Investigating student approaches to scenario-based assessments of systems thinking

Matthew B. Norris*, Jacob R. Grohs and David B. Knight

Department of Engineering Education, Virginia Tech, Blacksburg, VA, United States

The development of systems thinking is considered a critical skill set for addressing interdisciplinary problems. This skill set is particularly important in the field of engineering, where engineers are often tasked with solving socio-technical problems that often require knowledge beyond their original discipline and practice in unfamiliar contexts. However, existing assessments often fail to accurately measure teachable knowledge or skills that constitute systems thinking. To investigate this issue, we compared students' performance on two previously and independently peer-reviewed scenario-based assessments for systems thinking: The Village of Abeesee and the Lake Urmia Vignette. Twenty undergraduate engineering students participated in a multi-phase case study utilizing think aloud protocols and semi-structured interview methods to elicit the approaches students took thinking across the two instruments and past experiences that they felt prepared them to solve these ill-structured problems. We found that the way a scenario is presented to students impacts their subsequent problem-solving approach, which complicates assessment of systems thinking. Additionally, students identified only limited opportunities for the development of ill-structured problem-solving skills necessary for systems thinking. Our findings inform future work on improving systems thinking assessments and emphasize the importance of more intentionally supplying opportunities for students to practice solving ill-structured problems throughout the curriculum.

KEYWORDS

systems thinking, problem-solving, assessment, think-aloud interviews, engineering education, ill-structured problems

Introduction

An important yet often vague aspiration of higher education involves cultivating future leaders for a complex and unknown tomorrow. As problems continue to grow in complexity and connectivity, individuals in search of innovative ideas and solutions are being asked to exhibit systems-thinking capabilities, characterized by an ability to think and work across perspectives, contexts, and disciplines (ACED, 2019; ABET, 2021). Because of these complex challenges, governmental bodies, industry, and funding agencies have all pressed colleges and universities to graduate students who are capable of tackling broad problems and who can integrate and connect ideas and perspectives across multiple disciplines and content areas to discover and invent new solutions (e.g., Burrelli, 2010). Indeed, the National Academies of

Sciences, Engineering, and Medicine (2018), National Institutes of Health (2006), and National Research Council (2012) all have identified interdisciplinary issues as among the most pressing for society, and postsecondary institutions are being asked to promote a generation of "systems-thinkers" who can lead the search for solutions. A challenge for colleges and universities is to overcome discipline-based organizational structures (Warburton, 2003) that hinder student development of systems-thinking and interdisciplinary problem solving skills (Warburton, 2003; Svanström et al., 2008).

Fundamental to this challenge is that the assessment of systems thinking is difficult due to competing definitions and constructions that often vary across disciplines. Additionally, these assessments vary in usability and scalability for large-scale deployment to colleges and universities in a standardized manner. A recent systematic literature review by Dugan et al. (2022) mapped out existing tools for the assessment of systems thinking and their associated definitions of systems thinking. Many of these assessments (19/27) were identified as behavior-based, involving the performance of a specific task or the completion of ill-structured questions. These behavior-based assessments offer diverse content domains to situate ill-structured problems for evaluating systems thinking across interdisciplinary contexts. However, student performance on these assessments is often low, and interdisciplinary skills are often relegated to the peripherals of curriculum (Palmer et al., 2011; Lattuca et al., 2017). For this reason, a better understanding of how students are approaching these open-ended assessments and what portions of existing curriculum prepare them for systems thinking is needed. This improved understanding will allow for more targeted assessments of student learning and identify potential inadequacies in existing curriculum.

The overarching goal of this study is to *investigate student approaches to scenario-based assessments of systems thinking for the future improvement of teaching systems thinking and integration into engineering curriculum*. To achieve this the following sub-questions were developed to help investigate this broader goal:

Q1: How do students approach systems thinking tasks in open-ended scenario-based assessments?

Q2: How do students describe differences and similarities in their approaches to two open-ended scenario-based assessments of systems thinking?

Q3: Where in their educational experiences do students say they encounter ill-structured problems like those presented in open-ended scenario-based assessments?

Literature review

Systems thinking

"Systems thinking" is defined in numerous and occasionally contradictory ways. Senge et al. (1994) defined it as "a way of thinking about, and a language for describing and understanding,

the forces and interrelationships that shape the behavior of systems" (p. 6). Some perspectives of systems thinking describe the world as a series of interconnected physical and social systems that one can engage with to predict and direct outcomes (Meadows, 2008). These systems thinking perspectives are often called 'hard' systems methods and are associated with stock and flow diagrams, modeling, and feedback loops. In contrast, Checkland (1999) defines systems thinking as "applied to the process of our dealing with this world" (p. A10) rather than describing the world itself. This 'soft' systems methodology poses the world as complex and mysterious, and instead defines systems thinking as a systematic approach one can take to increase understanding. However, as emerging problems increase in complexity the lines between technical and social systems have blurred. The interdependencies between social, cultural, economic, and political contexts and potential technical solutions increase and further complicate problems.

Another popular method of conceptualizing systems thinking emphasizes solving problems that call for a flexible method of framing and understanding, acting across interdependent social and technical dimensions. This perspective of systems thinking emphasizes the analysis of an individual's reasoning through a posed dilemma, over whether their posed solution is 'right' or 'wrong'. The process by which an individual approaches an open-ended problem is more important to determining their expertise in systems thinking than their ability to produce a successful solution to the problem.

Scenarios and definitions of systems thinking

The assessment of complex skills like systems thinking can be attempted through a variety of methods. Self-report instruments have been widely used to assess constructs, but are often criticized for their ability to predict respondent behaviors or student performance (Peng et al., 1997). Alternatively, simulation-based assessments and capstone experiences can provide a more realistic environment and reliable results, however these often suffer from poor scalability. Scenario-based assessments offer a middle ground between self-report instruments and long-scale project simulations for assessing systems thinking (Mazzurco and Daniel, 2020). The key characteristic of scenario-based assessments is that they present a realistic problem followed by a series of open- or close-ended questions designed to elicit responses addressing particular knowledge ro skills.

A number of scenario-based assessments exist for measuring student knowledge and skills, including design skills (e.g., Atman et al., 2007) and knowledge of contextual contexts (McKenna et al., 2016). In this paper, the assessments we employ specifically measure systems thinking, though they conceptualize it through partially different constructs. In order to meaningfully discuss and compare students' approaches to these instruments the definition

of systems thinking underlying each instrument must be discussed.

In the first assessment, Grohs et al. (2018) describe a three dimensional framework for general usage in interdisciplinary contexts (Figure 1). The *problem dimension* considers technical elements broadly and the contextual environment that a given problem is situated within, including assumptions, prioritizations, and constraints. The *perspective dimension* considers the roles of diverse stakeholders in understanding and solving ill-structured problems. The *time dimension* considers the history of the given problem, stakeholders, and past attempted solutions and how these factors influence possible future solutions. This framework was developed from existing literature of systems thinking as well as problem-solving literature in engineering, the Dimensions of Systems Thinking Framework emphasizes the interconnectedness of technical and social aspects of modern problems and the importance of stakeholder perspectives.

The second assessment, the Lake Urmia Vignette, was framed using a definition of systems thinking grounded in system dynamics (Davis et al., 2020). From this definition, systems thinking is about examining physical and social systems as a whole and identifying connections between different parts of the system (Meadows, 2008). The way these system parts interact with one another over time to form feedback loops and flows is fundamental to several schools of thought on systems thinking (Sterman, 2010). The Lake Urmia Vignette is structured around students' ability to identify unique variables and the relationships between them, emphasizing the importance of feedback loops and more complicated webs of interconnections between variables. Davis et al.'s (2020) assessment associates more complex mental models with greater systems thinking ability.

