International Journal of Engineering Education Vol. 37, No. 3, pp. 672–689, 2021
Printed in Great Britain

0949-149X/91 $3.00+0.00

Assessing Students' Design Processes and Design Outcomes*

DOUGLAS B. CLARK**, CHRISTOPHER OSTROWDUN and STEFAN ROTHSCHUH
Werklund School of Education, University of Calgary EDT 1340, 2500 University Drive NW, Calgary, Alberta T2N1N4, Canada.
E-mail: douglas.clark@ucalgary.ca, chris.ostrowdun@ucalgary.ca, stefan.rothschuh@ucalgary.ca

SENAY PURZER
School of Engineering Education, Purdue University, 610 Purdue Mall, West Lafayette, IN, 47907, USA. E-mail: senay@purdue.edu

Classroom assessments can affect what is valued, taught, and learned by teachers and students. It is therefore of great importance to review the approaches proposed for assessing student design (processes and outcomes) given the increasing focus on design in STEM curricula. We reviewed 17 prominent journals that address STEM and design education. Inclusion criteria and search terms were identified through a systematized process. Database searches resulted in 2101 raw hits. Articles that straddled the borders of the inclusion criteria were first reviewed at the abstract level by the researchers. We then reviewed the main text of each article and evaluated whether the assessment instruments therein were described in sufficient detail to meet the criteria. The literature search resulted in a sample of 27 articles. Most of the articles (23 of 27) were in engineering education journals and four were in a design journal. We performed a content analysis of the final 27 articles with a goal of identifying key components of assessment in terms of design foci, student age, evaluator type, project type, and granularity of assessment. The majority (20) included a focus on performance of the design, and a substantial number included a focus on communication (15) and scoping (11). While less prevalent, divergent thinking (9), creativity (9), convergent thinking (8), and collaboration (8) were also broadly represented. Ethical considerations were not strongly represented in the reviewed assessments, although there were notable exceptions.

Keywords: design, assessment; design processes; design outcomes; review

1. Introduction

The structure and content of assessments drive what is valued by educators and learned by students [1–4]. According to Hattingh, Dison, and Woollacott [5], the influence of classroom assessment practices on student learning is particularly evident in engineering classrooms. Given that assessment drives significant aspects of learning and teaching, researchers argue that systematic approaches and clear assessment strategies are necessary to promote deep learning [5], especially when assessing constructs with multiple dimensions [6].

The multi-faceted nature of design is evident in Crismond and Adams's [7] definition of design as a goal-directed problem-solving activity that targets user needs by optimizing parameters and balancing trade-offs. First, assessment in design requires nuance due to the open-ended nature of design problems with multiple solutions [8, 9]. Second, design requires complex reasoning processes and switching between divergent and convergent ways of thinking [10]. Third, as a function of its pragmatic nature, design is inherently interdisciplinary, often requiring a combination of technical and scientific principles, business and economics, and human sciences [11–13].

There have been multiple efforts aimed at synthesizing assessment practices in engineering and design. For example, prior studies have compared metrics for assessing design creativity [14, 15]. Purzer, Fila, and Nataraja [6] examined classroom assessment practices related to the entrepreneurial aspects of engineering design. Cardella and colleagues [16] brought together engineering educators who use novel approaches that elicit meta-thinking about informed design practices such as problem scoping and iteration. Further, the increased importance of design in K-12 education prompted reviews of assessment practices [17, 18] and the addition of technology and engineering literacy to assessments by the National Assessment of Educational Progress (NAEP) [19]. It is therefore of great importance to review the approaches used for assessing design thinking and practices given the increasing attention to design in engineering and more broadly in STEM education.

To gauge the assessment of design, we performed a systematized review [20] and content analysis of literature from seventeen prominent journals that address STEM and design education. We selected journals based on Journal Citation Reports (JCR), Scientific Journal Rankings (SJR), and h-index statistics. Ultimately, only journal articles in engineering education journals and a design studies journal met the criteria for the search and inclusion.

** Corresponding author.

* Accepted 10 December 2020.

Table 1. Design Practices and Ways of Assessing Them

Design Stage	Purpose and Design Behaviors	Potential Ways of Assessing
Problem Scoping	Build knowledge: Treat design as problem framing as opposed to problem-solving. Gather information to learn about the need/problem, the system, and prior solutions. Talk with stakeholders and potential primary and secondary users. Determine design specifications: Take time to better explore implicit needs, articulate assumptions, and better understand the problem. Define design constraints, criteria, and metrics.	Evaluation of design reports for information gathering used for problem scoping. Performance tasks that ask students to frame a problem but not solve it.
Idea Generation	Generate ideas fluently: Generate alternative ideas fluently without judgment and fixation. Explore alternative solutions with an open mind, divergent thinking, and using deliberate ideation strategies. Represent ideas fluently: Model ideas with multiple representations (sketches, elaborations, and quick prototypes) to explore and share design ideas.	Ideation tasks with sketched solutions evaluated for utility, variety, novelty, elaboration.
Evidence Gathering	Gather evidence in a systematic, planned manner: Conduct experiments to test performance of solutions, gather user input, and explore ways to optimize prototypes.	Analysis of design reports for data use and in-text citations.
Evaluation and Communication	Compare design alternatives. Organize and synthesize evidence to allow systematic comparison of design options. Recognize necessity of tradeoff decisions. Communicate: Form arguments for recommending a specific design option/solution. Use data displays to convince clients and supervisors. Articulate novelties as well as limitations of proposed solutions.	Design review sessions that elicit evidence in decision-making. Evaluating the quality of selected design in meeting the targeted requirements.
Reflection and Revision with Feedback	Reflect on process: Practice reflective thinking at all stages of design. Revise and iterate: Seek feedback and embrace iteration as a necessary element of the design practice.	Reflection and evaluation tasks asking students to evaluate their or another designer's design processes.

In this paper, we share our review of these articles and consider implications for engineering, design, and STEM education more broadly. We analyzed how educators assess students' design processes and outcomes in STEM and design education courses. We explored the following specific questions: (a) what disciplines are involved, (b) what aspects of the design process and/or outcomes are assessed, (c) what approaches are used to assess design, (d) at what granularity are the assessments conducted, (e) who is involved in the assessment process, and (f) what age groups of students are the focus. This paper contributes to STEM education a synthesis and an overview of current design assessment approaches to help inform design educators in their assessment practices.

1.1 Importance of Quality Assessment in Design Education

Design is central to engineering education, but it is also challenging to teach and even more challenging to assess [16, 21]. Student design tasks are often open-ended with many viable solutions [22]. Classroom design tasks and assessments help new designers develop the necessary skills to manage such complexities. Quality assessments help offer valuable inferences on student learning and misconceptions of design, which can inform curriculum and instructional scaffolds to improve design solutions and design processes [16, 23]. However, as with any assessment, it is important to have a clear definition and understanding of the construct to be assessed. Assessments should also articulate the practices, knowledge, and competencies students should demonstrate. As well, research should inform assessment development, such as studies on design expertise and student designers.

1.2 Determining What Should and Could be Assessed in Design

Design is a core student outcome in engineering education (and similar fields) [24] and an explicit component for accrediting engineering degree programs (e.g., ABET). Scholars have also argued that design knowledge and knowing are essential elements of the epistemology of engineering [25]. While there are contextual differences across and within disciplines [26], many researchers [7, 27, 28] agree on a common set of principles that define design processes. These principles are based on research examining expert behaviors, typically through case studies, observations, and verbal protocol studies [27, 29, 30]. In addition, while the designed products (i.e., objects, artifacts, or systems) often take the center stage, product quality depends highly on process quality. Table 1 provides a synthesis of key design practices described in these prior studies [e.g., (7, 31)].

According to Mehalik and Schunn [32], design education must focus on key elements of design that have been documented to support achieving effective design, such as framing a design problem, searching for alternative solutions, and using an iterative strategy. First, informed design can entail substantial time in scoping a problem to build knowledge and understanding of users' needs. When problem scoping, designers frequently engage clients and users to understand their needs and lived experiences. Second, through divergent and generative thinking, designers generate alternative ideas (i.e., ideation) and model these ideas in multiple ways [7, 33]. Designers use deliberate ideation strategies to avoid premature decision-making. Third, designers also collect and analyze data, practice information literacy, and organize evidence as they transition from divergent thinking to convergent thinking (i.e., to evaluate design alternatives). Fourth, design often occurs in social spaces as designers communicate, convince, and negotiate with other designers and stakeholders. Designers often work in teams, where individuals bring their unique expertise to a project. In addition, designers commonly use design review sessions to facilitate a feedback exchange processes [24] to ensure they meet the needs of users and clients. Fifth, iteration is essential in the design process and quality of a design [27, 34]. The many facets of design problems necessitate ongoing examination and re-framing of problems. Given these multiple dimensions of design, our systematized review sought to determine what aspects of design are targeted and assessed in design education.

2. Methods

Our goal was to identify peer-reviewed articles about design assessment and to analyze how educators assess students' designs in STEM and design education. Within this broader goal, we explored the following questions:

- What disciplines are involved in the classroom assessment of design?
- What aspects of the design process and/or outcomes are assessed?
- What approaches are used to assess design (e.g., rubrics, metrics)?
- At what granularity are the assessments conducted?
- Who is involved in the assessment process?
- What age groups of students are the foci?

