# Learning by Evaluating: An Exploration of Optimizing Design-Based Instruction

***Scott Thorne, Nathan Mentzer, Greg Strimel, Scott Bartholomew, Jason Ware***

**Abstract**

Learning by Evaluating (LbE) is an instructional approach that involves students making comparative judgements of pairs of artifacts, such as student work, portfolios, prototypes, or curated images related to a topic of instruction to enhance critical thinking and decision-making skills. Situated as a primer for learning, the efficacy of LbE stems from actively engaging students in the evaluation process, scaffolding their learning, fostering self-reflection in decision-making, and facilitating the transfer of acquired skills and concepts to academic contexts and project work. However, there is an opportunity to gain deeper insights into classroom integration of LbE and the factors that may influence the student experience. This study adopts a design-based research approach to analyze LbE within a secondary STEM education setting, with the objective of optimizing classroom integration. By analyzing student comments generated during LbE, the research explores factors shaping the students' learning experiences, examining the extent to which students engage in informed decision-making, offer justifications, and express sentiments throughout the process. Additionally, the study explores how teachers strategically incorporate LbE into their classroom, aligning LbE sessions with curriculum objectives. Findings indicate a diverse pattern of student engagement, sentiments, and decision-making approaches across STEM classrooms. This study contributes to research on LbE by offering insights into the dynamics between teacher implementation and student engagement. The insights gained highlight the potential for refining the effectiveness of LbE within the classroom. Notably, the research emphasizes the significance of how LbE sessions are framed and strategically integrated to enrich the overall educational experience for both students and educators.

**Keywords** – Learning by Evaluating, instructional approach, student engagement, classroom integration, informed decision-making, design-based research

Thorne, S., Mentzer, N., Strimel, G., Bartholomew, S., & Ware, J. (2024). Learning by Evaluating: An Exploration of Optimizing Design-Based Instruction, *35*(2), 53-80. https://doi.org/10.21061/jte.v35i2.a.4

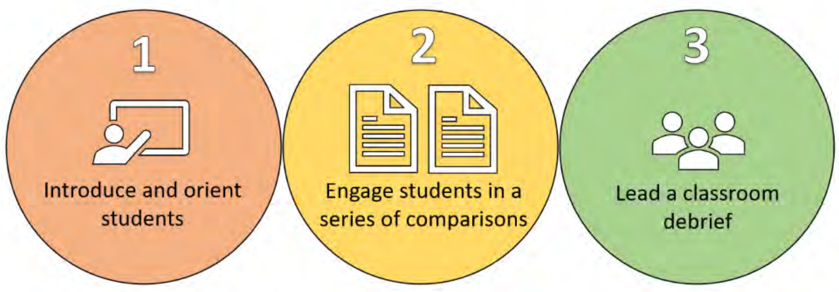
## Introduction

Learning by Evaluating (LbE) is a process or educational technique that applies the concept of comparative judgement to assist students in learning concepts and strategies related to open-ended design tasks by engaging them in making binary comparisons of student generated design artifacts (S. Bartholomew & Jones, 2022). LbE, as an educational practice, builds on work around the use of comparative judgment (CJ; Pollitt, 2012) and adaptive comparative judgment (ACJ; Kimbell, 2012, 2018) as assessment tools in educational settings. Both CJ and ACJ are approaches to evaluating items using paired comparisons; distinct approaches to evaluation which moves away from rubrics and other criterion-based approaches that rose out of efforts to improve the reliability of assessment outcomes in open-ended settings (Kimbell, 2007). CJ and ACJ—originally put forth as tools for use by teachers—have increasingly been used by students as a primer for later work; research in this vein with students engaging in ACJ as a part of their learning has shown promise in improving student capacity and attainment (Baniya et al., 2019; S. R. Bartholomew, 2017). As Bartholomew et al. (2022) points out, students benefit from this practice in four distinct ways: 1) exposure to new ideas, 2) critical comparison and evaluation of pairings, 3) providing and receiving feedback, and 4) an increased understanding of assignment criteria. LbE has been shown to help students who are engaged in resolving open-ended design challenges to define not only what a 'good' solution looks like, but to also encourage them to embrace challenges and learn from the mistakes of others to foster motivation toward project goals (Bartholomew, Mentzer, Jones, et al., 2022).

Teachers employing the LbE process (Figure 1) to help support their design-based instruction use a three-step process when implementing it into their classroom. Teachers first introduce and orient students to the topic/design challenge, identifying criteria for evaluation, often modeling one comparison before students start. Then students individually engage with comparative judgement software to systematically compare pairs of exemplars, judging which one better aligns with the given criteria. During these comparisons, students can submit comments to justify their judgements that can be consolidated and used as feedback. Finally, teachers lead the classroom in a debrief and a review of the comments to come to a consensus on criteria expectations and address any misconceptions.

**Figure 1**

*LbE Process*




This process may be situated as a primer before any step in the design process. In the context of design, exemplars represent samples of successful solutions or solutions that are typical student work of a given skill level. Integrating these exemplars into the instructional process allows educators to contextualize assignment criteria, enhancing students' comprehension of task expectations (Sadler, 2002). Research has indicated that by employing pairwise comparisons and making holistic judgements, students more reliably identify assignment criteria compared to subjective decisions based on a gallery walk-style evaluation (S. Bartholomew, Mentzer, & Jackson, 2022). In the STEM education setting, LbE is thought to be particularly beneficial for students new to the design process, aiding them in addressing challenges specific to design, such as 1) a lack of direction or steps to follow, 2) fear of failure, 3) limited prior experience to problem solving and collaboration, and 4) difficulty with ambiguity and interpreting the challenge (Crismond & Adams, 2012; Douglas et al., 2012; Ge et al., 2005; Jonassen, 1997).

Past studies on LbE have focused on the efficacy of comparisons as a treatment (Bartholomew, Mentzer, Jones, et al., 2022; Mentzer et al., 2021), the quality of items used for pairwise comparison (S. Bartholomew et al., 2021), and its adaptability across contexts (S. Bartholomew & Jones, 2022; Huber et al., 2021; Strimel et al., 2021). In LbE, exemplars are adapted to individualized judgment sessions using web-based software, randomizing samples for pairwise comparison in such a way that each student makes a different set of comparisons, followed by a whole-class discussion. This is a departure from the traditional use of exemplars in which all students see the same few samples to represent different qualities of student work. Pairwise comparison in LbE offers several advantages. Firstly, this approach allows students to make judgments within a specific timeframe, enabling faster students to complete more evaluations while ensuring that all students finish simultaneously. Secondly, teachers conduct group debriefings after students' judgments, providing an opportunity to recalibrate expectations and realign the focus based on their collective insights. Thirdly, by evaluating exemplars from a previous year, students receive feedback *before* initiating their drafts, which aids in refining their understanding and decision-making. Lastly, the use of flexible exemplar sessions empowers teachers to scaffold challenging topics at various stages throughout the design process, enhancing the overall learning experience. However, there is an opportunity to better understand how students engage with LbE in justifying and explaining knowledge-driven decisions through their evaluation comments to optimize the learning experience.

Employing a design-based research approach, this study investigates the initial implementation of LbE as a priming tool across multiple STEM classrooms, specifically within a secondary engineering setting. Situated in a three-year NSF funded project, this study aimed to identify a subset of specific classes and teachers to investigate with the goal of informing the next iteration of the LbE experience for both students and educators. To achieve this, we conducted a comprehensive directed content analysis, utilizing established theories for initial coding categories and expressing findings in terms of the percentage of codes for teacher participants (Hsieh & Shannon, 2005). This analysis was applied to 1012 justifications provided by students in the form of comments during pairwise comparative judgments within an LbE session. This analysis seeks to explore the student experience and its wider implications on classroom implementation. To ensure diverse perspectives, non-probability purposive sampling was utilized within a culturally and socio-economically diverse school district. This methodological choice aimed to capture a range of experiences and perspectives related to LbE in the specific context of secondary engineering education.

