Outcomes-Based Assessment Instrument for Engineering Problem-Solving Skills

Miss Bahar Memarian, University of Toronto

Bahar Memarian is a PhD candidate in Industrial Engineering and the Collaborative Specialization in Engineering Education at University of Toronto, Canada. Her primary research interests are in Human Factors and Systems Engineering, specifically their application in Education (e.g. learning outcomes assessment, engineering problem solving). Before that, she completed her MASc. (2015) and BASc. (2012) in Electrical Engineering from University of Toronto.

Dr. Susan McCahan, University of Toronto

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

Outcomes-Based Assessment Instrument for Engineering Problem-Solving Skills

Bahar Memarian, Susan McCahan
Department of Mechanical & Industrial Engineering, University of Toronto, Canada
bahar.memarian@utoronto.ca, susan.mccahan@utoronto.ca

Abstract

This poster paper provides an in-depth analysis of the design of a new assessment instrument. The instrument, Constructive Alignment Integrated Rating (CAIR), is a formative feedback scheme that facilitates the assessment of engineering problem-solving skills. Importantly, this instrument is designed to provide formative feedback on problem-solving skills that transcend the content of any particular problem. The CAIR instrument allows consistent feedback on problem-solving to cut across course and assignment boundaries, thus reinforcing the development of problem-solving skills. Our work is founded on a social constructivist conceptual framework, whereby assessment is viewed as a formative negotiation cycle between three main groups: the instructor, the assessors, and the students. The design of CAIR and how it is used at different points of the assessment cycle are discussed.

Keywords
Assessment, Assessment for Learning, Formative Feedback, Constructive Alignment

Introduction

This poster paper provides an analysis of the design of a new assessment instrument. The instrument was developed based on a social constructivist perspective on assessment and uses test-taking and grading opportunities to provide feedback on the quality of students' engineering problem-solving skills.

Presently, the assessment of engineering problem-solving skills is generally accomplished through grading individual assignments and then summing those grades at the end of the term [1]. As Carberry et al. point out, this type of summative grading does not provide formative feedback on key course learning outcomes and can mask the student's true demonstration of a skill, or lack thereof. Conceptually, assessment in engineering education has been shifting away from a measurement of knowledge perspective to an outcomes-based stance that is meant to measure and provide feedback on the development of transferable skills and the ability to apply those skills in the context of course content [2]. In this approach, assessment is intended to serve as an opportunity for students to practice a skill, demonstrate content proficiency, and to guide student learning and development of transferable skills. Assessment is no longer simply a summative evaluation of performance but instead an ongoing process anchoring student learning and, more specifically, problem-solving skill development.

Cyclic View of Assessment

Testing is an opportunity for students to practice a skill, demonstrate competency, and receive feedback to improve learning [3]. Effective communication between the student and the instructor is needed for the student to learn from their own problem-solving errors and to apply this learning to future assessment activities. In literature, using assessment to improve student learning is also referred to as *assessment for learning*.

Typically, an assessment cycle involves interactions between three main groups: the instructor, the assessors, and the students (see Figure 1). A cycle that follows the principles of constructive alignment ensures that the learning objectives and intended learning outcomes are linked at every step [4]. The cycle starts with the instructor identifying learning objectives for the course that align with the program of study level learning objectives. For a particular assignment, the instructor is responsible for selecting a type of assessment activity and creating the objectives and guides (e.g. the instructions, a grading scheme, and so on). The assessment is deployed, and the student is responsible for responding to the assignment instructions and communicating that response. In many engineering courses, the assessments require the students to demonstrate their ability to solve subject-matter specific closed-ended problems. The response from the student on this type of assignment or test is a written communication that states an answer to the problem and explains how that answer was obtained. Then, the assessors, who could be the instructor or teaching assistants (TAs), are responsible for evaluating the students' work by following the assessment guide developed by the instructor. The evaluation may, or may not, include providing formative feedback on the students' solutions.