Given the differences in foundational definitions of systems thinking and associated frameworks these scenario-based assessments differ across many constructs. However, both assessments attend to the problem definition phase of problem-solving. Grohs et al. (2018) explicitly define the problem definition phase and stakeholder perspectives as integral to systems thinking. Similarly, Davis et al. (2020) implicitly assess students' ability to define the problem through the identification of variables and agents. While the scoring scheme of Davis et al.'s assessment focuses primarily on the interconnectedness of different variables, both stakeholders and physical resources are taken into account. For the purposes of this study we will focus on how students approach the tasks of problem identification and the inclusion of contextual stakeholders.
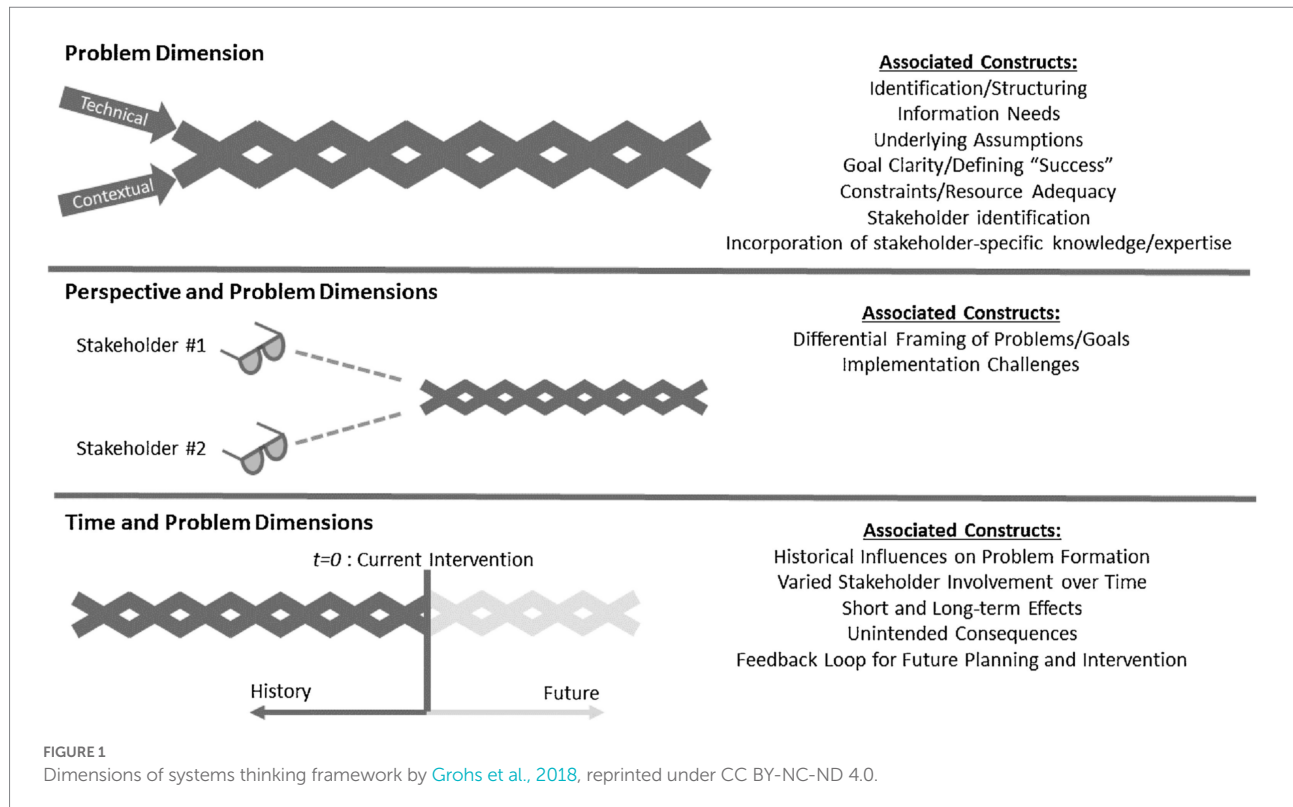
Problem solving and analogical reasoning

The idea of a 'problem' is core to our conceptualization of systems thinking and the assessment of students' systems-thinking ability. In problem-solving literature, it is the open-ended or close-ended nature of how the problem is framed that describes the problem's structuredness. Well-structured (or close-ended) problems are characterized by clear problem elements and defined goals leading to convergent absolute solutions through repeatable known algorithmic methods (Jonassen, 1997; Shin et al., 2003). In contrast, ill-structured problems are characterized by obscured or emergent problem elements, multiple or nonexistent solutions, no clear repeatable patterns for addressing the problem, and/or require interpretational judgments often based on personal subjective factors (Meacham and Emont, 1989; Shin et al., 2003). These ill-structured problems are the ideal use case for interdisciplinary skills like systems thinking. The scenarios selected for this study are open-ended and the problems contained within are ill-structured. The Village of Abeesee poses intentionally vague prompts after a brief vignette and does not value a specific problem approach over others (Grohs et al., 2018). The Lake Urmia Vignette presents a socially and politically complicated problem with decades of history that influences possible technical solutions (Davis et al., 2020).

Effective strategies for solving ill-structured problems are fundamentally different from the algorithmic methods used to solve well-structured problems. Solving ill-structured problems requires an iterative process of reflective design and argumentation in constructing the problem space (Schön., 1990; Jonassen, 1997), and problem solver success is related to epistemological beliefs (Jonassen, 2010) and advanced use of metacognitive strategies (Shin et al., 2003). However, the identification and transfer of problem schema—the underlying structure of the problem—can aid solvers in approaching seemingly intractable problems. One method of conceptualizing this transfer of problem schema is through analogical reasoning.

Analogical reasoning describes cognitive processes that conceptualize novel problems as modified instances of similar, previously-solved problems. By drawing such an analogy, the same schemas used to solve the analogous problem can then be applied in the novel solution. This sort of reasoning in-part explains differences between novices and experts. In "The Role of Attention in Cognition," Herbert Simon (1986) discusses "...when experts look at a problem situation in their domain of expertness, they immediately recognize familiar features in the situation, and these turn out to be the principal relevant features for correct handling of the situation" (p. 111). However, especially in the context of ill-structured problems, analogical reasoning must be multi-dimensional. Jonassen claims failed attempts to solve problems are often over-reliant on surface-level characteristics of a seemingly-similar problem causing an inaccurate analogy (Jonassen, 2010). He argues that expert problem-solvers focus on deeper-level structural similarities in order to draw upon multiple similar problems as possible analogues. Yet, for some problems, analogical reasoning alone is insufficient or even a hindrance as seen in some functional and design fixedness studies requiring creative thinking (Duncker, 1945; Adamson, 1952; Jansson and Smith, 1991; Chrysikou and Weisberg, 2005).

FIGURE 1
Dimensions of systems thinking framework by Grohs et al., 2018, reprinted under CC BY-NC-ND 4.0.

Materials and methods

This multi-phase qualitative study involved the completion of each scenario followed by an accompanying interview. The order students completed the two scenarios was staggered so that half would complete one scenario before the other. The first scenario was followed by a Think-aloud (TA) interview focusing on their approach to that scenario. One week after the first scenario participants completed the remaining scenario and metacognitive interview comparing the assessments. This study design is summarized in (Figure 2). Written responses to the systems thinking assessments were recorded using Qualtrics, and interviews were conducted and recorded over Zoom. A more detailed discussion of the assessment instruments and interview protocols is discussed in the following sections.

Systems thinking assessments

As discussed in the literature review above, two scenario-based assessments of systems thinking were selected to help investigate students' approaches. These instruments were administered through the online survey software Qualtrics and presented students with a scenario prompt and several questions corresponding to different systems thinking constructs. Details of each scenario are discussed in the following subsections:

Village of Abeesee: Problem identification and stakeholder involvement

The Village of Abeesee assessment is framed in a community setting, and is designed to be accessible to diverse populations of students by avoiding the necessity of domain-specific knowledge in the reasoning process. The scenario consists of three distinct phases: processing, where students are tasked with problem definition tasks, responses, where students are tasked with reasoning through a potential solution to their defined problem, and critique, where students are tasked with evaluating both an example plan and their own response (Grohs et al., 2018). The scenario prompt provided to students is as follows:

The Village of Abeesee has about 50,000 people. Its harsh winters and remote location make heating a living space very expensive. The rising price of fossil fuels has been reflected in the heating expenses of Abeesee residents. In fact, many residents are unable to afford heat for the entire winter (5 months). A University of Abeesee study shows that 38% of village residents have gone without heat for at least 30 winter days in the last 24 months. Last year, 27 Abeesee deaths were attributed to unheated homes. Most died from hypothermia/exposure (21), and the remainder died in fires or from carbon monoxide poisoning that resulted from improper use of alternative heat sources (e.g., burning trash in an unventilated space). (Grohs et al., 2018).

Student performance on the scenario is assessed using a detailed rubric across seven different constructs that support systems thinking. These constructs are then independently scored 0–3, with the score of 0 corresponds to no response or irrelevant responses,