## Literature and Theoretical Framework

LbE in practice builds on the use of CJ or ACJ in educational settings (S. Bartholomew, Mentzer, Jones, et al., 2022). Specifically, wherein the research around ACJ and CJ has primarily focused on educational evaluation and assessment *by teachers at the conclusion of a project*, LbE positions the use of ACJ and CJ *by students as an intentional primer for learning near the beginning of a project*. These differences position LbE as a distinct use of ACJ in educational settings with a focus on the student and future learning as opposed to a teacher performing an evaluation of work already completed. However, despite these differences, the core foundation for LbE as an effective educational tool rests on the body of work into CJ, and later ACJ as educational tools (S. Bartholomew & Jones, 2022). In addition to literature around ACJ and CJ, the theoretical framework guiding this study is at the intersection of the theory of cognitive apprenticeship (Collins et al., 1991), Crismond and Adams' (2012) informed design teaching and learning matrix, and Sadler's theoretical justification and structuring of exemplars (Sadler, 1987). Cognitive apprenticeship has roots in socio-cultural theories of learning, demonstrating that as students socialize with peers, teachers, and other adults they develop and refine cognitive and communicative skills (Ghefaili, 2003), specifically through interactions, observations, imitation, and modeling observed techniques. When making holistic, side-by-side comparisons of peer exemplars, students have the opportunity to not only evaluate, but to learn from and reflect on the examples they are assessing. As a pedagogical approach, cognitive apprenticeship has been found to be highly effective in helping students acquire new knowledge and skills as they are able to work alongside experts and learn from their experiences. When implementing a cognitive apprenticeship model to engineering curriculum across two courses, Poitras and Poitras (2011) found a number of benefits to students, including 1) better facilitated learning, 2) allowed better teamwork skills, 3) was more efficient at obtaining course objectives, and 4) helped develop critical analysis and logical reasoning skills. Additionally, they found that students asked more questions, consulted more sources when solving problems, and favored cognitive apprenticeship over the traditional model. Interestingly, grades obtained by students in treatment and control group did not significantly differ, indicating that grades may be independent of teaching method (Poitras & Poitras, 2011).

Similarly, Crismond and Adams (2012) emphasize the application of logical reasoning skills to key performance dimensions, or design strategies, that are fundamental to the act of "informed design" within K-16 STEM education and engineering contexts. Informed design can be described as the act of bringing together diverse knowledge sources through guided investigations and logical reasoning to identify intellectual patterns and skillsets that are needed for the design task, which then assists the students in learning during the design process. The LbE process can be supportive of the logical reasoning involved with informed design. For example, the LbE framework can be aligned with the key performance dimensions of 1) learning while designing, 2) making and explaining knowledge-driven decisions, 3) working creatively to generate design insights and solutions, and 4) perceiving and taking perspectives intelligently. Within the framework of LbE, students engage in *learning while designing* by critically evaluating the strengths and weaknesses of various exemplars, engaging in dialogue with their teacher and peers, and reflecting on their own criteria when developing plausible solutions. They *make and explain knowledge-driven decisions* as they justify the rationale behind their choices when transferring insights gained from judged exemplars to the project at hand. By exploring diverse exemplars, students are provided the scaffolding to not only generate potential solutions, but also derive valuable creative insights. Lastly, students *perceive and take perspectives intelligently* by discerning what is relevant for their specific project and using this understanding to define criteria for potential solutions.

The use of exemplars facilitates the transfer and application of tacit knowledge regarding criteria, standards, and the nature and quality of work

(Grainger et al., 2018; Hawe & Dixon, 2017; Sadler, 2002) provided that they are not too far below or above what the student is capable of, sometimes referred to as their current zone of proximal development (Vygotsky, 1987). Comparing items side by side in a LbE model allows students to engage in a form of teacher facilitated cognitive apprenticeship through prompts to think about their thought process as they discriminate what aspects elevate one example above another when engaged in a problem-solving process (Brown et al., 1989; Ghefaili, 2003). The utilization of information technology, specifically the software employed in the LbE approach, presents the opportunity to document feedback on student judgments. This practice not only motivates students to reflect on and critically analyze their decisions, but also permits their cognitive process to become more transparent. By pinpointing specific points in the problem-solving process where students may require assistance, the use of technology acts as a form of cognitive apprenticeship, a concept that has been shown to aid in scaffolding (Collins et al., 1991; Ghefaili, 2003). In LbE, this is done through the use of side-by-side comparison of exemplars. For purposes of this research, the term "exemplar" is used to mean a *typical* model, and not the popular definition of an excellent model (Carless & Boud, 2018; Carless & Chan, 2017; Chong, 2019; Dixon et al., 2020).

Much of the literature on exemplars stems from Sadler's early work on achievement standards (Sadler, 1987) and formative assessment (Sadler, 1989). It was found that rubrics and criteria are often too abstract and decontextualized for students, driving the need for exemplars to clearly convey teacher expectations (Sadler, 2002). The use of exemplars has been shown to help to reduce anxiety in students (Yucel et al., 2014), identify and focus on important aspects of the project (Aitken & Thompson, 2018; Bouwer et al., 2018; Knight et al., 2019), and raise overall student performance on tasks (Hendry & Tomitsch, 2014; Newlyn, 2013; Tam, 2021).

Concerns about showing examples (exemplars) relate to a concept called "design fixation," where students might get too stuck on existing ideas and struggle to be creative (Jansson & Smith, 1991). However, research suggests that when students see a variety of examples of different quality levels, it actually helps them be more creative and come up with new, original designs, improving their creative abilities and helping them generate unique ideas (Chong, 2019; Hendry & Tomitsch, 2014; Tam, 2021). Using software to engage students in multiple pair-wise holistic judgements through LbE allows for students to see a variety of interpretations of criteria, exposing them to anywhere from 10 to 20 exemplars in a single session (Canty et al., 2017). This allows for students to both gain tacit knowledge about the task at hand (Aitken & Thompson, 2018; Grainger et al., 2018; Rust et al., 2003) and potentially reduce the time to task completion (Grainger et al., 2018; Headley & Pittson, 2020; Knight et al., 2019).

## Method

The implementation and observation of a teaching strategy like LbE across multiple schools is ongoing, complex, and involves bridging the gap between theory and practice while introducing numerous variables. As part of an evolving NSF funded design-based research (DBR) project on LbE, the goal of this study was to better understand the student and teacher experience and develop a sustainable solution in authentic educational settings. In this setting, DBR provides several advantages over other models, including 1) the promotion of iterative design and improvement to refine and enhance methods based on feedback, 2) collaboration with teachers who are facilitating the LbE model to utilize their expertise and co-create meaningful sessions for the students, and 3) greater generalizability of research findings by exploring contextual factors that may influence the success of similar interventions in different educational contexts (Barab & Squire, 2004; Bell, 2004; Design-Based Research Collective, 2003). This allowed for a practical and collaborative approach to address complex educational problems, facilitated iterative design and improvement, and promoted the generalizability of research findings to a broader educational context.