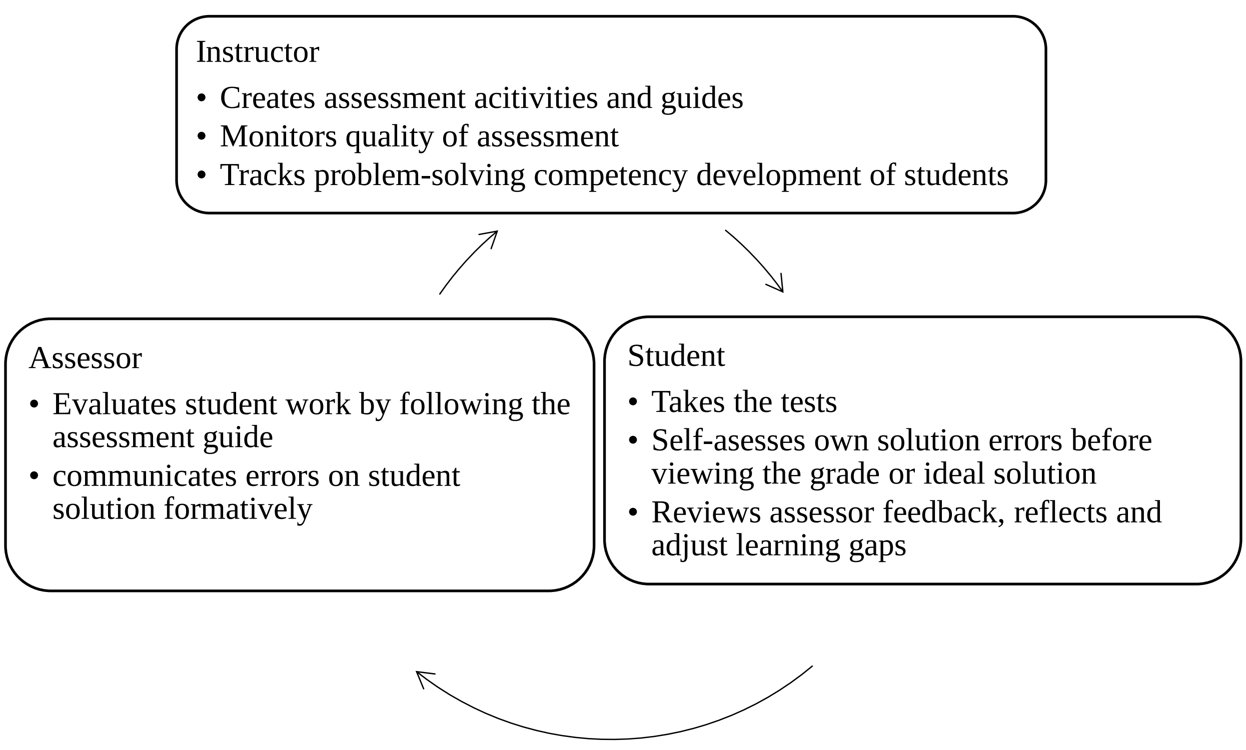

Figure 1 Cyclic view of assessment

Conceptual Framework

The interactions and the type of social negotiations between the groups in the assessment cycle are driven by the societal values and views about how assessment should be administered. It is typical for the emphasis on problem-solving activities to be on the content of the problem. For example, in developing a question about circuit analysis, the instructor may be looking for a correct application of Ohm's Law. This is certainly part of the demonstration of learning that is necessary in a discipline. However, it only captures the content-specific knowledge application. Ideally, problem-solving activities should also result in improved problem-solving competency that transcends the specific subject matter. Learning effective problem-solving in a way that is transferable to other problems, courses, and disciplines would broaden the outcomes from engineering "fact and principle" (i.e., engineering science) type courses.

Effectively, broadening the assessment of outcomes from a pure knowledge perspective to include problem-solving competencies is not simply a matter of changing the assessment approach from summative grading to one that includes feedback. The conceptual framework of assessment needs to change. This idea was captured by Jonassen nearly thirty years ago when he contrasted objectivism with constructivism [5]. As a pioneer in constructivism, Jonassen observed that every individual has a unique understanding of the world. The role of assessment then becomes both facilitating the development of that individual meaning-making and measuring whether the individual can apply that understanding to effectively solve problems. Epistemologically, we are still maintaining that there is a single, correct answer to a closed-ended problem that will align with our current understanding of physics. However, the way the student understands engineering science concepts and perhaps the way they get to the solution may be unique and this individuality should be supported in the assessment process.

Therefore, grading a problem in a rigid way that demands that the student replicate a set solution path to get full credit misses the point. The point is to facilitate the development of the student's problem-solving ability, not their ability to memorize a solution path and regurgitate it in a way that superficially replicates the instructor's thinking. Rust et al. [6] suggested a number of ways to engage a constructivist mindset in assessment including asking students to reflect on the feedback they receive; giving students feedback without a grade to increase engagement with the feedback; providing common feedback to the class and asking students to identify the feedback that applies to their work. The ultimate goal of engineering education is to teach the process rather than the product of engineering problem-solving [7], [8].

