From Problem Identification to Stakeholder Awareness: Systems Thinking Skill Assessment in Graduate Systems Engineering Course

Dr. Mengyu Li, University of Florida

Mengyu (May) Li is an Instructional Assistant Professor in the Department of Industrial and Systems Engineering at the University of Florida. She teaches a diverse range of undergraduate and graduate courses, including Inventory and Supply Chain Systems, Models for Supply Chain Management, Engineering Economy, Senior Design Projects, Simulation, Systems Design, and Systems Architecture.

Beyond her teaching, May serves as a faculty senator and member of the Engineering Faculty Council at UF, where she advocates for faculty representation and actively contributes to policy development. She is also the director of the UF Outreach Engineering Management (OEM) Program, currently serving her second term and overseeing all operational aspects while bridging academic and administrative domains.

Andrea Goncher P.E., University of Florida

Andrea Goncher is an Instructional Assistant Professor in Engineering Education at the University of Florida. She earned her PhD in Engineering Education from Virginia Tech and focuses on teaching and learning projects in human centered design and humanitarian engineering.

Dr. John Mendoza-Garcia, University of Florida

Dr. John Mendoza-Garcia is an Instructional Associate Professor in the Department of Engineering Education at the Herbert Wertheim College of Engineering, University of Florida. He received his Ph.D. in Engineering Education from Purdue University, and earned a Master's and a Bachelor's degree in Systems and Computing Engineering from Universidad de Los Andes and Universidad Nacional de Colombia, respectively. His teaching and research focus on developing professional competencies, including problem-solving, systems thinking, and design thinking. Before transitioning to academia, he worked in industry, gaining practical experience that informs his educational approach.

Paulina Gonzalez-Brito, University of Florida
Armeet Kaur Gahra, University of Florida

From Problem Identification to Stakeholder Awareness: Systems Thinking Skill Assessment in Graduate Systems Engineering Course

ABSTRACT

At the graduate level, students are expected to progress beyond component-level problem solving toward synthesizing knowledge across disciplines, recognizing feedback loops, and anticipating unintended consequences. This research builds upon and extends prior work on systems thinking development by exploring its evolution from undergraduate to graduate contexts. Specifically, it investigates how advanced academic and professional experiences shape key systems thinking dimensions, including problem identification and stakeholder awareness and involvement.

We investigate these dimensions through an analysis of student responses from a graduate-level Systems Architecture course. While open to all UF engineering graduate students, the course is predominantly composed of Industrial and Systems Engineering majors. Applying Groh's rubrics as an evaluative framework, we assessed responses collected over three consecutive years of course enrollments.

Our results indicate that, in general, graduate student participants demonstrated notable strengths in cross-disciplinary integration and consideration of long-term system impacts, moving beyond a narrow focus on discrete technical elements. However, certain aspects of systems thinking still require further development,

These findings underscore the influence of advanced education and professional experience in cultivating higher-order systems thinking and point to opportunities for introducing targeted scaffolding strategies earlier in the engineering curriculum to accelerate these capabilities. Insights from this analysis will inform the development of instructional strategies that explicitly help to structure systems thinking learning content at the graduate level.

1. INTRODUCTION

Engineering graduates are increasingly expected to reason for complex, socio-technical systems that extend beyond isolated technical components to include human, cultural, organizational, and environmental dimensions. Students are expected to progress beyond component-level problem solving toward synthesizing knowledge across disciplines, recognizing systems-level relationships, and anticipating unintended consequences [1].

Developing systems thinking in engineering education has emerged as a critical objective as educators seek to prepare students for the complexities of real-world problem-solving. Prior research has emphasized the need for engineering curricula to move toward holistic systems reasoning and problem integration [2,3,4,5].

As engineering problems grow in complexity and societal impact, the cultivation of systems thinking has become a core goal in engineering education [6]. Systems thinking entails the ability to integrate knowledge across domains, consider interactions among system components, anticipate unintended consequences, and engage multiple stakeholders in the decision-making process [2, 7]. This capability is increasingly viewed as essential for engineers working on open-ended, interdisciplinary problems that resist simple, component-level solutions. However, much of the research to date has emphasized undergraduate education, with fewer studies examining how systems thinking continues to evolve in graduate-level contexts, where students bring deeper technical expertise and often possess prior industry experience.

