Selecting and Designing Assessment to Measure Early Engineering Curriculum Impact on Application of Engineering Design and Attitudes Toward STEM

Dr. Kristin Kelly Frady, Clemson University

Kris Frady is an Assistant Professor in the Departments of Educational and Organizational Leadership and Development and Engineering and Science Education at Clemson University and Faculty Director for Clemson University Center for Workforce Development (CUCWD) and the National Science Foundation Advanced Technological Education Center for Aviation and Automotive Technological Education using Virtual E-Schools (CA2VES). Her research and experiences include implementation of digital learning solutions, development of career pathways including educator professional development, and analysis of economic development factors impacting education and workforce development. Kris earned an Ed.D. in Curriculum and Instruction in Education Technology from the University of Florida and also holds business and teaching degrees from Clemson University and the University of South Carolina. Kris has previously worked as a professional corporate trainer, as an instructor at a two-year institution, and as a career and technology education teacher at the high school level. Kris and the CUCWD/CA2VES team aim to develop an exchange which will help identify, understand, and elevate best practices; facilitate peer learning and deepen knowledge; spark replication and advancement across regions; and inform long-term public investment in talent development through research and educational resources.

Mr. Ikenna Q. Ezealah, Clemson University

Mr. Christopher Ciucia, SAE International

Chris Ciucia is the Director of Pre-Professional Education at SAE International. He oversees the strategic direction of SAE’s programmatic offerings at the K-16 level, including the National Science Board Award winning A World In Motion and Collegiate Design Series Programing. Chris leads numerous U.S.-based and global initiatives designed to increase science, technology, engineering and math (STEM) engagement, participation and achievement at the K-16 level.

As a recognized "STEM Expert", Chris serves in advisory roles for National Science Foundation, National Academy of Engineering and other private and publically funded STEM projects, is a published author on various educational strategies and techniques for delivering classroom-based instruction and has served as an expert on multiple STEM panels and through other media outlets.

As a former classroom teacher, Chris received peer and student-based recognition through honors such as school district-wide Teacher of the Year. Chris strives to bring his classroom experience into all efforts designed to deliver high-quality integrated STEM instruction.

Mrs. Meghan M. Stoyanoff, SAE International

Meghan Stoyanoff is a Pre-College Program Developer for the A Wold In Motion Program (AWIM). In this role, Meghan oversees several special initiatives such as Nissan and FCA and coordinates educational programming in schools in accordance with funds provided through these companies. Meghan is the lead on AWIM’s Pre-K curriculum development and is a member of NAEYC. Meghan has a background in education and has worked for several years inside and outside the classroom teaching, tutoring, and writing curriculum. Meghan’s background in science education and curriculum writing contribute to her ability to deliver effective professional development training to both teachers and industry professionals that volunteer in the classroom. Her goal is to provide educators with the tools they need to deliver effective, hands-on STEM instruction in the classroom.
**Selecting and Designing Assessment to Measure Early Engineering Curriculum Impact on Application of Engineering Design and Attitudes Toward STEM**

**I. Introduction**

STEM education is vital to the future of our country, state, and children. In fact, the United States Department of Commerce has reported that Science, Technology, Engineering and Math (STEM) jobs are growing at a rate of 17%, while other occupations are growing at only 9.8%. Engaging students early in STEM education is critical to building motivation and interest in STEM educational and career pathways. Programs providing early enrichment, engagement and support have been proven to be ideal gateways into interest and development of STEM career pathways.

SAE International’s “A World in Motion (AWIM)” program provides students and teachers with innovative experiential learning opportunities, in the classroom, that require students to transfer academic knowledge and experience to solving real-world engineering challenges. More than 75,000 curriculum challenge kits have been provided with over 5 million students and 35,000 volunteers having participated. The AWIM program, aligned with K-8 national standards, incorporates integrated STEM learning experiences through hands-on activities that reinforce classroom STEM learning.

