Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
Paper ID #48766

Towards a framework for assessing systems thinking in collaborative problem-solving in STEM

Dr. Lykke Brogaard Bertel, Aalborg University

Lykke Brogaard Bertel is Associate Professor in PBL and digital transformation at the UNESCO center for Problem-based Learning in Engineering, Science and Sustainability at Aalborg University.

Dr. Rea Lavi, Massachusetts Institute of Technology

Rea Lavi received his Ph.D. degree from the Faculty of Education in Science and Technology, Technion—Israel Institute of Technology, Haifa, Israel. He is Lecturer and a Curriculum Designer with the NEET program, School of Engineering, Massachusetts Institute of Technology, Cambridge, MA, where he teaches thinking skills to undergraduate students. His research interests in STEM education involve the fostering and assessment of thinking skills involved in complex problem-solving, with special focus on systems thinking, creative thinking, and metacognition. His doctoral research received several awards, including the Zeff Fellowship for Excelling First-year Ph.D. Students and the Miriam and Aaron Gutwirth Fellowship for Excelling Ph.D. Students. Rea is also the inventor of the SNAP Method® for structured creative problem-solving (US & UK trademarks).

Kamilla Rathcke, Aalborg University
Nelson F. Coelho, Aalborg University
Sun Cole Seeberg Dyremose, Aalborg University

Towards a framework for assessing systems thinking in collaborative problem-solving in STEM

Lykke Brogaard Bertel[1] [0000-0002-1460-2905], Rea Lavi[2] [0000-0002-0788-7236], Kamilla Rathcke[1],
Nelson F. Coelho[1] [0000-0001-5287-2538] and Sun Cole Seeberg Dyremose[1] [0000-0001-7971-8266]

[1] Aalborg University, Aalborg, Denmark
[2] Massachusetts Institute of Technology, Cambridge MA, USA
lykke@plan.aau.dk

Abstract

In this evidence-based practice research brief, we expand on previous work developing SAFO - a framework for teaching introductory systems thinking in first-year STEM education. We refine a rubric useful for assessing systems thinking, and present initial results from applying this rubric to structured case work involving collaborative problem-solving. Finally, we discuss the potential of applying SAFO as a research tool to compare variations of interdisciplinarity and complexity in collaborative problem-solving in STEM.

Introduction

Systems thinking is a higher order thinking skill important for addressing complex, real-world problems in STEM [1-3]. Systems thinking can be assessed in a multitude of ways, including rubrics, open- and close-ended tools, scenarios, mapping and coding schemes, and more, depending on the focus and field of study [3-6]. We have previously shown that an introductory systems thinking (IST) framework, System Architecture-Function-Outcome (SAFO), can be used to teach and assess first-year STEM students' systems thinking, and validated a rubric suitable for this purpose [7]. While the previous study focused on students' individual IST, we hypothesize that SAFO and the associated rubric are applicable to a variety of STEM education settings, including those that involve collaborative problem-solving.

In STEM, collaborative problem-solving is often facilitated through problem-oriented case- or project work, and, depending on the structuredness and complexity of the problem, may involve several iterations of problem design and analysis [8]. In this process, students inquire and explore the problem field and negotiate different perspectives and approaches to finally agree on a suitable solution. Thus, for systems thinking to be utilized in collaborative problem-solving, it needs to be integrated into and assessed formatively throughout these processes, to support communication and argumentation in the scoping and analysis of problems [7].

The purpose of our research is two-fold; to elaborate and employ SAFO to refine a rubric useful for teaching and assessing introductory systems thinking in collaborative problem-solving in STEM, and, to investigate its applicability as a research tool to compare IST across variations of collaborative problem-solving in STEM. In this research brief we focus primarily on how we refined the rubric and present initial results from applying the rubric to structured case-work involving collaborative problem-solving. We discuss the implications of this approach to teaching and assessing IST skills in collaborative settings and discuss the potential of applying the framework to compare variations of collaborative problem solving, i.e. structured, discipline-specific case-work compared to open-ended, student-led and interdisciplinary project work.

Systems thinking in STEM

The concept of systems thinking has been independently developed in various disciplines, such as natural science and social science, and is increasingly recognized as a crucial skill across STEM disciplines too due to its critical role in addressing complex, multifaceted problems [3, 6]. In STEM, systems thinking enables students to identify, understand, predict, and suggest improvements of every aspect of an *engineered* (artificial) system, including the way these aspects inter-relate within the system and with other natural or socio-technical systems [6, 9]. Introducing systems thinking to STEM novices can be challenging, primarily due to limited disciplinary knowledge and lack of prior exposure to complex systems analysis. Thus, in previous studies, IST constructs have been utilized to create structured assignments as mediums for teaching systems thinking in the initial learning of problem-solving [9-11]. The SAFO framework is one such construct, which, through its inclusion of 'Outcome' directs attention towards the *problem* that the system is intended to solve, as well as its *benefits* and *detriments* to *stakeholders* [7].

