Paper ID #20388

An Assessment Framework for First-Year Introduction to Engineering Courses

Dr. Senay Purzer, Purdue University, West Lafayette (College of Engineering)

Senay Purzer is an Associate Professor in the School of Engineering Education. Her research focuses on teaching and assessment associated with key aspects of engineering design such as innovation and decision-making.

Dr. Kerrie Anna Douglas, Purdue University, West Lafayette (College of Engineering)

Dr. Douglas is an Assistant Professor in the Purdue School of Engineering Education. Her research is focused on methods of assessment and evaluation unique to engineering learning contexts.

Jill Anne Folkerts, Purdue University

Mr. Taylor V. Williams, Purdue University, West Lafayette (College of Engineering)

Taylor Williams is a Ph.D. student in engineering education at Purdue University. He is currently on an academic leave from his role as an instructor of engineering at Harding University. While at Harding he taught undergraduate engineering in biomedical, computer, and first-year engineering. Taylor has also worked as a systems engineer in industry. Taylor received his master's in biomedical engineering from Tufts University and his bachelor's in computer engineering and mathematics from Harding University.

An Assessment Framework for First- Year Introduction to Engineering Courses

Abstract

In this evidence-based practice paper, we describe an assessment framework that applies to first-year introductory engineering courses. First-year engineering courses cover a variety of learning objectives that address both technical and professional outcomes outlined in ABET. These courses also often involve open-ended design and modeling projects. The assessment of multiple competencies along with open-ended design can be a challenging task for educators. In this paper, we describe a framework that guides instructional processes for effective assessment for student learning. This assessment-centered teaching and learning framework helps connect specific learning objectives to broader learning goals or competencies and on-going formative feedback targeting student progression on specific learning objectives. Our plan is to refine the framework using a design-based research approach. Following the description of the model and its development, we present results from the first cycle of implementation. We conclude by discussing hybrid ways for combining traditional methods of assessment with the ability to highlight performance expectations and the appropriate uses of the framework in the classroom.

Introduction

As a gateway to engineering, first-year engineering or introduction to engineering courses cover a variety of learning objectives. An important and common component of first-year courses in engineering programs is introducing students to engineering concepts, practices, and the engineering profession as well as motivating the students towards engineering.[1] According to a Delphi study by Reid and colleagues,[2] these courses cover four main areas: engineering skills (e.g., design process, programming), professional skills (e.g., teamwork, technical communication), orientation to the engineering program (e.g., discipline selection), and orientation to the engineering profession (e.g., professional societies). Hence, these courses address both technical and professional outcomes outlined by ABET as well as orientations to engineering school and the profession. Similarly, Gustafsson and colleagues analyzed first-year introductory courses at three universities in Sweden and MIT in the U.S.[1] They specifically used a Conceive, Design, Implement, and Operate (CDIO) model for engineering education as part of a reform effort. The four components of this model include technical knowledge and reasoning, personal and professional skills and attributes, interpersonal skills, and CDIO systems in the enterprise and societal context. While prior research studies on the classification and models of first-year engineering programs help frame program development and syllabi, there exists little research on the issue of assessment in first-year engineering programs with such rich and distinct competencies and objectives.

The assessment of multiple competencies along with open-ended design is a challenging task for educators.[3] In an ideal classroom setting, students demonstrate learning through a variety of means and multiple sources of evidence, yet there are practical challenges that can prevent rich assessment. Therefore, we need a practical assessment system where multiple forms of evidence can be used to assess student learning and inform instruction. The purpose of this paper is twofold: first, to describe the development of a framework that guides instructional processes for effective assessment of student learning, and second, to share our refinement efforts using a design-based research approach. This assessment-centered teaching and learning framework is designed to help connect specific learning objectives to broader learning goals (i.e., competencies) and to enable ongoing formative feedback targeting student progression on specific learning objectives.

Literature Review

Assessing Higher-Order Skills, Measuring Competencies Across Tasks

There is broad consensus within the engineering education community that students should actively use knowledge to develop skills rather than merely memorizing facts and theories.[4–6] Therefore, most of the tasks students encounter should tap into cognitive skills that are higher-order.[7] Higher-level or higher-order skills also support transferable learning into other contexts far better than lower-level skills. Hence, classroom assessment and practices should enable continuous assessment of high-level key competencies across multiple tasks, be comprehensive (based on evidence from multiple sources), and be coherent (in alignment with higher-order course goals and objectives).[8] Assessing higher-order skills is inherently more difficult than assessing rote learning or basic procedural knowledge and skills. While Bloom's taxonomy is widely known,[9] a more contemporary and useful model is Webb's Depth of Knowledge (DOK) taxonomy.[10] According to the DOK, higher-level and transferable skills go beyond recall skills such as the ability to list, calculate, and use (level 1) and target the abilities to predict, graph, and compare (level 2); to assess, revise, and investigate (level 3); as well as the abilities to analyze, synthesize, design, and create (level 4). Moreover, Darling-Hammond et al. state that "if assessments are to reflect and encourage transferable abilities, a substantial majority of the items and tasks (at least two-thirds) should tap conceptual knowledge and abilities (level 2, 3, or 4 in the DOK taxonomy)" (p. 5).[7]

