



An Assessment Tool to Evaluate Student Learning of Engineering (Fundamental)

Dr. Tamara J Moore, Purdue University, West Lafayette

Tamara J. Moore, Ph.D., is an Associate Professor in the School of Engineering Education and Director of STEM Integration in the INSPIRE Institute at Purdue University. Dr. Moore's research is centered on the integration of STEM concepts in K-12 and postsecondary classrooms in order to help students make connections among the STEM disciplines and achieve deep understanding. Her work focuses on defining STEM integration and investigating its power for student learning.

Dr. Siddika Selcen Guzey, Purdue University, West Lafayette

Dr. Guzey is an assistant professor of biology and biology education at Purdue University. Her research and teaching focus on integrated STEM Education.

Mr. James Holly Jr., INSPIRE Institute, Purdue University

James Holly Jr. is a Ph.D. Student in Engineering Education at Purdue University. He received a B.S. from Tuskegee University and a M.S. from Michigan State University, both in Mechanical Engineering. His research interest is exploring formal and informal K-12 engineering education learning contexts. Specifically, he is interested in how the engineering design process can be used to emphasize the humanistic side of engineering and investigating how engineering habits of mind can enhance pre-college students' learning abilities.

An Assessment Tool to Evaluate Student Learning of Engineering (Fundamental)

Strand: K-12 Engineering Resources: Best practices in curriculum design

While STEM subjects have traditionally been taught separately in K-12 schools the new initiatives share a focus on integrated approaches to teaching STEM. For example, the recently released *Next Generation Science Standards* (NGSS)¹ addressed the need for explicit integration of science with engineering. Science teachers are expected to teach intersecting concepts and core disciplinary science using scientific and engineering practices. One challenge that many teachers face is the lack of measurement instruments for student learning required in the NGSS. Measuring the student understanding of engineering, connections between science and engineering, and engineering practices requires the development of new assessments align to NGSS. The purpose of this study was to develop and validate a student assessment for engineering. The use of item response theory to assess item functioning was a focus of the study. The work is part of a larger project focused on increasing student learning in STEM-related areas in Grades 4 – 8 through an engineering design-based, integrated approach to STEM instruction and assessment.

The Review of the Literature

For more than a decade, the United States has been shifting K-12 curriculum to a focus on science, technology, engineering, and mathematics (STEM). Over time, the argument for STEM education has grown stronger, in hopes of establishing a citizenry that is literate in all four STEM areas regardless of one's occupation. However, until recently, many STEM educational programs tended to focus on the "S" and "M" of the acronym, with slight mention of technology and largely ignoring engineering. Nevertheless, the future of engineering instruction in K-12 learning settings is brightening as many educational standards and research initiatives are being enacted to address the problem of genuine and effective implementation of STEM-focused curricula.

This study highlights one aspect of furthering the development of true STEM education. Primarily, we focus on a component of engineering education to bring to the forefront the "E" that has long gone unnoticed. The U.S. national report *Engineering in K-12 Education: Understanding the Status and Improving the Prospects*² stated, "The presence of engineering in K-12 classrooms is an important phenomenon, not because of the students impacted, which is still small relative to other school subjects, but because of the implications of engineering education for the future of science, technology, engineering, and mathematics (STEM) education more broadly" (p. 1). Generally, the endorsement of K-12 engineering education is motivated by interests in improving the quantity and quality of domestic students pursuing STEM careers. The benefits of an engineering education at the K-12 level extend beyond the expansion of the engineering pathway, reaching further to provide students opportunities to authentically interact with subject matter from other subjects, and actively engage the world around them. As Author et al.⁴ explain, "Because engineering requires the application of mathematics and science through the development of technologies, it can provide a way to integrate the STEM disciplines meaningfully" (p. 2).

Engineering education, at any grade level, cultivates competences that are useful beyond the academic context. Ioannis N. Miaoulis⁵, founding director of the National Center for Technological Literacy (NCTL), writes “I use my engineering training constantly to solve problems far removed from engineering, such as dealing with personnel issues or fundraising” (p. 39). The content of engineering allows students to make connections between their academic studies and their daily lives. Engineering education trains students to think analytically, and to use their knowledge base to make improvements. As Author⁴ states “Engineering requires students to be independent, reflective, and metacognitive thinkers who can understand that prior experience and learning from failure can ultimately lead to better solutions” (p. 5). These skills, often referred to as engineering “habits of mind” – values, attitudes, and thinking skills associated with engineering – are transferrable to multiple contexts within our society.