Research question one was guided by the cognitive apprenticeship framework, exploring how, and at which stages of the design process, teachers implement LbE sessions in STEM classrooms. This was specifically focused on secondary engineering classrooms, that rely heavily on the use of open-ended design tasks for teaching a variety of concepts and skills. This examination was then positioned to help identify and better understand the factors that influence the student experience when engaging in LbE as a primer for open-ended design projects, ultimately with a goal to optimize the LbE pedagogical approach. To address research question two, the study focused on investigating how students made and explained knowledge-driven decisions based on their LbE experience. Additionally, the study sought to gain insights into the nature of student engagement during LbE sessions through the comments they made when justifying judgement decisions.

From a design-based research perspective, the study aimed to determine whether the student experience was consistent across different teachers. With the variations discovered across classrooms, the study sought to identify specific classes and teachers to further investigate and enhance the LbE experience for both students and teachers. To address these objectives, the following research questions (RQs) guided the study:

RQ-1: Following training on the LbE approach and use of software for conducting judgement, how and in what ways do secondary engineering teachers uniquely integrate Learning by Evaluating sessions into their curriculum?

RQ-2: During participation in the Learning by Evaluating process, how and in what ways do comments made by students during the comparative judgement sessions to justify their decisions vary across different teachers?

### Participants and Data

During the 2021-2022 academic year, data were gathered in both spring and fall from students who utilized LbE as an introduction to a range of projects within the first-year course, Foundations of Technology & Engineering (FoTE), in the Engineering by Design (EbD) program. This context was chosen for our study as FoTE frequently serves as students' initial exposure to open-ended design projects, offering support in a process that may be unfamiliar to them. The study included students from five distinct schools within the DeKalb County School District in Atlanta, Georgia. To ensure consistency, all teachers involved in the study received training on both the LbE approach and the relevant software required for conducting judgment sessions. The five teachers participating in the study were all FoTE educators with a range of teaching experience, with some having from less than five years and others with more than 20 years of experience in the classroom (Table 1). Although all students in the FoTE course (414) were exposed to the Learning by Evaluating primer, only 98 completed the necessary consent and assent forms, thereby fully consenting to participate in the study.

**Table 1**

*Distribution of Student Participants*

| Teacher Number | Years of Teaching Experience | Years Teaching the FoTE Course | Number of Student Participants |
|---|---|---|---|
| 1 | >10 | <5 | 17 |
| 2 | >20 | <5 | 10 |
| 3 | <5 | <5 | 23 |
| 4 | >10 | >5 | 10 |
| 5 | >20 | >5 | 38 |

### Implementation Approach

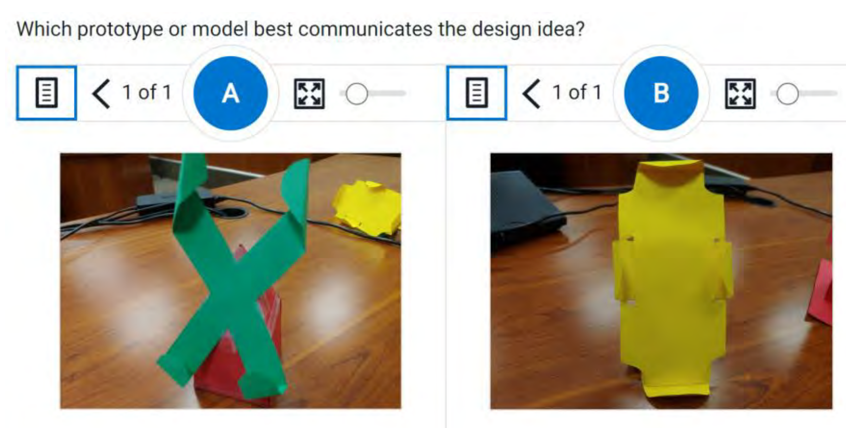
When discussing the implementation of a pedagogical strategy such as LbE, it is important to understand the role that teachers play in shaping the classroom experience. This section introduces the ways in which educators integrated LbE sessions into the teaching process, how these sessions are strategically aligned with classroom curriculum, and the preparation provided to educators for implementation. This provides the context necessary to understand the data utilized for this directed content analysis, and how it aligns with the study's methodology and how it may be used to optimize the experience in future iterations.

During class, teachers were encouraged to use examples of previous design solutions, such as the one presented in Figure 2, to introduce LbE to their students through a structured three-step process. Initially, the teacher would introduce the concept to be evaluated and help students understand the relevance to their ongoing project. Next, students would engage in a series of at least five comparisons using a *holistic statement* as a focal point to make each judgement, documenting evidence and reasoning for each decision when prompted. Finally, the teacher would lead a classroom debrief, focusing on the main concept being evaluated, technical insights derived from comparing exemplars, and ways students may transfer this knowledge to their own projects. The holistic statements used in LbE sessions are comprehensive prompts or questions to assist and focus students when making evaluations of various exemplars. It is designed to guide students to consistently evaluate each pair of exemplars on the same set of criteria, aligning with aspects and qualities valued by their respective classroom teacher. A well-structured holistic statement assists students in identifying criteria that are of most importance to the evaluation process, guiding students to make informed judgments or selections throughout the process.

During a session, a student would see a holistic statement across the top of their screen, such as “Which prototype or model best communicates the design idea” as well as two pictures of protypes from a previous year side by side (Figure 2).

**Figure 2**

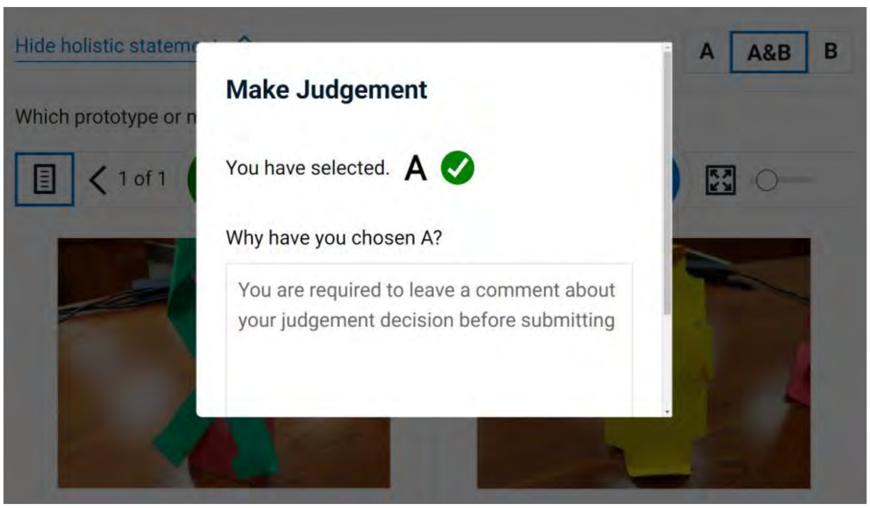
*LbE Judgement Interfaace*



Upon clicking one of the two images, the student would be prompted to provide an explanation for their decision (Figure 3). The screen would then change, replacing only the two images for comparison while maintaining the holistic statement. The student would continue to independently make comparisons and provide explanations for each judgment until the teacher ends the session, at which point the teacher would debrief with the students.

**Figure 3**

*Prompt for Justification of Decision*




Teachers were encouraged to facilitate these debriefs with conceptual questions such as “Why are quick prototypes important?”, technical questions such as “What made the good prototypes good?”, and transfer questions such as “What testing is important for your design?”.

Each LbE session is designed to align with the classroom curriculum with the flexibility to focus on various steps within the design process where additional scaffolding may be beneficial. By providing students with exemplars relevant to the specific stage of their design process, LbE primes and hones students' focus before continuing their ongoing design projects. Notably, we have observed that the most effective exemplars are those that lack a clear right answer, challenging students to engage in profound reflection about what truly matters and to establish their own criteria.