Literature Review

Outcomes-based assessment research in engineering education has been mainly focused on problems that are open-ended in nature (e.g. design problems) or enabling skills more broadly (i.e. ethics, communication). For an open-ended problem, multiple viable and correct solutions exist. Students' writings, portfolios, or design-based projects, laboratories, or fourth-year capstone projects are areas in which outcomes-based research has been extensively investigated [9]–[12].

Most of the work done on closed-ended problem solving is related to aiding students with self-regulation and building their problem-solving capability, rather than aiding the feedback process.

Examples of the former include models of problem-solving in engineering and information processing [13]–[17]. These models provide guiding principles and strategies for students on how to approach engineering problem solving and communicate their thought processes. These models, however, were not intrinsically developed for formative assessment purposes.

In engineering education literature, the latest work on the constructivist-based assessment of problem-solving skills can be seen in the development of rubrics and coding schemes created for the assessment of universal engineering problem-solving skills. Two examples are the VALUE rubrics and the PROCESS coding schema [18], [19]. The VALUE rubrics follow the typical models of engineering problem solving closely and further provide descriptors to categorize the quality of student problem-solving skills at each step. However, VALUE rubrics are designed for program-level use in evaluation and discussion around student learning and not for grading per se [18]. The PROCESS rubric, in contrast, evaluates the tasks, errors, strategies, and accuracy of a student's solution via a 54 item coding scheme. Like the problem-solving models, the coding scheme has been investigated from the student, rather than the assessor perspective. A study by Call et al. [20], for example, tested the PROCESS rubric qualitatively to investigate the cognitive strategies used by engineering students, and gaps in their learning, on 2-D and 3-D force equilibrium concepts in engineering statics. Both of these approaches support outcomes-based learning strategies.

This work of Jonassen, Woods and others points to the need for assessment tools that can differentiate between critical thinkers and memorization of a solution path [14], [21], [22] There is a concern that current tools do now allow the assessor to identify whether the student has grasped the conceptual goal of a problem or has just stated key principles governing the problem from memorizing a similar solution from examples. Grading identifies points of error in a solution, but does not provide formative feedback on the error usually [23]. Rubrics identify, in some sense, what is right about a solution by showing students how they did relative to an optimal (e.g. exceeds expectations) performance, but do not provide explicit feedback on points of error. We are seeking to design a tool that works at this intersection: one that explicitly identifies points of error and provides feedback on the nature of errors in a student's solution relative to engineering problem-solving competencies.

Methods

To address this gap, we have designed a new formative feedback instrument for the assessment of engineering problem-solving skills that takes a constructivist perspective. Our work aims to bridge features of grading and feedback schemes by explicitly identifying errors and providing feedback relative to optimal performance. Our area of interest is the assessment of engineering problem-solving skills in the context of paper and pencil activities; well structured and information-rich problems [24]. Such problems are frequently used in engineering courses through tests and problem sets. The design of the new instrument was inspired by the concept of Work Domain Analysis (WDA) from the domain of human factors engineering [25], [26]. The method of WDA has been used to support engineering problem-solving and troubleshooting in industry [27]. Here we use WDA to assist the work of assessors because our goal is to facilitate the identification and classification of errors in engineering problem solutions.

WDA proposes that the scope for the design of a system should be shaped by the underlying goals and constraints in the environment (i.e., the work domain). In this case, the work domain is the work done to assess and provide feedback on problem solutions submitted by students. Gathering the needs from assessment literature informed the scope of the design. This information, classified using a WDA centered lens, is presented in Table 1. The primary function of the new tool is to facilitate the outcomes-based assessment of problem-solving skills on each solution to support both student learning and institutional assessment administration. Shaeiwitz et al. [28] suggest that a well-grounded outcomes-based assessment should: 1) offer a statement of educational goals, 2) offer multiple measures of achievement of these goals and 3) use the information gathered to correct and improve the educational process. For these functions, we gather the key objectives expressed in the literature and offer our design approach to achieve these objectives.