We conducted a systematized [20] literature review of articles from seventeen prominent journals that address STEM and design education. The review was systematized [20] in the sense that it adhered to most criteria of a systematic review [35, 36] but diverged in terms of (a) our specific selection of prominent journals from which to draw rather than opening the search to broader databases and (b) the fact that our coders separately coded a subset of the data, worked to consensus on that subset, and then separately coded the remainder of the data. It is a "state-of-the-art" style of review [20] in the sense that it focuses on the "[c]urrent state of knowledge and priorities for future investigation and research" [20, p. 95]. The following sections detail our procedures for the literature search, article selection, and analysis.

2.1 Journal Selection

Our search focused on leading journals in STEM education and design studies disciplines to understand prominent approaches to assessing students' designs across the STEM education and design studies disciplines. We therefore purposefully focused our search on prominent journals in these fields/disciplines. We selected journals based on Journal Citation Reports (JCR), Scientific Journal Rankings (SJR), and h-index statistics. Table 2 lists the journals and databases searched for each discipline. Multiple databases were required because not all journals were indexed in the same databases. We bounded our review to articles published from 2002 through August 2018. This timeframe approximated the introduction of the Accreditation Board for Engineering and Technology (ABET) Engineering Criteria 2000 (www.abet.org), which spurred a surge in publications about engineering and design education. We aimed to review the current, rather than historical, state of design assessment.

2.2 Search Terms

We searched across four databases (i.e., ASC, ERC, ERIC, and AA) (Table 2) and used the following search terms and Boolean operators at the abstract and title level: (Design* OR Ideation OR Creativity) AND (Assess* OR Rubric OR Framework OR (Coding Scheme) OR Measure OR Metric OR Instrument OR Method). This scan resulted in 2101 raw hits across the 17 journals.

2.3 Selection Criteria

Through iterative discussions about the scope of our review and our research questions, we developed the following inclusion criteria at the abstract level, which we applied to the 2101 raw hits:

Proposing/presenting/validating design assessment. Abstracts must have explicitly stated the article was presenting/proposing/validating an approach to assessing design processes, products, or outcomes. The abstract must have explicitly communicated that the manuscript shared or proposed a specific assessment approach and indicated that assessment was a focus of the article. It was insufficient to only mention assessment without any details or indication of an intentional approach to assessment. Our rationale was that while most design education articles involved assessment, we wished to focus on articles where assessment itself was a key aspect.

Assessing students. Articles must have explicitly stated in the abstract that the assessment is for assessing students in an education setting (e.g., courses, classes). The intended users of the assessment tool may have been instructors of students, students assessing themselves, or others assessing students. The assessment approach must be feasible to use in a real-world learning environment, not just as a research instrument or only feasible in a controlled laboratory environment. Our rationale was to understand how assessment of students' designs in instructional settings is currently envisioned in the literature.

Assessment of design product/process. The assessment must have been specifically applied to design processes or products and not just to personality or cognitive traits of individuals that might be associated with design knowledge or ability (i.e., students' understanding of design process or critique of design process in absence of students engaging in design). Assessments of creativity, self-efficacy, or design knowledge alone were not included. The rationale was to focus on the assessment of designs and design processes specifically and not solely on a students' capacity for design.

Theoretical warranting. Theoretically warranting of an assessment tool was sufficient. Data to support conceptual assessment tools was not required provided the assessment approach was intended for real-world application in a teaching and learning context. We did not want to limit the search solely to validated measures to allow for newer approaches to be included.

Abstract/title. All of the criteria must have been met by the description in the abstract and/or title. The rationale was that the criteria outlined above needed to be salient enough to the purpose of the article.

2.4 Selection Process

We divided the 2101 raw results among three of us (co-authors) for a first-pass culling of the articles based on the selection criteria at the abstract level. Any articles that straddled the borders of the inclusion criteria were reviewed at the abstract level by the group for a final decision. This reduced the sample to 30 articles, all of which were located in engineering education or design journals. We then reviewed the main-text of each article and evaluated whether the assessment instruments therein were described in sufficient detail to meet the criteria listed above.

The literature search resulted in a sample of 27 articles, with 23 of the articles in engineering education journals and 4 articles in a design studies journal.

2.5 Content Analysis and Coding

We performed a content analysis of the final 27

Table 2. Journals Searched

Journal	**Field**	**Database**	**Raw**	**Match**	**Years***
International Journal of Engineering Education	Eng Ed	ERC	510	10	2005+
European Journal of Engineering Education	Eng Ed	ASC	141	6	2002+
Journal of Engineering Education	Eng Ed	ASC	257	4	2002+
Advances in Engineering Education	Eng Ed	ERIC	66	3	2007+
Journal of Pre-College Engineering Education Research	Eng Ed	ERIC	21	0	2011+
Design Studies	D Studies	ASC	193	4	2002+
Design Issues	D Studies	AA	84	0	2002+
International Journal of Science Education	Sci Ed	ASC	207	0	2002+
Journal of Science Education and Technology	Sci Ed	ERIC	147	0	2003+
Journal of Research in Science Teaching	Sci Ed	ERIC	111	0	2002+
Science Education	Sci Ed	ASC	78	0	2002+
Research in Science Education	Sci Ed	ERIC	73	0	2004+
ZDM (Zentralblatt für Didaktik der Mathematik)	Math Ed	ERIC	90	0	2007+
Educational Studies in Mathematics	Math Ed	ASC	55	0	2002+
Journal of Research in Mathematics Education	Math Ed	ASC	11	0	2002+
Journal of the Learning Sciences	Lrn Sci	ASC	42	0	2002+
Cognition and Instruction	Lrn Sci	ERIC	15	0	2002+

Note. Columns describe the journals searched, disciplinary field, search database, initial raw hits from search terms, criteria matches, and years considered.

* Earliest year of publication or availability through our library database from 2002 onwards.

ERC = Education Research Complete, ASC = Academic Search Complete, ERIC = Education Resource Information Center, AA = Arts Abstracts H.W. Wilson, Eng Ed = Engineering Education, D Studies = Design Studies, Sci Ed = Science Education, Math Ed = Mathematics Education, Lrn Sci = Learning Sciences.

Table 3. Coding Categories

Focal Aspects of Design Assessed	
Problem Scoping	Problem analysis, information gathering, objectives, requirements, constraints
Creativity and Divergent Thinking	Idea generation, fluency, brainstorming, creativity, novelty, originality
Design Performance and Functionality	Problem solving, usefulness, client satisfaction
Decision Making and Convergent Thinking	Evaluation, critical analysis, merit of concepts
Collaboration	Group dynamics, collaborative experience, team member citizenship
Communication	Communication skills, delivery and articulation, presentation skills
Assessment Conceptualization and Operationalization	
Education Level Targeted by the Assessments	Third-year engineering students, design students, undergraduate engineering students
Evaluator Categories	Instructor, self-assessment, peer-assessor external assessor (e.g., industry expert or external guest rater)
Nature of Design Task	Large-scale problem, proposed challenge, hands on design activities
Granularity of Assessment	Rubrics, assessment scheme, assessment criteria

articles to identify key components of assessment of students' designs. Two of us divided the 27 articles to code for demographic information (e.g., age, discipline), who did the assessment (e.g., instructor, peer), the nature of the design task, and assessment approaches for various aspects of design. We used a grounded approach [37, 38] to derive the coding categories for the assessment approaches based on the terms and definitions described in the articles wherever possible (e.g., divergent thinking, collaboration, creativity; see Table 3 for the coding categories and descriptions).

Following the initial coding, we reviewed and revised our codes to be inclusive of articles that addressed the same concepts with similar or synonymous terminology, such as information gathering and problem definition under the broader umbrella of design scoping. After revising our coding protocol, we divided the 27 articles among us and conducted a second pass of coding to catch anything missed in the initial coding. A subset of articles was double-coded to strengthen reliability, and articles that did not readily fit into the categories were reviewed and discussed by the group.

Our discussions typically centered around whether an article described an assessment as a practical application of an assessment in real-world contexts or was more oriented as a research instrument in a controlled environment. We also discussed which categories to combine and which to keep separate (e.g., convergent thinking and performance where the latter assesses a final design rather than decisions made to arrive at the final design).

We initially attempted to group codes under the umbrella categories of design product and design process but found there was often too much overlap between the two to cleanly distinguish, with many assessments having subcategories that overlapped across multiple codes (e.g., convergent thinking as a process and in terms of the nature of the final product). This became particularly challenging for assessments of intermediary design artifacts or for assessments of communication and collaboration. Rather than artificially forcing product or process categories, we opted to retain categories we could apply more reliably. To apply a code, an article had to explicitly describe an assessment for a given category. For example, some papers [e.g., (39)] described divergent thinking as part of the design process but only assessed the final design.

3. Results

Within engineering education, design assessments are especially emphasized in first-year cornerstone and final-year capstone projects in undergraduate and graduate courses [e.g., (40)]. Cornerstone projects give students a taste of engineering design, typically with simplified design challenges, and focus on assessments of the general design process [21]. When working on capstone projects, students typically use content and concepts from prior courses to find solutions to design problems. Capstone projects are typically aligned with formal professional accreditation requirements, such as those outlined by the Accreditation Board of Engineering and Technology (ABET). In turn, these requirements often guide assessment strategies in engineering design courses [40].

In the following sections we analyze these assessments in terms of: (a) problem scoping, (b) idea generation / divergent thinking, (c) design creativity, (d) design performance, (e) decision making / convergent thinking, (f) collaboration, and (g) communication. These categories include both process and product where applicable, with many assessments corresponding to multiple categories.