The ability to engage in systems thinking is increasingly recognized as a core competency for engineering graduates navigating complex, interdisciplinary challenges in professional practice [8]. Engineering education researchers have emphasized the importance of cultivating this capability early in students' academic journeys, particularly as engineering problems often involve dynamic systems, multiple stakeholders, and long-term consequences, which are necessary for the future workforce [9,10]. At the graduate level, however, there remains limited understanding of how systems thinking develops, particularly as students can apply advanced disciplinary knowledge and real-world experience to their studies. This study examines how graduate engineering students demonstrate systems thinking through written reflections in a Systems Architecture course. Specifically, we assess key dimensions such as stakeholder engagement, feedback loop recognition, and problem framing, using Groh's [11] systems thinking rubrics.

This research builds upon and extends prior work on systems thinking development by exploring its evolution from undergraduate to graduate contexts. Specifically, it investigates how advanced academic and professional experiences shape key systems thinking dimensions, including problem identification and stakeholder awareness and involvement. We aim to inform both instructional design and program-level structures that prepare graduate engineers to lead in the complex systems they will encounter in various contexts, including industry and academia.

2. METHODS

2.1 Context

Graduate students enrolled in a systems engineering course participated in a systems thinking activity as part of an individual assignment. The course focuses on system architecture and incorporates human-centered design processes to support the development of functional prototypes. Core topics include systems thinking, feedback loops, and the integration of technical and contextual factors in system design.

Although the course is open to all engineering majors, it primarily attracts students majoring in industrial and systems engineering. It serves as a capstone project course typically taken in the final year of study.

Data for this research were collected over three consecutive years, comprising responses from a total of 38 graduate engineering students (Year 1: 11 students; Year 2: 19 students; Year 3: 8 students), all of whom consented to participate in the study. All procedures adhered to institutional review board (IRB) protocols for ethical research involving human subjects.

2.2 Systems Thinking Assessment Tool

To assess systems thinking competencies, a problem-based scenario and accompanying rubric were used. This assessment moved beyond self-reported perceptions to evaluate demonstrated systems thinking behaviors.

The scenario presented a complex, ill-structured problem that required students to consider multiple dimensions, including technical, economic, ethical, and cultural factors, when formulating potential solutions. The scenario described a community facing severe challenges related to energy access and safety during harsh winter conditions [11]:

> "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 Yakutia residents. In fact, many residents are unable to afford heat for the entire winter (5 months). A Northeastern Federal University 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 Yakutia 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)."

2.3 Data Collection

For the purposes of data collection, the researchers slightly modified the scenario by renaming the village "Yakutia" to enhance realism. Students were informed that, while the region described was real and faced severe winters, the details provided were fictional and intended solely for the assignment.

Student responses were evaluated using the established systems thinking assessment rubric. Analysis was grounded in a conceptual framework comprising three interrelated dimensions: problem, perspective, and time. Each dimension included specific constructs that enabled the researchers to examine how students demonstrated systems thinking competencies, namely, their ability to integrate multiple perspectives, address complex interactions, and consider longer-term implications in their reasoning [11].

This study examined participants' performance across four prompts designed to assess different dimensions of systems thinking. In particular, we analyzed responses related to problem identification, information needs, and stakeholder awareness. Prompts 3 and 4 specifically focused on participants' ability to recognize relevant stakeholders and outline processes for engaging them. For clarity, we refer to these components as *stakeholder identification* (Prompt 3) and *stakeholder involvement* (Prompt 4).

Participants responded to the following prompts:

Prompt 1 - Problem identification: Given what you know from the scenario, please write a statement describing your perception of the problems and/or issues facing.

Prompt 2 - Information needs: What additional information do you need before you could begin to develop a response in Yakutia? Consider both detail and context of the problems/issues you identified.

Prompt 3 - Stakeholder identification: What groups or stakeholders would you involve in planning a response to the problems/issues in Yakutia?