While the current national focus on K-12 engineering education is encouraging, it is critical to consider the components of engineering literacy for pre-college students. For the purpose of clarification, a definition of engineering literacy developed by Purzer, Strobel, & Cardella⁶, is provided below:

Engineering literacy is the ability to solve problems and accomplish goals by applying the engineering design process – a systematic and often iterative approach to designing objects, processes, and systems to meet human needs and accomplish goals. Students who are able to apply the engineering design process to new situations know how to define a solvable problem, to generate and test potential solutions, and to modify the design by making tradeoffs among multiple considerations in order to reach an optimal solution. Engineering literacy also involves understanding the mutually supportive relationship between science and engineering, and the ways in which engineers respond to the interests and needs of society and in turn affect society and the environment by bringing about technological change (p. 8-9).

Engineering pedagogy in the K-12 setting should be centered on the engineering design process². However, already demanding academic standards place a constraint on the flexibility of pre-college curricula. STEM integration is an effective method to establish engineering literacy without compromising scientific and mathematical literacy. As Author et al.⁷ stated, STEM integration is defined to be “the blending of science, technology, engineering, and mathematics content and context into one learning environment for the purpose of (1) deepening student understanding of each discipline by contextualizing concepts, (2) broadening student understanding of STEM disciplines through exposure to socially and culturally relevant STEM contexts, and (3) increasing interest in STEM disciplines to broaden the pipeline of students entering STEM fields” (p. 2). With the STEM integration concept, educators are able to use the engineering design process as content knowledge to familiarize students with engineering big ideas, or as a context to deepen student’s learning of math and science content. Since engineering makes use of the other STEM disciplines, it allows for an organic implementation of STEM instructional strategies in a relevant and purposeful manner.

The increasing recognition of K-12 engineering education establishes a need for quality assessment to effectively measure students’ learning of engineering. Arguments have been made about the importance of pre-college engineering education, but as pedagogies and standards are designed to satisfy this objective, sufficient measurement tools are required as well. Educators

must be careful in creating and/or performing evaluations to determine whether the desired engineering outcomes have been realized by their students. Furthermore, it is essential that the instructional activities educators use appropriately prepare students for the engineering concepts that are assessed². According to Olds et al.⁸, assessment is “the act of collecting data or evidence that can be used to answer classroom, curricular, or research questions” (p. 13). Whereas, the authors⁸ define evaluation as “the interpretations that are made of the evidence collected about a given question” (p.13). The differentiation between these terms is essential to accurately determine the role of each in the educational process, and delineate the next steps for their continued advancement.

This paper focuses on the design and development of an assessment to measure student learning of engineering in grades 4-8. There are few assessment tools available to measure student learning of engineering. An Engineering Concept Assessment designed by Daugherty et al.⁹ focuses on engineering education at the secondary school level. Dyehouse et al.¹⁰ generated student knowledge tests with similar intentions of measuring the science and engineering comprehension of elementary students (grades 2-4). Recently, National Assessment of Educational Progress (NAEP) also developed Technology and Engineering Literacy (TEL)¹¹ assessment that aims to measure how well students apply their understanding of technology and engineering principles in to real life situations. TEL assessment is computer-based and uses short-answer and multiple-choice questions. In 2014, the test was administered to a national sample of eight graders in 2014. The test takes approximately 120 minutes to complete and questions are not available to the public. These existing tools are not appropriate for grades 4-8 so this paper reports the first step in research on the development of an engineering assessment appropriate for these levels.

Methodology

Context

The engineering assessment was developed as a part of a larger study investigating the role of engineering design-based science curriculum units on student learning of STEM subjects. 15,000 students will participate in the study during the course of the five-year project. The purpose of the National Science Foundation funded project is to support teachers (grades 4-8) through professional development in developing and implementing effective integrated STEM curricula. To assess the engineering literacy of students as a result of STEM curriculum units, we developed an engineering assessment since no such assessment is currently available. In doing so, we took the view that our work was the first phase of research in what we expect to be an ongoing process of test development. As we collect additional student assessment data during the project and data for external variables such as state-mandated mathematics test scores, we will modify the assessment accordingly. Our research team consisted of a statistician and his graduate students, as well as two STEM education researchers, one with expertise in science and the other in engineering.

Assessment Development

Development of the engineering assessment was predicated on the goals and definitions underlying the STEM curriculum units described earlier, which were embedded in state-level

engineering standards for students. We began with an assessment development team of academic researchers with collective expertise in engineering, science, and mathematics. Researchers wrote 21 multiple-choice items, and content and face validity of the items were established by expert review. Part of this expert review was to have researchers who were experts in engineering literacy evaluate whether the assessment addressed the tenants of engineering literacy^{4,6}. The 21 items were used to develop engineering assessments for elementary and middle school students—what is described here focuses on the middle school assessment but the assessment development for elementary students went through the same steps. The final version of the middle school engineering assessment had 15 items so that it could be integrated with science content questions (not reported here) and completed by students within one school period (50 minutes).