Table 4. Coding Results Overall and by Article with Counts Representing Number of Articles Including that Explicit Focal Outcome, Target Group, Evaluator Group, and Assessment Approach

	Totals/AVG	Acar (2004)	Azmy & Mokhtar (2017)	Bar-Eli (2013)	Bartholomew et al. (2018)	Berry & Carlson (2010)	Charyton & Merrill (2009)	Charyton et al. (2011)	Chiaradia, et al. (2017)	Dancz et al. (2017)	Davis et al. (2002)	Davis et al. (2010)	Demirkan & Afacan (2012)	Guzzomi et al. (2015)	Jaeger & Adair (2015)	Lans & Verkroost (2004)	Mccormack et al. (2015)	Oehlers (2006)	Ozaltin et al.(2015)	Reid & Cooney (2008)	Ringwood et al. (2005)	Sherrett et al. (2013)	Sluis-Thiescheffer et al. (2016)	Steiner et al. (2011)	Sung-Hee et al. (2015)	Thompson et al. (2013)	Watson et al. (2017)	Welch et al. (2009)
Design Focus																												
Scoping	11		1				1	1			1					1	1		1	1		1		1	1			
Divergent Think	9			1			1	1			1				1	1			1				1		1			
Creativity	9						1	1			1		1					1	1		1		1		1			
Convergent Think	8		1				1	1			1					1	1					1			1			
Performance	20		1	1	1		1	1	1	1	1			1	1	1	1	1	1	1	1	1		1	1		1	
Collaboration	8										1	1			1		1			1				1	1			1
Communication	15	1	1			1				1	1	1			1	1	1		1	1				1	1	1		1
Total Design Foci	*2.96*	*1*	*4*	*2*	*1*	*1*	*5*	*5*	*1*	*2*	*7*	*2*	*1*	*1*	*4*	*5*	*5*	*2*	*5*	*4*	*2*	*3*	*2*	*4*	*7*	*1*	*1*	*2*
Students																												
K12	1																						1					
Undergraduate	25	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1
Graduate	1								1																			
Evaluators																												
Instructor/TA	25	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		1
Peer	12	1	1			1			1			1		1	1		1			1	1			1	1			
Self	9	1				1					1	1			1		1			1				1	1			
External	9	1			1		1	1	1				1				1		1								1	
Total Evaluators	*2.04*	*4*	*2*	*1*	*2*	*2*	*2*	*2*	*3*	*1*	*2*	*3*	*2*	*2*	*3*	*1*	*4*	*1*	*2*	*3*	*2*	*1*	*1*	*3*	*3*	*1*	*1*	*1*
Assessment Approach																												
Formative	12	1			1	1			1		1	1		1			1		1	1				1	1			
Summative	25	1	1	1	1	1	1	1		1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	1	1	1
Detailed Criteria	20		1		1		1	1	1	1	1	1			1*	1	1	1		1	1	1		1*	1	1	1	1

Table 4 presents the list of the 27 articles matching the search criteria with descriptive coding from the analysis.

3.1 Design Scoping

Before engaging in idea generation, students are typically expected to define the scope and parameters of the design or problem to be solved. Davis et al. [41] referred to this as information gathering and problem definition, where students collect and synthesize relevant information about the problem at hand, such as constraints, client requirements, technical or regulatory requirements, and design goals. Information gathering and problem definition were common to multiple articles as the first step in the design process to identify known and unknown variables that could impact the design [42, 43]. Similarly, problem finding has been described as "the ability to identify problems or be able to foresee potential problems that may occur, but have not occurred yet" [44, p. 785]. Sung-Hee et al. [45] explained that "students look for underlying needs by observations, interviews, or survey methods and search for sufficient information on the initial problem, such as patents, products, or professional knowledge to analyze reasons and identify design requirements and constraints" [45, p. 1008]. We found assessment of design scoping in 11 of the articles. Please refer to Table 4 for an overview of the occurrence of design scoping in our sample of 27 articles, as well as all other assessment criteria.

3.2 Divergent Thinking and Creativity

Divergent thinking and creativity were most closely related of the coding categories, which led us to discussing them together within the same section. Divergent thinking and idea generation are a key part of the design process in coming up with multiple solutions to address a problem and selecting among those ideas [41, 45]. Creativity is often encouraged, but the emphasis is on a feasible and effective solution. For example, Davis et al. and Sung-Hee et al. [41, 45] assessed idea generation in terms of students' ability to stimulate and support creativity, use varied methods to generate ideas, and the quantity of ideas. Sung-Hee et al. [45] considered not only the number of ideas generated but also the team dynamics during the generation process. Students were assessed on whether they "actively suggested their own creative ideas and motivated and stimulated other members' opinions" [45, p. 1011]. Similarly, Jaeger and Adair [46] assessed how ideas were developed and required students to show critical reflection and support from external sources to demonstrate idea development using peer and self-assessments. In Bar-Eli's [47] framework, design students' sketches were evaluated according to characteristics such as detail or scale, with the intent to identify certain behavioral patterns. These patterns were used to determine sketching profiles. Each of the three distinct profiles represent specific actions and approaches to the divergent thinking phase during design [47]. Articles assessing divergent thinking tended to take a pragmatic approach to design prioritizing function over form.

Intertwined with the idea of divergent thinking is creativity. We found that assessments of divergent thinking generally focused on the number of novel or unique design ideas [46, p. 643], whereas creativity assessments incorporated additional aspects or considerations. For instance, creativity has been assessed on metrics of originality/novelty (terms often used interchangeably; [e.g., (48)]), the ability to solve problems (usefulness/performance), affective characteristics, or by combining all of the above. The conceptual overlaps and blending of categories for creativity assessment in the literature prompted our coding to incorporate foci on novelty, originality, and innovativeness into an overarching creativity code.

Charyton and Merrill [49], for example, positioned creativity as a pillar of engineering design. They developed the Creative Engineering Design Assessment (CEDA) as an overarching assessment tool that includes diverse elements. CEDA combines measures of divergent thinking, convergent thinking, constraint satisfaction, problem finding, and problem solving to assess the overall creative process. In the revised CEDA [44], students were assessed for fluency (number of ideas), flexibility (number of types of ideas), originality/novelty, and usefulness. Fluency and flexibility were scored by summing the total number of designs, descriptions provided, materials used, problems solved, and end-users identified for each component. Originality was scored on an 11-point scale (0 = dull, 1 = commonplace, to 10 = genius). While usefulness would more typically be considered in terms of convergent thinking and design performance (see sections 3.3 and 3.4, CEDA includes it with a scoring a 5-point scale (0 = not useful to 4 = indispensable). These subscores were weighted and combined into a total CEDA score.

Demirkan and Afacan [50] presented another example with their 3-fold model for assessing creativity. They implemented their model in the context of an undergraduate interior architecture and environmental design course, wherein the categories assessed an artifact's creativity, specific design elements, as well as the assembly of those design elements. Demirkan and Afacan [50] outlined measurement items for each dimension including lists of illustrative adjectives for *artifact creativity* ("integrated, coherent, detailed, refined deliberate, polished, balanced, significant, adequate, sensible, different, unconventional, infrequent, extraordinary, exciting, zippy, fresh, eccentric, new, novel, unusual, unique, original, pleasant, good, delighted, appealed", p. 265), descriptive nouns defining *design elements* ("shape, colour, size, proportion, number, geometric relations, figure-ground relation" p. 265) and the *assembly of design elements* ("harmony, rhythm, unity, variety, repetition, balance, order", p. 265). After expert ratings for all 41 items were collected, statistical analyses validated the model. Since some items were inherently subjective, a creative design was the combination of elements rather than any one element on its own. The use of synonymic adjectives is notable, as well as the lack of antonymic adjectives describing poor or lacking originality.

Some frameworks included creativity assessments in a less central role. For example, Ringwood et al. [51] incorporated creativity, but treated it as more of an aspiration than a requirement to reduce copying of ideas by student teams. Others included ambitious mathematical models to determine the overall innovativeness of design [e.g., (40, 49)]. Ozaltin et al. [40], for example, described economist Joseph Schumpeter's innovation theory as a basis for their assessment framework. They expanded the notion of originality (or novelty) by including the frequency of specific design products recorded in pre-defined conceptual design categories. For instance, if several engineering students designed products of the same pre-defined category, the occurrence value for the category would be high. Threshold values for high, medium, and low occurrences were predetermined by considering the total number of students in the class or the total number of conceptual categories [40]. High numbers of design products in one category meant that many engineering students made use of designs or aspects of designs in that category. By dividing the innovation scores of design products in one category by the usage levels and expressing it in terms of probability, a measure of utilization for each category was calculated [40]. The resulting probability was conditional, thus enabling learners to further explore the innovativeness with regards to the usage: "given that a product is innovative, which usage level for conceptual design (i.e., low, medium, or high) has the highest probability of having occurred and should therefore be emulated?" [40, p. 11]. The utilization score could provide members of an engineering class or a design team with more balanced feedback on a design's originality because innovation could be described in terms of usage or frequency and vice versa.

Similarly, Sluis-Thiescheffer et al. [52] assessed novelty by measuring the frequency of a design, where less frequent designs were considered novel (we consider "novel" to be relatively synonymous with "creative" in this context). Before conducting the assessment, similar or identical designs and/or features would be grouped together and ranked by their frequency [52]. The authors set an (arbitrary) .75 percentile cut-off to distinguish between the frequency of novel and non-novel designs, noting, it was logical "to choose a value that disregards at least 50% of the most frequently generated solutions" [52, p. 62]. In other words, the 25% least frequent designs were considered novel. Sluis-Thiescheffer et al. [52] explained that setting the 'expectation threshold' marked an important part of their assessment procedure, which had to be negotiated for each novelty-assessment. Moreover, the approach could be used for different levels of granularity, for instance by focusing on the overall design, or on specific design components such as the wheels of a car versus the whole car. This allowed the same group of design artifacts to be evaluated repeatedly, whereby different aspects of the design were assessed each time [52].