Prompt 4 - Stakeholder involvement: Please briefly describe the process you would use to plan a response to the problems/issues in Yakutia.

The rubrics developed by Grohs et al. [11] were used to evaluate participants' written responses across multiple prompts in this study. Figure 1 presents the rubric used to assess students' performance on problem identification, focusing on their ability to recognize and articulate the core issues embedded in the scenario. Figure 2 displays the rubric applied to the information needs prompt, which evaluates how effectively participants identified the additional data and context required to develop an informed response. Figure 3 shows the rubric for evaluating participants' responses to both the stakeholder identification and stakeholder involvement prompts, assessing their understanding of relevant stakeholder groups and their ability to describe an appropriate engagement process.

Rubric – Operationalizing Systems Thinking: Problem Identification.

Construct	Criteria and Rating Guide
Problem Identification *Prompt 1*	**0** No response was provided or respondent was unable to identify a relevant problem **1** The problem statement identified is ***only technical*** or ***only contextual*** (economic, political, environmental, social, time, etc) in scope **2** The problem statement a) dentified both ***technical*** and ***contextual aspects*** but did not acknowledge ***interaction*** and complexity between issues b) identified ***technical aspect*** or ***contextual aspect only***, and acknowledges ***interactions*** and complexities between issues **3** The problem statement identified both ***technical*** and ***contextual aspects*** and acknowledges ***interactions*** and complexity between issues

Figure 1 - Rubrics by Grohs et al. [11] for evaluating problem identification responses.

Rubric – Operationalizing Systems Thinking: Information Needs.

Construct	Criteria and Rating Guide
Information Needs ***Prompt 2***	**0** No response was provided, or respondent sought information that was not relevant to the scenario **1** The response identified information needs focused only on one ***aspect:*** either ***technical*** only or ***contextual*** (economic, political, environmental, social, time, etc) only **2** The response 1. identified several relevant information needs addressing both ***technical*** and ***contextual aspects***, but these ***aspects*** are not specifically ***integrated*** 2. identified several relevant information needs addressing ***technical aspect*** or ***contextual aspect only***, and there is acknowledgment of ***integration*** within information needs of the ***aspect in focus*** **3** The response identified several relevant information needs that address both ***technical*** and ***contextual aspects*** and ***integrates*** these ***aspects***

Figure 2 - Rubrics by Grohs et al. [11] for evaluating information needs responses.

Construct	Criteria and Rating Guide	Rating
Stakeholder Awareness ***Prompts 3 and 4***	**0** No response was provided, or respondent only provided a list of stakeholders but no discussion on the role that the stakeholders will play in identifying and implementing possible solutions **1** The response includes a list of stakeholders; discussion of role of stakeholders is limited only to ***one group of stakeholders*** (community, or power/politics, or experts) providing ***input*** in discussions to identify possible solutions **2** The response lists an array of various stakeholders (community, power/politics, experts). Discussion of the role of stakeholders includes: 1. ***one group of stakeholders*** being ***engaged*** in activities to identify and implement possible solutions; or 2. ***more than one group of stakeholders*** providing ***input*** in discussions to identify possible solutions **3** The response lists an array of various stakeholders (community, power/politics, experts). Discussion of the role of stakeholders includes all stakeholders ***iteratively giving input and engaging with each other*** to identify and implement possible solutions. The discussion explicitly includes listening to the community voice and getting buy-in from the community	

Figure 3 - Rubrics by Grohs et al. [11] for evaluating stakeholder awareness responses

3. RESULTS

3.1 Data Analysis

Thirty-eight graduate students from three different cohorts participated in the study. The data were analyzed using Groh's systems thinking assessment tool, treating all participants as if they had taken the course in the same semester, since the same instructor delivered identical content across the three years. Two undergraduate researchers, supervised by one of the authors, conducted the analysis. After training on the evaluation and grading procedures, each undergraduate researcher independently assessed participants' responses to Prompts 1, 2, and 4. For Prompt 3, the two graders evaluated the responses jointly.