FIGURE 2
Scenario and interview administration plan.
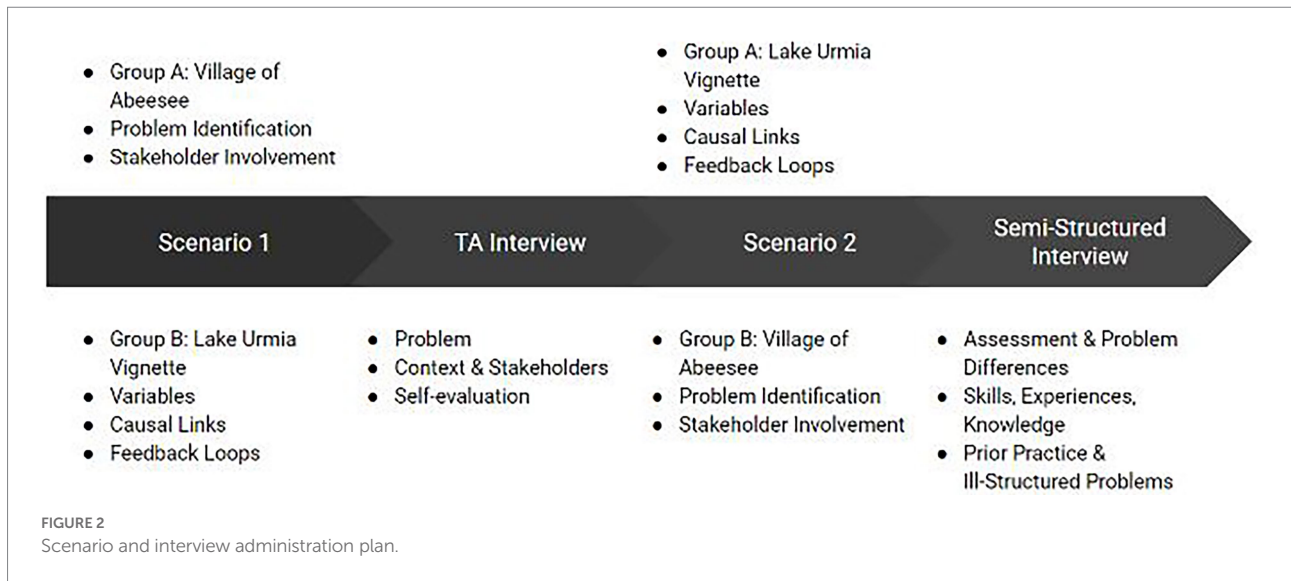
and the score of 3 is characterized by the ideal qualities of the response. Of the seven constructs involved in the Village of Abeesee assessment, in this study interviews primarily focused on the constructs of *problem identification* and *stakeholder involvement*. The scoring process is completed by a human rater using detailed examples alongside the provided scoring rubric. The full assessment instrument can be found in Supplementary materials.

Lake Urmia vignette: Variables, causal links, and feedback loops

The Lake Urmia Vignette assessment was designed as an assessment tool for systems thinking in a socio-environmental context. The scenario consists of several paragraphs describing the real-world shrinkage of Lake Urmia in north-west Iran and details a number of the reported environmental impacts and social unrest surrounding the lake's ecological problems (Davis et al., 2020). Participants read the text and then describe the source of the problem and important stakeholders in their own words. A portion of the vignette is included below:

Several reports show that the lake suffers from serious ecological problems, and many of the indicators are easily observable from the lake itself. Bexperts look at a problem situation etween 1972 and 2014, the area of the lake shrank by 88%. The evaporation of the water has exposed the lakebed and caused windblown salt, which may lead to environmental health crises, including increase in infant mortality, cancer, and liver, kidney, and respiratory diseases. This phenomenon is similar to what happened after the death of the Aral Sea. In addition, it will increase unemployment by reducing tourism and shrinking the fertility of the land in the region. Fortunately, the public awareness about the lake has increased… (truncated for clarity; Davis et al., 2020).

Student performance on the scenario is assessed by transforming student text into word and arrow diagrams consistent with systems dynamics modeling methods. Responses are scored on three items: the number of main variables identified,

the number of causal relationships between variables, and the number of feedback loops explicitly or implicitly mentioned. These items are assigned point values according to a provided scoring rubric. The responses are scored by two trained raters, and student scores are capped at 10 points per item for a maximum score of 30. Direct scores are then compared as a measure of a student's systems thinking ability. The full assessment instrument can be found in Supplementary materials.

Study participants

The participants for this study consisted of 20 undergraduate students in mechanical and civil engineering. These disciplines were expected to have relevant background knowledge that would potentially help students address the scenarios presented in the assessment instruments. While the fundamental knowledge base and coursework for both majors overlaps, the civil engineering program has an increased emphasis on involving stakeholders in the engineering process. In accordance with approved protocol for human subjects research students were recruited through advertisement in their respective departments' weekly informational email and compensated $50 for their participation after the completion of all study parts.

The 20 participants were evenly split between mechanical and civil engineering, with 7 s-year, 7 third-year, and 6 fourth-year students. First-year engineering students were ineligible to participate in the study as they do not become associated with a specific engineering discipline until their second year.

Think aloud protocol

Think-aloud (TA) protocol analysis is a research method originating in psychology that has been used to understand

participants' cognitive processes and discover how participants interact with knowledge (Krahmer and Ummelen, 2004). Originally built on the information processing model of human cognition, Ericsson and Simon's (1980) model assumes that participants transform knowledge from input states to output information. However, TA protocols can also be used for accessing participant-generated knowledge and beliefs as part of completing an ill-structured task (Koro-Ljungberg et al., 2013). Examples of prior applications of TA protocols in engineering education include assessing student design processes (Atman and Bursic, 1998), determining students' level of reflective practice (Adams et al., 2003), and the development of self-directed learning skills (Marra et al., 2022).

Ericsson and Simon (1980) originally described two types of verbal protocols: concurrent verbalization, where participants articulate their thoughts while simultaneously completing a specific task, and retrospective verbalization, where a participant describes cognitive processes that occurred at a prior point in time. Previous work by our team indicated that participants feel overloaded when attempting concurrent think alouds with these scenarios and have led us to choose retrospective think aloud protocols in this and future research. The retrospective TA protocols in this study were conducted immediately following the scenarios to improve the reliability of students' memory recall (Charters, 2003).

The TA protocol for both assessment instruments consisted of three broad questions targeting students' thought processes during the scenario. These questions were tailored to each scenario and worded to elicit detailed recollections of what students were thinking while answering scenario questions. The first prompting question focused on students' understanding of the problem, their reasoning, and the source of their conclusions. The second prompting question emphasized how participants described the potential stakeholders and broader contextual aspects, and their reasoning behind the inclusion or exclusion of specific stakeholder groups. The third prompting question asked participants to engage in a self-evaluation of their understanding of the scenario and reflect on why they made certain decisions not covered in the first two questions. This third question aimed at engaging participants in metacognitive processes regarding their approaches to the assessment scenarios.

This retrospective TA protocol also allowed for active engagement from the interviewer to encourage more detailed responses through acknowledgements (e.g., "mm-hmm," "go on") and prompt participants to clarify on comments (e.g.: participant: "I thought that was important," interviewer: "why so?"; Boren and Ramey, 2000). This allowed the interviewer to prompt participants to elaborate on short comments and their thought processes. The full details for the TA interview protocols can be found in Supplementary materials.

Semi-structured interview

The second interview following participants' final scenario assessment followed a semi-structured interview model. This model of interviews allows for a more natural flow of conversation between participant and researcher. Semi-structured interviews are also ideal when asking "probing, open-ended questions and [we] want to know the independent thoughts of each individual" (Adams, 2015, p. 494).

These interviews consisted of several parts, each with a different emphasis. The interviews began with a brief recollection of the first instrument from the prior session and associated TA interview. Then came an overall discussion on the differences between the two scenarios and assessment processes, beginning with an explicit question comparing the two scenarios (e.g., How would you describe the difference between the "Lake Urmia" and "Village of ABC" scenarios?).

Following this comparative component, participants discussed a more detailed evaluation of what skills, experiences, or knowledge they found necessary to complete each assessment, and where they gained those skill sets. Questions included a detailed breakdown of the needed skills for the problems presented in each instrument, as well as a metacognitive evaluation of the skills needed to complete the assessments themselves (e.g., What types of skills or abilities were needed to address the 'Village of ABC' scenario?, What about the 'Lake Urmia' scenario?).

The final portion of each interview discussed the comparative difficulty students had with each scenario and where they had attempted to solve ill-structured problems prior to completing the scenarios. These included personal and academic experiences, with the interviewer specifically prompting participants to consider areas not initially mentioned by the student. Throughout this process the interviewer asked probing questions (e.g., "Can you tell me more about what you mean by that?") when responses to a particular question were vague or failed to address the question.

Coding and analysis

Following data collection interviews were transcribed and collected across participants. We then conducted a first cycle holistic coding pass of both interviews following standard analysis practices (Miles et al., 2014). The goal of this coding pass was to familiarize the research team with students' responses and gain a sense of possible categories. Given the varied length and detail of participant responses, the flexibility of a holistic coding pass allows for more variability in the size of the text assigned a code. Using these developed codes and the initial research questions, a structure for analytical memos was developed to organize the analysis of interview transcripts. This structured memo targeted the initial research questions and attempted to categorize student responses accordingly. An example of this structured memo can be found in Supplementary materials.

From these analytical memos a second cycle of pattern coding, or themeing, was conducted to condense the first-cycle codes and structured memos into higher level constructs (Saldaña, 2015). These codes represented the types of approaches students took in

each scenario assessment and the ways in which they discussed differences between the scenarios.

While the assignment of codes to interview transcripts was completed by a single member of the research team, the process of code development and theming involved the entire team in a process of collaborative development and peer auditing. Due to the shorter length of some interviews and the vague nature of those students' responses, some interviews were characterized as having no discernable approach with sufficient evidence for second-cycle coding.