Assessment for Learning

The focus of the framework we have developed is on student learning. Often assessment in the classroom is equated with exams, quizzes, and grades rather than emphasizing ways that assessment can be useful in support of teaching and student learning. Moreover, adding to the confusion, in higher education the term assessment has many disparate uses referring to measures of institutional effectiveness, the appraisal associated with accrediting programs, the evaluation of faculty and staff, and finally the appraisal of student learning.[11]

It is important that assessment practices target higher-order skills (e.g., level 2, 3, or 4 in Webb's DOK taxonomy[10]). It is also important to note that classroom assessment practices can either help support or hinder student learning.[12–15] The types of assessments educators use in their classroom instruction, as well as the manner in which they use assessment, can have an influence on the learning outcomes of their students.[16] In a model that focusses on measuring core competencies across multiple tasks and targeting higher-order skills, it is not only important to help educators move towards evidence-based, assessment-centered teaching but also to challenge educators' perceptions of assessment as only consisting of grades and exams towards a belief with assessment as an integral part of teaching.[17]

According to York, who wrote about the roles of formative assessment in higher education, effective assessment relies on educators who are not only aware of the epistemology of the discipline and stages of student development but also the psychology of giving and receiving feedback.[12] Hence, we argue that effective assessment and formative feedback does not focus on what students are not able to do (negative or backward feedback) but is instead forward-looking and progressive.

Assessment Frameworks

Various forms of assessment frameworks exist in the literature targeting assessment broadly[11,18] or specific content areas such as science, mathematics, and engineering. In science education, the Task Analysis Guide in Science (TAGS) framework by Tekkumru-Kisa and colleagues specifically focuses on two components of assessment: cognitive demand (memorized to practiced) and aspects assessed (content, practices, or integration of content and practices).[19] In other models content and skills are prioritized. For example, Bleiler and Thompson propose a multidimensional approach to assessing students' mathematical understanding across four dimensions (SPUR): skills, properties, uses, and representations, which provides educators useful information about the depth of their students' understanding of a mathematical topic.[20,21] The examples in engineering education target interpretation and feedback processes that takes into account student cognition. Diefes-Dux et al.'s framework is on feedback in model-eliciting activities.[22] In engineering design, Beyerlein and colleagues present an assessment framework for capstone design courses[23] building on the assessment triangle model by Pellegrino, Chudowsky, and Glaser.[18] To our knowledge, there are no assessment frameworks that specifically target first-year introductory engineering courses.

Competencies and Learning Objectives

Higher-level skills encompass multiple distinct components (i.e., core competencies) associated with each skill. Table 1 presents five competencies among a total of ten with a sample set of associated learning objectives. Note that higher-level engineering skills are not equivalent to task specific skills but rather are expansive skills such as design or problem solving.[24] Each of these higher-level skills encompasses a variety of distinct, broad competencies (e.g., the high-level skill of design contains the competencies of evidence-based decision making, EB, and solution quality, SQ). While these competencies identify the broad components of the skill, they are still too broad to assess directly, so each goal is further subdivided into measurable learning objectives (e.g., the design competency of evidence-based decision making, EB, contains the learning objectives EB04, justify the metrics chosen for evaluating potential solutions, and SQ01, solutions are technically accurate). As another example, Sorby writes about engineers' need for spatial visualization skills.[25] The associated competencies would include improved communication and augmented creativity. Specific learning objectives might include demonstrating proficiency in computer-aided design (CAD) or drafting and freehand sketching.