Sessions were summarized comprehensively on a “live” shared document that allowed editing and modifications throughout the year. This document provided teachers with links to more than 10 LbE sessions on various topics, each containing essential concepts, potential overarching statements, the specific location of curated images for exemplars, and reflective questions (Figure 4).

**Figure 4**

*LbE Prototyping Session Overview*

| | |
|---|---|
| **Key Concepts:** | Prototyping allows users and designers to test an idea.<br>Prototyping or modeling allows designers and users to communicate design intentions with physical representations. |
| **Potential Holistic Statements:** | Which prototype is most likely to be functional?<br>Which prototype or model best communicates the design idea? |
| **Link to Artifacts Folder:** | [url link to cloud storage] |
| **Debrief Questions:** | |
| **Conceptual:** | Why are quick prototypes important?<br>How do prototypes help with design iteration?<br>Why is it important to communicate your ideas with a model or prototype? |
| **Technical:** | What made the good prototypes good?<br>How were the good prototypes built?<br>Do you think you would learn from interacting with these prototypes? Why? |
| **Transfer:** | What testing is important for your design?<br>How can you test that aspect of your idea?<br>How might you best show your ideas using a prototype or model? |

While the research team provided initial holistic statements, additional statements were generated collaboratively with the classroom teachers. Across the five teachers, the DBR approach was used to co-create 14 unique holistic statements for LbE sessions, with consenting students generating 1012 comments on individual judgements. In partnership with the participating teachers, the LbE sessions were developed to address different aspects of the design process. These sessions encompassed activities such as documenting, formulating problem statements, brainstorming, sketching, conducting research, and prototyping.

As part of the professional development provided to the teachers, we highlighted the importance of anticipating that students' conclusions may not always align with their teachers. This emphasized the significance of conducting debriefing sessions to facilitate constructive discussions and gain deeper insights into students' perspectives. These debriefs enabled teachers to recalibrate student understanding through open dialogues. During these discussions, students had the opportunity to share the evidence and reasoning they utilized during their individual comparison sessions and to reconsider their viewpoints based on the insights and reasoning of their peers. Through deliberating on what aspects are better and why, students, together with their teacher, can foster a deeper understanding of project expectations in preparation for their design work. The LbE sessions employed by teachers in their classrooms were analyzed and coded, alongside the responses students provided for each judgment.

### *Data Analysis*

The software utilized for conducting LbE sessions gathered data across multiple dimensions of the student experience. In the context of this study, the data of particular relevance included the 1) classroom teacher, 2) holistic statement, 3) judgement items, and 4) judgement comments. Initially, these data were transferred to a spreadsheet for preliminary analysis before being further processed and examined using NVivo software. Once in NVivo, student comments were coded by attributes, a recommended first cycle analysis in several passes (Saldaña, 2013), to identify emergent themes. The second cycle of analysis made sense of the data shared in the discussion section.

When analyzing and categorizing holistic statements in the first pass, one finding that emerged was that teachers were using LbE sessions with both *convergent* and *divergent* approaches. In the convergent approach, LbE helped students narrow in on what makes something 'good'. On the other hand, the divergent approach involved students using LbE for ideation, expanding their approaches to problem-solving. While there are several studies focused on comparative judgement for convergent tasks (Bartholomew, Mentzer, Jones, et al., 2022; Bartholomew & Connolly, 2017; Potter et al., 2017; Seery et al., 2019), in the literature review conduced for this study no sources could be found exploring their use for divergent thinking. For example, the statement "*Which of these resources will help you create an informative website about your topic?*" was coded as divergent, as students are using the judgment process to broaden their approach to their project and help with brainstorming. However, the statement "*Which problem statement is the best?*" was coded as convergent, as all students in the class would narrow in what makes a good problem statement.

Once themes were identified for each holistic statement, a second pass was conducted to develop rules for coding student responses for a thematic analysis for each decision made in a LbE session (Saldaña, 2013). With inductive coding we discovered that some of the students' comments emphasized positive attributes for the item they selected, while others focused on negative attributes on the item that was not selected. For this paper we will call this the *sentiment* of each judgment and categorize each as either a *positive sentiment* or *negative sentiment*. *Positive sentiment* was used to code judgments where it was clear the justification was for the item the student picked, such as "*I choose this statement* *because it talks about the material they will use to create the backpack out of and it also says who the intended audience for this product will be.*" While a *negative sentiment* was used to code justifications that were made because they did not like the other option such as "*I just don't like A at all, it's too complex.*" During this analysis, open coding was used deductively to identify judgements that used both examples, A and B, in their justification. One such example was "*A is more readable and in order compared to B's writing being everywhere*". Finally, there were several judgements that did not seem to fit any of these categories, in which case they were coded as *unclear*. This could be because the student did not put effort into making a comment with nonsensical responses such as "*lk;hgjktfrewq*" or "*n/a*", but also included responses indicating a glitch in the software "*There was no project b"* or when students were confused and used the justification section to describe what they were seeing such as "*Its showing that the tooth paste is a product that builds strong teeth.*"

Driven by insights from Crismond and Adams (2012), a third analysis pass was conducted to explore how each student makes and explains knowledge driven decisions. Inductive coding was initially used to identify students who provided explanations for their decisions as *engaged* while those that typed random keystrokes were identified as *disengaged*. However, it was apparent during analysis that not all students engaged to the same extent. Therefore, each justification made when deciding between two options in the LbE session was analyzed using a claims, evidence, and reasoning framework. To reflect this, codes were revised to focus on *claims with evidence*, *claims without evidence*, *neither claims or evidence*, and *other*. For an item to be coded as *claims with evidence,* it had to specify *why* the decision was made. This resulted in both lengthy statements such as "*I liked how the animation demonstrated lab incidents that could not be represented in real life without serious injuries*" and decisions justified with short statements such as "*Easier to carry.*" Student responses identified as *claims without evidence* showed that they were still attempting to participate in the activity but did not demonstrate a rationale in their response. Examples include "*this one is better",* or "*I chose this design because the design in the other group looks like it will not go as fast as the one in this group.*" Even though the second example demonstrates a longer response, the student does not specify *why* one design may not be as fast as another design. Student responses coded as *neither claims or evidence* consisted of "*n/a"* and "*htrshtrs*" style comments, but also included were copy/paste responses that at first glance might be coded as *high-* or *low-engagement* such as "*I love the design and how clean and nice it is and it looks just like the design.*" This is a valid response by itself, however this particular student used the same response for 16 consecutive judgements over a timespan of about five minutes. The category of *other* was used to identify student comments in which they did not seem to understand the activity, or they were reporting a problem with the software. Examples include a description of what they were seeing on the screen, such as "*the picture is showing you that there is something to eat which shows you the restaurant is trying to advertise their food,*" confusion "*I don't know what archived data is or how it would be used…,*" and problems "*photo b isn't loading.*"

During the final coding passes, attribute coding was again utilized to align each comment with the student's classroom teacher and associate the holistic statement with a specific step in the design process. This second coding cycle, following Saldaña's methodology (2013), allows for the development of themes and patterns across different teachers. The analysis aims to identify effective teaching methods and styles, providing valuable insights for further investigation in understanding teaching practices' impact on the student experience.

### *Trustworthiness and Credibility*

Several strategies outlined in *Qualitative Inquiry and Research Design* (Creswell & Poth, 2017) were utilized to build trustworthiness and credibility of data analysis. The researcher spent time engaging with each of the teachers by co-delivering professional development, making observations of classrooms, and reviewing interviews of both teachers and students, which demonstrates prolonged engagement and persistent observation. Findings were reviewed with co-authors and co-investigators on the research team as an external check of the research process, and an external consultant was utilized throughout the study to examine the research process and assess the accuracy of findings. It is suggested that a researcher engage in at least two strategies as a means of building trustworthiness and credibility (Creswell & Poth, 2017, p. 253).