Limitations

One limitation of this study arises from the differing definitions of systems thinking across assessment scenarios. The Village of Abeesee scenario was developed to assess systems thinking as "a metacognitive strategy for flexibility and iteratively considering problems" (Grohs et al., 2018, p. 113). In contrast, the Lake Urmia Vignette defines systems thinking as the ability to understand the complexity of complex systems and recognize the interconnections between system variables. These differing perspectives of systems thinking, and the lack of a definitive definition of systems thinking, make more detailed comparisons across instrument constructions difficult.

Another limitation of this study lies in the data collection methods. Students' written responses and interview explanations varied in detail and complexity across both instruments. This discrepancy in between the written responses and interview discussion could be attributed to social-desirability bias (Krumpal, 2013), with participants expanding on their initial written answers to be viewed more favorably by the researcher conducting the interviews. Despite combining written and verbal responses and problem reasoning through think aloud protocols we cannot ensure that students' approaches to systems thinking have been captured completely. The approaches students used for these scenarios may not reflect their approaches while encountering authentic ill-structured problems in practice. The format of the scenario-based assessments limits the authenticity and complexity of the presented problem, potentially limiting the transferability of our results to professional practice.

A final limitation of these scenarios in measuring systems thinking is that they are designed to assess individual students' reasoning through a single pass of the problem. Authentic systems thinking practice in engineering takes place in team settings *via* iterative design processes, an aspect of systems thinking that most current instruments fail to assess (Dugan et al., 2022). In attempting to decrease time and resource investment into the assessment process these scenario-based assessments limit analysis to a single iteration of the problem solution process and require students to interact only with the scenario or provided materials. Given systems thinking's emphasis on bigger-picture

holistic evaluations of problems, this limits the applicability of these scenario assessments in terms of disciplinary scope.

Results

Student scores on both the Lake Urmia Vignette and Village of Abeesee scenarios closely resemble results from prior separate administrations. Counter to initial expectations, we observed no differences across class standing for this cross-sectional data collection. Given the sample size, we do not intend to make strong claims about this observation, but consistent findings with prior administrations does raise questions for further research. Additionally, we expected civil engineering students to perform better than mechanical engineering students on stakeholder awareness constructs given civil engineering's increased emphasis on public engagement. In actuality students from both majors performed equivalently on stakeholder awareness constructs, raising further questions for future investigation. A more detailed summary table of instrument scores can be found in Appendix A. However, our research questions are more appropriately addressed through the analysis of the follow-on interviews.

Approaches to problem solving

Thematic analysis of the retrospective think aloud interview data produced five initial codes defining approaches to the tasks of problem identification and awareness of stakeholder considerations. Consolidation of codes generated four distinct themes characterizing the underlying structure of students' approaches. These problem solving themes were developed to answer *Q1: How do students approach systems thinking tasks in complex open-ended scenario-based assessments?* (Table 1) provides a brief overview of these four themes.

Analysis of students' thought processes during the initial retrospective think aloud interview demonstrated different approaches to problem solving. These coded approaches occurred across both *problem identification* and *stakeholder considerations* portions of each assessment. Given that only a small number of students exhibited the *No Structure* approach in their interviews and the inherent variance of responses with no consistent structures this approach will not be discussed in detail. While the remaining approaches follow common themes, there are some variations within each that deserve further description.

Reductionist approach

The first theme, the Reductionist approach to the scenarios, where students would fixate on a specific aspect of the problem and develop a solution focusing only on that identified aspect, occurred across both scenarios as well as problem identification and stakeholder consideration components of each scenario. This approach was the most commonly used in TA responses about the Village of Abeesee scenario. The method students used to

TABLE 1 Summary of problem solving themes.

Approach	Summary	Example
Reductionist	Characterized by simplifying the scenario, explicitly focusing on a single dimension, or building a detailed response on a small subset of the provided problem.	"So it was the problem that I felt was most trying to be solved. There was a lot of information there where you could kind of solve some of the problems, but I was trying to figure out what most were we trying to reduce or improve..The root problem I felt was that people were dying." (ME9, Abeesee TA)
Assessment Prompt Driven	Characterized by using the written problem prompt to guide students' approach without the use of external information or making connections not present in the text.	"I think it's a combination of both, but I think that the main issue is human, mostly due to, given the article. The article listed a natural cause once, which was the drought, which I believe does have an effect. However, just reading the rest of the article, the population expansion, the starting of the pipe upstream…" (ME6, Lake Urmia TA)
Prior Experience	Characterized by the use of a prior experience or applicable knowledge to direct the process of identifying the problem or relevant stakeholders.	"First I had flashbacks for all my engineering classes in engineering, like general engineering, stakeholders [inaudible]. And, just tried to think, "Well, who is actually affected?" I mean, obviously it's the people who cannot heat their homes, but, okay, what else plays into that?" (ME1, Abeesee TA)
No Structure	Characterized by the failure to articulate any particular structure to students' thought processes. Non-systematic or erratic responses that jump between problem dimensions without clear connections.	"I do not know. There wasn't any stuff out there. Just of how I thought I would try to solve it, and the things I knew I needed to get to those points." (CE9, Abeesee TA)

conceptually structure their understanding of the problem varied, but the fundamental simplification of the problem remained present. For example, ME5 identified carbon monoxide poisoning as the primary problem present in the scenario prompt:

> So, this is one of the concerning issues that should be addressed because a lot of people got killed because of hypothermia and carbon monoxide poisoning. So we should definitely address that problem. (ME5, Abeesee TA).

The student then continues to discuss the ramifications of carbon monoxide and the complexities raised by these unsafe heating and burning methods, building their response to the problem on this single aspect of the presented problem:

> So I thought that since people were not able to pay for the heating they were cutting woods out in the forest and they were bringing it home, they were burning it to keep their house warm, which produces carbon monoxide because when we burn the word, it is not very effective. It does not produce completely carbon dioxide. Most of it is carbon monoxide, as well people said that if you are burning some kind of food in the house and you should open your windows, so that carbon monoxide escapes. But since the weather is very brutal outside, most of the people would not open their windows because they want the heat to be there within the house. And if they open the windows the heat is going to get out. So they are out of options. (ME5, Abeesee TA).

Participant ME5 fixates on a problem understanding centered on carbon monoxide poisoning and the current unsafe heating situation presented in the assessment scenario. Other responses

following this approach followed similar patterns and expanded on limited aspects of the problem in great detail, but left large segments of the problem context unaddressed.

The other common *Reductionist* approach was characterized by participants discussing the problem from the perspective of being present in Abeesee and personally tasked with solving the problem they identified. This approach was characterized by students focusing on immediate concerns and prioritizing the most obvious aspects of the problem instead of addressing broader elements of the problem in a more systematic process. CE5 discussed identifying the problem as imagining themselves present in the scenario, asking questions to try and identify a possible source of the problem:

> Yeah. So being in the village, I guess, taking a physical survey of the area, maybe you could harness some thermal energy from the earth, or I guess you'd have to take a ground survey for that. And also studying the soil to see what that's like, if there's a permafrost layer or anything that might prevent heating pipes underground or something like that. And then also just maybe looking at the cover, if there's a lot of trees nearby that might be preventing sunlight and natural heat from getting in. That could be useful to identify. And then just physically looking at the homes and the different buildings there, just to get a general sense of their material of construction. (CE5, Abeesee TA).

This demonstrates a more mature approach than artificially focusing on specific aspects of the scenario, but is still motivated by the perceived urgency of specific aspects of the problem. This caused students to fixate on what they perceived to be the most immediate or potentially harmful portions of the scenario to the

exclusion of broader contexts. When discussing stakeholders, students discussed imagining themselves in the assessment scenarios and attempting to determine who would be most immediately impacted. ME10 talked about how they felt that imagining themselves in the scenario would make it easier and more impactful to think through who should be included:

> I think when people see wildlife statistics and a bunch of statistics and numbers, it may not connect with them as much. But when there's people in the street and there's actually action being done and that's still not solving the problem, I think that's more of a tangible, visible problem that people can relate to more and it should have led to more action than it did. (ME10, Lake Urmia TA)

In these cases, the 'reduction' occurred from students limiting their consideration of potential stakeholders to those that would be immediately and obviously impacted by the problem. This prevented these students from discussing more high-powered or indirectly impacted stakeholders that could be engaged to help address the problem. Both of the *Reductionist* types of responses are characterized by limiting the scope of the presented situation to a more directly manageable solution. This practice is in direct opposition to the desired holistic examination of a problem that systems thinking principles attempt to achieve.

Assessment prompt driven approach

The next most common approach, *Assessment Prompt Driven*, is characterized by explicit reference to the information provided in the scenario prompts as the source of students' understanding of the problem. Reference to the text alone was not sufficient to categorize a response as being driven by the prompt, students needed to articulate the prompt as the origin of what they considered to be the problem in the scenario. While this approach was demonstrated across both scenarios, it was more prevalent for the Lake Urmia Scenario, which contains a more detailed prompt. When asked what approach they took to identify the problem in the Lake Urmia Vignette CE8 responded with:

> Right **when I first started reading**, I thought the problem was that the lake had decreased in size. It had dried up. So in the very first thing that really stood out to me was the decrease in certain organisms that had been there in the past. And then the second thing that popped out right away, **I think it was in the first or second paragraph**, was that because the lake had dried up there was salt blowing in the region (CE8, Lake Urmia TA, emphasis added).