Rubrics and Rubric Development

One challenge inherent in assessing higher-order skills is rater subjectivity (real or perceived). Rubrics, also known as scoring guides, are one option to increase rater reliability.[26] Rubrics provide a systematized method for raters to judge the quality of student responses against a set of established measurement criteria and are particularly useful for evaluating significant tasks, like performance tests (i.e., authentic tasks for assessing high-level skills). In addition to being an assessment tool for instructors, rubrics can assist planning of appropriate instruction[26,27] as well as promote student learning.[26,28,29] If the students have access to the rubric while completing their task they are better able to self-assess their own progress and learning and can even provide peer feedback.[26,30] To achieve these benefits, rubrics must be high quality and raters must be trained in their use. Lovorn and Rezaei found that low-quality rubrics and lack of training can result in assessments that are just as subjective as if no rubric had been used at all.[31]

Table 1. *Sample core competencies and objectives*

CORE COMPETENCIES		LEARNING OBJECTIVES
Data Visualization and Analysis (DV) Visually represent data and derive meaningful information from data.	DV01	Use built-in cell referencing and functions for efficiency of calculations.
	DV04	Prepare chart or table for technical presentation with proper formatting (headers, units, meaningful decimal points, appropriately scaled axes, appropriately sized marker and axis labels).
	DV05	Describe, with calculations, the central tendency of data using descriptive statistics (mean, median, and mode).
Evidence-Based Decision Making (EB) Use evidence to develop and optimize the solution. Evaluate solutions, test and optimize the chosen solution based on evidence.	EB02	Identify assumptions made in cases when there are barriers to accessing information.
	EB04	Justify chosen metrics and the corresponding assigned weights to evaluate potential solutions, based on stakeholder needs.
Information Literacy (IL) Seek, find, use and document appropriate and trustworthy information sources.	IL03	Support all claims made with evidence that is either generated or found.
Professional Communication (PC) Communicate engineering concepts, ideas, and decisions effectively and professionally in diverse ways such as written, visual and oral.	PC02	Make clear and complete arguments or statements by fully addressing all parts of the assignment.
	PC03	Present all visuals with captions (e.g., figure number, table number, and brief description).
Solution Quality (SQ) Design final solution to be of high technical quality. Design final solution to meet client and user needs.	SQ01	Use accurate scientific, mathematical, and/or technical concepts, units, and/or data in solutions.
	SQ02	Justify design solution based on how well it meets criteria and constraints.

Rubric development is an iterative process that begins with defining the competencies to assess, identifying teachable scoring criteria addressing each goal (i.e., learning objectives), and decomposing each learning objective into clearly identified and described gradations of quality or levels of mastery (e.g., strong, middling, and problematic student work).[29,32,33]

Scoring strategies for rubrics can be divided into analytic or holistic types.[32,34–38] Holistic rubrics assess multiple criteria within a single score. These are most appropriate when the assessment criteria have significant overlap or when making broad judgments of quality. While holistic rubrics are generally less time consuming than an analytic rubric, they do provide limited feedback to the student, restricting its educative value. Rater bias is also a concern when using holistic rubrics as broad judgments are inherently more prone to variations in rater judgment. Analytic rubrics, on the other hand, have the rater assign as a separate score for each criterion. This type of rubric helps instructors and students better identify areas needing improvement but is more time consuming than a holistic rubric.

Popham discusses three types of rubrics: task-specific, hypergeneral, and skill-focused, and argues that the first two of which should be avoided.[27] The first type of rubric Popham discourages is the task-specific rubric. This kind of rubric links its evaluative criteria directly to specific tasks called for in the assignment; thus, it is unable to help instructors plan teaching to promote generalizable skill development and transfer in their students. On the other extreme, hypergeneral rubrics are equally problematic in that they use general and poorly defined terms in their criteria's quality descriptions. Without meaningfully clarified descriptions of performance criteria, these rubrics also provide little to no guidance in instructional planning or student learning.

Popham does recommend a third type of rubric—skill-focused. These rubrics are designed specifically to access higher-order skills. Instructors familiar with a skill-focused rubric's key features will usually plan better instruction (as they will know what to emphasize). Popham provided five guidelines for designing a skill-focused rubric: (1) ensure the assessed skill is significant, (2) ensure each of the rubric's evaluative criteria can be taught, (3) minimize the number of evaluative criteria to around three or four, (4) include a brief label for each criterion, and (5) constrain the rubric to a usable length (one or two pages for most people). Additional suggestions for rubric design are available in the literature.[34,39,40] In Table 2, we present an example for an analytic rubric that targets five evaluative criteria across four levels of competencies.