New curricula in STEM education, such as AWIM, provide opportunities to examine an experiential learning curriculum's impact on early elementary through middle school students specifically in areas of student (1) application of engineering design as an engineering approach in STEM education and (2) knowledge of and attitudes towards STEM-related fields, careers, and educational opportunities.

This paper examines a process by which assessment tools were selected and designed to measure early engineering curriculum impact on application of engineering design, attitudes, and motivations toward STEM.

**II. Literature Review**

In an era of increased complexity where industry necessitates the synthesis of various disciplines to continually advance, STEM (Science, Technology, Engineering, and Math) stands at the forefront. The naturally expected involvement of STEM causes two great needs: 1) increased demand for workers, and 2) more programs that awaken interest in the profession. However, the academic achievement of students in elementary and middle school as it relates to STEM disciplines through integrated methods is a growing concern. To support the impetus for heightened achievement, the NGSS (Next Generation Science Standards) were developed. The NGSS are a framework for K-12 education that is based on the most current research, and it defines the three dimensions of what it means to be proficient in science: science and engineering practices, crosscutting concepts, and core ideas.

The NGSS standards along with the effort of educators require the partnership of STEM organizations to address the ultimate need of more highly-qualified workers. There are a growing
number of organizations creating STEM curriculums that aim to increase interest within youth about identifying and pursuing STEM as a career. One of these organizations is SAE International. SAE International originally developed the AWIM (A World in Motion) program in 1990. Over the past 27 years, the AWIM program has evolved to meet growing and future research-based educational needs. AWIM is benchmarked to the NGSS and aims to bring STEM education to life in the classrooms. In order to select and design assessments which are able to measure STEM related experiential learning programs such as AWIM through the framework of NGSS, it was necessary to critically examine literature and prior research of evidence-based STEM assessments utilized in elementary and middle school environments. This literature review provides an overview of elementary and middle school engineering design curriculum, the AWIM program, NGSS, and grounds those concepts in an empirical framework, building on the successes of previous STEM program assessments and evaluations.

**Engineering Design Curriculum in Elementary and Middle School**

The increased involvement of STEM literacy in all industries makes its influence in the future world economy significant. To accommodate this inevitable expansion, schools need to help students increase their performance in certain subjects, as well as expose children to engineering curriculum. If performance is key, 2015 data from the National Assessment of Educational Progress (NAEP) indicates that STEM student achievement in K-12 is underwhelming. Approximately 60% of 4th graders and 77% of 8th graders were not proficient in math. In science, 62% of 4th graders and 66% of 8th graders were not proficient. Abel (2015) mentions that, recognizing the shortage of STEM workers and the underperforming K-12 students, the federal budget of 2011 included $3.7 billion for STEM education policy and $4.3 billion for Race-to-the-Top programs. Against this backdrop, there is increased effort across the board to prepare students for STEM majors.

A recent report of the NAE and the NRC confirms the major endeavor: “Today, several dozen different engineering programs and curricula are offered in school districts around the country, and thousands of teachers have attended professional development sessions to teach engineering-related coursework” (NAE & NRC, 2009, p. 1). Numerous programs are developing curricula focused on engineering to address future workforce need. Some of the more notable programs are A World in Motion (AWIM), Engineering is Elementary (EiE), Project Lead the Way (PLTW), and Full Option Science System (FOSS).

**Overview of A World in Motion**

AWIM was initially developed as a collection of teacher lesson plans and has expanded into today’s program which is a K-8 STEM education solution that is a teacher-administered, industry volunteer-assisted program that brings science, technology, engineering and math (STEM) education to life in the classroom for students. Benchmarked to the national education standards, the AWIM program incorporates integrated STEM learning experiences through hands-on activities that reinforce classroom STEM learning. The program serves K-8 schools, teachers, and students nationwide in STEM related classrooms. More than 75,000 curriculum/challenge kits have been provided with over 5 million students and over 35,000 volunteers having participated.
Quality STEM learning incorporates facets from multiple disciplines. By design, AWIM integrates STEM subjects through an approach identified as the Engineering Design Experience (EDE). The EDE places students in the roles of scientists as they explore an AWIM Engineering Design Challenge. The EDE provides a problem-solving context in which students design a product or devise a solution to a problem throughout the teacher-facilitated challenge. Teams of students examine who the product is for and what must be accomplished; gather and synthesize information; design, develop, and test a prototype design; and prepare a presentation of their design ideas. The specific detail of each step of the EDE is outlined below in Table 1.