They begin by referencing the act of reading the provided prompt, and then go on to discuss specific paragraphs from memory. Students taking this approach often referenced the provided text multiple times and mitigated their statements about the problem. While students would often cite the prompt multiple

times during their interviews, their discussions tended to center around only a subset of the information provided.

Responses to the Lake Urmia Vignette also saw some students use *Assessment Prompt Driven* approaches specifically while discussing stakeholder engagement. These students tend to reference the local community and government protestors described in the prompt. Participant CE4 expands upon these, including protestors to the local community and describes what portion of the prompt text led them to include these stakeholders:

> And also, based on the paragraph with the protests and stuff going on and how people were dedicated to the cause, I figured you are going to want to get that group of people who are so devoted that they want to go and protest and want to try to say that, you want to include them. And also, just including people from the area because they tend to know a lot more about the general thing going on. Because I can be an engineer and read the description of what's going on and maybe look for solutions, but usually there are a lot more smaller, fine tuned details that go into it. And so I figured that the local community would be good. (CE4, Lake Urmia TA).

While CE4 does recognize the limitations of the scenario prompt's information on Lake Urmia's situation, it still served as the guide for what stakeholder groups they discussed. Earlier in their interview they also discuss the emotional impact on possible stakeholders, echoing the *Reductionist* approach's emphasis on the more directly impacted and disadvantaged stakeholders.

Prior experience approach

The last major approach to problem solving that students exhibited was *Prior Experience*, where students discussed specific experiences in their past they utilized to help understand the scenario problems. Several students expressed prior experience with including stakeholders in the problem solving process from academic and industry experiences.

In particular, participant ME2 references classes from a minor in Green Engineering as the source of their prior experience considering what stakeholders need to be involved, "My minor is green engineering. In all my classes for that, we always have to do case studies and the first thing they always ask us [about] is the stakeholders." (ME2, Lake Urmia TA). These students tended to have more complete and detailed discussions of potential stakeholders and methods to leverage those stakeholders to address the problem they defined. However, not all student experiences discussed directly relate to solving the problems within the scenarios. Some students followed an approach based on preconceptions about the problem:

> And then in addition, obviously I think, as in most issues, the government would be a stakeholder because the government is the one that is most likely going to be providing the funds in creating the plan. So I put government and then I put fishermen, which I was just

thinking about the lakes that were feeding into… (CE8, Lake urmia TA).

Participant CE8 makes assumptions about the source of funding in the Lake Urmia scenario based on prior experience with governments funding large-scale ecological projects. Several other students built their approaches on the assumption of certain external resources, political organization, or stakeholder motivations that were not provided within the content of the scenario prompts. These *Prior Experience* approaches are built on knowledge, assumptions, or experiences from outside the assessment scenarios and directly impacted the ways in which students attempted to answer the assessments.

Metacognitive comparisons

Relevant data to answer *Q2: How do students describe differences and similarities in their approaches to two open-ended scenario-based assessments of systems thinking?* was drawn from the metacognitive interviews after students had completed both scenarios. Analysis of students' explicit comparisons of the systems thinking assessments identified two primary types of comparisons across instruments, with a third subset of students making both types of comparisons. The most common comparison made by students consisted of discussion centered on the **domain-content** of the scenarios. The least detailed of these comparisons is limited to identifying similarities between the general knowledge domains of the problems presented in each assessment. Participant ME4 is one such who made this type of assessment when asked to compare the two scenarios and their purpose:

"At the core, the issues are slightly different. They're both environmental issues with the resources, but you'd be solving two different issues or two different problems, at least." (ME4, Meta Interview)

Here ME4 only compares the environmentally situated content of the assessments, focusing on the domain of the problems presented in the scenarios over the structure of either. The majority of these types of comparisons focused on the environmental or economic aspects of the problems described in the assessments. Some students interpreted the assessments as addressing two different problem domains, but their comparisons remain limited to the domain content of the problems presented.

A small number of students (4/20 participants) discussed the two assessments beyond the content provided in the prompts of each. These students recognized differences in the **underlying structure** of the two scenarios and discussed the similarities between each instruments' intended assessment. The following is an excerpt of participant CE8's explanation of the structure or the Lake Urmia and Village of Abeesee scenarios:

I think the lake situation seemed to have a bigger effect overall on so many different things. It was going to probably affect the economy because farmers weren't able to grow crops, and then it was also causing health problems and it was decreasing tourism. So it was having a big economic effect. Whereas this problem [Abesse] seems more condensed and centralized. (CE8, Meta Interview)

So I think both problems, in my opinion, I felt like I was asked to look at the problem holistically and not just solve one part of the problem, if that makes sense." (CE8, Meta Interview)

Participant CE8 then goes on to discuss both similarities in the problem solving process that was needed to address the presented problems as well as differences in how the problems are structured to provide information, of the Village of Abeesee's case, withhold it. Students who made these comparisons also tended to recognize and discuss the cognitive skills needed to complete both assessments.

Another third grouping of students make **both domain-content and structural comparisons**. These students spent the majority of their time discussing similarities in problem content, but briefly discussed aspects of the scenarios that point toward clearer understanding of the structural differences between scenarios. Participant ME2 begins their discussion of the assessments with:

They both were largely affecting the people that lived in the area and were putting people at risks to their health. They were doing that for different reasons, but, in both scenarios, you're looking at people that are reliant on an area and they can't either, well, in the Abeesee village, they couldn't afford to keep their homes. Whereas in the Lake Urmia, they were at risk of living, losing their livelihood due to the lake disappearing and both scenarios. There were ecological factors." (ME2, Meta Interview)

As seen in the majority of responses they begin by focussing primarily on the environmental and ecological setting of both problems. However, they then go on to discuss how they understand the structure of the tasks each assessment instrument asked them to complete:

The first scenario [Lake Urmia] was definitely asking. It was definitely a lot more of a web of issues. I think. You had the lake disappearing, and then you had the various things that they'd listed, all the details about things that had been done to the lake over the years … (ME2, Meta Interview)

This is a particularly insightful observation, as the Lake Urmia Vignette is scored based on the complexity of the 'web of issues' or causal map that students are able to articulate in their written response to the prompt. In addition to explicit

comparisons of the assessment instruments, students discussed what relevant skills, knowledge, and experiences they believed were necessary to address (1) the questions the assessment instruments were asking, and (2) the problem presented in each scenario. When discussing the questions asked by the assessment instruments the majority of students identified various cognitive skills, such as general problem solving, critical reasoning, and reading comprehension. Participant CE9 described both reading comprehension and general problem solving skills when discussing both instruments.

> "I need to be able to pull information out from the text that was important and be able to analyze it. Analyze and see what issues that really need to be solved. And then I needed to know the general steps of what I think would need to happen, the process of like a situation…" (CE9, Meta Interview)

While some students made distinctions between the skills needed to address one instrument or another, this was uncommon and multiple cognitive skills were identified for each scenario. When it came to identifying necessary skills, knowledge or experiences for addressing the problems presented in the assessment scenarios students mainly identified various domains of background knowledge as relevant. The specific domains of background knowledge varied with students' interpretation of the problem, though many discussed ecological or environmental knowledge as being important. Participant ME7 breaks down the different pieces of knowledge they felt would be necessary to address the different scenario problem is detail:

> I think it was just entirely critical thinking based and just having some type of background knowledge about things that could be potential solutions. (ME7, Meta Interview) Knowing the rough cost of a liquid natural gas storage tank and knowing more or less what those things would do and what they would solve and what their intended consequences would be. Whereas, something like build a dam. I've absolutely no idea how much that might cost. What that might do other than block water. But what does that mean for the environment? What does that mean for downstream? And so on and so forth. You need to have some kind of idea of both the very literal upfront effect of whatever piece of the puzzle you're referring to and also its consequences. (ME7, Meta Interview)

They then end their response by generalizing the importance of having specific technical knowledge to be able to provide well thought out solutions. Given that systems thinking is often situated as a socio-technical skill, the emphasis on both similar cognitive skills and background information helpful to solving the presented problems is promising for assessment.

When comparing how they approached both systems thinking assessments, ME2 described the Lake Urmia Vignette as, "a lot more just throwing stuff that I'd read down on paper." (ME2, Meta Interview). The detailed background provided by the vignette was

often referenced as the main feature of the assessment, and students' approaches would center around relating that information back as their understanding of the problem.

Source of relevant experiences

As part of the Metacognitive interview students were asked to discuss where they had previously solved ill-structured problems to help answer *Q3: Where in their educational experiences do students say they encounter ill-structured problems like those presented in open-ended scenario-based assessments?* The majority of students provided primarily academic experiences that they considered helpful or relevant to solving ill-structured problems. The academic experiences that are referenced as useful fell across students' academic careers from high school to undergraduate capstone, but with a notable exception of core technical classes. These core technical courses include fundamental engineering concepts covering content in materials science, thermodynamics, and engineering mechanics. These courses are primarily taken during students' second and third years as prerequisites for more discipline-specific engineering courses.