Table 2. *Sample scoring guide and criteria.*

	LEARNING OBJECTIVE	Proficient	Developing	Emerging	No evidence
DV05	**Describe, with calculations, the central tendency of data using descriptive statistics (mean, median, mode).**	Statistical calculations are correct and used to describe multiple aspects of the central values of the distribution of data (mean, median, mode).	Statistical calculations are correct but are not used (or used incorrectly) to describe the central tendency of the data.	Statistical calculations are incorrect, preventing proper interpretations of central tendency.	No evidence found related to the learning objective
EB02	**Identify relevant assumptions made in cases when there are barriers to accessing information.**	Identified relevant assumptions needed to be made in cases when there are barriers to accessing information but there is also the need to move forward.	Identified relevant assumptions needed to be made although a few of these are not relevant or could have been easily resolved with additional information gathering.	Assumptions are made but were not relevant or explicitly recognized.	No evidence found related to the learning objective
IL03	**Support all claims made with evidence that is either generated or found.**	Supported all claims made with evidence that is either generated or found.	Some of the claims made have evidence that was generated/found, but one or more claims missing substantive evidence.	Claims are made with little supporting evidence.	No evidence found related to the learning objective
PC03	**Present all visuals with captions (e.g., figure number, table number, and brief description).**	Presented all visuals with captions (e.g., figure number, table number, and brief description).	Used captions that do not clearly summarize the visual.	Repeatedly presented visuals without captions.	No evidence found related to the learning objective
SQ01	**Use accurate, scientific, mathematical, and/or technical concepts, units, and/or data in solutions.**	Correctly used scientific, mathematical, and/or technical concepts, units, and data in solution.	There are minor scientific, mathematical, and/or technical errors (typically calculation errors) with minor impact on the solution.	There are major scientific, mathematical, and/or technical errors with conceptual flaws or errors that have a significant impact on the solution (e.g. unit conversion error).	No evidence found related to the learning objective

Note. *DV (Data Visualization and Analysis), EB (Evidence-Based Decision Making), IL (Information Literacy), PC (Professional Communication), and SQ (Solution Quality).*

Assessment Framework for First-Year Engineering Courses

In light of the aforementioned reasons and literature, we developed the *Assessment Framework for First-Year Engineering Courses.* The purpose of this framework is to provide educators a roadmap for:

- determining core competencies targeting higher-order abilities;
- decoupling interconnected aspects of competencies by connecting specific learning objectives to broader competencies;
- valuing transfer of learning by measuring common competencies across multiple tasks;
- collecting evidence of student learning through a variety of means;
- providing a consistent approach to assessment in the classroom with scoring guides that are task-independent;
- guiding instructional processes with forward-looking, formative feedback that targets student progression on specific learning objectives; and
- producing interpretations that are usable, shareable, and educative.

To accomplish these design features, we used the three components of the assessment triangle (cognition, observation, and interpretation)[18] as a canvas for developing an actionable roadmap (see Figure 1). Our framework's roadmap is driven by research on student cognition and epistemological understanding of the discipline, with tools and strategies to support observations of student learning, and efforts that enable interpretations that are usable, shareable, and educative.