Table 1
The AWIM Engineering Design Experience (EDE)

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set Goals</td>
<td>Students are introduced to a challenge scenario. They review a toy company’s Request for Proposal letter, investigate and discuss what is asked of them, and share ideas on how to go about solving the problem. Students set goals and begin to work in teams, recording their work in design logs.</td>
</tr>
<tr>
<td>Build Knowledge</td>
<td>Students work through a variety of structured and age-appropriate activities to build knowledge and skills they will need to design their own solutions or products. First, students build a physical model and test how well it works. In the next several activities, teams change variables on the model, record observations, and discuss results as a team and as a class. They move from simple explorations and opinions to controlled experiments and performance predictions based on the graphs or tables of results they create. Design Student teams design their own toy to meet the requirements stated in the toy company’s letter. They determine the values of variables, plan construction, and predict performance based on knowledge from previous activities and data collected.</td>
</tr>
<tr>
<td>Build &amp; Test</td>
<td>Student teams build and test their design to see how well it meets the performance criteria. Design teams optimize performance based on previous knowledge gained. Students make certain their solution meets the design criteria outlined in earlier phases of the EDE. Students begin to build a case for why they made their design decisions in preparation for the team presentation of their solution.</td>
</tr>
<tr>
<td>Present</td>
<td>Student teams present their work to an audience (i.e., their class, whole school, parents, and/or community) and provide arguments for why their solution/design meets the design criteria.</td>
</tr>
</tbody>
</table>

Students complete the EDE for each design challenge; all are aligned to specific age/grade-appropriate math and science standards. This structure is important because it facilitates teachers bringing a real-life inquiry-based approach into the classroom setting.
Next Generation Science Standards

To help create a standard of high achievement, the National Research Council (NRC), the staff arm of the National Academy of Sciences, developed a Framework for K-12 Science Education, which is based on the most current research on science and science learning that identifies the science all students should know. The NRC Framework that describes what it means to be proficient in science has three dimensions: science and engineering practices, crosscutting concepts, and core ideas in each science discipline (Drape, Lopez, Radford, 2016). States subsequently used this framework to develop K-12 science standards that are rich in content. Dimension 1 describes the practices scientists employ as they investigate and build models and theories about the natural world, and it also describes the key set of engineering practices that engineers use as they design and build models and systems (Quinn, Schweingruber, & Keller, 2012).

Crosscutting concepts apply across all domains of science and are thus a way of linking them. Concepts that include “patterns, similarity, and diversity; Cause and effect; Scale, proportion and quantity; Systems and system models; Energy and matter; Structure and function; Stability and change” (Duschl, 2012, p. 1). Disciplinary Core Ideas are meant to focus K-12 science curriculum, instruction, and assessment on the most important aspects of science. The ideas are to meet at least two of the following four criteria to be considered core: 1) Have broad importance across multiple sciences or engineering disciplines; 2) Provide a key tool for understanding or investigating more complex ideas and solving problems; 3) Relate to the interests and life experiences of students or be connected to societal or personal concerns that require scientific or technological knowledge; 4) Be teachable and learnable over multiple grades at increasing levels of depth and sophistication.