## Results

During an inductive coding analysis of student comments and LbE sessions, three key areas emerged. Firstly, we observed how teachers are integrating LbE into their classrooms, noting the frequency of its usage and the step in the design process that aligns with each session. The second area of focus relates to the claims, evidence, and reasoning expressed by students to support their judgments during LbE sessions. Lastly, we examined the sentiments of students when making judgements and the specific aspects they prioritized in each decision. Each of these aspects will be discussed in detail in the following sections.

### *RQ-1: How and in What Way do Teachers Integrate LbE*

To examine how secondary teachers integrated LbE sessions into their curriculum, each session was aligned with specific steps in the curriculum and design process, and a frequency count of each topic was generated for each teacher (Table 2). Among various design process models, the 12-step process from the International Technology and Engineering Educators Association

(ITEEA, 2011) was chosen for analysis as it aligns with the design process used by students in the course. This approach offers insights into the topics that teachers emphasized more frequently, indicating potential areas of effectiveness or the need for revision to enhance the impact of certain sessions. Given the iterative nature of design practices and the repetition of steps across multiple projects in the course, many teachers had repeated LbE topics. This repetition may provide valuable information on which topics are particularly influential for LbE.

**Table 2**

*Alignment of LbE Session with Design Process*

| Topic | Teacher Number | | | | | Total |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| **Course Orientation** | | 2 | | | | **2** |
| **Define Problem** | | | 1 | | 1 | **2** |
| **Brainstorm Possible Solutions** | 2 | 1 | 2 | | 2 | **7** |
| **Research Ideas/Explore Possibilities** | 1 | 2 | | | | **3** |
| **Consider Alternative Solutions** | 1 | | 1 | | 2 | **4** |
| **Develop Written Design Proposal** | 1 | | | | | **1** |
| **Test and Evaluate** | | | | 2 | | **2** |
| **Create/Make Product** | | | | 1 | | **1** |
| **Communicate Results** | 1 | 2 | 2 | 1 | | **6** |
| **Total Counts** | 6 | 7 | 6 | 4 | 5 | 28 |

Visualizing the data in this way offers several insights. First, it became evident that each teacher implemented a relatively similar number of LbE sessions throughout the school year. Although the majority of these were focused on steps in the design process, one teacher, Teacher 2, utilized LbE as an ice breaker during course orientation with two different sessions, having students converge on what makes a good student and what makes a good teacher. Furthermore, the data demonstrates that LbE was used throughout the design process, indicating that LbE was utilized to support students throughout their projects, not just at the initial or final stages. Not every step in the design process was represented, steps that were not included in LbE by the teachers this year included 1) specify constraints and identify criteria, 2) select an approach, 3) make model/prototype, and 4) refine/improve. Additionally, we can also see that four out of five teachers in the study incorporated brainstorming and communicating results as crucial and often recurring elements in their LbE sessions.

A unique feature of the brainstorming category is the discovery of teachers using LbE as a tool to foster divergent thinking. While the original purpose of LbE in this study was to assist students in refining their understanding and converging on the concept of what makes something "good," we observed that this was not the case for all teachers. Specifically, two teachers, teacher 1 and teacher 5, framed the holistic statement in a way that encouraged students to explore a broad range of divergent ideas. For instance, in one LbE session, students were prompted to find inspiration from items that were not directly related to the given problem, with the holistic statement, "How could each item inspire you to solve the problem? Which item best helps you with your design and why?" Another instance of promoting divergent thinking occurred in a session where students were presented with websites showcasing new technologies and were asked, "Which of these resources will help you create an informative website about your topic?" These examples showcase how LbE sessions were tailored to foster students' creativity and exploration of diverse ideas in their design projects.

It's worth noting that the number of students consenting to be part of the study varied significantly among teachers, despite each teacher having a similar number of students in each classroom. By using a frequency count of LbE sessions per teacher, researchers can gain insights when analyzing comment data and investigating the student experience, accounting for potential variations in the number of students participating from different classrooms.

### *RQ-2: Student Justifications: Claims, Evidence, and Reasoning*

Student responses were analyzed using attribute coding and quantified to identify themes and patterns across teachers. A matrix of student responses was created, using number identifiers for each classroom teacher and the appropriate category using a claims, evidence, and reasoning framework. The teacher identifiers in this matrix correspond with the ones presented in previous tables. For every teacher, the matrix contains the count of student comments coded for each engagement level, along with the corresponding percentage. This arrangement allows for a comparison of engagement levels across different teachers (Table 3).

**Table 3**

*Matrix of Claims Evidencing Reasoning Findings*

| Making and explaining knowledge-driven decisions | Teacher Number | | | | | | | | | | Total Counts | Total (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | 3 | | 4 | | 5 | | | |
| | # | % | # | % | # | % | # | % | # | % | | |
| Claim and Evidence | 31 | 30.4 | 104 | 48.4 | 4 | 100 | 156 | 26.7 | 61 | 57.5 | 356 | 35.2 |
| Only Claim | 64 | 62.7 | 68 | 31.6 | 0 | 0 | 295 | 50.4 | 26 | 24.5 | 453 | 44.8 |
| Neither Evidence nor Claim | 0 | 0.0 | 31 | 14.4 | 0.0 | 0.0 | 123 | 21.0 | 6 | 5.7 | 160 | 15.8 |
| Unclear | 7 | 6.9 | 12 | 5.6 | 0 | 0 | 11 | 1.9 | 13 | 12.3 | 43 | 4.2 |
| **Total Counts** | 102 | | 215 | | 4 | | 585 | | 106 | | 1012 | |

In analyzing this data, it appears that student decisions varied across different teachers in regard to providing evidence for claims when making knowledge-driven decisions. With teachers 1 and 4, the most common approach among students was to make decisions, or claims, without offering supporting evidence, such as one student who made the following comment while comparing advertisements "*it seems like something people would enjoy*." However, most of the students in teacher 2 and 5's classes provided both a claim and the corresponding evidence to substantiate it. An example of this is a student comparing hand-drawn sketches who stated, "*B looks more professional with shading and hatching*". Comments identified as 'Neither Evidence nor Claim" lacked substantive information, often stating "*n/a*" or containing random strings of text. Additionally, comparing tables 2 and 3 reveals that the number of student comments available for analysis is not proportional to the number of LbE sessions run by teachers. For example, teacher 3 conducted six LbE sessions, however, the limited number of fully consented students and the student's participation in only one session resulted in only four available student comments for analysis.

***Student Sentiment***

Following the inductive coding process, an additional measure of student responses was explored which focused on capturing the sentiment of each student comment. The majority of students across all teachers' classes emphasized the positive aspects of the item they selected saying things like "*This image shows more uses for the product in a visual way*" when brainstorming backpack designs, or "*I chose this one because it looks more symmetrical than the other group (option)*" when evaluating CO2 dragster designs. Similarly consistent across teachers, a small percentage of student comments focused on the negative aspects of the unchosen item with remarks such as "*Option B had no links to navigate you while this did*" when comparing website designs for their portfolio. This represented about 5% of all student comments, and less than 9% for any one teacher. Even fewer student comments addressed both the positive aspects of the item they selected as well as aspects they did not like about the unselected item as seen in the following comment where a student was comparing lab safety videos, "*Video A is more effective in demonstrating lab safety. Video B contained too many loud clips which made it hard to watch*." Similar to the previous table, the teacher identifiers in this matrix correspond with those presented in Table 1. For each teacher, the matrix displays the count of student comments coded for each sentiment level, along with the corresponding percentage. This approach facilitates a comparison of sentiment levels across different teachers (Table 4).