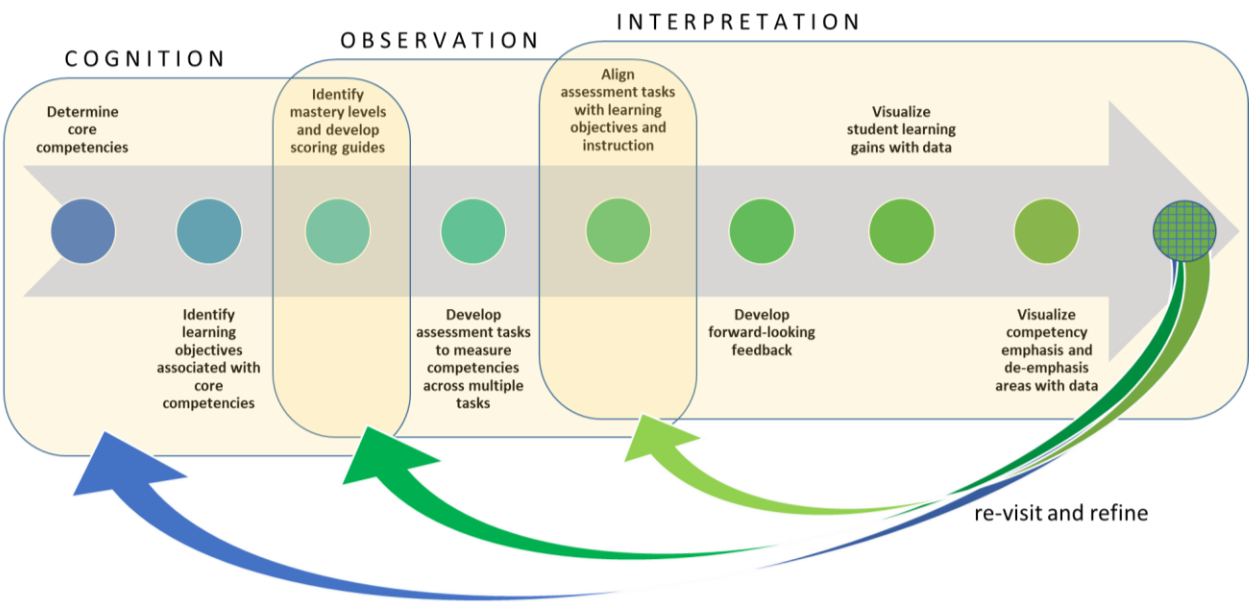

Figure 1. Assessment Framework for First-Year Engineering Courses

(1) *Determine core competencies*: Core competencies are high-level learning goals that an educator identifies as critical, broad aspects of learning expected in a course. These are typically written as learning goals or learning outcomes in a syllabus.

(2) *Identify learning objectives associated with core competencies*: Because core competencies are high-level and broad, they entail sub-aspects that we layout as learning objectives. It is important to align each learning objective to core competencies so, when aggregated, the learning objectives provide information on student performance on the higher-level competencies.

(3) *Identify mastery levels, develop scoring guides*: Given that learning objectives target areas for new learning, and that students come to the classroom with diverse levels of prior knowledge, we would expect students will meet each objective at varying levels. Hence, the framework suggests determining three levels of competency for each objective: proficient (high), developing (medium), and emerging (low).

(4) *Develop assessment tasks to measure common competencies across multiple tasks*: Higher-order competencies require transferable concepts and skills. Therefore, assessment of these competencies needs to occur multiple times in a semester to allow the student to practice such transfer, but also to allow the educator to assess the student's ability to transfer these competencies across tasks.

(5) *Align assessment tasks with learning objectives and instruction*: Once core competencies and specific learning objectives are developed, the next step is reviewing the assessment tasks alongside the planned instruction so that all of these components coalesce toward coherent and common goals.

(6) *Develop forward-looking feedback*: While assessment is often thought as a cognitive endeavor, there are psychological factors that can hinder the effectiveness of the approach even when all other aspects are in place. It is hence important to ensure that assessment motivates the student toward higher performance by emphasizing what the student can improve rather than focusing on errors and mistakes related to a past task.

(7) *Visualize competency emphasis and de-emphasis areas with data*: While much care might have been given to the allocation of learning objectives across various tasks, it is still possible that some areas might be over or under emphasized unintentionally. A mapping and visualization of core competencies and learning objectives can help visualize these emphasis areas. These data can then help enhance instruction and assessment.

(8) *Visualize student gains with data*: Perhaps the most important and challenging part of the framework is visualizing student progress over time. The ability to do that depends on our ability to develop assessment tasks to measure common competencies across multiple tasks and the relative similarity or difficulty of these tasks.

(9) Revisit 1 through 8 and refine.

Research Methods

Context of the Study

We implemented the *Assessment Framework for First-Year Engineering Courses* in the context of a specific first-semester introductory course as part of a large first-year engineering program. The course covered topics in the following main content areas: mathematical modeling, data analysis with Excel, design, teamwork, technical communication, ethics, sustainable energy concepts, information literacy, and information on engineering programs.

Methodological Framework: Design-based Research

The design and development of the framework is an iterative process. Hence, we are using design-based research (DBR)[41] to systematically study the framework and its appropriate uses. The DBR approach involves iterative cycles of testing and research-informed revisions which are especially suitable for studying novel educational products and processes.[42] In this study, the assessment framework, as well as its components and artifacts, are examined by defining a problematic situation, establishing conceptual foundations, developing initial product design and process for users, and iteration.[43]

Results

Defining the Problematic Situation

The problematic situation we started with is the need to address diverse goals of first-year engineering courses in ways that provide useful information to support student learning in higher-order competencies. These issues are discussed in detail in the introduction and literature review sections.

Forming Conceptual Foundations

The conceptual foundations that informed the design of the framework are outlined in earlier sections that lay out the framework in Figure 1.

Determine Core Competencies and Identify Learning Objectives Associated with Core Competencies

The initial product included 10 learning goals (core competencies) and 45 learning objectives. Each objective is noted with an abbreviated description connecting it to the core competency. Several examples from this initial product are presented in Table 1.

Identify Mastery Levels and Develop Scoring Guides

Keeping the forward feedback in mind, we developed three levels of competencies (proficient, developing, and emerging) along with a “no evidence” category. We then wrote behaviors, or performances that are aligned with these levels. Table 2 presents a subset of the resulting scoring guide with examples of assessment criteria and levels of competency as they relate to five learning objectives.