Once the assessment development team agreed that the assessment addressed comprehensive engineering literacy consistent with the goals and definitions underlying the STEM curriculum engineering units, the assessment was piloted in two waves. In the first wave, the assessment was administered to a small group of about 10-20 middle school students (grades 6-8). The purpose was to obtain and analyze preliminary item response data as well as obtain feedback on characteristics of the assessment and the testing environment that affected performance (e.g., readability, clarity of items). This small group of students was interviewed after they completed the assessment (40% were students of color and 44% students in the free and reduced lunch program). Based on information provided by wave 1 piloting, the assessment were modified by the assessment development team and then piloted a second time using a larger sample of students that allowed item analyses to be performed.

In the second wave of piloting the revised assessment was administered to approximately 170 middle school students. A diverse group of students completed the assessment (approximately 70% students of color and 75% students in the free and reduced lunch program). Data from wave two were used to conduct extensive item analyses of the assessment including a factor analysis of the data. The goal of these analyses was to assess the likelihood that items represented a single underlying construct (engineering literacy). The results reported in this paper come from this second wave of piloting.

Results

Factor-analytic results suggested a single (major) factor underlies the assessment (unidimensionality), a finding that was robust to different methods of factoring (principal axis, maximum likelihood) and factor rotation (varimax, oblique). Theoretically, the idea of a single factor underlying engineering literacy makes sense because by definition engineering literacy is an integrated and interdependent way of thinking⁶. Our claim of unidimensionality relied on four sources of information. First, the eigenvalue for factor 1 was approximately 5 and that for factor 2 was approximately 1.4 suggesting a single underlying factor. Second, the corresponding scree plot showed an almost ideal pattern supporting unidimensionality of a steep curve followed by a bend and then an essentially horizontal line. Third, the Tucker-Lewis reliability coefficient¹² was estimated using maximum likelihood and the AMOS software in which models for 1 vs 2 factors were fitted and the corresponding chi-square fit statistics used to compute this coefficient. The resulting value of .85 supports the argument that a single factor underlies the middle school engineering assessment. Fourth, an examination of the size, pattern, and meaningfulness of factor

loadings for the 1 factor model provided support for unidimensionality. These results collectively suggests that the goal of constructing items for the assessment that reflected a single factor was generally met, and helped to justify the use of item response theory (IRT) to generate proficiency scores for students.

Next the Rasch IRT model was fitted to the data for each item on each assessment to estimate that item’s ability to contribute to estimates of student proficiency. In order to perform the item analyses for the test, we estimated the coefficient alpha, item mean, item discrimination and the frequencies of the items response options. To perform the analyses, we used Winsteps software. Several items statistics were estimated. However, the most important statistics to look at are the item location (item difficulty) and the item infit and outfit values. For the item location the interpretation is the lowest the value the easiest the item and the highest the value the hardest the item; in the case of infit and outfit values, they help to determine whether or not any item is productive for measurement purposes. Table 1 is the Winstep output. In Table 1, ENTRY represent items and MEASURE represents item difficulty. The number of students taking the test is represented by the COUNT column. IN.MSQ is the “Infit Mean Square, IN.ZSTD and OUT.ZSTD are simply standardized versions of IN.MSQ and OUT.MS, respectively. PTME represents the point-biserial correlation for each item. The Rasch model adequately fits the data of all 21 items; however, we eliminated 6 items to reduce the tests to the desired length of 15 items. We removed items 5, 12, 17, 18, and 21 because of their large IN.ZSTD standardized residuals and item 1 because of its negative biserial correlation. The final version of the test includes 15 questions.