**Table 4**

*Matrix of Sentiment Findings*

| | Teacher Number | | | | | | | | | | Total Counts | Total (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | 3 | | 4 | | 5 | | | |
| **Sentiment** | **#** | **%** | **#** | **%** | **#** | **%** | **#** | **%** | **#** | **%** | | |
| Positive Sentiment | 81 | 79.4 | 165 | 76.7 | 4 | 100 | 524 | 89.6 | 93 | 87.7 | 867 | 85.7 |
| Negative Sentiment | 9 | 8.8 | 12 | 5.6 | | | 26 | 4.4 | 6 | 5.7 | 53 | 5.2 |
| Both | 3 | 2.9 | 6 | 2.8 | | | 19 | 3.2 | 3 | 2.8 | 31 | 3.1 |
| Unclear | 9 | 8.8 | 32 | 14.9 | | | 15 | 2.6 | 4 | 3.8 | 60 | 5.9 |
| **Total Counts** | 102 | | 215 | | | | 585 | | 106 | | 1012 | |

Our analysis suggests that the teaching method has minimal influence on student sentiment when students justify their decisions. The majority of students tended to emphasize the positive aspects of the item they chose, while the smallest percentage of students discussed both positive and negative aspects of the two choices presented. Additionally, a higher number of students did not participate in justifying their decisions compared to those who provided both positive and negative evidence for making informed design choices. These findings suggest current teaching methods might not significantly influence how students express their sentiments and justifications during the decision-making process.

The findings from this study reveal several significant observations. Teachers incorporated LbE not just at the start of a project, but consistently throughout the design process, serving as primers for the entire course, individual projects, and specific stages of the design process, showing that it can be adaptable for any or all of these three aspects. Additionally, all teachers employed LbE as a tool for brainstorming and envisioning the final projects. Interestingly, teachers also utilized LbE in unexpected ways, such as using it as an icebreaker to commence the academic year and to foster divergent thinking among students. Delving into the student experience, the study highlights the variability in student justifications with the inclusion of evidence when making claims across different teachers. This is in contrast to the sentiment used when providing evidence, placing the most emphasis on positive aspects of the item chosen while refraining from commenting on the item that was not chosen. These insights shed light on the diverse and dynamic ways in which LbE is implemented and experienced in the educational setting.

## Discussion

We began this study with the assumption that engaging students in learning through evaluating exemplars is an effective scaffolding strategy that may be utilized at any point in the design process. From analyzing LbE sessions utilized in the Foundations of Technology & Engineering course, we see that when left up to teachers to decide when to provide scaffolding, they provided opportunities for students to evaluate exemplars in eight of the 12 steps of the design process. The majority of these were in the brainstorming phase (7 sessions) followed closely by communicating results (6 sessions). Four out of five teachers implemented these sessions in their classrooms, often more than once. This may indicate that LbE is particularly well suited for sessions focused on these topics or these are easily accessible topics for which to create sessions. Other sessions that teachers utilized more than once include 1) research ideas/explore possibilities, 2) consider alternative solutions, and 3) test and evaluate. As this was a pilot year for the program, it may be informative to investigate what teachers implement in the second year, especially after conferring with their colleagues in the program.

Among the twelve steps in the design process, participating teachers covered eight steps at various points throughout the year, while four steps were notably absent. The missing steps include 1) specify constraints and identify criteria, 2) select an approach, 3) make model/prototype, and 4) refine/improve. Possible reasons for this could be that 1) teachers were uncertain about how to create side-by-side comparison clips for these topics, 2) they opted for topics that were already started as a time-saving measure, or 3) they might feel that students do not need additional support in these particular areas. An area for potential exploration in future research when optimizing this process is whether or not these topics are well-suited for the application of LbE sessions. It is worth noting that exemplars have been shown to assist students in recognizing and synthesizing constraints and criteria (Grainger et al., 2018; Hawe & Dixon, 2017; Sadler, 2002). Consequently, students may gain a tacit knowledge of what is required through earlier LbE sessions, thus facilitating the process of identifying constraints and criteria and selecting an approach based on informed decisions. The step of refining and improving a design is often unique to each design itself, making it a task that may not be as suited for LbE sessions due to its specialized nature. Additionally, we discovered that in the researcher created library of sessions, we did not offer any specific LbE session on strategies students could utilize when creating models or prototypes, an aspect that warrants further investigation for future research.

According to Crismond and Adams (2012), good design work is characterized by thoughtfulness, rationalization, and insight. In our study, we sought to observe students demonstrating these qualities by providing both claims and evidence in their comments, justifying each of their design decisions. However, our findings indicated that students did not consistently offer evidence to support their reasoning across classrooms. On average, 44.8% of comments contained claims without evidence, while only 35.2% of comments included both claims and evidence. Interestingly, teachers 2 and 5 had more students providing both claims and evidence than claims alone. For instance, teacher 5's students demonstrated evidence to justify decisions in 57.5% of their comments. These findings suggest the presence of other factors that may influence student success, prompting the need for closer examination of how these teachers facilitate LbE sessions in future studies.

While an initial analysis might suggest that 4-6% of students whose comments were categorized as 'unclear' were not actively participating in LbE sessions, additional insights were gained from classroom observations. As LbE sessions were conducted using students' laptops or school desktops computers, it was observed that some students displayed limited proficiency in keyboard typing. These students employed various strategies when attempting the in-class task: 1) typing slowly using one or two fingers to search for each letter to compose short responses, 2) crafting what they considered to be one 'good' response, to then copy and paste it as justification for each decision, or 3) resorting to random keystrokes to move on to the next comparison. This highlights a limitation in our study, as students' typed responses may not accurately reflect their cognition or engagement and could be an unreliable predictor of informed design (Crismond & Adams, 2012). Alternative approaches to capture decision-making processes, such as speech-to-text software, individual interviews, or think-aloud protocols (Ericsson & Simon, 1998), could be considered. However, an equally effective and less intrusive strategy might involve assessing engagement through the outcomes of students after an LbE session when compared to a control group.

When looking at student sentiment while analyzing designs, we found that overall students commented on the item they picked, however, they often did not comment on the item that was not picked. This is likely due to the phrasing of the holistic statement, often asking students to choose the “best” exemplar and then justifying why they made that decision. However, upon examining the 3% of student responses that emphasized advantages to both items before justifying their selection, it was evident that these students exhibited a stronger engineering habit of mind (Katehi et al., 2009). Termed as optimism, the ability to perceive possibilities and opportunities within designs, even those not chosen, is believed to foster students' innovation and resilience in overcoming setbacks in design.

In future iterations of the study, the framing of the holistic statement could be adjusted to prompt students to identify positive aspects of both options A and B before choosing the best exemplar and justifying their decision. Additional areas of future research may explore how LbE enhances STEM learning and technology education by actively engaging students, promoting reflection, and exposing them to diverse problem-solving strategies. Additionally, investigations into LbE's impact on cross-disciplinary integration in STEM, including critical evaluation and consideration of various factors and perspectives in decision-making, offer promising directions for future exploration.

## Conclusion

This study has provided valuable insights into the implementation and impact of LbE sessions within secondary engineering classrooms with the goal to improve and optimize the experience for both students and educators. Through a directed content analysis of student comments generated during LbE sessions, we have explored the diverse ways in which students engage with this pedagogical approach, the dynamics of their justifications for each decision made in the comparison session, and how this may vary across classrooms.