To offer multiple measures of a learning outcome, as Shaeiwitz [28] suggests, an assessment instrument must work across different tests and/or problem sets. Therefore, the new tool needs to use criteria universal to generic engineering problem-solving skills so that it can be applied to a variety of engineering problem-solving instances. Doing so can allow collecting multiple measures of the skill over time and checking to see if students are on the intended learning trajectory. This can also provide feedback to students that can guide their self-directed learning.

A tool that uses the information gathered to correct and improve the educational process needs to effectively communicate relevant information to each user group, in particular, the students. The feedback produced by using the tool should be consistently formative, allowing engineering problem-solving skill development feedback to be concretely linked to information on the student's solutions. We decided to use a coding convention (e.g. letter/symbol) in the instrument so feedback delivery is not time-intensive and learning outcomes information can be consistently communicated across problems. Further, using a tagging approach to capture errors on the student's solution can help guide and inform students about the location and nature of their errors.

Table 1. Scope of the design. The primary function is to facilitate assessment for learning.

Function:	Objectives	The design approach
Offer Statement of educational goals	Should follow the criterion-based perspective	Use feedback rather than grading scheme to communicate
Offer multiple measures of achievement	Should be applicable to a variety of engineering problem-solving instances	Use criteria universal to generic engineering problem-solving skill
Use the information gathered to correct and improve the educational process	Should be consistently formative, allowing engineering problem-solving skills development to be directly traceable from student solutions	Use letter/symbol coding to reduce marking time and learning outcomes information can be consistently communicated across assessments

The development of the CAIR instrument

Using the principles that inform effective formative assessment design, we create the organization of the tool and its content. To create the instrument, we borrowed another concept from industrial engineering; the abstraction decomposition space (ADS) which can be a table. An ADS can be used to communicate the structure of processes, such as engineering problem solving, that are hierarchical by nature.

In engineering problem solving the student ideally: 1) has a deep understanding of a problem's goal at a high level, 2) has a correct understanding of the theoretical principles needed to solve the problem, and 3) must apply a computational approach correctly to arrive at an end result. These three things must occur concurrently, not necessarily serially, for a successful result. Basically, we can evaluate the problem solution from each of these three perspectives. We utilize this idea to classify the errors that may be present in a student's solution (see Figure 2).

There are two dimensions to the ADS structure; the levels of abstraction (shown in rows of the table) and the levels of decomposition (shown in columns of the table). The levels of abstraction offer different representations of the same entity, in this case, the problem solution. The levels of decomposition examine the quality of the problem-solution from a particular abstraction level perspective. The levels of decomposition are used to distinguish between errors that stemmed from lacking fundamental understanding (deep errors) versus superficial or careless mistakes (surface errors).

To reduce marking time, and ensure a consistent degree of formative feedback delivery, an alphabetic tagging convention was developed. Assessors are asked to tag each error in a student's solution with an alphabetic symbol (see Figure 2). The alphabetic symbols, though somewhat arbitrary, loosely align with the abstraction and decomposition level. A different symbol set could be chosen by the instructor to fit their course.

CAIR is used in a way that is fundamentally different than a rubric or traditional grading scheme alone. The instructor provides an ideal solution along with the total mark for the problem to the assessors (e.g. Teaching Assistants). Assessors compare the student's solution to the ideal solution to identify, circle, and then tag errors on the student's solution using the alphabetic tagging system to indicate the nature of the error. The tag on the student's work serves as formative feedback that carries more information than a mark deduction alone. In addition, the type and number of errors in each category can be used to determine the performance of the student at each abstraction level: fails to meet expectations, below expectations, meets expectations or exceeds expectations. In this way, the instrument acts as a learning outcomes-based rubric. The instructor could, if they wished, use the type and number of errors to determine a grade formulaically.

Tag errors on student solution based on an alphabetic convention		
	Deep Decomposition	Surface Decomposition
Problem-solving Goal Abstraction:	**U**nknown variables *E.g. Found variables problem had not asked for.*	**K**nown variables *E.g. Did not use the necessary variables problem provided.*
Fails○ Below○ Meets○ Exceeds○		
Engineering Principles Abstraction:	**T**heoretical model *E.g. Wrote an incorrect expression/formula for theory.*	**D**isciplinary standards within the model *E.g. Did not follow assumptions, conversions of the theoretical model.*
Fails○ Below○ Meets○ Exceeds○		
Mathematics Abstraction:	**C**omputational technique *E.g. Rules of standard computational technique violated.*	**A**rithmetic work within the technique *E.g. Arithmetic calculations made incorrectly.*
Fails○ Below○ Meets○ Exceeds○		

Figure 2. The CAIR feedback instrument

Work to date

Our work to date has been primarily focused on electrical engineering problem-solving activities [29], [30]. We have examined the distribution and types of feedback assessors most commonly provide on graded student solutions. We have also investigated the assessors' self-reported experiences toward assessment when using grading schemes to evaluate student work [31], [32] Our findings suggest that the efficiency of grading schemes make is attractive for assessors. However, the feedback yielded from grading schemes are mostly symbolic and minimally descriptive. The findings support the design decisions that led to the development of CAIR. Preliminary testing of an early version of the instrument suggests that it has merit as an alternative to traditional marking or rubrics [29].