Table 1: *Winsteps results for the engineering assessment for grades 6-8*

ENTRY	MEASURE	COUNT	IN.MSQ	IN.ZSTD	OUT.MS	OUT.ZSTD	PTME
1	3.104	168	1.23	1.1	7.68	7.07	-0.12
2	-0.561	168	1.08	0.84	1.01	0.14	0.41
3	-0.257	168	1.02	0.25	1.11	0.97	0.43
4	-1.321	168	1.03	0.26	1.01	0.11	0.42
5	0.659	168	1.17	2.63	1.35	2.77	0.28
6	-1.23	168	0.82	-1.49	0.83	-0.85	0.56
7	0.778	168	0.93	-1.04	1	0.05	0.47
8	-0.975	168	0.94	-0.5	0.96	-0.16	0.48
9	0.242	168	1.01	0.22	1.03	0.36	0.44
10	0.391	168	1.08	1.21	1.14	1.29	0.38
11	-1.23	168	0.76	-2.06	0.62	-2.18	0.61
12	-0.935	168	0.74	-2.61	0.63	-2.54	0.64
13	-0.895	168	0.78	-2.15	0.68	-2.14	0.61
14	1.111	168	1.14	2.06	1.3	1.95	0.28
15	0.51	168	1.04	0.7	1.08	0.79	0.4
16	0.778	168	1.01	0.23	1.08	0.67	0.4
17	0.302	168	1.17	2.39	1.15	1.39	0.33
18	-0.526	168	0.8	-2.33	0.69	-2.6	0.61
19	-0.596	168	0.8	-2.24	0.72	-2.26	0.6
20	-0.78	168	0.83	-1.75	0.76	-1.71	0.58
21	1.432	168	1.2	2.46	1.52	2.62	0.2

Conclusion

The aim of the present study was to develop and provide phase 1 evidence of validity of a simple instrument to assess student learning of engineering literacy. The results provide evidence that the assessment represents the first research step in developing a psychometrically sound instrument. The short length of the test and the fact that all items are multiple choice items suggest that teachers or researchers can easily administer the test to evaluate student learning of engineering. Additional student responses to the assessment will continue to be collected as part of the larger research project, as will external variables such as a state-mandated assessment that may provide validity evidence. We will also continue to probe more specific features of the assessment such as the pattern of items on the IRT difficulty scale and whether the assessment performs similarly for student subgroups (e.g., gender).

References

1. NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.
2. National Academies of Engineering and National Research Council. (2009). *Engineering in K-12 Education: Understanding the Status and Improving the Prospects*. Washington, D.C.: National Academies Press.
3. Strobel, J., Mendoza Diaz, N. V., & Diaz, N. V. M. (2012). Exploration of NSF-ATE projects approaches in the integration of technology and engineering education at the K-12 levels. *Advances in Engineering Education*, 3(2), 1–23.
4. Moore, Tamara J.; Glancy, Aran W.; Tank, Kristina M.; Kersten, Jennifer A.; Smith, Karl A.; and Stohlmann, Micah S. (2014) "A Framework for Quality K-12 Engineering Education: Research and Development," *Journal of Pre-College Engineering Education Research (J-PEER)*: Vol. 4: Iss. 1, Article 2. <http://dx.doi.org/10.7771/2157-9288.1069>
5. Miaoulis, I. (2010). K-12 Engineering – the missing core discipline. In Grasso, D. and Burkins, M. B. (Eds) *Holistic Engineering Education*, pp. 37-51. Dordrecht:Springer.
6. Purzer, S., Strobel, J., & Cardella, M. E. (Eds.). (2014). *Engineering in pre-college settings: synthesizing research, policy, and practices*. Purdue University Press.
7. Moore, T.J., Stohlmann, M.S., Wang, H.-H., Tank, K.M., Glancy, A.W. & Roehrig, G.H. (2014). Implementation and integration of engineering in K-12 STEM education. In Ş. Purzer, J. Strobel, & M. Cardella (Eds.), *Engineering in precollege settings: Research into practice* (pp.35-60). West Lafayette, IN: Purdue Press.
8. Olds, B. M., Moskal, B. M., & Miller, R. L. (2005). Assessment in engineering education: Evolution, approaches and future collaborations. *Journal of Engineering Education*, 94(1), 13-25.
9. Daugherty, J., Custer, R. L., Brockway, D., & Spake, D. A. (2012). Engineering concept assessment: Design & development. Proceedings of the 119th American Society for Engineering Education Annual (ASEE) Conference and Exposition, San Antonio, TX.

10. Dyehouse, M., Diefes-Dux, H. A., & Capobianco, B. (2011). Measuring the effects of integrating engineering into the elementary school curriculum on students' science and engineering design content knowledge. Proceedings of the 118th Annual American Society for Engineering Education (ASEE) Conference and Exposition, Vancouver, BC, Canada.

11. National Assessment of Educational Progress, Technology and Engineering Literacy Assessment (TEL)
<http://nces.ed.gov/nationsreportcard/tel/moreabout.aspx>

12. Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1-10.