To evaluate inter-rater consistency in applying the grading rubrics, a quadratic weighted Cohen's kappa (kw) was calculated [12,13]. This statistic accounts for agreement due to chance and applies greater penalties to larger discrepancies between raters. Table 1 presents the kw results for Prompts 1, 2, and 4. A summary of the kw values can be found on Table 1

Prompt	Inter-rater reliability score using Quadratic weighted Cohen's kappa (kw)
Prompt 1	0.92
Prompt 2	0.42
Prompt 3	Raters worked together
Prompt 4	0.48

Table 1 - Inter-rater reliability kw scores for prompts 1 to 4

The kw results indicate almost perfect agreement for Prompt 1 (kw = 0.92) and moderate reliability for Prompts 2 (kw = 0.42) and 4 (kw = 0.48). These values suggest that inter-rater agreement was well above chance overall and that Prompt 1 was more straightforward to evaluate than Prompts 2 and 4.

When disagreements occurred for Prompts 1 and 2, the instructor provided additional guidance to help the undergraduate raters re-examine the responses. This was followed by a second review in which the two raters discussed their interpretations and reached consensus on previously disputed scores. For Prompt 3, responses were evaluated jointly by the two undergraduate raters. For Prompt 4, when the undergraduate raters were unavailable to reconcile disagreements, the instructor reviewed the participant's response and either selected one of the initial scores or assigned a new score as appropriate.

Examples of participant responses that initially resulted in rater disagreement are provided below for each prompt.

Specifically, for Prompt 1, Participant SP2301's response illustrates a case in which the two undergraduate raters did not initially agree:

> *"The issue is simply that the people in this region cannot afford heating [sic] their homes due to how expensive fossil fuel is. This is even a bigger issue as the winter climate is extremely harsh and cause [sic]several deaths."*

The two raters provided the following scores for this answer and also had a second pass in which they both agreed on a score, as shown in Table 2.

Prompt 1: Problem Identification						
Participant	Rater 1	*Notes by Rater 1*	Rater 2	*Notes by Rater 2*	Final	*Notes by Rater 3*
Sp2301	2.1	Context given, but lacks explanation	3	Good concise.	**2.2**	**Needs slightly more detailed explanation**

Table 2 - Rating of participants' answers for Participant Sp2301 for Prompt 1.

For Prompt 2, we present Participant Sp2206's response as an example where the two raters did not fully agree:

"Heating costs per person. Insulation costs"

This answer was scored by the raters as shown on Table 3.

Prompt 2: Information Needs						
Participant	**Rater 1**	***Notes by Rater 1***	**Rater 2**	***Notes by Rater 2***	**Final score**	***Notes by Rater 3***
Sp2206	1	Lacks explanation	3	Answered thoroughly	2.1	Identified contextual and technical, but no interaction.

Table 3 - Rating of participants' answers for Participant Sp2206 for Prompt 2.

As noted above, for Prompt 3 the two undergraduate raters worked together, whereas for Prompt 4 any disagreements between the undergraduate raters were resolved by the instructor, who reviewed the participant's response and determined the final score (undergraduate students were no longer involved in the project since the semester had ended). One example of a disagreement and the decision made by the instructor is the response provided by Participant Sp2105, as it can be seen on Table 4:

"I would meet with the village leaders to discuss needs and budgets that the town could afford for its people. I would also need to meet with the heads of each family to discuss their budget or what they could contribute to their own homes."

Prompt 4: Stakeholder involvement						
Participant	**Rater 1**	**Notes by Rater 1**	**Rater 2**	**Notes by Rater 2**	**Final score**	**Notes by Rater 3**
Sp2105	2.2	Steps mentioned, but not in a strategic manner	3	Good answer, paragraph approach	2.2	More than one group of stakeholders was identified, but no mention of their involvement or the community involvement

Table 4 - Rating of participant's answers for Participant Sp2105 for prompt 4.