Evidence-Based STEM Assessment Frameworks

It is encouraging that evidence of the impact of structured informal learning environments’ ability to positively influence learning outcomes, attitudes towards STEM and likelihood that students will pursue STEM academic programs or careers is beginning to develop (Bell, 2009). However, there is limited data which empirically links the impact of engineering integrated education with improved learning and achievement in STEM (Yoon, Lucietto, Capobianco, Dyehouse, & Diefes-Dux, 2014; Wendell & Rogers, 2013). Coupled with the lack of data is the additional challenge that there are a limited number of research validated assessment tools designed to measure the learning outcomes of these types of environments. Of few the existing tools, they are specifically content and program aligned and not well suited for use in assessing learning gains of other STEM programs nor are they specifically aligned with NGSS in engineering design. Additionally, consideration of grade-level-appropriate contexts requires assessments across multiple levels. Classroom-embedded assessment tools based on NGSS have the potential to significantly impact assessment systems designed to evaluate student knowledge and performance following instruction (National Research Council, 2014).

Prior research studies which have been conducted in the broader STEM and NGSS assessment realms have produced suggested frameworks and practices for development of high-quality STEM assessments. Assessments in structured informal STEM environments should not contain only factual recall or other low cognition measures but should address a broad range of competencies and should provide authentic evidence of learning and abilities (Bell, 2009).
Specifically, NGSS assessments should reflect the three dimensions: disciplinary core ideas, crosscutting concepts, and science and engineering practices (Bell, 2009). To demonstrate competence of NGSS performance expectations, students must combine disciplinary core knowledge, crosscutting concepts and application of science and engineering practice. Further, no assessment tools exist which identify assessment strategies aligned specifically with NGSS engineering design standards K-8 and broader NGSS core ideas of engineering design which include: (1) defining and delimiting engineering problems, (2) designing solutions to engineering problems and (3) optimizing the design solution (NAP, 2013). The combination of these recommendations and frameworks suggest that development of robust assessment of engineering design NGSS in structured informal STEM environments must be multidimensional, span multiple levels of cognition and incorporate various STEM assessment strategies.

Defining and delimiting engineering problems.

Students who have mastered this core engineering design idea have the ability to (K-2) “identify situations that people want to change as problems that can be solved through engineering”, (3-5) “specify criteria and constraints that a possible solution to a simple problem must meet” and (6-8) “attend to precision of criteria and constraints and considerations likely to limit possible solutions” (NAP, 2013).

Designing solutions to engineering problems.

To experience and learn about scientific and engineering practices students must actively engage in practices which are used by scientists and engineers and can include modeling and developing explanations or solutions (National Research Council, 2014). Images are created to make sense of everyday experiences (Wilson & Wilson, 1977) and for a traditional pencil-and-paper assessment, asking students to draw and provide a written explanation of a model solving an engineering problem can provide both qualitative and quantitative insight into student thinking. Previously, assessments such as the “Draw a Scientist Test” or “Draw an Engineer Test” have been effectively used to quantify student understanding and perceptions (Barman, 1999; Diefes-Dux, 2015; Finson, 2002). Assessments which incorporate elements of drawing and explaining a model, and explaining the process for testing and improving the model help to meet the core engineering design competencies where students are able to from K-2 convey possible solutions through visual or physical representations; in grades 3-5 research and explore multiple possible solutions; and for grades 6-8 combine parts of different solutions to create new solutions.

Optimizing the design solution.

Students’ ability to interpret information for making claims about data and to reflect on the impact of different variables may be measured through use of a task set. This task set may include a variety of types of assessment questions ranging from discrete to short-answer to constructed-response all which lead a student to developing an extended response for a complex task or question (National Research Council, 2014). Further, assessments which challenge students to reason about data representation (Lehrer et al., 2013) provide insight into the student’s ability to interpret and manipulate data. Students who have mastered the core engineering design idea of optimizing the design solution are able to at K-2 compare solutions, test them, and evaluate each; in grades 3-5 improve a solution based on results of simple tests, including failure points; and in grades 6-8 use systematic processes to iteratively test and refine a
solution. Further, in order to make claims about student learning no single assessment strategy or piece of evidence is sufficient where instead assessment should incorporate a pattern of evidence across multiple components (National Research Council, 2001).