Overall, of all our codes divergent thinking and creativity were the most intertwined. We identified divergent thinking in 9 of 27 articles and creativity in another 9 articles (please refer to Table 4). Twelve of the articles assessed at least one of the two. Six measured both [40, 41, 44, 45, 49, 52], three measured only creativity [48, 50, 51], and three measured divergent thinking [46, 47, 53]. Of the six that measured both, two of the assessment frameworks evaluated divergent thinking and creativity using separate scores without combining the two into a mutual result, [i.e., (41, 45)], and we identified four papers that merged divergent thinking and creativity into a combined assessment framework, [i.e., (40, 44, 49, 52)].

3.3 Design Performance and Functionality

A design is typically assessed on whether it meets requirements, solves the problem posed, and/or how well it performs [44]. Unsurprisingly, design performance was the most common assessment category of the reviewed literature (20 of the 27 articles, please refer to Table 4 for an overview). We found variations in terminology to assess performance such as usefulness, practicality, or appropriate choice and use of resources, but the foci were similar overall [43, 49, 54]. Some of the characteristics assessed include:

- Evaluating design sketches in terms of their practical applicability, for instance by identifying features such as *detail* or *scale* [47].
- Expert ratings on which design product is *holistically better* [54].
- Measuring the *usefulness* (i.e., practicality for functionality) [44].
- Determining the *quality of the work* [55].
- Identifying whether a design includes elements that situate it above or below a *base line* (or a design according to the book) [48].
- Focusing on aspects such as technical elements, or overall complexity of the project [43].
- Measuring the *sustainability* of the design product [56].

In the following, we describe selected frameworks of the design performance category in more detail to exemplify the range of assessments in this section. Most articles included rubrics or rating scales to evaluate design performance. Watson et al.'s [56] sustainability design rubric, for instance, was developed for a civil and environmental engineering capstone design course. Through expert consultation, 16 sustainability criteria were established, where each was rated on a four-point scale. To achieve exemplary scores, students had to develop designs with characteristics such as "Minimizes natural resource depletion," "Uses renewable energy sources," "Protects human health and well-being," or "Incorporates environmental impact assessment tools" [56, p. 6].

Some articles relied on approaches other than rubrics to determine design performance (also see section 3.10). Batholomew, Strimel, and Jackson's [54] Adaptive Comparative Judgement (ACJ), for instance, stands out as a unique assessment strategy. ACJ relies on expert judges who compare design products against each other *holistically* using their professional opinion: "The judges are not asked to provide a grade for each piece of work but rather asked to make a holistic decision as to which artifact is better" [54, p. 31]. As judgements on designs are made, each artifact attains a "win-loss" result. Eventually, designs are ranked according to their scores. Batholomew, Strimel, and Jackson [54] highlighted the reliability and validity of ADJ resembles traditional assessment methods because a panel of independent judges will generate a more balanced overall review than an individual rater.

In terms of self-evaluation of functionality, Chiaradia et al.'s [57] self-assessment for urban planning concentrated on property appraisals in the design process, thus assessing the functionality of the design. Master's students were given the task to re-plan an inner-city area and had to include a self-evaluation of their design products with a "scorecard." By using the scorecard, students assessed their property development plan in terms of retail, office, and housing value, wherein each component was aggregated into a gross development value appraisal [57]. This pedagogical approach aimed at bridging students' design ideas with considerations and/or limitations of development appraisals so that students learned to critique their own propositions and modify their designs accordingly.

Overall, we found that 20 articles included a focus on the performance of design, making it the most frequently assessed aspect of design (see Table 4).

3.4 Convergent Thinking

As a stepping stone toward finalizing a design, designers must often strategically select one solution from multiple options [44, 45, 49]. Convergent thinking is part of the evaluation and decision-making process to determine how well a proposed solution meets design requirements and any given constraints, such as materials, cost, or time. Curiously, few articles explicitly assessed the decision-making process; instead, most articles focused on how well the final design solved the problem. Perhaps how a design is selected is implicitly baked into the assessments of the final design, such as through students' presentations of their designs or written documentation with design rationale, but these assessments generally happen after-the-fact, when the design is already finished.

Sung-Hee et al. [45] was one of the few articles that explicitly assessed convergent thinking, which they termed "optimal solution selection." In solution selection, to achieve the highest score of 'excellent,' students had to evaluate ideas in consultation with outside experts using metrics predetermined by the design team. The selected solution also had to be considered 'innovative' and 'indispensable.' Conversely, only relying on team discussions for evaluating an idea was scored as marginal.

Davis et al. [41] similarly outlined specific assessment criteria for evaluating the process of convergent decision making. For a maximum score in evaluating ideas, students had to consider "technical, financial, system, life-cycle, [and] failure" [41, p. 216] aspects in their analysis, as well as, use the "best methods" for doing the analysis. As with Sung-Hee et al. [45], students were expected to seek outside consultation in evaluating ideas. In decision making, the full design team had to participate in reviewing, refining, and weighting design solutions [41]. Assessing convergent thinking was a way of demonstrating due diligence in the design process to justify what decisions informed the final design.

Whereas we found that 20 articles focused on the performance of design, only 8 out of the 27 articles focused on the convergent thinking process involved in selecting that design (please refer to Table 4).

3.5 Collaboration

Several articles integrated teamwork/collaboration as part of the overall design process assessment, which are undoubtedly typical expectations of real-world design work. Collaboration focused on interactions between people and how this shaped the design process. The educational outcome for teamwork in engineering design was described in Davis et al. [41] as "organizing, performing, and refining member actions that capitalize on capabilities and resources of all team members to achieve collective goals" [41, p. 213]. Teamwork was assessed in terms of a team's identity and purpose, roles and responsibilities, attitude and climate, resource management, operating procedures, and rewards for achievement. Davis et al. [58] extended this in later work to address team member citizenship by focusing on attributes such as member contributions, effectiveness, member strengths, and areas for improvement.

Reid and Cooney [43] developed a teaming rubric that included three levels (excellent, average, poor) and six criteria (contributions, division of labour, communication, professional conduct, group discipline, group dynamics). This was one of several rubrics intended to assess the non-technical aspects of engineering design and aligned with accreditation criteria outlined by ABET. Instructors used the rubric to evaluate teams, while students used it to evaluate their team members. Students also had access to summaries of the results.

Steiner et al. [59] took a subjective approach to assessing team effectiveness, relying on the opinion of a project mentor and evaluator regarding the team's interactions during class and the quality of posts on a project forum.

Overall, collaboration was assessed in 8 articles (see Table 4). Collaboration was generally interwoven as part of the broader design process. The collaboration assessments tended to focus on interactions between people and how this shaped the design process.

3.6 Communication

Assessment of communication was interwoven as part of the broader design process and focused on how ideas, designs, and technical information were communicated to other people. For instance, Berry and Carlson [60] concentrated on the practice and development of writing skills as part of engineering design. Students completed a series of writing tasks aligned with ABET criteria for effective communication. The assessments were done using a calibrated peer review tool, where students were first trained on using a rubric to assess writing samples. Second, students did peer reviews on three anonymous essays. Third, students self-assessed using the same rubrics to prompt reflection. The writing tasks were structured as scaffolds to help students develop senior design project proposals. Students had to develop competency in writing about both the technical aspects of design as well as their ethical stance and the social impact of their projects.

Thompson et al. [61] developed a rating scale to assess technical posters produced by first-year engineering design students in South Korea. The posters were assessed primarily in terms of how information was communicated rather than the technical correctness of students' designs. The posters had to clearly communicated the design problem, the design process, and the final design while also being persuasive and convincing. The use of visuals, space, text, and writing mechanics were also assessed.

Welch et al. [62] took a comprehensive approach in assessing oral presentations of students in capstone engineering design courses. In their final presentations, students were assessed individually and as a group in terms of their participation in the presentation, audience engagement, technical design, responses to questions, and being able to communicate a cohesive message.

Overall, communication was assessed in 15 articles, as outlined in Table 4. This was the second most frequently common focus, after design performance.

3.7 Education Level Targeted by the Assessments

We coded for the education level targeted by the assessment to provide insight into the educational degree participants were pursuing when the assessment data was collected. We coded the education level of participants including elementary school, middle school, high school, undergraduate, and graduate level. In cases where only general descriptors were given (e.g., design students), we categorized these papers as being undergraduate level.

Overall, our review indicates that the majority of assessment frameworks are implemented at the undergraduate level (25 out of 27, please refer to Table 4), and only very few articles focus on a different student demographic. We identified one article that included participants at the graduate level [57] and one other article that assessed students in the elementary grade level [52].

3.8 Evaluator Categories

Most of the articles provided demographic information on the evaluators, i.e., the individuals who assessed design processes and products. Please see Table 4 in terms of which articles employed which forms of assessment. We differentiate between the following evaluator categories:

Instructor. Unsurprisingly, instructors/teaching assistants were involved in assessment in nearly all (25) of the articles. The course instructor uses the proposed assessment framework to evaluate student progress or design results, such as in engineering capstone projects. Other terms to refer to instructors include: member of an academic facilitator team [63], examiner [64], teacher [53], team instructor [59], and faculty member [61]. We also consider teaching assistants to be part of the instructor category.