3.2 Findings

3.2.1 Prompt 1: Problem identification

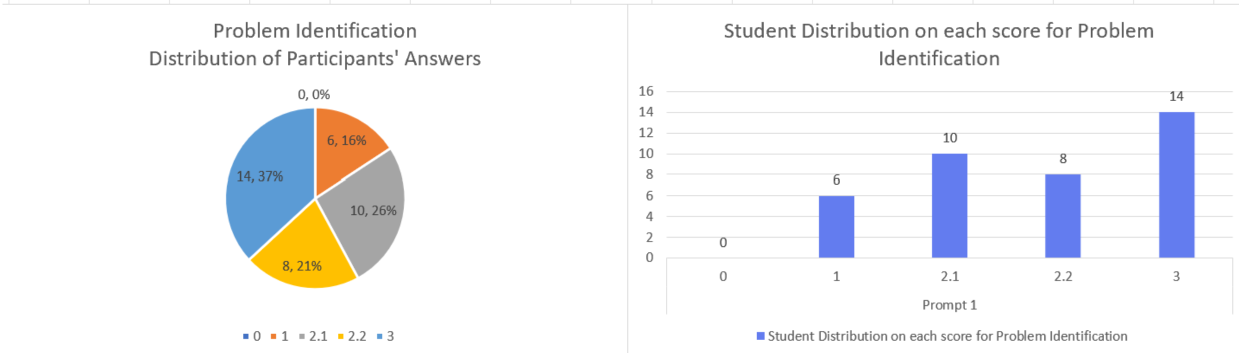

Figure 4 - Figure: Problem identification prompt 1, pie chart, and histogram analysis

Examining the results for Prompt 1 (Figure 4), the distribution of scores is illustrated using both a pie chart and a histogram. Only 37% of students achieved the highest score of 3, indicating that they correctly identified the problem while adequately addressing both technical and contextual aspects. For the intermediate scores of 2.1 and 2.2, 26% and 21% of students, respectively, demonstrated a partial but generally sound understanding, though their responses did not fully address all required elements. Approximately 16% of students received a score of 1, reflecting minimal explanation and the omission of key points. No students received a score of 0, suggesting that all participants were able to identify the problem to some extent.

3.2.2 Prompt 2: Information needs identification

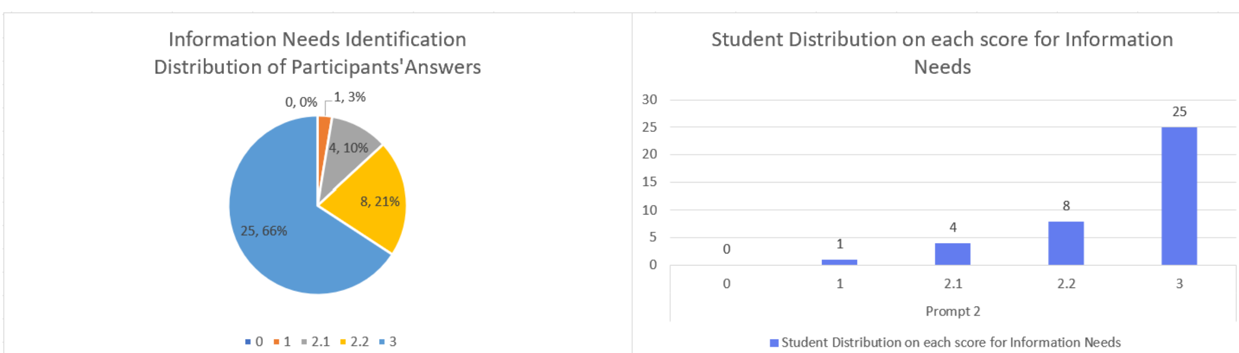

Figure 5 - Information Needs identification prompt 2, pie chart, and histogram analysis

Examining the score distribution for Prompt 2 (Figure 5), approximately two-thirds of the students (66%) achieved the highest score of 3, indicating that they provided a complete and well-explained response. For the intermediate scores of 2.1 and 2.2, 10% and 21% of students, respectively, received these ratings. In total, about 31% of students demonstrated a partial understanding of Prompt 2 by identifying some of the information needed to analyze the village, but their responses did not address all required aspects. Only one student received a score of 1, reflecting an incomplete or missing response.