Our investigation also focused on the ways in which teachers integrated LbE into their classrooms, emphasizing different stages of the design process and creatively adapting LbE for both convergent and divergent thinking. During ideation, teachers utilized LbE as a primer, presenting students with a diverse range of exemplars to spark creativity and encourage innovative ideas. Some teachers revisited LbE for ideation in subsequent projects, fostering an environment conducive to exploring multiple perspectives and generating novel solutions.

Teachers also integrated LbE into subsequent stages of the design process, utilizing it as a tool to converge on solutions, refine designs, establish testing procedures, and develop effective communication strategies. Additionally, several teachers incorporated LbE multiple times throughout the subsequent stages of the design process, recognizing its value in fostering collaborative decision-making among students. By leveraging peer feedback through comparisons and critical evaluation, students were empowered to make informed decisions, enhancing the quality of their designs.

These findings highlight the adaptability of LbE across the curriculum, demonstrating its application in both creative and analytical thinking skills. As a dynamic pedagogical approach, LbE offers educators a flexible framework to help students navigate the design process while promoting critical thinking skills such as formulating claims, providing evidence, and reasoning through a variety of problems. Notably, students from certain teachers continued to utilize both claims and evidence in their decision-making throughout the design process. This finding suggests that there may be other instructional strategies or factors contributing to students' critical thinking and reasoning abilities that, if utilized, may enhance the application of LbE.

By focusing on claims, evidence, and reasoning expressed by students during LbE sessions, we observed variations in how students rationalized their decisions. One recommendation for future refinement may be for educators to place more emphasis on how to make claims prior to LbE sessions as a way to scaffold students into an informed design mindset. The presence of both positive and negative sentiments among students when making comments further highlighted the ways in which students may engage. Modifications of holistic statements and prompts for comments may additionally assist students with seeing value in various designs to work toward an engineering habit of mind.

Situated in design-based research methodology, we bridged the theoretical and practical realms, enabling a deeper understanding of the nuances associated with LbE implementation in a high school setting. This iterative and collaborative approach facilitated refinement, utilizing the insights of teachers with the creation and delivery of LbE sessions to inform future LbE sessions and our collective understanding of the learning process. Additionally, the directed content analysis of student comments provided context to understand how students perceived and engaged with each session.

However, our study is not without limitations. If capturing the thought process of the students when making decisions is of value, the reliance on typed responses and potential discrepancies between written responses and true cognitive engagement may justify further consideration. Future investigations could explore alternative methods to better capture the way in which students interpret prompts and rationalize decisions during LbE sessions.

In conclusion, this study has highlighted the ways in which teachers use LbE sessions to facilitate the design process and how students engage with them. By delving into the students' perspectives and experiences, we have gained valuable insights that can contribute to the ongoing evolution of LbE as an effective pedagogical tool. As we continue to refine and adapt the approach, we hope that our findings will empower educators to create enriched learning environments that foster decision-making skills, informed design thinking, and engineering habits of mind among their students.

**Funding**

This material is based upon work supported by the National Science Foundation under Grant 2101235. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

## Ethics Approval

This research was approved through the Human Research Protection Program housed in the Office of Research at Purdue University.

## References

Aitken, A., & Thompson, D. G. (2018). Using software to engage design students in academic writing. *International Journal of Technology and Design Education*, *28*(3), 885–898. https://doi.org/10.1007/s10798-017-9413-4

Baniya, S., Mentzer, N., Bartholomew, S., Chesley, A., Moon, C., & Sherman, D. (2019). Using Adaptive Comparative Judgment in writing assessment: An investigation of reliability among interdisciplinary evaluators. *The Journal of Technology Studies*, *45*(1), 24–35.

Barab, S., & Squire, K. (2004). Design-Based Research: Putting a Stake in the Ground. *Journal of the Learning Sciences*, *13*(1), 1–14. https://doi.org/10.1207/s15327809jls1301_1

Bartholomew, S. (2017). Assessing open-ended design problems. *Technology and Engineering Teacher*, *76*(6), 13–17.

Bartholomew, S., & Connolly, P. E. (2017, June 24). *Adaptive Comparative Judgment in Graphics Applications and Education*. 2017 ASEE Annual Conference & Exposition. https://strategy.asee.org/adaptive-comparative-judgment-in-graphics-applications-and-education

Bartholomew, S., & Jones, M. D. (2022). A systematized review of research with adaptive comparative judgment (ACJ) in higher education. *International Journal of Technology and Design Education*, *32*(2), 1159–1190. https://doi.org/10.1007/s10798-020-09642-6

Bartholomew, S., Mentzer, N., & Jackson, A. (2022). Lessons from Dilbert: Clarifying design expectations. *Technology & Engineering Teacher*.

Bartholomew, S., Mentzer, N., Jones, M., Sherman, D., & Baniya, S. (2022). Learning by evaluating (LbE) through adaptive comparative judgment. *International Journal of Technology and Design Education*, *32*(2), 1191–1205. https://doi.org/10.1007/s10798-020-09639-1

Bartholomew, S., Yauney, J., Mentzer, N., & Thorne, S. (2021). *Investigating the impacts of differentiated stimulus materials in a Learning by Evaluating activity*. Mississippi Valley Technology Teacher Education Conference (MVTTEC), Nashville, TN, USA.

Bell, P. (2004). On the theoretical breadth of design-based research in education. *Educational Psychologist*, *39*(4), 243–253. https://doi.org/10.1207/s15326985ep3904_6

Bouwer, R., Lesterhuis, M., Bonne, P., & De Maeyer, S. (2018). Applying criteria to examples or learning by comparison: Effects on students' evaluative judgment and performance in writing. *Frontiers in Education*, *3*. https://www.frontiersin.org/articles/10.3389/feduc.2018.00086

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32–42. https://doi.org/10.3102/0013189X018001032
Canty, D., Seery, N., Hartell, E., & Doyle, A. (2017). *Integrating peer assessment in technology education through adaptive comparative judgment*.
Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354
Carless, D., & Chan, K. K. H. (2017). Managing dialogic use of exemplars. *Assessment & Evaluation in Higher Education*, *42*(6), 930–941. https://doi.org/10.1080/02602938.2016.1211246
Chong, S. W. (2019). The use of exemplars in English writing classrooms: From theory to practice. *Assessment & Evaluation in Higher Education*, *44*(5), 748–763. https://doi.org/10.1080/02602938.2018.1535051
Collins, A., Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *AMERICAN EDUCATOR*, *6*, 38--46.
Creswell, J., & Poth, C. (2017). *Qualitative inquiry and research design* (4th ed.). SAGE Publications Inc. https://us.sagepub.com/en-us/nam/qualitative-inquiry-and-research-design/book246896
Crismond, D. P., & Adams, R. S. (2012). The Informed Design Teaching and Learning Matrix. *Journal of Engineering Education*, *101*(4), 738–797. https://doi.org/10.1002/j.2168-9830.2012.tb01127.x
Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, *32*(1), 5–8. https://doi.org/10.3102/0013189X032001005
Dixon, H., Hawe, E., & Hamilton, R. (2020). The case for using exemplars to develop academic self-efficacy. *Assessment & Evaluation in Higher Education*, *45*(3), 460–471. https://doi.org/10.1080/02602938.2019.1666084
Douglas, E. P., Koro-Ljungberg, M., McNeill, N. J., Malcolm, Z. T., & Therriault, D. J. (2012). Moving beyond formulas and fixations: Solving open-ended engineering problems. *European Journal of Engineering Education*, *37*(6), 627–651. https://doi.org/10.1080/03043797.2012.738358
Ge, X., Chen, C.-H., & Davis, K. A. (2005). Scaffolding novice instructional designers' problem-solving processes using question prompts in a web-based learning environment. *Journal of Educational Computing Research*, *33*, 219–248. https://doi.org/10.2190/5F6J-HHVF-2U2B-8T3G
Ghefaili, A. (2003). Cognitive apprenticeship, technology, and the contextualization of learning environments. *Journal of Educational Computing, Design & Online Learning*, *4*.