Peer. Peer assessment was involved in 12 of the articles. Students in the assessments engage in peer-assessment during or after a design activity. They may use a specific assessment framework, such as calibrated peer review, where students are trained on how to do peer assessments [60], or they may be handed an additional rubric/evaluative sheet that focuses on team assessment and team members' contributions [43, 58].

Self-assessment. Self-assessment was involved in nine of the articles. Design students evaluate their own work by submitting an evaluative report [63] or using scorecards [57]. We also assigned individuals' self-assessment of group work to this category, for example, when group members reflect on their own design contributions [58].

External. External assessment was involved in nine of the articles. In this category, one or more external reviewers (outside experts) are introduced to the evaluative procedure. This may include guest reviewers from industry [63], faculty members from departments other than the one where the assessment is implemented [49], or unspecified individuals such as field experts [50], "trained judges" [56, p. 10], or "independent judges" [54, p. 23].

In summary, many of the assessments relied on assessment by multiple categories of evaluators. Two articles relied on all four categories [42, 63], six articles relied on three categories [43, 45, 46, 57–59], ten articles relied on two categories, and eleven articles relied on only one category. Our review indicates that most articles relied on instructor assessment (25 out of 27). Self-assessment was used nine times, and we were able to identify peer review in twelve cases. External evaluators were consulted in nine articles.

3.9 Nature of Design Task

Most of the articles described the intended design tasks which students were to carry out and accomplish. In the following, we illustrate the range and diverse nature of design tasks assigned, and briefly outline how assessment frameworks accommodated these different outputs. Some of these tasks were part of cornerstone or capstone design projects [e.g., (56)]. We noticed a wide range in design tasks' scope, difficulty level, and detail; ranging from tangible, small-scale artifacts such as "a mobile robot to collect three drink cans in a square area" [51, p. 98] to abstract refined conceptions of urban housing and retail development.

Our synthesis shows that fully functional prototypes were the objective of design activities in fewer articles, whereas the majority of frameworks relied on unfinished and/or conceptual design documents. In cases where participants were asked to design one or more functional prototypes, a given real-world problem would usually be addressed, for instance the design of bridges or roadways in engineering courses [56]. Sometimes, constraints on resources or materials that could be used were also part of the task. Examples of functional prototypes include the design of "a water purification system using low cost, readily available materials" [54, p. 22] or "a truss structure made out of drinking straws and wire" [48, p. 493].

Besides functional prototypes, design briefs often required the conceptualization of designs without creating a finished product. These design tasks called for conceptual representations such as illustrations, sketches, written descriptions, or digital models. Examples include the design of an "interior space to be used as a place for older people, which that surpasses conventional expectations" [47, p. 480], or the creation of sketches and journal posts for a functional roller coaster without building a model [49].

For the functional prototype design tasks, such as the ones mentioned above, the assessment procedure normally required students to hand in their completed artifacts for evaluation. Design tasks that focused on conceptual representations required different kinds of submissions for assessment, including illustrations, sketches, or drawings (digital or on paper), digital representations of design products (models or photos), written descriptions of design products, or unfinished design drafts. In some cases, the exact artifacts for assessment were not fully specified in the articles, as the descriptions were either general [e.g., (62, 63)] or were not included in the article [e.g., (61)].

3.10 Granularity of Assessment

To understand how the assessments were scored, we explored the criteria used for determining high and low performance. Most assessments used scaled assessments that were often operationalized as a rubric. The number of assessment items and sophistication of criteria varied from a few items with condensed scales to dozens of items with detailed descriptions on 7-point scales. For example, Davis et al. [41] used separate rubrics for design process, teamwork, and communication. Each rubric included approximately six subcategories, and each subcategory listed multiple items assessed on a 7-point scale. A low score of 1, for example, included the descriptor: "Climate: Critical, stifling; Methods: Single effort; Types: Solution ideas" whereas a high score of 7 included the descriptor: "Climate: Stimulating, supportive of creativity; Methods: Methods varied, revised, used multiple times; Types: Solution and process ideas." Thus, a fair amount of detail is included to anchor the numerical score to a detailed rubric. Conversely, Charyton and Merrill [49] assessed four main categories using a 5- to 10-point scale with simple descriptors, such as a low score of 1 for the originality category having the simple descriptor "dull" while a 10 included the descriptor "genius." Most assessments tended towards the former with several assessment items and brief descriptions to distinguish low to high performance. Lans and Verkroost [53] took a more mixed approach, providing substantial detail for a given category (e.g., "Identifying, designing and screening the alternatives: Theoretical insights and/or precedents are used in the design. Different alternatives have been identified. The alternatives represent the space of possible solutions", p. 279) but employing a scale system ranging simply from "– –" to "+ +".

Overall, similar assessment items were used in most articles, with slight variations in the purpose of a given rubric. For example, Davis et al. [41] described the rubric as a wholistic instrument to map the expected performance of an engineer at various stages of their education and professional practice. They noted a score of three out of seven was the target for engineering students after two years in their degree program, while a score of seven represented a practicing, professional engineer.

Most other assessment scales were bound to specific courses or years of an engineering or design program. For engineering design, the assessment items were often linked to learning outcomes and program objectives, which were linked in turn to accreditation criteria given by professional engineering bodies (e.g., ABET). Azmy and Mokhtar [64] likewise assessed technical aspects of design but also described the evaluation process as a way to promote quality teaching and guide student learning. Bartholomew et al. [54] took a hybrid approach by using a traditional, scaled rubric as a basis for judges to give holistic assessments of student design projects. Rather than giving a final score, design projects were comparatively ranked.

In addition to the scaled rubric approaches, which represented the majority of the assessments, some assessments pursued alternative mechanics. Alternative approaches focused on similar design criteria to scaled rubric assessments, and were sometimes used in tandem with them, but took different approaches in doing so, such as the win-loss comparisons used by Bartholomew et al. [54]. Motivations for alternative approaches included increased efficiency in grading, greater depth of learning, greater engagement of students, and greater authenticity with professional practice.

Sherrett, Nefcy, Gummer, and Koretsky [65] proposed another alternative approach, assessing information gathering, formulation of the problem, iterative modeling, and experimentation using model-maps as a visual way to track design process and assess decision making competency. Sluis-Thiescheffer et al. [52] represent yet another category of alternative approaches based on calculating rarity across an overall group. They used the novelty metric proposed by Shah et al. [15] to calculate the frequency of an approach in the solutions generated across the entire group, where infrequent solutions received novelty scores calculated based on the relative rarity.

These non-scaled approaches show promise and power, and they provide interesting foundations on which to build in the future, but ultimately most of the assessments focused on scaled rubrics.

4. Discussion

4.1 What Focal Aspects of Design are Assessed?

In our grounded coding, categories arose around multiple aspects of design, including: (a) problem scoping, (b) divergent thinking, (c) creativity, (d) design performance and functionality, (e) convergent thinking, (f) collaboration, and (g) communication. While these seven categories were generally distinct, the clearest distinctions were between performance, communication, and collaboration, while the distinction between divergent thinking and creativity was the least clear in many cases. This was complicated by terminology that was often used interchangeably in the literature (e.g., originality and novelty).

Most of the assessment approaches focused on multiple areas (2.96 out of 7 categories on average). Among the 27 articles, the majority (20) included a focus on performance assessment, and a substantial number included a focus on communication (15) and scoping (11) (please refer to Table 4 for details). While less prevalent, divergent thinking (9), creativity (9), convergent thinking (8), and collaboration (8) were also broadly represented (Table 4), and if divergent thinking and creativity were merged as a single category then 12 of the articles would include that focus. Notably, most articles assessed the performance or function of a final design outcome. We speculate the logistics of assessing a final design are often more manageable for instructors than assessing intangible or real-time aspects of the design process in the midst of running a class. As well, we speculate that the assessment of design processes is done in implicit ways with a formative intent. Undoubtedly, the real-time processes of design are as important as final projects as argued in prior research [see (27)], but we recognize that the pragmatic realities of facilitating and teaching a class may challenge the bandwidth to conduct simultaneous real-time assessments of process. There is work underway to collect real-time process data digitally to automate assessment of process [66, 67], which may change the nature of assessment of design radically in the future, but these approaches are not yet scalable. These digital tools can help complement, extend, or add depth to what is currently feasible.

Another major influence on assessment foci are professional engineering organizations, particularly the Accreditation Board for Engineering and Technology (ABET) based in the United States but also similar organizations internationally (e.g., Engineers Canada, United Kingdom Engineering Council, Japan Accreditation Board for Engineering Education). ABET, for example, is a non-governmental organization that accredits post-secondary education programs in applied and natural science, computing, engineering, and engineering technology. According to the ABET website (abet.org/about-abet/), 4,005 programs are accredited, distributed over 793 universities and colleges in 32 countries. Several of the assessments are conceived in explicit relationship extensively articulated to ABET broadly, [e.g., (43, 60)], or to specific components of ABET, [e.g., (62)], in terms of communication. The ABET 2019–20 Criteria for accrediting Engineering programs at the baccalaureate level in terms of student outcomes focus on:

(1) an ability to identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics;
(2) an ability to apply engineering design to produce solutions that meet specified needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors;
(3) an ability to communicate effectively with a range of audiences;
(4) an ability to recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts;
(5) an ability to function effectively on a team whose members together provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives;
(6) an ability to develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions;
(7) an ability to acquire and apply new knowledge as needed, using appropriate learning strategies.