3.2.3. Stakeholder identification (stakeholder awareness)

Figure 6 - Stakeholder identification prompt 3, pie chart, and histogram analysis

With respect to stakeholder identification (Figure 6), 14 participants (37%) achieved the highest score. An additional 14 participants (37%) were rated at level 2.2, while 7 participants (18%) received a score of 2.1 and 3 participants (8%) were scored at level 1. Overall, 74% of participants (levels 2.2 and 3) identified more than one stakeholder in their responses.

3.2.4. Stakeholder involvement

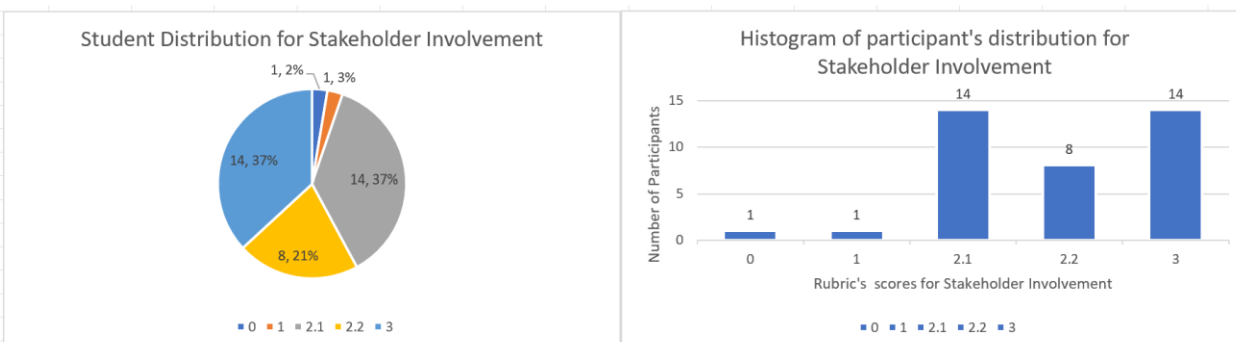

Figure 7 - Stakeholder identification prompt 3, pie chart, and histogram analysis

For Prompt 4 (Figure 7), the data indicate that when describing how stakeholders would be involved in defining the solution, 14 participants (37%) achieved the highest score. Twenty-two participants (58%) were rated at level 2 (21% at level 2.2 and 14% at level 2.1). One participant provided a response that did not mention stakeholders, and one participant did not provide an answer.

4. DISCUSSION AND IMPLICATIONS

The findings reveal varied levels of student performance across problem identification, information needs identification, and stakeholder awareness and involvement, highlighting uneven development of early-stage systems thinking competencies.

For **problem identification (Prompt 1)**, fewer than half of the students (37%) achieved the highest score, indicating that while all students could recognize the problem to some extent, many struggled to fully articulate both the technical and contextual dimensions. The sizable proportion of students scoring at intermediate levels (47% combined at levels 2.1 and 2.2) suggests partial understanding, with students recognizing surface-level technical issues but failing to integrate broader contextual or systemic considerations. This aligns with prior observations in engineering education that novice problem solvers often frame problems narrowly and prematurely converge on solutions without fully exploring the problem space.

In contrast, **information needs identification (Prompt 2)** showed notably stronger performance, with 66% of students achieving the highest score. This suggests that once a problem context is presented, students are more comfortable identifying data and information required for analysis than articulating the problem itself. However, roughly one-third of students scored at intermediate levels, indicating gaps in comprehensiveness and prioritization. This pattern may reflect the traditional emphasis in engineering curricula on data-driven analysis rather than upstream problem framing.

Results for **stakeholder identification (Prompt 3)** further demonstrate moderate to strong awareness, with 74% of participants identifying more than one stakeholder. This suggests that students are generally capable of recognizing multiple actors within a system. Nevertheless, the distribution across levels indicates variability in depth, with some students listing stakeholders without clearly articulating their roles or relationships. This pattern becomes more pronounced in **stakeholder involvement (Prompt 4)**, where only 37% achieved the highest score, and the majority remained at intermediate levels. While most students acknowledged stakeholder participation, fewer could meaningfully describe how stakeholders would influence or co-define solutions. This suggests that stakeholder consideration is often treated descriptively rather than as an integral component of decision-making and solution development.