Many students reference high school environmental or earth science classes as providing them with both background knowledge as well as practical experience in managing more ill-structured problems. Participant ME10 in particular talked about how high school still features prominently when they think about ill-structured problems, "I would say some of the bigger ill-structured problems are projects from my high school. I was in an engineering specialty, so I think I mentioned it last time. And so we had a lot of projects." (ME10, Meta Interview).

When discussing experiences from their undergraduate career students often referenced introductory engineering courses, with those in their final year highlighting capstone. Participant CE4 discusses both their capstone project as well as general projects from their introductory course 4 years earlier:

> And we did have to perform due diligence studies to understand what was going on in the background for that project. And then the other project I'm doing for our senior design is a roadway design and it was just mitigate the problems of traffic on [Road] in [University Town]. And that was really all that was given. So it was a lot more of having to figure out what those problems were and where some of the major sources of traffic came from and then proposing new roadways in that regard. (CE4, Meta Interview).. And then I guess also in the foundations of engineering classes too, with those projects (CE4, Meta Interview)

While Capstone class experiences were often referenced by senior students as being more similar to the ill-structured problem that both systems thinking assessments attempt to emulate, ME1 compares the open-ended scenarios to a well defined problem:

> Something like engineering classes, even the senior design I'm doing now. It's similar. I know similar to what I was talking about last session and that, even though that [inaudible 00:12:17], the engineering one's a lot more hard set. Because it is an actual engineering thing. We have to solve some very distinct or defined goal…" (ME1, Meta Interview, emphasis added)

This demonstrates that the relevance of capstone experiences to addressing these systems thinking scenarios may be dependent on the structure of the capstone project. Additionally, these instruments were developed in part to serve as formative assessments for improving systems thinking education, so capstone projects being the considered relevant experiences for these instruments would be of limited use.

Students do discuss some core technical engineering courses, though these almost all come from classes specific to a minor in green engineering, and are therefore not courses that all students will be exposed to as part of the curriculum. The green engineering minor courses are discussed as relevant from both a domain knowledge and ill-structured problem solving perspective. Participant ME2's view of their major discipline and green engineering minor emphasizes the sentiment found across most students:

> Probably not in my mechanical engineering classes, for sure. Like I said, in my green minor classes, I've had projects that were they would give us a set of numbers and the first half of the project would be like, for example, last year in one of my intro, one of my green engineering classes, we looked at the project they were doing to convert to add additional ability for the [University] power plant to do more natural gas energy production, and the cost of switching. (ME2, Meta Interview, emphasis added).

They even go so far as to explicitly state that their mechanical engineering courses are not providing opportunities to solve ill-structured problems. Several other students failed to provide examples of relevant experiences from core technical courses even when explicitly prompted to do so. This exclusion of core technical engineering courses wasn't total, with a single student making multiple connections to different courses, but collectively, significantly more introductory and high school experiences were discussed than was expected from undergraduate students.

Discussion

Anchoring and satisficing problem solving approaches

Students' approaches to open-ended scenarios appears to be critical for their understanding of the problem. Their development of a problem schema—the knowledge structure developed to represent the problem—is an important part of

solving that problem. The structural and situational characteristics of the problem schema define the solver's understanding of the problem and guide their development of potential solutions (Blessing and Ross, 1996; Jonassen, 2010). In some problems the amount of information available may be too great for the student to attend to every aspect equally. In these instances problem solvers are forced to 'satisfice' their understanding of the problem. Introduced by Herbert Simon (Simon, 1956), individuals satisfice their decisions by either reducing their understanding of the problem until they can find an optimal decision or accept satisfactory decisions in more complex situations. For example, participant ME9 discussed only being able to satisfy a portion of the problems they identified due to information overload and their own limitations ("There was a lot of information there where you could kind of solve some of the problems, but I was trying to figure out what most were we trying to reduce or improve..," Abeesee TA). In the context of these scenarios students enacted bounded rationality by attempting to reduce the complexity of the problems before developing a solution. In the case of information rich scenarios students may engage in this process to limit what they have to attend to, impacting potential systems thinking assessments that value holistic evaluations of contextual problems.

In other cases, problem solvers anchor their approaches based on similarities to prior experiences as a method of understanding and solving the problem. Multiple relevant prior experiences provide students with the opportunity to analogically compare multiple schemas with their current problem (Catrambone and Holyoak, 1989). Students with relevant experiences from which to draw were able to identify how relevant stakeholders tie in to understanding the problem, whereas others only made shallower content-domain analogies. For example, student ME4 identified only the environmental context of the scenarios ("They're both environmental issues with the resources…," Meta Interview), whereas student CE8 identified that both scenarios were assessing their ability to address the broader problem contexts holistically ("So I think both problems, in my opinion, I felt like I was asked to look at the problem holistically and not just solve one part of the problem..," Meta Interview). Current literature suggests that having multiple experiences across a variety of domains can be helpful for improving students' ability solving ill-structured problems analogical comparison (Gentner et al., 2003). Participants were often unable to identify multiple experiences to draw upon while developing their solutions, suggesting that the importance of providing students with multiple domain contexts for learning has not extended to teaching practice.

In addition, students' approaches to these open-ended problems proved to be susceptible to the framing of the scenario, in some cases directly guiding students' responses. We observed this phenomenon more frequently for students' approaches to the more detailed Lake Urmia scenario, where explicitly provided information was frequently used to guide students' understanding

of the problem ("And then the second thing that popped out right away, I think it was in the first or second paragraph…," CE8, Lake Urmia TA). However, unlike well-structured assessments, the organization of information and application of domain knowledge alone are insufficient in describing or solving the problem. This anchoring and dependence on the scenario vignette highlights a potential barrier to assessments of systems thinking that value problem considerations outside those provided. However, for systems thinking perspectives that value the *identification* of provided variables and interconnections a more detailed problem framing within vignette prompt may be desirable–individuals choosing between available assessments should consider these tradeoffs based on their own goals.

Analogical reasoning and developing expertise

In considering the ways in which students initially approach problems, it is clear that it is not the approach itself that is right or wrong, but how the approach aligns with the demands of any given problem. Taking this idea further, Simon (1986) argues that "a large component of expert skill resides in the ability to attend, upon seeing a stimulus in the domain of the skill, to the relevant parts of the stimulus; and, through that attention and the resultant recognition, to get access in long-term memory to the information that is required or executing the skill at that point" (p. 112). As discussed in the literature review, a broader cognitive skill at work here is analogical reasoning and describes one's ability to reason through a novel situation or problem as an analog of another previous situation or problem. Jonassen (2010) argues a critical factor in making these analogies is that expert problem solvers rely on structural similarities rather than over-reliance on surface-level characteristics. As an example of the foundational work leading to such claims, a classical study in physics education noted that novice problem solvers focused on problem features (e.g., both problem figures involve circular discs or blocks on an inclined plane), whereas experts focused on the underlying similar physics principles (e.g., two problems that look different on the surface both can be solved by the Law of Conservation of Energy; Chi et al., 1989).

However, what constitutes "surface" compared to "structural" similarities in ill-structured problems such as the Abeesee and Lake Urmia vignettes is not so obvious as with physics problem figures or the underlying governing equations. When students in our study were asked to describe their thought processes and comparisons across the two problem scenarios, students focused on domain-level ("environmental" or "ecological") and/or structural-level similarities (e.g., interconnected "web of issues") between the two scenarios. At first glance, anchoring on domain-level similarities alone would be problematic since it is expected such problems require interdisciplinary solutions and reducing to one domain perspective or one aspect of the problem alone would not

engage appropriate complexity and interconnectedness needed in a proposed solution. However, even the structural similarities noted in our data seem relatively simple— recognizing the problem posed as an interconnected web of issues is a great start but does not really help unless the problem-solver can make sense of the interconnectedness in parseable ways.

Evidence of this higher order expertise and analogical reasoning was sparse in our data which could be a limitation of our data collection as well as somewhat expected given that systems thinking is a complex competency and undergraduate students would be expected to have only emergent knowledge and skills which will be further improved in professional practice. That said, there were some implicit examples such as when a participant described their rationale with respect to unintended consequences: "You need to have some kind of idea of both the very literal upfront effect of whatever piece of the puzzle you are referring to and also its consequences." Considering the need to more explicitly understand reasoning processes prompts the question of what are key "structural" features in these contexts that problem-solvers should attend to and how can they be used to approach new problems?

Summarizing their own work and that of colleagues in the area of analogy generation and expertise, Goldwater et al. (2021) highlight the importance of understanding relational connections among objects and events and the explanatory processes that describe these relationships. From this perspective, it is the deeper relational connections and explanatory processes that affirm why "structural" similarity over "surface" similarity is associated with expertise. That is, an expert can reason best through a novel, ill-structured problem when they have a wide range of possible underlying causal relationships and potential theories that can be mixed and matched to make sense of the new problem.