Develop Assessment Tasks to Measure Common Competencies Across Multiple Tasks

Given the emphasis we made earlier that competencies would be assessed across multiple tasks and assessment guides or rubrics should address core competencies rather than being task-specific, our framework requires a careful development of assessment tasks.

Align Assessment Tasks with Learning Objectives and Instruction

Once assessment tasks are developed, it is important to ensure they are aligned with the learning objectives. Moreover, the instruction addresses these objectives, giving students the opportunity to learn the competencies they are asked to perform in an assessment task. . . . . .

Develop Forward-Looking Feedback

As discussed earlier in the literature review, not all feedback is helpful in promoting student learning. Table 3 presents examples of good insufficient and backward feedback as well as high-quality forward feedback.

Table 3. *Examples of insufficient (backward) and quality (forward) feedback*

Learning Objectives	**P, D, or E?**	**Insufficient (backward) Feedback**	**Quality Feedback (forward feedback)**
DV01: Use built-in cell referencing and functions for efficiency of calculations.	PROFICIENT	Good job.	Your spreadsheet demonstrated good use of absolute and relative cell-referencing practices.
DV04: Prepare chart or table for technical presentation with proper formatting (headers, units, meaningful decimal points, appropriately scaled axes, appropriately sized marker and axis labels).	EMERGING	Wrong decimal places. Chart missing units and labels	(1) Check that the values from your calculations are meaningful. In Column F, why report 10 significant digits? To determine the reasonable and appropriate number of decimal places refer to your input values. (2) Remember to format your columns so values are under the title cell. See values 0-340 in relationship to header cell "Time (mins)" (3) One could imply from the graph that two factors are being compared but it is not clear what the series represent? (4) Also, what are the units for y-axis and x-axis? Please prepare your charts for technical presentation with proper units and axis labels.
PC02: Make clear and complete arguments or statements by fully addressing all parts of the assignment.	DEVELOPING	Explain your outputs	How did you arrive at 3.2 cm and 0.92? While the spreadsheet addresses all aspects of the problem, the outputs (answers to part 4a and 4b) need further elaboration.
SQ01: Use accurate, scientific, mathematical, and/or technical concepts, units, and/or data in solutions.	EMERGING	Wrong answers for question4	You multiplied your terminal velocity by Pi twice and you forgot to account for the density of air. These errors led to an incorrect answer for part 4 (a).

Visualize Competency Emphasis and De-Emphasis Areas with Data
The visualization of competency focus areas helps determine the alignment between competency emphasis and de-emphasis areas and the intended goals of a course. Figure 2 shows that the 10 targeted competency areas were not emphasized equally.

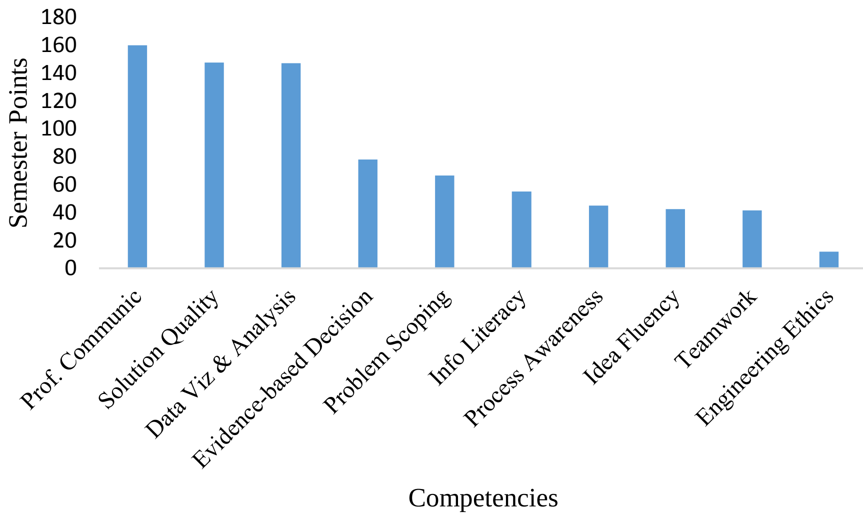

Figure 2. The available grading points for each competency showing the respective emphasis and de-emphasis areas

Iteration and Improvement
The project team gathered input from instructors who implemented the framework. Instructor reactions included the challenges of cognitive load when assessing for multiple learning objectives within an assignment with multiple components. The educators also wanted more specific guidance and professional development on how to assess learning objectives with overlapping components and strategies for giving quality feedback.