Taken together, these results suggest a developmental progression: students demonstrate stronger performance in identifying information needs and recognizing stakeholders than in fully framing complex problems or integrating stakeholders into solution processes. This imbalance highlights a common challenge in systems-oriented education. Students may understand *what* elements exist within a system, but struggle to articulate *how* these elements interact and shape problem definitions and outcomes.

The findings have several important implications for engineering education, particularly for courses aiming to foster systems thinking and socially informed design:

1. **Strengthening problem framing:** The relatively low proportion of students achieving the highest score in problem identification underscores the need for explicit instructional support. Structured frameworks, guided prompts, or iterative refinement activities can help students consider technical, social, environmental, and contextual dimensions before proposing solutions.
2. **Leveraging strengths in information analysis:** Students' relative proficiency in identifying information needs provides a foundation for scaffolding. Activities that connect data requirements back to problem definitions and stakeholder concerns can reinforce the link between analysis and system-level understanding.
3. **Enhancing stakeholder engagement:** While students recognize multiple stakeholders, the weaker performance in stakeholder involvement highlights the need to move beyond

identification. Role-playing exercises, stakeholder mapping, and scenario-based decision-making tasks can help students understand how stakeholders influence problem definition, trade-offs, and solution viability.

4. **Using detailed rubrics for targeted feedback:** Graded levels (e.g., 2.1, 2.2) are valuable for diagnosing partial understanding and guiding instruction. They can help identify specific gaps, such as missing contextual factors or superficial stakeholder treatment, and support tailored feedback.

Overall, these findings support the integration of more intentional, early-stage design instruction that emphasizes problem framing and stakeholder integration as foundational skills, rather than assuming they will naturally emerge through technical analysis alone.

4.1. Effectiveness of the Assessment Approach

The assessment tool employed in this study has demonstrated strong validity and reliability for evaluating systems thinking competencies in graduate engineering students. As shown in Table 1, the inter-rater reliability scores, measured using quadratic weighted Cohen's kappa (κw), indicate almost perfect agreement for Prompt 1 (κw = 0.92) and moderate reliability for Prompts 2 (κw = 0.42) and 4 (κw = 0.48). These scores confirm that the grading process is fair and that the rubrics can be applied consistently across different raters, which is essential for the credibility and replicability of the assessment. By measuring these competencies at key points in the curriculum, educators can assess growth over time and evaluate the effectiveness of specific instructional interventions. This aligns with program-level assessment requirements and supports continuous improvement efforts.

4.2. Limitations

This study was conducted with students from three different cohorts on a systems engineering graduate course offered by the Department of Industrial and Systems Engineering at our University. Still, several factors lead us to believe the results are limited and cannot be generalized to other institutions, disciplines, or graduate populations. Such factors include the sample size (n=38) and the fact that the participants were primarily industrial and systems engineering graduate students, who may have had prior theoretical or practical exposure to systems thinking concepts and/or practices, which limits generalizability to other graduate engineering programs. Additionally, participants were assessed using a single, written scenario-based instrument, which may not fully reflect students' performance in a team-based or applied design context.

5. FUTURE DIRECTIONS FOR RESEARCH

Future research should explore instructional interventions specifically designed to strengthen students' problem-framing and stakeholder-integration skills and empirically evaluate their impact across multiple course offerings. One promising direction would be to implement a pre/post assessment design in which students complete the same systems thinking task at the beginning of the course and again after receiving instruction on system-level thinking. This approach would enable direct comparison of their performance before and after targeted pedagogical interventions, providing stronger evidence of instructional effectiveness and skill development within a single semester. Such a design could help isolate the impact of specific course components on systems thinking competencies.

Longitudinal studies could examine how these competencies develop over time and whether early improvements in problem identification translate into higher-quality design outcomes. Additionally, qualitative analyses of student responses may provide deeper insight into how students conceptualize system boundaries and stakeholder roles. Expanding the study across institutions and disciplines would further support the generalizability of these findings and inform best practices for embedding systems thinking in engineering curricula.