Thinking about the Abeesee and Lake Urmia instruments in the context of this literature, the established scoring of the instruments gives insight into what "structural" (underlying relationships and explanatory processes) elements of these interdisciplinary problems might be associated with expertise. For example, the highest score of the problem identification construct in the Abeesee rubric involves students not only identifying several technical and contextual components of the problem (e.g., "web of issues") but further articulating interrelations among them (e.g., "relational connections"). Or, for the Stakeholder Awareness construct, expertise is about identifying stakeholders across sectors and power differentials and recognizing that stakeholder perspective is needed in a regular engagement process. In this construct, experts would be recognizing that many interdisciplinary ill-structured problems require considering the problem-solving process as iterative, cross-sector, and collaborative (e.g., awareness of theories of collaboration, change, and/or organizations).

Our investigation supports the notion that the field should continue to more explicitly define and operationalize the

knowledge and skills that are part of broader rhetorical arguments for interdisciplinary systems thinkers. In particular, from the perspective of furthering analogical reasoning, more attention should be focused on how expert ability to consider underlying relationships and reasoning principles or theories from a range of domains might influence problem-solving ability in novel situations. Additionally, although much remains to be learned about the underlying cognitive science around analogical reasoning, research also suggests that problem-solvers need practice with contrasting case examples and consistent use of relational language within those examples to help facilitate abstraction of the underlying reasoning skills across domains (e.g., Jamrozik and Gentner, 2020).

Limited opportunities to explicitly learn and practice systems thinking

In the case of building expertise in ill-structured problem-solving, our findings suggest that understanding the problem features students attend to, and the resultant activation of relevant past problems, experience, or learned skills they use to reduce ill-structuredness into manageably structured problems, is particularly important. However, teaching such skills is not trivial. Much of the engineering curriculum our engineering student participants engage in are about complicated yet well-structured problem-solving where identifying appropriate equations and performing the necessary challenging mathematical calculations leads to singular correct answers. Our analyses highlighted a variety of academic experiences in which students described opportunities to develop their ill-structured problem solving skills. However, many of these experiences were limited to high school settings, at the bookends of the postsecondary curriculum–either during a cornerstone experience in their first year or a capstone experience in their final year–or within a set of courses that comprised a focused minor. If developing systems thinking and/or related ill-structured problem solving skills is viewed as one of the more important goals of undergraduate engineering programs, our findings pose the question of whether the current structure of the curriculum is designed optimally to support that goal. Moreover, because minors are only accessed by a fraction of enrolled students, our findings point to questions regarding equitable access to potentially pivotal educational experiences that can help position a student for continued future success in their career.

These results focused on systems thinking development within undergraduate engineering join findings in the literature that pose such questions regarding the equitable and optimal structure of the engineering curriculum. Indeed, several other prior studies of a variety of competencies and skill sets that are purportedly deemed incredibly valuable by industry (e.g., National Academy of Engineering, 2004, 2005) seem to only appear to be intentionally developed at a few touch points within the curriculum or out-sourced to the co-curriculum. For example,

Palmer et al. (2011) describe how students' opportunities to develop design thinking skills tend to be limited to certain experiences within the curriculum or to co-curricular opportunities that are not accessed broadly. Lattuca et al. (2017) drew a similar conclusion for the development of students' interdisciplinary competence, as did Knight and Novoselich (2017) for students' leadership skills. In Daly et al.'s (2014) analysis of teaching creativity within engineering, the exemplar courses in the curriculum tended to be limited to design-focused courses, and Colby and Sullivan (2008) noted how ethics education tends to be sporadic and unintentional throughout engineering students' curriculum. Shuman et al.'s (2005) broad discussion of development of engineering students' professional skills describes how such skills require pedagogies and assessments that traditionally have not been embedded throughout the engineering curriculum but had an optimistic view about progress on that front. Although we have certainly seen some fine examples of progress over the past few decades since that article, our findings in this analysis of systems thinking skills joins the conclusions of several other studies of related skills and competencies that there is much more work to be done with respect to large-scale curriculum change if we are to meet the goal of integrating equitable opportunities for skill and competency development throughout the curriculum.

Implications for educators

Our findings have several implications for the education of systems thinking in higher education, and engineering education in particular. The assessment of skills like systems thinking using scenario-based assessments offers a promising avenue of assessment, but greater attention must be given designing these assessments to target the aspects of systems thinking that educators hope to measure. Educators should prioritize creating more and varied opportunities for learning systems thinking across curriculum and ensure that these opportunities are explicitly identified to students as fundamental for developing ill-structured problem solving abilities. Educators can discuss similarities and differences between these opportunities to help students learn to identify deeper domain independent skills and knowledge that can be applied more easily to future problems.

Future work

Our results have several implications for the development of skills like systems thinking and our methods for assessing these skills. The framing of assessment can influence students' approach to the assessment alongside the domain content the instrument is assessing. The way these assessments are framed can help or hinder our evaluation of students' skills as they prompt different approaches to ill-structured problems. The development of future scenarios must therefore include

intentional design considerations for the presentation of the scenario alongside the scenario's content. For example, future work may present the same problem scenario in different ways (i.e., text, images, systems map, news briefing) and analyze how student approaches vary across scenario framing. The behavior-based scenarios in this study offer potential opportunities for reflective evaluation of student's systems thinking competency in a way that balances the difficulties of ill-structured problems with the need for timely feedback in instruction. However, further study into dynamic assessments including iterative problem solving or systems thinking in groups can address specific limitations of these scenarios. Having initiated investigation into student approaches to systems thinking assessments, our subsequent research will explore how students develop their systems thinking skills in academic and co-curricular environments.

Conclusion

The purpose of this study was to investigate student approaches to scenario-based assessments of systems thinking for improved teaching and curriculum integration. We found that students utilize the assessment provided to them alongside their own past experiences to satisfice and anchor their understanding of ill-structured problems. The development of scenario-based assessments of systems thinking must more intentionally elicit the constructs of systems thinking (i.e., variables, problem identification) that are being targeted. Students primarily compare these systems thinking scenarios using shallow domain-specific context, and fail to identify more structural differences that would indicate expertise with systems thinking. Our findings also contribute to a growing body of work identifying the need for valued skills and competencies to be systematically integrated into existing curriculum. These gaps in curriculum highlight the need for increased opportunities for students to engage in ill-structured open-ended problem solving in varied authentic environments.

Data availability statement

Requests to access raw, de-identified data can be made to the authors, who will work with the university institutional review board to grant access on a case-by-case basis.

Ethics statement

This study involving human participants was reviewed and approved by Virginia Tech's Institutional Review Board (18-942). The participants provided their oral informed consent to participate in this study.

Author contributions

Funding

Acknowledgments

Conflict of interest

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2022.1055403/full#supplementary-material

References

ABET (2021). *Criteria for Accrediting Engineering Programs 2022–2023*. Baltimore: Engineering Accreditation Commission.

ACED (2019). *Engineering Futures 2035: A Scoping Study*. Washington: Australian Council of Engineering Deans.

Adams, W. C. (2015). "Conducting semi-structured interviews," in *Handbook of Practical Program Evaluation*. eds. K. E. Newcomer, H. P. Hatry and J. S. Wholey (Hoboken: John Wiley & Sons, Inc.)

Adams, R. S., Turns, J., and Atman, C. J. (2003). Educating effective engineering designers: the role of reflective practice. *Des. Stud.* 24, 275–294. doi: 10.1016/S0142-694X(02)00056-X

Adamson, R. E. (1952). Functional fixedness as related to problem solving: a repetition of three experiments. *J. Exp. Psychol.* 44, 288–291. doi: 10.1037/h0062487

Atman, C. J., Adams, R. S., Cardella, M. E., Turns, J., Mosborg, S., and Saleem, J. (2007). Engineering design processes: a comparison of students and expert practitioners. *J. Eng. Educ.* 96, 359–379. doi: 10.1002/j.2168-9830.2007.tb00945.x

Atman, C. J., and Bursic, K. M. (1998). Verbal protocol analysis as a method to document engineering student design processes. *J. Eng. Educ.* 87, 121–132. doi: 10.1002/j.2168-9830.1998.tb00332.x

Blessing, S. B., and Ross, B. H. (1996). Content effects in problem categorization and problem solving. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 792–810. doi: 10.1037/0278-7393.22.3.792

Boren, T., and Ramey, J. (2000). Thinking aloud: reconciling theory and practice. *IEEE Trans. Prof. Commun.* 43, 261–278. doi: 10.1109/47.867942

Burrelli, J. S. (2010). "Ch. 2: Higher education in science and engineering," in *Science and Engineering Indicators: 2010*. eds. R. F. Lehming, L. T. Carlson, M. J. Frase, M. P. Gutmann, J. S. Sunley and D. W. Lightfoot (National Science Foundation), 2.1–2.48.