In addition to the specified student outcomes, ABET also defines Engineering Design:

> "Engineering design is a process of devising a system, component, or process to meet desired needs and specifications within constraints. It is an iterative, creative, decision-making process in which the basic sciences, mathematics, and engineering sciences are applied to convert resources into solutions. Engineering design involves identifying opportunities, developing requirements, performing analysis and synthesis, generating multiple solutions, evaluating solutions against requirements, considering risks, and making trade-offs, for the purpose of obtaining a high-quality solution under the given circumstances. For illustrative purposes only, examples of possible constraints include accessibility, aesthetics, codes, constructability, cost, ergonomics, extensibility, functionality, interoperability, legal considerations, maintainability, manufacturability, marketability, policy, regulations, schedule, standards, sustainability, or usability." [68]

While the ABET specified student outcomes and their definition of Engineering Design continue to evolve [69], we see a strong alignment with the aspects of design processes and outcomes highlighted by our coding of the assessments in terms of project scoping (e.g., identifying opportunities and developing requirements), creativity (e.g., creative decision-making processes), design performance and functionality (e.g., meeting needs and specifications), ideation (e.g., generating multiple solutions), convergent thinking (e.g., evaluating solutions against requirements). Collaboration (e.g., ability to function effectively on a team) and communication (e.g., ability to communicate effectively with a range of audiences) are key ideas represented from the broader outcomes.

Aspects of ABET that are less represented in the reviewed assessments include conducting experiments, applying science and mathematics, and information literacy/life-long learning, except when integrated into a larger category of solution performance. Engineering ethics and ethical considerations seem to be the least represented, although there are exceptions. Berry [60], for instance, focused heavily in their approach to Calibrated Peer Review on having the students write a social impact document using the IEEE Code of Ethics as the rubric. Other assessments may include some aspect of attending to ethical issues as part of the task but not as a major aspect of the assessment rubrics themselves. While beyond the scope of this review, we find such omissions troubling given the very real impacts design can have on people's lives. For example, designers can face heavy ideological and ethical dilemmas when engineering products that are used in warfare [70] or to marginalize specific populations [71]. This seems to be an area for future growth in the assessments. While not an assessment of students' actual designs, the approach outlined by Christensen, Hjorth, Iversen, and Blikstein [72] assessed students' stance toward inquiry in a way that attends to the complexity inherent in wicked problems in terms of social and ethical dimensions among others.

One final consideration is that many of the assessments synthesized here often prioritize the role of the designer(s) for the accomplishment of the design processes and project goals with less emphasis being placed on engaging actual stakeholders or actual end-users in more participatory design processes [73]. An exception to this trend involves programs focused on industry-sponsored design projects, [e.g., (74)], where students engage with more real-world audiences. Consequently, skills of interaction and true stakeholder participation are not typically represented in design assessment tasks beyond high-level design scoping or constraint compliance. This could be partly a pragmatic result of engaging large numbers of students in coursework simultaneously, but also signals a divide between design processes and broader stakeholder and social ramifications. To advance more realistic and participatory design scenarios, efforts to involve authentic audiences and consider broader societal impacts of design activities should be strengthened.

4.2 How is Assessment Conceptualized and Operationalized

As outlined in Table 4, while some articles addressed both summative and formative evaluation (9 of 27), a larger percentage of articles overall described summative (25 of 27) rather than formative assessment procedures (12 of 27). We surmise this stems from a combination of formative assessments often being less formal (e.g., undocumented verbal commentary from an instructor circulating among students), traditional preferences to summative assessments in STEM education, and an interest in supporting accreditation processes (e.g., ABET).

For focal student level, we anticipated undergraduate education would be the predominant focus of the assessments but were surprised by the limited focus at the K-12 level given the increasing interest and emphasis in K-12 education on design, design thinking, and engineering. There are many excellent studies of K-12 student design that are not focused on "assessment" or where design-based learning uses design as a pedagogy and hence their assessment focuses on student learning of science concepts [e.g., (75–77)]. While there is an increasing emphasis on design in K-12 STEM domains, explicit systematic development of approaches to assessment is less prevalent in journals.

In terms of evaluators, multiple types of assessor were often used (2.04 on average). Instructors and TAs were involved in almost all of the assessments (25 of 27), but most (18) of the assessments involved other people in the assessments, including peer assessment (12), self-assessment (9), and external disciplinary expert scoring (9) (please see Table 4). Essentially, most scoring was done by instructors and TAs but self, peer, and external scoring was more prevalent than anticipated.

Primarily, the assessments depend on the application of rubrics. Some articles do not provide enough information about their rubrics to be reliably applied by others, but many of the articles engage in a high degree of granularity with highly descriptive rubrics. These efforts toward systematicity are laudable and desirable, but it is important that the field be realistic about the degree of objectivity that can be achieved in a process that involves substantial subjectivity. Some of the assessments leverage complex approaches and mathematics within the assessment in an effort to create a final numerical score with great precision, but this apparent precision is necessarily limited because the input numbers are based at times on subjective or underspecified criteria. Another group of the assessments are more subjective in their orientation from beginning to end. Sometimes this results in advantages in terms of not artificially systematizing the process, but greater specification within the levels of the rubrics themselves would likely be the most powerful approach to improving the actual rigor and reliability of the assessment processes across the assessments we reviewed.

Related to this, we noticed a considerable range in the granularity of assessments. For instance, Charyton and Merrill [49] used single-words to describe originality criteria, while Davis et al. [41] used detailed criteria and sub-criteria in their assessment. We find benefits and drawbacks to both approaches. The simplified or open-ended descriptors afford flexibility and space for amorphic aspects of design such as idea generation and creativity, but they also limit consistency in how to carry out assessment procedures and how students understand what is expected of them. In contrast, detailed criteria reduce ambiguity in what is being assessed and how but also makes design more rigid and systematic, which can run against more freeform or fluid notions of innovation and novelty.

4.3 Limitations of Our Approach

There are several limitations and caveats to our study. First, the key search terms, "assessment" and "design," are common terms in many contexts beyond our scope. This was one of the reasons why we examined specific journals rather than library databases. The common nature of these terms results in many false positive search results and can make inclusion decisions complex. Teasing apart this distinction required multiple readings of the articles and multiple discussions among us to achieve consensus.

Second, the bounds of the inclusion criteria themselves could be envisioned differently. We ultimately decided to set the criteria to include only those assessments that focused on assessing students' design products and processes but not to include more general cognitive inventories that might be relevant to students' abilities to engage in various aspects of design, such as creative ability [e.g., (78)], or other cognitive capacities [e.g., (39)]. We acknowledge inventories of skills and creativity are a rich area of research and assessment, but we chose to limit to assessments of students' actual design processes partly because drawing bounds around some but not all cognitive and creative inventories created too subjective a criterion.

Third, the stringency of the search terms and inclusion criteria resulted in a primary emphasis on engineering journals (23 of 27 articles). There is certainly critical work in K-12 on integrating design into the curriculum, such as Kolodner and Hmelo-Silver's seminal work on Learning by Design [e.g., (76, 77)] and other core work in STEM education [e.g., (75)]. We also recognize our focus on articles that propose assessment approaches, as opposed to studies of students' designs focused on student learning of science or math concepts, limited the appearance of K-12 research in our search results. This is likely a function of the more nascent emphasis on design in K-12 compared to undergraduate engineering education. Similarly, research studies may have been left out where different terms were used instead of assessment [e.g., design review sessions; e.g., (24)]. Another area for increased attention is design in the context of K-12 STEM education, where further assessment is needed, building on the existing ones such as by Alemdar, Lingle, Wind, and Moore [79] and Cardella, Hsu, and Ricco [80].

Fourth, some articles did not fully describe all aspects of their assessment practices or only focused on specific components of a broader assessment scheme, resulting in some codes not being applied. While some authors may have considered or used other assessment practices, we chose to limit our coding to clear evidence as presented in the articles rather than speculate on what authors/researchers might have done.

5. Conclusions

To gauge the assessment of design, we performed a systematized review and content analysis of literature from seventeen prominent journals that address STEM and design education. We selected journals based on Journal Citation Reports (JCR), Scientific Journal Rankings (SJR), and h-index statistics. Inclusion criteria and search terms were identified through a systematized process. The literature search resulted in a sample of 27 articles. Most of the articles (23 of 27) were in engineering education journals and four were in a design journal. We performed a content analysis of the final 27 articles with a goal of identifying key components of assessment in terms of design foci, student age, evaluator type, project type, and granularity of assessment. The majority (20) included a focus on performance of the design, and a substantial number included a focus on communication (15) and scoping (11). While less prevalent, divergent thinking (9), creativity (9), convergent thinking (8), and collaboration (8) were also broadly represented. Most of the assessment approaches focused on multiple of these areas (2.96 out of 7 categories on average).

Overall, there is a strong alignment of the assessments we have identified with ABET's definition of Engineering Design and associated student outcome criteria, particularly in terms of design performance and requirement satisfaction. There is also substantial emphasis on communication. Emphasis is also evident on problem scoping, ideation, creativity, convergent thinking, and collaboration. There are also opportunities for enriching the assessment practices in design. Aspects of ABET that are less represented in the reviewed assessments include conducting experiments, applying science and mathematics, and information literacy/life-long learning, except when integrated into a larger category of solution performance. Furthermore, many of the assessments prioritize the role of the designer(s) for the accomplishment of the design processes and project goals with less emphasis being placed on engaging actual stakeholders or actual end-users in more participatory design processes. Finally, engineering ethics and ethical considerations seem to be the least represented, although there are exceptions. While beyond the scope of this review, we find such omissions troubling given the very real impacts design can have on people's lives.