Another important direction for future work involves expanding the sample size and diversity of test cases. Since this course is offered only once per year, there is an inherent limitation on the number of participants that can be recruited within a short time interval. Future iterations could explore multi-institutional collaborations to increase participant diversity and sample size more rapidly. Additionally, incorporating multiple ill-structured scenarios with varying levels of complexity and different contextual factors (e.g., urban vs. rural settings, different engineering domains, or varying scales of impact) would strengthen the robustness of the assessment tool and provide a more comprehensive understanding of how systems thinking skills transfer across different problem contexts.

REFERENCES

[1] S. Fukushige, T. Nonaka, H. Kobayashi, and C. Tokoro, *EcoDesign for Circular Value Creation: Volume I*. Singapore: Springer, 2025.

[2] M. Li, J. A. Mendoza-Garcia, and A. Goncher, "Assessing students' stakeholder awareness skills in an introductory engineering design course through systems thinking scenarios," in *Proc. 2023 ASEE Annu. Conf. & Expo.*, Baltimore, MD, USA, Jun. 2023, doi: 10.18260/1-2--42307.

[3] J. Mendoza-Garcia, A. Goncher-Sevilla, and M. Li, "Assessing systems thinking skills in engineering education: Addressing implementation challenges and unintended consequences in ill-structured problems," in *Proc. 2025 ASEE Annu. Conf. & Expo.*, Montreal, QC, Canada, Jun. 2025, doi: 10.18260/1-2--55473.

[4] A. Goncher-Sevilla, J. A. Mendoza-Garcia, and M. Li, "Student goal formulation in an introductory engineering design course through systems thinking scenarios," in *Proc. 2024 ASEE Annu. Conf. & Expo.*, Portland, OR, USA, Jun. 2024, doi: 10.18260/1-2--48011.

[5] J. A. Mendoza-Garcia, A. Goncher, and M. Li, "WIP: Assessing baseline systems thinking in an introductory engineering design course," in *Proc. 2021 ASEE Virtual Annu. Conf.*, Virtual Conference, Jul. 2021, doi: 10.18260/1-2--38069.

[6] B. Richmond, "Systems thinking: Critical thinking skills for the 1990s and beyond," *Syst. Dyn. Rev.*, vol. 9, no. 2, pp. 113–133, 1993. [Online].
Available: https://doi.org/10.1002/sdr.4260090203

[7] B. Blanchard, W. Fabrycky, and W. Fabrycky, *Systems engineering and analysis*. 1990.

[8] S. N. Ankrah and O. Al-Tabbaa, “Universities–industry collaboration: A systematic review,” *Scand. J. Manag.*, vol. 31, no. 3, pp. 387–408, 2015. [Online]. Available: https://doi.org/10.1016/j.scaman.2015.02.003

[9] J. P. Monat, T. F. Gannon, and M. Amissah, “The Case for Systems Thinking in Undergraduate Engineering Education,” *Int. J. Eng. Ped.*, vol. 12, no. 3, pp. pp. 50–88, May 2022.

[10] C. Frey and M. Osborne, “The future of employment: How susceptible are jobs to computerisation?,” *Technol. Forecast. Soc. Change*, no. 114, pp. 254–280, Jan. 2017, Accessed: Jan. 09, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0040162516302244.

[11] J. R. Grohs, G. R. Kirk, M. M. Soledad, and D. B. Knight, “Assessing systems thinking: A tool to measure complex reasoning through ill-structured problems,” Think. Ski. Creat., vol. 28, no. September 2017, pp. 110–130, 2018, doi: 10.1016/j.tsc.2018.03.003.

[12] J. L. Fleiss and J. Cohen, “The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability,” Educ. Psychol. Meas., vol. 33, no. 3, pp. 613–619, 1973.

[13] J. L. Fleiss, J. Cohen, and B. S. Everitt, “Large sample standard errors of kappa and weighted kappa,” Psychol. Bull., vol. 72, no. 5, pp. 323–327, 1969.