Catrambone, R., and Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 1147–1156. doi: 10.1037/0278-7393.15.6.1147

Charters, E. (2003). The use of think-aloud methods in qualitative research an introduction to think-aloud methods. *Brock. Educ. J.* 12, 68–82. doi: 10.26522/brocked.v12i2.38

Checkland, P. (1999). *Systems Thinking, Systems Practice* Hoboken: John Wiley.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., and Glaser, R. (1989). Self-explanations: how students study and use examples in learning to solve problems. *Cogn. Sci.* 13, 145–182. doi: 10.1016/0364-0213(89)90002-5

Chrysikou, E. G., and Weisberg, R. W. (2005). Following the wrong footsteps: fixation effects of pictorial examples in a design problem-solving task. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 1134–1148. doi: 10.1037/0278-7393.31.5.1134

Colby, A., and Sullivan, W. M. (2008). Ethics teaching in undergraduate engineering education. *J. Eng. Educ.* 97, 327–338. doi: 10.1002/j.2168-9830.2008.tb00982.x

Daly, S. R., Mosyjowski, E. A., and Seifert, C. M. (2014). Teaching creativity in engineering courses. *J. Eng. Educ.* 103, 417–449. doi: 10.1002/jee.20048

Davis, K., Ghaffarzadegan, N., Grohs, J., Grote, D., Hosseinichimeh, N., Knight, D., et al. (2020). The Lake Urmia vignette: a tool to assess understanding of complexity in socio-environmental systems. *Syst. Dyn. Rev.* 36, 191–222. doi: 10.1002/sdr.1659

Dugan, K. E., Mosyjowski, E. A., Daly, S. R., and Lattuca, L. R. (2022). Systems thinking assessments in engineering: a systematic literature review. *Syst. Res. Behav. Sci.* 39, 840–866. doi: 10.1002/sres.2808

Duncker, K. (1945). On Problem Solving. *Psychol. Monogr.* 58, 1–113. doi: 10.1037/h0093599

Ericsson, K. A., and Simon, H. A. (1980). Verbal reports as data. *Psychol. Rev.* 87, 215–251. doi: 10.1037/0033-295X.87.3.215

Gentner, D., Loewenstein, J., and Thompson, L. (2003). Learning and transfer: a general role for analogical encoding. *J. Educ. Psychol.* 95, 393–408. doi: 10.1037/0022-0663.95.2.393

Goldwater, M. B., Gentner, D., LaDue, N. D., and Libarkin, J. C. (2021). Analogy generation in science experts and novices. *Cogn. Sci.* 45:e13036. doi: 10.1111/cogs.13036

Grohs, J. R., Kirk, G. R., Soledad, M. M., and Knight, D. B. (2018). Assessing systems thinking: a tool to measure complex reasoning through ill-structured problems. *Think. Skills Creat.* 28, 110–130. doi: 10.1016/j.tsc.2018.03.003

Jamrozik, A., and Gentner, D. (2020). Relational labeling unlocks inert knowledge. *Cognition* 196:104146. doi: 10.1016/j.cognition.2019.104146

Jansson, D. G., and Smith, S. M. (1991). Design fixation. *Des. Stud.* 12, 3–11. doi: 10.1016/0142-694X(91)90003-F

Jonassen, D. H. (1997). Instructional design models for well-structured and III-structured problem-solving learning outcomes. *Educ. Technol. Res. Dev.* 45, 65–94. doi: 10.1007/BF02299613

Jonassen, D. H. (2010). *Learning to Solve Problems: A Handbook for Designing Problem-solving Learning Environments*. Milton Park: Taylor & Francis.

Knight, D. B., and Novoselich, B. J. (2017). Curricular and co-curricular influences on undergraduate engineering student leadership. *J. Eng. Educ.* 106, 44–70. doi: 10.1002/jee.20153

Koro-Ljungberg, M., Douglas, E. P., Therriault, D., Malcolm, Z., and McNeill, N. (2013). Reconceptualizing and decentering think-aloud methodology in qualitative research. *Qual. Res.* 13, 735–753. doi: 10.1177/1468794112455040

Krahmer, E., and Ummelen, N. (2004). Thinking about thinking aloud: a comparison of two verbal protocols for usability testing. *IEEE Trans. Prof. Commun.* 47, 105–117. doi: 10.1109/TPC.2004.828205

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Qual. Quant.* 47, 2025–2047. doi: 10.1007/s11135-011-9640-9

Lattuca, L. R., Knight, D. B., Ro, H. K., and Novoselich, B. J. (2017). Supporting the development of engineers' interdisciplinary competence. *J. Eng. Educ.* 106, 71–97. doi: 10.1002/jee.20155

Marra, R. M., Hacker, D. J., and Plumb, C. (2022). Metacognition and the development of self-directed learning in a problem-based engineering curriculum. *J. Eng. Educ.* 111, 137–161. doi: 10.1002/jee.20437

Mazzurco, A., and Daniel, S. (2020). Socio-technical thinking of students and practitioners in the context of humanitarian engineering. *J. Eng. Educ.* 109, 243–261. doi: 10.1002/jee.20307

McKenna, A. F., Hynes, M. M., Johnson, A. M., and Carberry, A. R. (2016). The use of engineering design scenarios to assess student knowledge of global, societal, economic, and environmental contexts. *Eur. J. Eng. Educ.* 41, 411–425. doi: 10.1080/03043797.2015.1085836

Meacham, J. A., and Emont, N. M. (1989). "The interpersonal basis of everyday problem solving," in *Everyday Problem Solving: Theory and Applications*. ed. J. D. Sinnott (Santa Barbara: Praeger).

Meadows, D. H. (2008). *Thinking in Systems: A Primer*. Vermont: Chelsea Green Publishing.

Miles, M. B., Huberman, A. M., and Saldaña, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook* (*4th edn*). London: SAGE Publications, Inc.

National Academies of Sciences, Engineering, and Medicine, Committee on How People Learn II: The Science and Practice of Learning, Board on Behavioral, Cognitive, and Sensory Sciences, Board on Science Education, and Division of Behavioral and Social Sciences and Education (2018). *How People Learn II: Learners, Contexts, and Cultures*. Washington: National Academies Press.

National Academy of Engineering (2004). *The Engineer of 2020: Visions of Engineering in the New Century*. Washington: National Academies Press.

National Academy of Engineering (2005). *Educating the Engineer of 2020: Adapting Engineering Education to the New Century*. Washington: National Academies Press.

National Institutes of Health (2006). *Summary of the President's FY 2006 Budget*. Bethesda: National Institutes of Health.

National Research Council (2012). *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st century*. Washington: The National Academies Press.

Palmer, B., Terenzini, P. T., McKenna, A. F., Harper, B. J., and Merson, D. (2011). "Design in context: Where do the engineers of 2020 learn this skill?" in *2011 ASEE Annual Conference & Exposition*, 22–43.

Peng, K., Nisbett, R. E., and Wong, N. Y. C. (1997). Validity problems comparing values across cultures and possible solutions. *Psychol. Methods* 2, 329–344. doi: 10.1037/1082-989X.2.4.329

Saldaña, J. (2015). *The Coding Manual for Qualitative Researchers*. London: SAGE.

Schön, (1990). "The design process," in *Varieties of Thinking: Essays from Harvard's Philosophy of Education Research Center*. ed. V. A. Howard (London: Routledge).

Senge, P. M., Kleiner, A., Roberts, C., Ross, R. B., and Smith, B. J. (1994). *The Fifth Discipline Fieldbook: Strategies and Tools for Building a Learning Organization*. New York: Crown Business.

Shin, N., Jonassen, D. H., and McGee, S. (2003). Predictors of well-structured and ill-structured problem solving in an astronomy simulation. *J. Res. Sci. Teach.* 40, 6–33. doi: 10.1002/tea.10058

Shuman, L., Besterfield-Sacre, M., and McGourty, J. (2005). The ABET "professional skills"—can they be taught? Can they be assessed? *J. Eng. Educ.* 94, 41–55. doi: 10.1002/j.2168-9830.2005.tb00828.x

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychol. Rev.* 63, 129–138. doi: 10.1037/h0042769

Simon, H. A. (ed.) (1986). "The role of attention in cognition: a cultural-historical, interpersonal, and individual approach to development," in *The Brain, Cognition, and*

Education. eds. S. L. Freidman, K. A. Klivington and R. W. Peterson (Cambridge: Academic Press).

Sterman, J. (2010). "Causal Loop Diagrams," in *Business dynamics: Systems thinking and modeling for a complex world* (New York: McGraw-Hill).

Svanström, M., Lozano-García, F. J., and Rowe, D. (2008). Learning outcomes for sustainable development in higher education. *Int. J. Sustain. High. Educ.* 9, 339–351. doi: 10.1108/14676370810885925

Warburton, K. (2003). Deep learning and education for sustainability. *Int. J. Sustain. High. Educ.* 4, 44–56. doi: 10.1108/14676370310455332