Discussion

The goal for the design of CAIR is to facilitate access to rich data relevant to each user group. Aligned with a social constructivist perspective, we designed a feedback scheme that can allow: the students to receive formative information from test-taking opportunities, the assessors to provide formative feedback directly relatable to solution errors in a time-efficient manner, and the instructors to track the quality of feedback from the assessors. In addition, it allows instructors to track the problem-solving skill development of the students across problems, assignments, and courses.

CAIR facilitates outcomes-based assessment by using feedback rather than a grading scheme and communicating students' engineering-problem solving error types. The structural layout of CAIR and, particularly, the levels of abstraction of CAIR enable differentiating between problem solvers who understand the goal of a problem in context, versus those who memorize solution profiles. The levels of decomposition of CAIR, on the other hand, enable categorizing the severity of errors.

The assessment cycle, shown in Figure 1, is repeated multiple times in engineering courses and across the curriculum. CAIR can contribute to achieving high-quality outcomes-based assessment of engineering problem-solving activities at the course and curriculum levels. Using a consistent system that measures problem-solving outcomes, such as CAIR, allows every assessment activity to contribute evidence on engineering problem-solving competency development. For example, analytics on assessment data using this tool can help the instructor check for the quality of the marking activity; e.g., the reliability (i.e. inter-rater, intra-rater) with which assessors provide feedback across different solution qualities (e.g. Fails, Below, Meets and Exceeds expectations). Also, if the instructor finds that many students in the class exhibit one or a combination of errors frequently, this may suggest that the instructor needs to reinforce a particular concept or problem-solving skill.

When using CAIR, the nature of the error is clear from the tag. It is not implicit in a mark deduction. This can have two benefits. First, more formative feedback is provided making the thought process of the assessor more transparent. Second, the assessor does not directly decide a grade and instead, the system can be used to automatically calculate the grade based on the nature of errors captured. We speculate that this could reduce re-mark requests from students and arguments concerning mark deductions. If mediated by a computer, the entire evaluation activity could be simplified to clicks on the student solution (to circle an area) and selection of a tag.

With CAIR, students are able to reflect on the assessor's feedback and self-assess their own problem-solving. Best practices would suggest giving the students the ideal solution after the test and asking them to self-assess their own work using the CAIR feedback scheme. The students then receive their marked paper and can compare and contrast their own evaluation with that of the expert assessor. Or, alternatively, providing the CAIR marked paper to the student (tagged and with a performance assessment at each abstraction level) and then asking them to reflect on this formative information before they receive the numerical grade.

Self-assessment, when it happens before receiving assessor feedback, can build metacognition and critical thinking capacities, yet it is often left out of the assessment picture. With grading schemes, because feedback is often implicit (e.g. check marks, cross marks, grade deductions, text, and so on), self-assessment can devolve into grade disputes. CAIR encourages the student to reflect on the nature of the error, not just whether the point deduction is justified. Put another way, both the assessor and student may agree on the number and location of errors in a solution, but not how many points should be deducted – a conversation which should be secondary to the learning outcomes.

The design and use of the CAIR tool aligns with a social constructivist perspective. Reality is an internally developed set of constructs [5], which are only apparent when they are effectively communicated to an external audience. CAIR is designed to facilitate this communication – to make the thinking processes of the assessors transparent to the students. Contrary to the traditional measurement of knowledge perspective, learning is not achieved by accepting assessors' grading as the absolute truth, but rather learning happens when the student compares their self-assessment with that of the assessors. The work by VanLehn and colleagues on student and tutor interactions showed that students are more likely to gain knowledge at what they call an "impasse point", that is a point of error in the solution [33]. The use of CAIR encourages the student to examine their work and obtain information about their impasse points during the assessment cycle.