III. Theoretical Framework

Learning environments are complex and have many different dimensions. Evidence-based learning environments and curricula, designed with authentic learning experiences and workforce outcomes in mind, utilize contemporary educational theories of learning. Assessment of specific outcomes across multiple diverse learning environments can be meaningfully done by integrating theoretical frameworks aligned with the specific aims. Two leading theories frame the development of assessment and evaluation tools utilized in this study: Kolb’s Experiential Learning Theory and Lent, Brown, and Hackett’s Social Cognitive Career Theory.

Assessing Application of Engineering Design Approaches

The AWIM curriculum emphasizes the importance of active learning experiences paired with goal setting, observations and reflections to build understanding, testing, application of ideas, and presentation. This structure suggests an experiential learning framework where students take an active role in applying their knowledge to authentic, real-world learning experiences. Experiential learning is “the process whereby knowledge is created through the transformation of experience. Knowledge results from the combination of grasping and transforming experience” (Kolb, 1984, p.41). This theory also has strong foundations in cognitive learning (Piaget, 1964) and pragmatism (Dewey, 1933). Kolb’s (1984) learning cycle explains learning as a cycle of concrete experiences, reflective observations, abstract conceptualization (analysis and generalization), and active experimentation which then result in new experiences. An application of experiential learning is problem-based learning (PBL) where PBL initiates student learning through the need to solve an authentic, real-world, and ill-formed problem and facilitates student problem-solving and self-directed learning skills (Hung, Jonassen, & Liu, 2008) as students research, analyze, and synthesize ideas to solve the problem.

Experiential learning through the lens of a PBL framework is well suited to support assessment of student awareness of using engineering design as an approach to STEM education. Key factors for assessment of this outcome with respect to experiential learning theories are opportunities for students to demonstrate ability to ask questions, make observations, gather information, and define ill-formed problems; analyze materials, data and variables; and develop models and tests of solutions. Not only do these factors align well with AWIM curriculum and experiential learning frameworks but also facilitate seamless integration of NGSS engineering design standards.

Assessing Student Knowledge of, Attitudes Toward, and Motivations to Pursue STEM

Prior studies have demonstrated that student knowledge of, attitudes toward, and motivation to pursue STEM education can not only positively impact academic performance but also indicate future career interest and participation in the STEM workforce (Unfried, Faber, Stanhope, & Wiebe, 2015). Numerous studies have measured and quantified attitude and motivation integrating frameworks of self-efficacy theories with outcome expectancy theories. Self-efficacy is social cognitive approach with roots in self-determination theory and describes a person’s
perception or beliefs about their capabilities to produce effects (Bandura, 1986). This is closely tied with outcome expectancy which is a person’s expectations about the consequences of an action (or outcome of a task). The combination of these frameworks, in the context of STEM education, have been shown to influence motivation and persistence in an academic track (Unfried, Faber, Stanhope, & Wiebe, 2015). However; in order to situate student attitudes, interests, and experiences within a larger career context, the social cognitive career theory provides a three-way lens linking self-efficacy beliefs, outcomes expectations (a broader interpretation of outcome expectancy), and personal goals (Lent & Brown, 1996). Social cognitive career theory aims to “understand the processes through which people form interests, make choices, and achieve varying levels of success in educational and occupational pursuits” (Lent & Brown, 2000, p. 36).

Social cognitive career theory has strong roots in constructivism and is an important complementary framework to build on experiential learning. There is a strong overlap between academic and career factors which is recognized by this theory and helps to explain multiple constructs of STEM attitudes and career interests (Unfried, Faber, Stanhope, & Wiebe, 2015).