Grainger, P. R., Heck, D., & Carey, M. D. (2018). Are Assessment Exemplars Perceived to Support Self-Regulated Learning in Teacher Education? *Frontiers in Education*, *3*. https://www.frontiersin.org/articles/10.3389/feduc.2018.00060

Hawe, E., & Dixon, H. (2017). Assessment for learning: A catalyst for student self-regulation. *Assessment & Evaluation in Higher Education*, *42*(8), 1181–1192. https://doi.org/10.1080/02602938.2016.1236360

Headley, J., & Pittson, H. (2020). Bottom Up! Establishing a University Wide Community of Practice for Exemplars and Exploration of Its Affect in Enhancing the Quality of Student Assessment. *Practitioner Research in Higher Education*, *13*(1), 18–26. eric.

Hendry, G. D., & Tomitsch, M. (2014). Implementing an exemplar-based approach in an interaction design subject: Enhancing students' awareness of the need to be creative. *International Journal of Technology and Design Education*, *24*(3), 337–348. https://doi.org/10.1007/s10798-013-9256-6

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Huber, S., Bosman, L., & Bartholomew, S. (2021). Library instruction and adaptive comparative judgment to foster visual literacy skills. *Portal: Libraries and the Academy*, *21*(1), 149–169. https://doi.org/10.1353/pla.2021.0001

ITEEA. (2011). *Engineering Design Process (EDP) grades 6-12*. Retrieved August 10, 2023, from https://www.iteea.org/engineering-bydesign

Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design Studies*, *12*(1), 3–11. https://doi.org/10.1016/0142-694X(91)90003-F

Jonassen, D. H. (1997). Instructional design models for well-structured and III-structured problem-solving learning outcomes. *Educational Technology Research and Development*, *45*(1), 65–94. https://doi.org/10.1007/BF02299613

Katehi, L., Pearson, G., & Feder, M. (2009). The status and nature of K-12 engineering education in the United States. *The Bridge on K-12 Engineering Education*, *3*, 5–10.

Kimbell, R. (2007). E-assessment in project e-scape. *Design & Technology Education: An International Journal*, *12*(2).

Kimbell, R. (2012). The origins and underpinning principles of e-scape. *International Journal of Technology and Design Education*, *22*(2), 123–134. https://doi.org/10.1007/s10798-011-9197-x

Kimbell, R. (2018). Constructs of quality and the power of holism. *Perspectives on Human Capacity and Development*, 181–186.

Knight, S., Leigh, A., Davila, Y. C., Martin, L. J., & Krix, D. W. (2019). Calibrating assessment literacy through benchmarking tasks. *Assessment &*

*Evaluation in Higher Education*, *44*(8), 1121–1132. https://doi.org/10.1080/02602938.2019.1570483
Mentzer, N., Lee, W., & Bartholomew, S. (2021). Examining the validity of adaptive comparative judgment for peer evaluation in a design thinking course. *Frontiers in Education*, *6*. https://www.frontiersin.org/articles/10.3389/feduc.2021.772832
Newlyn, D. (2013). Providing Exemplars in the Learning Environment: The Case For and Against. *Universal Journal of Educational Research*, *1*(1), 26–32.
Poitras, G., & Poitras, E. (2011). A cognitive apprenticeship approach to engineering education: The role of learning styles. *Engineering Education*, *6*(1), 62–72. https://doi.org/10.11120/ened.2011.06010062
Pollitt, A. (2012). The method of Adaptive Comparative Judgement. *Assessment in Education: Principles, Policy & Practice*, *19*(3), 281–300. https://doi.org/10.1080/0969594X.2012.665354
Potter, T., Englund, L., Charbonneau, J., MacLean, M. T., Newell, J., & Roll, I. (2017). ComPAIR: A new online tool using adaptive comparative judgement to support learning with peer feedback. *Teaching and Learning Inquiry*, *5*(2), Article 2. https://doi.org/10.20343/teachlearninqu.5.2.8
Rust, C., Price, M., & O'Donovan, B. (2003). Improving Students' Learning by Developing their Understanding of Assessment Criteria and Processes. *Assessment & Evaluation in Higher Education*, *28*(2), 147–164. https://doi.org/10.1080/02602930301671
Sadler, D. R. (1987). Specifying and Promulgating Achievement Standards. *Oxford Review of Education*, *13*(2), 191–209. https://doi.org/10.1080/0305498870130207
Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144. https://doi.org/10.1007/BF00117714
Sadler, D. R. (2002). Ah!… so that's 'quality'. In *Assessment*. Routledge.
Saldaña, J. (2013). *The Coding Manual for Qualitative Researchers* (Second). SAGE Publications Inc. https://us.sagepub.com/en-us/nam/the-coding-manual-for-qualitative-researchers/book243616
Seery, N., Buckley, J., Delahunty, T., & Canty, D. (2019). Integrating learners into the assessment process using adaptive comparative judgement with an ipsative approach to identifying competence based gains relative to student ability levels. *International Journal of Technology and Design Education*, *29*(4), 701–715. https://doi.org/10.1007/s10798-018-9468-x
Strimel, G. J., Bartholomew, S., Purzer, S., Zhang, L., & Ruesch, E. Y. (2021). Informing engineering design through adaptive comparative judgment. *European Journal of Engineering Education*, *46*(2), 227–246. https://doi.org/10.1080/03043797.2020.1718614

Tam, A. C. F. (2021). Undergraduate students' perceptions of and responses to exemplar-based dialogic feedback. *Assessment & Evaluation in Higher Education*, *46*(2), 269–285. https://doi.org/10.1080/02602938.2020.1772957

Vygotsky, L. S. (1987). *The Collected Works of L.S. Vygotsky: The Fundamentals of Defectology*. Springer Science & Business Media.

Yucel, R., Bird, F. L., Young, J., & Blanksby, T. (2014). The road to self-assessment: Exemplar marking before peer review develops first-year students' capacity to judge the quality of a scientific report. *Assessment & Evaluation in Higher Education*, *39*(8), 971–986. https://doi.org/10.1080/02602938.2014.880400

## About the Authors

**Scott Thorne** (sthorne@purdue.edu) is a doctoral student in the department of Technology Leadership and Innovation at Purdue University, West Lafayette, IN. Correspondence concerning this manuscript should be addressed to Scott Thorne via email. ORCID# 0000-0002-3742-8903

**Nathan Mentzer** (nmentzer@purdue.edu), Ph.D., is a professor of Engineering Technology Teacher Education in the department of Technology Leadership & Innovation and Curriculum & Instruction at Purdue University, West Lafayette, IN. ORCID# 0000-0001-9788-7901

**Greg J. Strimel** (gstrimel@purdue.edu), Ph.D., is an associate professor of Engineering Technology Teacher Education in the department of Technology Leadership & Innovation at Purdue University, West Lafayette, IN. ORCID# 0000-0002-4847-4526

**Scott Bartholomew** (scottbartholomew@byu.edu), Ph.D., is an assistant professor of Technology and Engineering Studies at Brigham Young University, Provo, UT. ORCID# 0000-0002-1680-2433

**Jason Ware** (jaware@purdue.edu), Ph.D., is a clinical associate professor in the John Martinson Honors College at Purdue University, West Lafayette, IN. ORCID# 0000-0003-4615-3451