Conclusion

There are three key contributions made through this work. First, CAIR demonstrates an approach that enables a focus on the process rather than the product of engineering problem-solving. Grading schemes, when used alone, emphasize the numerical end result [34]. However, using this alternative approach, we can examine whether the student has understood the overarching goal of a problem, utilized appropriate engineering principles, and carried out computational work completely. Second, this information is valuable for outcomes-based assessment of a course and a program. The consistent use of CAIR could allow aggregation of outcomes-based data at the course and the program levels. Third, the formative nature of CAIR using tags opens up the opportunity for the students to self-assess their own work without focusing solely on grades. While rubrics also allow this type of reflection, they do not explicitly allow the student to connect the rubric criteria to an identified point in their work that illustrates the criteria. By using tagging, CAIR concretizes the feedback.

Future work could include testing CAIR in large classes and could examine how grading is coordinated with respect to CAIR dimensions across different conceptual domains as well as how the instructors make use of data at the end of the assessment cycle. CAIR could also be examined longitudinally to identify trends. In addition, direct data collection, the assessors' thought process and judgment need to also be collected and analyzed through qualitative means. Thorough benchmarking could be provided to support inter-rater reliability. To get a deeper understanding of the evaluation trajectory, a computer-mediated version of CAIR could track and record the type of tags assessors select when evaluating each solution. When testing CAIR with the student group, future work could study how students approach self-assessment using CAIR. Along with the outcomes-based performance analysis, it would be beneficial to examine how CAIR influences student problem-solving strategies over time.

References

[1] A. Carberry, M. Siniawski, S. A. Atwood, and H. A. Diefes-Dux, "Best practices for using standards-based grading in engineering courses," in *Proceedings of the ASEE Annual Conference & Exposition*, 2016.

[2] R. L. Olds, B. M., Moskal, B. M., & Miller, "Assessment in engineering education: Evolution, approaches and future collaborations," *J. Eng. Educ.*, vol. 94, no. 1, pp. 13–25, 2005.

[3] R. Stiggins, "From formative assessment to assessment FOR learning: A path to success in standards-based schools," *Phi Delta Kappan*, vol. 87, no. 4, pp. 324–328, 2005.

[4] J. Biggs and C. Tang, "Applying constructive alignment to outcomes-based teaching and learning," *In Training material for "quality teaching for learning in higher education" workshop for master trainers*. Ministry of Higher Education, Kuala Lumpur, pp. 23–25, 2010.

[5] D. H. Jonassen, "Objectivism versus constructivism: Do we need a new philosophical paradigm?," *Educ. Technol. Res. Dev.*, vol. 39, no. 3, pp. 5–14, 1991.

[6] C. Rust, B. O'Donovan, and M. Price, "A social constructivist assessment process model: how the research literature shows us this could be best practice," *Assess. Eval. High. Educ.*, vol. 30, no. 3, pp. 231–240, 2005.

[7] D. Jonassen, J. Strobel, and C. B. Lee, "Everyday problem solving in engineering: Lessons for engineering educators," *J. Eng. Educ.*, vol. 95, no. 2, pp. 139–151, 2006.

[8] NAE, "The engineer of 2020: Visions of engineering in the new century," National Academies Press, DC, 2004.

[9] C. Plumb and C. Scott, "Outcomes assessment of engineering writing at the University of Washington," *J. Eng. Educ.*, vol. 91, no. 3, pp. 333–338, 2002.

[10] A. C. Heinricher, J. E. Miller, L. Schachterle, N. K. Kildahl, V. Bluemel, and V. Crawford, "Undergraduate learning portfolios for institutional assessment," *J. Eng. Educ.*, vol. 91, no. 2, pp. 249–253, 2002.

[11] J. H. Hanson and J. M. Williams, "Using writing assignments to improve self-assessment and communication skills in an engineering statics course," *J. Eng. Educ.*, vol. 97, no. 4, pp. 515–529, 2008.

[12] B. Frank, N. Simper, and J. Kaupp, "Formative feedback and scaffolding for developing complex problem solving and modelling outcomes," *Eur. J. Eng. Educ.*, vol. 43, no. 4, pp. 552–568, 2018.

[13] G. Polya, "How to solve it," New Jersey: Princeton university press, 1957.

[14] D. R. Woods, "An Evidence-Based Strategy for Problem Solving," *J. Eng. Educ.*, vol. 89, no. 4, pp. 443–459, 2000.