As a result of the first implementation several activities and revisions took place:

- Emphasis and de-emphasis areas have been identified and discussed among instructional leadership team.
- Several learning objectives have been revised to improve clarity. In addition, we added new learning objectives that were relevant but left out in the first iteration.
- Mastery levels for several learning objectives have also been refined.
- The learning objectives associated with each assignment have been carefully reviewed to reduce the number of learning objectives targeted in order to reduce the cognitive load on assessors and time for quality feedback.
- A new competency area will be added to cover engineering concepts that are found to be important by faculty but were not included in the initial list.
- Assignments (assessment tasks) are refined to better align with learning objectives.
- A multi-phase professional development for effective assessment and feedback has been developed to support quality feedback given to students.

Future Research

In this paper, we presented an alpha version of the *Assessment Framework for First-Year Engineering Courses*. The refinement of the framework will include an iterative process of testing and research-informed revisions using design-based research (DBR)[41] to systematically study the framework and its appropriate uses. Specific pieces of evidence will address the following questions:

- Does the use of the *Framework* produce actionable information that informs instructional decisions?
- Do the instructors and graders see value in the *Framework* and are they able to use it effectively?
- Has the use of the *Framework* helped advance student learning and students' perceptions of learning?

Discussion

First-year engineering courses cover a range of competencies and objectives. The assessment of such a diverse set of goals and objectives can be a daunting task for educators. To address this challenge, we developed the *Assessment Framework for First-Year Engineering Courses* to help guide instructional processes to more effectively assess student learning. This assessment-centered teaching and learning framework aims to connect specific learning objectives to broader competencies through on-going formative feedback targeting learning transfer and progression on higher-order abilities. The development of such a framework is an iterative, long-term task. In this paper, we presented an initial framework and results from its first implementation. Our plan is to refine the framework using a design-based research approach. We found that the framework is useful in the first-year engineering courses in which we have implemented it. Future research will focus on further refinement of the framework and research on its broader utility in first-year engineering programs.

Acknowledgements

We would like to thank the team of instructors who significantly contributed to the writing of competencies, objectives, and rubrics in this first-year engineering course.

References

1. Gustafsson, G., Malmqvist, J., Newman, D. J., Stafström, S. & Wallin, H. P. Towards a new model for first-year introductory courses in engineering education programmes. in *NordDesign 2002* (2002).
2. Reid, K., Hertenstein, T. J., Fennell, G. T. & Reeping, D. Development of a first-year engineering course classification scheme. *Am. Soc. Eng. Educ. Annu. Conf. Expo.* (2013).
3. Trevisan, M., Davis, D., Beyerlein, S., Thompson, P. & Harrison, O. A review of literature on assessment practices in capstone engineering design courses: Implications for formative assessment. in *American Society for Engineering Education Annual Conference & Exposition* (2006).
4. ABET. Criteria for accrediting engineering programs 2015-2016. (2014).

5. Committee on the Engineer of 2020 Phase I. *The engineer of 2020: Visions of engineering in the new century*. (The National Academies Press, 2004). doi:10.17226/10999
6. Froyd, J. E., Wankat, P. C. & Smith, K. A. Five major shifts in 100 years of engineering education. *Proc. IEEE* **100,** 1344–1360 (2012).
7. Darling-Hammond, L. *et al.* Criteria for High Quality Assessment. 25 (2013).
8. Ruiz-Primo, M. A., DiBello, L. & Solano-Flores, G. Supporting the implementation of the Next Generation Science Standards (NGSS) through research: Assessment. (2014). Available at: https://narst.org/ngsspapers/assessment.cfm. (Accessed: 10th February 2017)
9. Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H. & Krathwohl, D. R. *Taxonomy of educational objectives, Handbook I: The cognitive domain*. (David McKay, 1956).
10. Webb, N. L. Depth-of-knowledge levels for four content areas. *Unpublished* 1–9 (2002).
11. Heywood, J. *Assessment in higher education: Student learning, teaching, programmes and institutions*. (Jessica Kingsley Publishers, 2000).
12. Yorke, M. Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *High. Educ.* **45,** 477–501 (2003).
13. Hattie, J. in *Tertiary Assessment & Higher Education Student Outcomes: Policy, Practice & Research* (eds. Meyer, L. H. et al.) 259–275 (Ako Aotearoa, 2009).
14. Lidstone, M.-L. & Ammon, P. A key to successful teaching is understanding and focusing on student learning; implications for teacher development. *ERS Spectr.* **20,** 27–37 (2002).
15. Hattie, J. & Jaeger, R. Assessment and classroom learning: A deductive approach. *Assess. Educ. Princ. policy Pract.* **5,** 111–122 (1998).
16. Stiggins, R. J. The unfulfilled promise of classroom assessment. *Educ. Meas. Issues Pract.* **20,** 5–15 (2001).
17. DiRanna, K., Barakos, L. & Carnahan, D. *Assessment-centered teaching: A reflective practice*. (Corwin Press, 2008).
18. National Research Council. *Knowing what students know: The science and design of educational assessment*. *Issues in Science and Technology* **19,** (National Academies Press, 2001).
19. Tekkumru-Kisa, M., Stein, M. K. & Schunn, C. A framework for analyzing cognitive demand and content-practices integration: Task analysis guide in science. *J. Res. Sci. Teach.* **52,** 659–685 (2015).
20. Bleiler, S. K. & Thompson, D. R. Multidimensional Assessment of CCSSM. *Teach. Child. Math.* **19,** 292–300 (2012).
21. Usiskin, Z. The case of the university of chicago school mathematics project—secondary component. *Perspect. Des. Dev. Sch. Math. curricula* 173–182 (2007).
22. Diefes-Dux, H. A., Zawojewski, J. S., Hjalmarson, M. A. & Cardella, M. E. A framework for analyzing feedback in a formative assessment system for mathematical modeling problems. *J. Eng. Educ.* **101,** 375–406 (2012).
23. Beyerlein, S., Davis, D., Trevisan, M., Thompson, P. & Harrison, O. Assessment framework for capstone design courses. in *American Society for Engineering Education Annual Conference and Exposition* (2006).
24. Mann, L., Alba, G. D. & Radcliffe, D. Using phenomenography to investigate different ways of experiencing sustainable design. in *American Society for Engineering Education Annual Conference & Exposition* (2007).
25. Sorby, S. A. Educational research in developing 3-D spatial skills for engineering students. *Int. J. Sci. Educ.* **31,** 459–480 (2009).

