juvera, and P. Correll, "Operationalizing culturally responsive instruction: Preliminary findings of CRIOP research," *Teachers College Record,* vol. 118, no. 1, pp. 1-46, 2016.