What is assessed greatly influences what is taught and learned. Ongoing research and development on approaches to design assessment will therefore be critical for design education. As opposed to the traditional heavier focus on convergent aspects of constraint satisfaction and performance, this research and development should increase attention to the divergent aspects of the design process and the social and ethical ramifications of design.

References

* = *Reviewed Article (All 27 cited in the text)*

1. G. J. J. Biesta, Good education in an age of measurement: On the need to reconnect with the question of purpose in education, *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, **21**(1), pp. 33–46, 2009.
2. G. J. J. Biesta, *Good Education in an Age of Measurement: Ethics, Politics, Democracy*, Paradigm Publishers, Boulder, CO, 2010.
3. N. J. Entwistle and E. R. Peterson, Conceptions of learning and knowledge in higher education: Relationships with study behaviour and influences of learning environments, *International Journal of Educational Research*, **41**(6), pp. 407–428, 2004.
4. H. Jürges, K. Schneider, M. Senkbeil and C. H. Carstensen, Assessment drives learning: The effect of central exit exams on curricular knowledge and mathematical literacy, *Economics of Education Review*, **31**(1), pp. 56–65, 2012.
5. T. Hattingh, L. Dison and L. Woollacott, Student learning behaviours around assessments, *Australasian Journal of Engineering Education*, **24**(1), pp. 14–24, 2019.

6. S. Purzer, N. Fila and K. Nataraja, Evaluation of current assessment methods in engineering entrepreneurship education, *Advances in Engineering Education*, **5**(1), pp. 1–27, 2016.
7. D. P. Crismond and R. S. Adams, The informed design teaching and learning matrix, *Journal of Engineering Education*, **101**(4), pp. 738–797, 2012.
8. N. Cross, Designerly ways of knowing: Design discipline versus design science, *Design Issues*, **17**(3), pp. 49–55, 2001.
9. E. Dringenberg and Ş. Purzer, Experiences of first-year engineering students working on ill-structured problems in teams, *Journal of Engineering Education*, **107**(3), pp. 442–467, 2018.
10. G. Goldschmidt, Linkographic evidence for concurrent divergent and convergent thinking in creative design, *Creativity Research Journal*, **28**(2), pp. 115–122, 2016.
11. A. D. d. Figueiredo, Toward an epistemology of engineering (November 1, 2008), Workshop on Philosophy and Engineering, The Royal Academy of Engineering, London, 2008.
12. D. Radcliffe, *A tale of two STEMs*, PRISM, Available: http://www.asee-prism.org/last-word-dec-3/, Accessed: January 31, 2020.
13. S. Purzer and J. P. Quintana-Cifuentes, Integrating engineering in K-12 science education: Spelling out the pedagogical, epistemological, and methodological arguments, *Disciplinary and Interdisciplinary Science Education Research*, **1**(1), p. 13, 2019.
14. P. Sarkar and A. Chakrabarti, Assessing design creativity, *Design Studies*, **32**(4), pp. 348–383, 2011.
15. J. J. Shah, S. M. Smith and N. Vargas-Hernandez, Metrics for measuring ideation effectiveness, *Design Studies*, **24**(2), pp. 111–134, 2003.
16. M. E. Cardella, W. C. Oakes, C. B. Zoltowski, R. Adams, Ş. Purzer, J. Borgford-Parnell, R. Bailey and D. Davis, Special session – Assessing student learning of engineering design, in *IEEE Frontiers in Education (FIE) Conference Proceedings*, Rapid City, SD, October 2011, pp. T4A1–T4A3, 2011.
17. National Academy of Engineering, National Research Council, Committee on Assessing Technological Literacy, E. Garmire, and G. Pearson, Tech Tally: Approaches to Assessing Technological Literacy, The National Academies Press, Washington, DC, 2006, Available: https://www.nap.edu/catalog/11691/tech-tally-approaches-to-assessing-technological-literacy.
18. J. W. Pellegrino, M. R. Wilson, J. A. Koenig and A. S. Beatty (Eds.), Developing assessments for the next generation science standards, The National Academies Press, Washington, DC, 2014.
19. NAEP, 2014 Technology & Engineering Literacy (TEL) Overall Results, 2014, Available: https://www.nationsreportcard.gov/tel_2014/#results/overall.
20. M. J. Grant and A. Booth, A typology of reviews: An analysis of 14 review types and associated methodologies, *Health Information & Libraries Journal*, **26**(2), pp. 91–108, 2009.
21. C. L. Dym, A. M. Agogino, O. Eris, D. D. Frey and L. J. Leifer, Engineering design thinking, teaching, and learning, *Journal of Engineering Education*, **94**(1), pp. 103–120, 2005.
22. S. Purzer, T. J. Moore, and E. Dringenberg, Engineering cognition: A process of knowledge acquisition and application, in Y. Dori, Z. Mevarech, and D. Baker (eds), *Cognition, Metacognition, and Culture in STEM Education. Innovations in Science Education and Technology*, Springer, Cham, pp. 167–190, 2018.
23. D. Crismond, L. Gellert, R. Cain and S. Wright, Minding design missteps, *Science and Children*, **51**(2), pp. 80–85, 2013.
24. R. S. Adams and J. A. Siddiqui (eds), *Analyzing Design Review Conversations*, Purdue University Press, West Lafayette, IN, 2016.
25. L. L. Bucciarelli, *Engineering Philosophy*, Delft University Press, Delft, The Netherlands, 2003.
26. N. D. Fila, Ş. Purzer and R. Chakroun, Cultures of innovation among chemical, civil, and mechanical engineering students: A qualitative study, in *IEEE Frontiers in Education (FIE) Conference Proceedings*, Madrid, Spain, October 2014, pp. 1–7, 2014.
27. C. J. Atman, R. S. Adams, M. E. Cardella, J. Turns, S. Mosborg and J. Saleem, Engineering design processes: A comparison of students and expert practitioners, *Journal of Engineering Education*, **96**(4), pp. 359–379, 2007.
28. N. Cross, The expertise of exceptional designers, presented at the Design Thinking Research Symposium (DTRS) 6, Sydney, Australia, pp. 79–103, 2003.
29. D. P. Dannels and K. N. Martin, Critiquing critiques: A genre analysis of feedback across novice to expert design studios, *Journal of Business and Technical Communication*, **22**(2), pp. 135–159, 2008.
30. M. M. Mehalik, Y. Doppelt and C. D. Schunn, Middle-school science through design-based learning versus scripted inquiry: Better overall science concept learning and equity gap reduction, *Journal of Engineering Education*, **97**(1), pp. 71–85, 2008.
31. D. Missingham, S. Shah, F. Sabir and J. Willison, Student engineers optimising problem solving and research skills, *Journal of University Teaching & Learning Practice*, **15**(4), pp. 1–17, 2018.
32. M. Mehalik and C. Schunn, What constitutes good design? A review of empirical studies of design processes, *International Journal of Engineering Education*, **22**(3), pp. 519–532, 2006.
33. Y. Luo, Design fixation and cooperative learning in elementary engineering design project: A case study, *International Electronic Journal of Elementary Education*, **8**(1), pp. 133–146, 2015.
34. A. Looijenga, R. Klapwijk and M. J. de Vries, The effect of iteration on the design performance of primary school children, *International Journal of Technology and Design Education*, **25**(1), pp. 1–23, 2015.
35. M. Borrego, M. J. Foster and J. E. Froyd, Systematic literature reviews in engineering education and other developing interdisciplinary fields, *Journal of Engineering Education*, **103**(1), pp. 45–76, 2014.
36. J. P. T. Higgins and S. Green (eds), *Cochrane Handbook for Systematic Reviews of Interventions, version 5.0.1 (updated September 2008)*, the Cochrane Collaboration, 2008.
37. K. Charmaz, *Constructing Grounded Theory*, Sage, London, 2006.
38. J. Corbin and A. Strauss, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 3rd ed., Sage, Thousand Oaks, CA, 2014.
39. C. L. A. Dancz, K. J. Ketchman, R. D. Burke, T. A. Hottle, K. Parrish, M. M. Bilec and A. E. Landis, Utilizing civil engineering senior design capstone projects to evaluate students' sustainability education across engineering curriculum, *Advances in Engineering Education*, **6**(2), pp. 1–27, 2017.*
40. N. Ozaltin, M. Besterfield-Sacre and R. Clark, An engineering educator's decision support tool for improving innovation in student design projects, *Advances in Engineering Education*, **4**(4), pp. 1–20, 2015.*
41. D. C. Davis, K. L. Gentili, M. S. Trevisan and D. E. Calkins, Engineering design assessment processes and scoring scales for program improvement and accountability, *Journal of Engineering Education*, **91**(2), pp. 211–221, 2002.*