26. Jonsson, A. & Svingby, G. The use of scoring rubrics: Reliability, validity and educational consequences. *Educ. Res. Rev.* **2,** 130–144 (2007).
27. Popham, W. J. *Classroom assessment: What teachers need to know*. (Pearson, 2014).
28. Reddy, Y. M. & Andrade, H. A review of rubric use in higher education. *Assess. Eval. High. Educ.* **35,** 435–448 (2010).
29. Andrade, H. G. Using rubrics to promote thinking and learning. *Educ. Leadersh.* **57,** 13–18 (2000).
30. Kellogg, R. S., Mann, J. A. & Dieterich, A. Developing and using rubrics to evaluate subjective engineering laboratory and design reports. in *ASEE Annual Conference & Exposition* 1–10 (2001).
31. Lovorn, M. G. & Rezaei, A. R. Assessing the assessment: Rubrics training for pre-service and new in-service teachers. *Pract. Assessment, Res. Eval.* **16,** (2011).
32. Coso, A. E. & Pritchett, A. The development of a rubric to evaluate and promote students' integration of stakeholder considerations into the engineering design process. *ASEE Annu. Conf. Expo. Conf. Proc.* (2014).
33. Hansen, E. J. *Idea-based learning: A course design process to promote conceptual understanding*. (Stylus Publishing, 2012).
34. Goldberg, G. L. Revising an engineering design rubric: A case study illustrating principles and practices to ensure technical quality of rubrics. *Pract. Assessment, Res. Eval.* **19,** 7714 (2014).
35. Moskal, B. M. Recommendations for developing classroom performance assessments and scoring rubrics. *Pract. Assessment, Res. Eval.* **8,** 1–5 (2003).
36. Moskal, B. M. & Leydens, J. A. Scoring rubric development: Validity and reliability. *Pract. Assessment, Res. Eval.* **7,** 1–10 (2000).
37. Mertler, C. A. Designing scoring rubrics for your classroom. *Pract. Assess. Res. Eval.* **7,** 1–10 (2001).
38. Sadler, D. R. Indeterminacy in the use of preset criteria for assessment and grading. *Assess. Eval. High. Educ.* **34,** 159–179 (2009).
39. Popham, W. J. What's wrong—and what's right—with rubrics. *Educational Leadership* **55,** 72–75 (1997).
40. Huba, M. E. & Freed, J. E. *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning*. (Allyn and Bacon, 2000).
41. Kelly, A. E., Lesh, R. A. & Baek, J. Y. *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching*. (Taylor & Francis, 2014).
42. Shavelson, R. J., Phillips, D. C., Towne, L. & Feuer, M. J. On the science of education design studies. *Educ. Res.* **32,** 25–28 (2003).
43. Hjalmarson, M. A. & Lesh, R. Design research. Engineering, systems, products, and processes for innovation. *Handb. Int. Res. Math. Educ.* **2,** (2008).