Engineering Assessment- 6th-8th Grade

- Which of the following activities is most important to the work of an engineer?
 - Using power tools to fix broken things.
 - Using power tools to build things.
 - Developing understanding about what makes things break.
 - Fixing broken things for people.
- Engineers are designing a bridge across a river. Some of the constraints that the engineers will consider are the time it takes to build the bridge, the cost of the construction of the bridge, and how much traffic will cross the bridge. Which of the following is also a very important constraint that the engineers would need to consider?
 - The population of fish in the river.
 - The types of plant life living in the river.
 - The pollution levels in the river.
 - The erosion of the river.
- Which of these statements describes something that an engineer would do as part of his or her job?
 - Figure out what materials to use to make bridges strong.
 - Operate cranes.
 - Build chimneys out of bricks.
 - Pour concrete or cement for new roads.
- When engineers test their designed product, which of the following is something they must consider in order to know if the product works?
 - How many team members were needed to design the product.
 - How much of the allowable money to design the product was used.
 - How many ideas they came up with during the brainstorming phase.
 - How well the product meets the client's needs.
- Which of the following would NOT be a part of the job of an environmental engineer?
 - Design a technology to clean up a pollutant.
 - Investigate the source of underground soil pollutants.
 - Pollute a stream to test the effect on the environments.
 - Make recommendations to a community on how to clean up river pollution.
- You are working as an engineer in a food storage container company and find out that one of the company products contains a potentially harmful chemical. What should you do?
 - Tell your boss so that the product can be taken out of stores and customers can return them for full refund of their money.
 - Change the name and packaging of the product so that no one knows it is the same.
 - Test the products on people to see if they get sick.
 - Sell the products that are already on the shelves, but stop making new ones.

7. Which of the following statements represents something that an engineer would do as part of his or her job?
- A. Repair the engine in a car that will not start.
 - B. Improve your truck by putting new wheels on it.
 - C. Figure out how to improve the safety of cars.
 - D. Drive cars in racing competitions.
8. When designing a bridge, which of the following measures will help the engineers make a decision about the safety of the bridge?
- A. The color of the paint on the bridge.
 - B. The brands of cars that will cross the bridge.
 - C. The amount of weight the bridge can hold.
 - D. The final destinations of the people crossing the bridge.
9. A team of engineers are designing an earthquake resistant building. They make a list of construction materials and build a prototype of their earthquake resistant building. They decide to use a “shake table” to help them with their design. Shake tables such as the one shown below provide conditions representative of actual earthquakes.

Engineers will use the shake tables to:

- A. test the color of the materials used to build their earthquake resistant prototype.
 - B. find the cost of the materials to build the actual earthquake resistant building.
 - C. test their prototype earthquake resistant building.
 - D. display the prototype of the earthquake resistant building.
- 
10. Kai is working on a solar oven design. A solar oven is a device that uses the energy from direct sunlight in order to cook food. Kai’s main goal is to cook the food quickly. Which of the following should Kai focus on when designing his solar oven?
- A. How to design a solar oven to heat the air inside the oven.
 - B. How to design the solar oven so it is easy to carry.
 - C. How to design the least expensive solar oven.
 - D. How to design the solar oven to be the smallest.
11. Ava is a geotechnical engineer who investigates soil and rock properties on and below planned construction areas. Today, she is investigating an area where a chain grocery store is to be built. What should she do if she finds the planned construction area has a high risk of earthquakes?
- A. Approve the design of the grocery store as originally designed.
 - B. Deny construction of any building on the area.
 - C. Suggest a redesign of the building that can tolerate movement.
 - D. Approve the construction of a building and then investigate the soil and rock properties of the area again after the building is built.

12. Nina needs to design and build a water filter for her science fair project. She brainstormed different filter materials that she could use in her design, tested them, and decided on which materials she would use. What should she do next?
- A. Build her filter design and test it.
 - B. Keep brainstorming different filter materials.
 - C. Ask questions of her teacher about water pollution.
 - D. Draw a plan of her filter design.
13. Jayson is working on his design of a video game. He has redesigned his video game several times. His problem is that the game stops working in level 3 each time he tests it. In order for Jayson to think creatively his next redesign, he should:
- A. Test his current design again to see if the problem goes away.
 - B. Use an old way of solving the problem.
 - C. Look at his problem in only one way.
 - D. Look at his problem in a new or different way.
14. Which of the following statements about engineers is NOT true?
- A. Engineers usually work in teams but they are also independent thinkers.
 - B. Engineers use only mathematics to solve problems.
 - C. Engineers manage constraints, risk, and safety factors.
 - D. Engineers use a variety of tools, skills, and processes at work.
15. Kellie needs to design and build a speaker for a class project. After completing the construction of her speaker, she tested it. She found a failure in her design. What does she need to do next?
- A. She needs to start working on a completely different design.
 - B. She should evaluate her design, identify strengths and weaknesses, and then use this feedback in redesign.
 - C. She does not need to make any changes on her design.
 - D. She should stop working individually and start working with a friend on a new design.