42. J. McCormack, S. Beyerlein, P. Brackin, E. Odom, and D. Cordon, Snapshot style poster sessions for formative inter-team design feedback in capstone courses, *International Journal of Engineering Education*, **31**(6B), pp. 1736–1747, 2015.*
43. K. Reid and E. Cooney, Implementing rubrics as part of an assessment plan, *International Journal of Engineering Education*, **24**(5), pp. 893–900, 2008.*
44. C. Charyton, R. J. Jagacinski, J. A. Merrill, W. Clifton and S. DeDios, Assessing creativity specific to engineering with the revised creative engineering design assessment, *Journal of Engineering Education*, **100**(4), pp. 778–799, 2011.*
45. J. Sung-Hee, K. I. Song, D. Shin and S. Shin, A performance-based evaluation rubric for assessing and enhancing engineering design skills in introductory engineering design courses, *International Journal of Engineering Education*, **31**(4), pp. 1007–1020, 2015.*
46. M. Jaeger and D. Adair, Using an evidential reasoning approach for portfolio assessments in a project-based learning engineering design course, *European Journal of Engineering Education*, **40**(6), pp. 638–652, 2015.*
47. S. Bar-Eli, Sketching profiles: Awareness to individual differences in sketching as a means of enhancing design solution development, *Design Studies*, **34**(4), pp. 472–493, 2013.*
48. D. J. Oehlers, Sequential assessment of engineering design projects at university level, *European Journal of Engineering Education*, **31**(4), pp. 487–495, 2006.*
49. C. Charyton and J. A. Merrill, Assessing general creativity and creative engineering design in first year engineering students, *Journal of Engineering Education*, **98**(2), pp. 145–156, 2009.*
50. H. Demirkan and Y. Afacan, Assessing creativity in design education: Analysis of creativity factors in the first-year design studio, *Design Studies*, **33**(3), pp. 262–278, 2012.*
51. J. V. Ringwood, K. Monaghan and J. Maloco, Teaching engineering design through Lego® Mindstorms™, *European Journal of Engineering Education*, **30**(1), pp. 91–104, 2005.*
52. W. Sluis-Thiescheffer, T. Bekker, B. Eggen, A. Vermeeren and H. De Ridder, Measuring and comparing novelty for design solutions generated by young children through different design methods, *Design Studies*, **43**, pp. 48–73, 2016.*
53. W. Lans and M.-J. Verkroost, Bottom-up: Teachers' contribution in developing uniform criteria to assess design products, *European Journal of Engineering Education*, **29**(2), pp. 275–282, 2004.*
54. S. R. Bartholomew, G. J. Strimel and A. Jackson, A comparison of traditional and adaptive comparative judgment assessment techniques for freshmen engineering design projects, *International Journal of Engineering Education*, **34**(1), pp. 20–33, 2018.*
55. A. L. Guzzomi, S. A. Male and K. Miller, Students' responses to authentic assessment designed to develop commitment to performing at their best, *European Journal of Engineering Education*, **42**(3), pp. 219–240, 2017.*
56. M. K. Watson, E. Barrella, T. A. Wall, C. Noyes and M. Rodgers, A rubric to analyze student abilities to engage in sustainable design, *Advances in Engineering Education*, **6**(1), pp. 1–25, 2017.*
57. A. J. F. Chiaradia, L. Sieh and F. Plimmer, Values in urban design: A design studio teaching approach, *Design Studies*, **49**, pp. 66–100, 2017.*
58. D. Davis, M. Trevisan, R. Gerlick, H. Davis, J. Mccormack, S. Beyerlein, P. Thompson, S. Howe, P. Leiffer and P. Brackin, Assessing team member citizenship in capstone engineering design courses, *International Journal of Engineering Education*, **26**(4), pp. 771–783, 2010.*
59. M. Steiner, J. Kanai, C. Hsu, R. Alben and L. Gerhardt, Holistic assessment of student performance in multidisciplinary engineering capstone design projects, *International Journal of Engineering Education*, **27**(6), pp. 1259–1272, 2011.*
60. F. C. Berry and P. A. Carlson, Assessing engineering design experiences using calibrated peer review, *International Journal of Engineering Education*, **26**(6), pp. 1503–1507, 2010.*
61. M. K. Thompson, L. K. H. Clemmensen and B.-u. Ahn, The effect of rubric rating scale on the evaluation of engineering design projects, *International Journal of Engineering Education*, **29**(6), pp. 1490–1502, 2013.*
62. H. Welch, D. Suri and E. Durant, Rubrics for assessing oral communication in the capstone design experience: Development, application, analysis and refinement, *International Journal of Engineering Education*, **25**(5), pp. 952–961, 2009.*
63. B. S. Acar, Analysis of an assessment method for problem-based learning, *European Journal of Engineering Education*, **29**(2), pp. 231–240, 2004.*
64. A. Azmy and H. Mokhtar, A project-based framework for teaching and assessment of design modules, *International Journal of Engineering Education*, **33**(2), pp. 629–638, 2017.*
65. B. U. Sherrett, E. J. Nefcy, E. S. Gummer and M. D. Koretsky, An expert solution to assess an industrially situated, computer-enabled design project, *Journal of Engineering Education*, **102**(4), pp. 541–576, 2013.*
66. M. Worsley and P. Blikstein, Reasoning strategies in the context of engineering design with everyday materials, *Journal of Pre-College Engineering Education Research (J-PEER)*, **6**(2), pp. 58–74, 2016.
67. C. Xie, Z. Zhang, S. Nourian, A. Pallant and E. Hazzard, A time series analysis method for assessing engineering design processes using a CAD tool, *International Journal of Engineering Education*, **30**(1), pp. 218–230, 2014.
68. ABET, *Criteria for Accrediting Engineering Programs, 2019–2020*, Available: https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2019-2020/, Accessed: November 4, 2019.
69. ABET, *Mapping Changes in Criteria for Accrediting Engineering Programs Between 2018–19 and 2019-20*, Available: https://www.abet.org/wp-content/uploads/2018/03/C3_C5_mapping_SEC_1-13-2018.pdf, Accessed: November 4, 2019.
70. T. M. Philip, A. Gupta, A. Elby and C. Turpen, Why ideology matters for learning: A case of ideological convergence in an engineering ethics classroom discussion on drone warfare, *Journal of the Learning Sciences*, **27**(2), pp. 183–223, 2018.
71. C. Chellew, Defending suburbia: Exploring the use of defensive urban design outside of the city centre, *Canadian Journal of Urban Research*, **28**(1), pp. 19–33, 2019.
72. K. S. Christensen, M. Hjorth, O. S. Iversen and P. Blikstein, Towards a formal assessment of design literacy: Analyzing K-12 students' stance towards inquiry, *Design Studies*, **46**, pp. 125–151, 2016.
73. R. E. Ladner, Design for user empowerment, *Interactions*, **22**(2), pp. 24–29, 2015.
74. M. Keefe, J. Glancey and N. Cloud, Assessing student team performance in industry sponsored design projects, *Journal of Mechanical Design*, **129**(7), pp. 692–700, 2007.
75. D. Crismond, Learning and using science ideas when doing investigate-and-redesign tasks: A study of naive, novice, and expert designers doing constrained and scaffolded design work, *Journal of Research in Science Teaching*, **38**(7), pp. 791–820, 2001.

76. C. E. Hmelo, D. L. Holton and J. L. Kolodner, Designing to learn about complex systems, *Journal of the Learning Sciences*, **9**(3), pp. 247–298, 2000.
77. J. L. Kolodner, P. J. Camp, D. Crismond, B. Fasse, J. Gray, J. Holbrook, S. Puntambekar and M. Ryan, Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design(tm) into practice, *Journal of the Learning Sciences*, **12**(4), pp. 495–547, 2003.
78. S. A. Atwood and J. E. Pretz, Creativity as a factor in persistence and academic achievement of engineering undergraduates, *Journal of Engineering Education*, **105**(4), pp. 540–559, 2016.
79. M. Alemdar, J. A. Lingle, S. A. Wind and R. A. Moore, Developing an engineering design process assessment using think-aloud interviews, *International Journal of Engineering Education*, **33**(1), pp. 441–452, 2017.
80. M. Cardella, M. C. Hsu and G. D. Ricco, Analysis of design process knowledge task responses: Statistical approaches to uncover patterns, in *Proceedings of the 121st ASEE Annual Conference and Exposition*, Indianapolis, IN, February 2014, pp. 24.180.1–24.180.18, 2014.

Doug Clark is the Werklund Research Chair of Design-Based Learning at the University of Calgary's Werklund School of Education. His research focuses on the design of digital learning environments, students' learning processes, and the opportunities of design for teachers and students. His earlier work has focused specifically on students' conceptual change processes, computational thinking, argumentation, modeling, and use of representations in digital and non-digital inquiry and game environments. Doug has been awarded multiple funded grants from the US National Science Foundation, Department of Education, and the US National Academy of Education/Spencer Foundation.

Christopher Ostrowdun is a PhD candidate in educational research (Learning Sciences) within the Werklund School of Education at the University of Calgary. He has a BSc in Physics and Astrophysics and spent some time working in professional engineering environments before pursuing graduate studies in education. His research focuses on supporting educators in fostering inclusive learning environments that address and recognize learners of marginalized and equity-seeking groups.

Stefan Rothschuh is a PhD Student in mathematics education at the University of Calgary. He received a B.Ed. and a M.Ed. in mathematics education from the University of Potsdam, Germany. Stefan is a former high school mathematics teacher. His research investigates the interconnectedness of action-, image-, and symbol-based mathematics learning, embodied learning designs for mathematical functions and calculus, as well as technology integration in secondary school settings. Stefan teaches STEM education in the undergraduate education program.

Senay Purzer is a Professor in the School of Engineering Education at Purdue University and the Director of Assessment Research at the INSPIRE Institute for Pre-college Engineering Research. Her research focuses on learning and assessment in complex learning environments with a specific focus on engineering design. Her research in design education is at the intersection of critical thinking (trade-off decisions) and creative thinking (design innovation). She conducts research at college and pre-college levels. She is the chief editor of the Journal of Pre-College Engineering Education Research (J-PEER) and an editorial board member of Science Education.