IV. Selecting and Designing Instruments for Assessment

Selection and design of instruments for assessment was directly aligned with research questions and assessment objectives. The primary research questions examining the curriculum’s impact on early elementary through middle school students specifically in areas of student (1) application of engineering design as an engineering approach in STEM education and (2) knowledge of and attitudes towards STEM-related fields, careers, and educational opportunities. To answer these questions, the team will utilize mixed methods in a pre-post with comparison group framework. The research instruments being utilized are (1) the are the Student Attitude Toward STEM (S-STEM) Survey developed at North Carolina State University by the Friday Institute and (2) a custom designed tool to assess the students’ knowledge of the Engineering Design Experience and their confidence in using it as an approach in STEM education.

Selection of Instruments

A literature review and extensive research of existing assessment tools used to measure learning of engineering design and attitudes toward and motivations to pursue STEM education and careers revealed a variety of instruments and approaches. Many of these instruments however were designed to measure these constructs in either high school or post-secondary students which, although they provided key insight into the construction, measurement, and use of the instrument, were ill-suited for early elementary through middle school aged students. Strong consideration was given the Engineering Identity Development Scale (EIDS) which has been validated by numerous research studies however, its focus on the field of engineering and student development of identity was problematic since the researchers were interested in evaluating the broader fields of STEM and student interests and motivations, not identity development (Capobianco, French, & Diefes-Dux, 2012). Also considered was the Hopes and Goals Survey (Douglas & Strobel, 2015) however this instrument did not have the broad focus on STEM and had a large construct designed around the concept of hope which was not of primary interest in answering the specific research questions of this study. Ultimately the Student Attitude Toward STEM (S-STEM) Survey developed at North Carolina State University by the Friday Institute was selected for use in this study (Friday Institute for Educational
Innovation, 2012). The S-STEM survey aims to measure changes in students’ confidence and efficacy in STEM subjects, 21st century learning skills, and interest in STEM careers and has been designed to support program coordinators in making decisions about program improvement and impact. Since their development in 2012, the S-STEM surveys have been given to over 21,000 students and the validity of the survey has been rigorously tested (Unfried, Faber, Stanhope, & Wiebe, 2015). This strong empirical base and the direct focus on attitudes, motivations, and interests STEM broadly made this tool and ideal choice for this study.

Additionally, the S-STEM survey has been developed in multiple formats which include upper elementary, middle school, and high school. The ability to use a similar instrument for students across all challenge levels was an attractive feature of utilizing this survey. However, since an early elementary version of the instrument did not exist the researchers employed tactics utilized in prior studies with early elementary students (Capobianco, French, & Diefes-Dux, 2012) to adapt another version for younger students. This adapted version contained the same statements and questions as the upper elementary version but instead of asking students to respond to a 5-point Likert scale (which has been shown to be problematic with younger students) it instead utilized a graphic depicting smiley faces (Figure 1).

![Figure 1: Adapted S-STEM survey for early elementary students (page 1 only)](image)

**Development of the Instruments**

After reviewing multiple tools, no assessment tools exist which identify assessment strategies aligned specifically with NGSS engineering design standards K-8 and broader NGSS core ideas of engineering design which include: (1) defining and delimiting engineering problems, (2) designing solutions to engineering problems and (3) optimizing the design solution (NAP, 2013). In order to understand the extent to which AWIM changes students’ awareness of engineering design as an engineering approach in STEM education it was necessary to design a new instrument. Using engineering design NGSS standards provides several advantages. First, it
increases internal validity by reducing researcher bias by ensuring that students were being evaluated on nationally accepted standards and not standards developed by or used only within AWIM. Use of NGSS also provides a more universal instrument which can be used in a variety of engineering education research studies to assess student changes in approach to engineering design.

Three total instruments were developed, one for each challenge being evaluated in this study: Straw Rocket - early elementary (2-3rd grades); JetToy - upper elementary (4-5th grades); and Gravity Cruiser - middle school (5-7th grades). All instruments were designed through four stages with similar formats but altered for appropriateness and reading level at each level of implementation.