[15] J. Pérez, C. Vizcarro, J. García, A. Bermúdez, and R. Cobos, "Development of Procedures to Assess Problem-Solving Competence in Computing Engineering," *IEEE Trans. Educ.*, vol. 60, no. 1, pp. 22–28, 2016.

[16] J. H. Larkin and F. Reif, "Understanding and teaching problem-solving in physics," *Eur. J. Sci. Educ.*, vol. 1, no. 2, pp. 191–203, 1979.

[17] M. T. Chi, R. Glaser, and E. Rees, “Expertise in problem solving,” in *R. Sternberg (Ed.), Advances in the psychology of human intelligence*, NJ: Erlbaum, 1982.

[18] AAC&U, “Problem Solving VALUE Rubric,” 2009. [Online]. Available: https://www.aacu.org/value/rubrics/problem-solving. [Accessed: 01-Jan-2017].

[19] S. J. Grigg, “A process analysis of engineering problem solving and assessment of problem solving skills.,” Dissertation. Clemson University, Clemson, SC., 2012.

[20] B. J. Call, W. H. Goodridge, and C. Green, “Strategy, Task Performance, and Behavioral Themes from Students Solving 2-D and 3-D Force Equilibrium Problems,” in *2015 ASEE Annual Conference & Exposition*, 2015, pp. 26–1405.

[21] J. M. Woods, D. R., Kourti, T., Wood, P. E., Sheardown, H., Crowe, C. M., & Dickson, “Assessing Problem-Solving Skills Part 1: The Context for Assessment,” *Chem. Eng. Educ.*, vol. 35, no. 4, pp. 300–7, 2001.

[22] J. M. Woods, D. R., Kourti, T., Wood, P. E., Sheardown, H., Crowe, C. M., & Dickson, “Assessing Problem Solving Skills Part 2: Assessing the Process of Problem Solving,” *Chem. Eng. Educ.*, vol. 36, no. 1, pp. 60–67, 2002.

[23] B. Memarian and S. McCahan, “Analysis of Feedback Quality on Engineering Problem Solving Tasks,” in *Proceedings of the American Society of Engineering Education (ASEE)*, 2019.

[24] R. Taraban, C. Craig, and E. E. Anderson, “Using Paper-and-Pencil Solutions to Assess Problem Solving Skill,” *J. Eng. Educ.*, vol. 100, no. 3, pp. 498–519, 2011.

[25] K. J. Vicente, *Cognitive Work Analysis - Toward Safe, Productive, and Healthy Computer-Based Work*. CRC Press, 1999.

[26] J. Rasmussen, A. M. Pejtersen, and K. Schmidt, *Taxonomy for cognitive work analysis*. Roskilde, Denmark: Risø National Laboratory, 1990.

[27] N. Naikar, R. Hopcroft, and A. Moylan, “Work domain analysis: Theoretical concepts and methodology (No. DSTO-TR-1665),” *Def. Sci. Technol. Organ. Victoria Air Oper. Div*, 2005.

[28] J. A. Shaeiwitz, “Outcomes assessment in engineering education,” *J. Eng. Educ.*, vol. 85, no. 3, pp. 239–246, 1996.

[29] B. Memarian and S. McCahan, “Performance Analysis: Feedback Checklist versus Traditional Grading,” in *Proceedings of the Human Factors and Ergonomics Society (HFES)*, 2019.

[30] B. Memarian and S. McCahan, “Systematic Analysis of Formative Feedback, Focus on Electrical Engineering Assessments,” in *Proceedings of the American Society of Engineering Education (ASEE)*, 2019.

[31] B. Memarian and S. McCahan, “Exploring Challenges of Assessing Students’ Errors in Engineering Problem Solving; a Thematic Analysis,” in *Proceedings of the Canadian Engineering Education Association Annual Conference*, 2018.

[32] B. Memarian and S. McCahan, “Usability Analysis of Freeform Marking on Engineering Problem Solving,” in *Proceedings of the Human Factors and Ergonomics Society (HFES)*, 2018.

[33] K. VanLehn, S. Siler, C. Murray, T. Yamauchi, and W. B. Baggett, “Why do only some events cause learning during human tutoring?,” *Cogn. Instr.*, vol. 21, no. 3, pp. 209–249, 2003.

[34] J. Docktor and K. Heller, “Robust assessment instrument for student problem solving,” in *Proceedings of the NARST 2009 Annual Meeting*, 2009, pp. 1–19.