*Pre-pilot stages.*

Stage 1: Research of elementary and middle engineering and science assessments

The first step in the process of developing a new assessment was to establish an empirical foundation by conducting a review of elementary and middle school engineering and science assessments. As previously mentioned, few assessment instruments are available for elementary and middle school students that focus specifically on students’ understanding and application of engineering design as an overall approach to STEM education. Some research studies have explored assessing STEM and/or engineering design. Many examples of these assessments can be found on the Engineering is Elementary Research Instruments site and a variety of other STEM education focused websites. However, the difference in specific assessment outcomes made these tools impractical for evaluating the research questions of study. Construction of engineering design-based activities and assessments used in these studies were important in informing development of the new instruments in areas of concept to task alignment, use of combinations of objective and constructive-response types of questions, use of drawings as an idea-eliciting task, and use of task-design approaches (Capobianco, Diefes-Dux, Mena, & Weller, 2011; Capobianco, Yu, & French, 2015; NAEP, 2015; Yoon, Lucietto, Capobianco, Dyehouse, & Diefes-Dux, 2014).

Stage 2: Identification, breaking down, and mapping standards

To create a strong alignment between AWIM challenges and STEM learning outcomes, overall design of the AWIM curriculum program aligns with a variety of standards ranging from Common Core Mathematics to Common Core English and Language Arts (ELA) to NGSS standards in forces and interactions, energy, and engineering design. Since one of the primary research questions of this study is exploring students’ application of engineering design as an approach in STEM education the team of researchers chose to focus only on the engineering design standards. All NGSS engineering design standards, for each appropriate level, were measured in the assessment. An issue in developing an assessment measuring each standard however was that each individual standard was quite dense and spanned multiple learning objectives, thus it became necessary to deconstruct each objective into more finite and measurable pieces. Once the standards were deconstructed they were then mapped to a core idea for assessment and general assessment plans and objectives could be identified. Figure 2 provides an overview of this mapping for each level of standards. This mapping document provided a roadmap for constructing an engineering design assessment for three different age
levels and helped to provide content validity to ensure that assessment questions linked directly to identified learning objectives and standards.

Stage 3: Development of instruments for each grade level

Two types of assessment questions were designed for use in this research study; objective questions and constructed-response questions. The assessment is divided into two or three parts with ten total questions for early and upper elementary school assessments and fifteen for middle school. Part 1 begins with a PBL type of scenario where students are asked to investigate and ill-formed problem to ask questions, make observations, and define a problem, criteria, and constraints. The questions and prompts related to these scenarios have been designed as constructed-response items to allow the researchers to more appropriately evaluate high-level cognitive objectives and also provide to considerable insight into the learner’s understanding and thinking processes (Morrison, Ross, & Kemp, 2007). Figure 3 provides an example of the early elementary (2nd-3rd grade Straw Rockets) PBL scenario with related prompts and questions and shows the researchers’ alignment to the NGSS standard mapping document developed in the previous stage.

Next, students across all levels of the assessments are also asked problem-solving questions where they are given a scenario, a list of materials (including images of the materials), and asked to design and draw a solution (see Figure 4 for an example of this at the upper elementary - 4th-5th grade Jet Toy). Teachers will collect Part 1 from students prior to distributing the next part.
Part 2 of the assessment builds on the same scenario presented in Part 1 however; it is given separately because it contains drawings, variables, suggested solutions, and data which may impact the way in which the students completed the PBL scenario-related questions, specifically those where they are asked to develop and draw an original design. Part 2 contains a mixture of constructed-response questions and objective questions. In this part of the assessment, students are asked to compare, analyze, and assess designs and data of two or three different objects. Figure 5 provides an example of the types of questions and formats presented in Part 2 of the elementary level assessments.

Part 3 of the assessment is only for middle school students who are participating in the Gravity Cruiser challenge. This part focuses on providing deeper opportunities for students to identify and distinguish between variables and characteristics of different design solutions as well as analyze related data. Figure 6 illustrates an example of these types of questions.

Stage 4: Expert review, editing, and revision

After development of initial assessment instruments, multiple different types of experts were contacted for their review of the assessments. Experts included three research faculty in P-12 STEM education and two elementary educators. After expert review the following suggestions were incorporated: instructions were included for 2nd and 3rd grade students to complete the assessments as a class with the teacher reading the questions, concrete examples and explicit definitions were provided for career descriptions, and the level of the language and vocabulary was reduced to make it more appropriate for elementary school students at each level. All other expert feedback was
compiled, reviewed, and incorporated where possible. All documents were then proofed by the research team and appropriate revisions were made.

**Planned use of instruments.**

Both instruments, the Friday Institute S-STEM surveys and the custom designed engineering design instrument, will be given to all students as a pre-survey and pretest prior to beginning the AWIM challenge and as a post-survey and post-test at the conclusion of their AWIM experience. The purpose of using a pretest-posttest design is to determine the degree to which the learners are improving as a result of participation in the engineering curriculum. Many experts in instructional design recommend using the actual evaluation test for both pretesting and posttesting to ensure that the instructional objectives and aligned assessment items are consistent between testing (Morrison, Ross, & Kemp, 2007). All pre-post information will be analyzed for changes to individual students. Further, after all stages of implementation and data collection have been complete, a comparison will be made between classes using AWIM and those not using AWIM.

**Pilot and Next Steps**

As of March 2017, the research team is engaging in a pilot where the custom designed tool to assess the students’ knowledge of the Engineering Design Experience and their confidence in using it as an approach in STEM education is being given to students in each level of the AWIM program being evaluated; 2-3rd grade, 4th-5th grade, and 5th-7th grade. Five STEM teachers in two different school districts are participating in this pilot. Each teacher teaches between 85-150 students and it is anticipated that there will be much student feedback and data.

This pilot serves several important purposes: (1) to determine appropriateness of question for age of the student and ability of the instrument to capture learning data robust enough to address both research questions, (2) utilize a combination of evidence-based practices and the results of the student data to develop scoring keys and rubrics, (3) to collect initial teacher feedback on the instruments and the process of delivering them in a classroom with students, and (4) as a trail run for the researchers to develop best practices for the printing, organizing, shipping, and receiving of the instruments.

Also, following the collection of pilot results initial construct and content validity of the assessments will be conducted. Providing tasks which incorporate methods such as analysis of
reasons, where students provide rationale and insight into how they solve the problem through their responses, is a key factor in establishing validity of an assessment tool (National Research Council, 2001). Further, early reliability will be established by (1) providing explicit instructions to teachers to ensure that all assessments are being administered in a standardized fashion and (2) construction of specific rubrics to ensure standardization among the scoring of the instruments (Morrison, Ross, & Kemp, 2007).

Upon completion of the pilot and following necessary revisions of the instruments, a broader scale analysis will be completed in Spring 2017. The team plans to evaluate 24 classrooms, 12 where AWIM programs are being implemented and 12 similar classrooms where the challenges are not being implemented. The results of this work will focus on supporting development of tools and findings which will enable data-driven decision making in selection of educational pathways supporting the K-12 STEM education ecosystems in establishing a STEM pipeline of future innovators.

V. Conclusion

Strong STEM education and a highly-qualified STEM workforce is vital to the health and security of our nation. Preparation of tomorrow’s STEM workers begins with early exposure and evidence-based STEM educational experiences however; it is critical that STEM programs and curricula be rigorously evaluated to maximize impact. Evaluation of assessment must be built on a framework of research, standards and learning objectives, and best practices from the field of STEM education. Selection and design of assessment instruments which can measure early engineering curriculum’s impact on application of engineering design, attitudes, and motivations toward STEM is the result of a systematic process integrating best practices from the fields of learning science, instructional design, and STEM education. Support of existing tools and introduction of these new instruments into the field of early STEM and engineering education has the potential to significantly impact the ability of related programs to use national standards, such as NGSS, to measure impact and effectiveness of STEM programs and curricula.

References