Teaching thinking skills and problem-solving in STEM is often guided to some degree through cases and case-based learning (CBL) with well-defined problems to scaffold the development of problem-solving competencies [12]. These can later be further developed through increasingly open-ended and ill-defined problem- and student-led project-based project work (PBL) that may eventually involve multi- or interdisciplinary communication and collaboration [8, 13]. As such, PBL can be considered a type of CBL with high learner autonomy, leaving more of the preparation stage of problem-solving to students. This may point to simpler CBL types being suitable for the very first introduction of systems thinking, e.g. to first-year students, and with a potential to scaffold progression in systems thinking by gradually increasing complexity. Thus, a rubric for assessing IST would need to be flexible enough to be applicable in first-year STEM while also allowing for a gradual increase of complexity in the problem-solving process.

Assessing introductory systems thinking in collaborative problem-solving

When developing the initial rubric for assessing IST based on SAFO, we focused specifically on enabling formative assessment by separating the two criteria, i.e. 'adherence' (framework understanding) and 'correctness' (content knowledge according to case description). Having a high level of systems thinking skill requires both, but separating them makes it possible to detect issues in the students' IST more precisely and thus provide targeted feedback. The assessment rubric included scoring guidelines for both criteria, which were refined iteratively and finally tested for inter-rater agreement. The original rubric can be found in [7].

The previous study focused on individual assessment and took point of departure in a lecture-based and highly structured assignment with a written case description serving as a pre-defined 'source of truth'. However, follow-up discussions with students revealed a promising potential for the framework to be applied in other ways too, e.g. to support collaborative problem-solving processes within teams or interdisciplinary communication between teams. Thus, in the current study we wanted to expand on these initial findings, applying the framework in a collaborative problem-solving setting in STEM with a more open-ended case description and no pre-defined 'source of truth', and through this explore possible scenarios for future applications and research.

Study design

In this study, discipline-specific student teams were given a case and a structured approach through which to collaboratively identify a system of interest from the perspective of their discipline and map it according to the SAFO framework. A total of 93 students (17 teams of 5-7 students) participated in the study, in the fall of 2023 as part of an introductory course on PBL for students from biology, biotechnology, chemistry, chemical engineering and environmental science. Over the course of four weeks, the students were introduced to a deep seabed mining case, systems thinking and SAFO as a structured approach through which to identify, analyze and describe a system of interest (SoI) and associated problem through the lens of their specific discipline. Team responses and individual reflections were collected using SurveyXact.

Five evaluators, of which two are experts on IST and three are content knowledge experts in deep seabed mining, assessed the student team responses. Each response received a score of 3 (strong – no changes required), 2 (moderate – few and/or minor changes required), 1 (weak – many and/or major changes required) or 0 (must be redone). This process served to both provide formative feedback to student teams on their IST, and to further develop and refine the rubric itself with detailed assessment guidelines for adherence (framework understanding) and correctness (content knowledge) within the context of collaborative problem-solving.

Results and discussion

Student teams were generally able to understand and apply SAFO to the deep seabed mining case and identify associated problems related to a chosen SoI. Based on the assessment results (Table 1 and 2 below) it is clear, that students needed more feedback on system *function*, i.e., its interaction with other boundary systems through input and output) as opposed to its *architecture* (structure and behavior between parts) and *outcome* (problem, stakeholders, benefit/detriment), which received higher scores across all disciplines. Describing the *benefit* of a system seemed easier compared to describing its *detriment*. The relative lower score on system *output* compared to system *input*, could indicate a potential confusion of 'output' and 'benefit', i.e. describing a key positive intended benefit of the system of interest rather than its interaction with a boundary system which receives output (material, energy or information) from said SoI.

Table 1. Median Team scores for adherence.

Discipline	Structure	Behavior	Input	Output	Stake-holders	Problem	Benefit	Detriment
All (17)	3	2	2	1	2	2	2	1
Biology (4)	3	2	2	0.5	1	2	2.5	1.5
Biotechnology (3)	3	3	2	0	2	2	3	1
Chemistry and chemical engineering (6)	3	2.5	2.5	1.5	2.5	2	2	1
Environmental science (2)	3	2	2	0.5	2	1.5	1.5	2

Table 2. Median Team scores for correctness.

Discipline	Structure	Behavior	Input	Output	Stake-holders	Problem	Benefit	Detriment
All (17)	2	3	2	0	2	2	2	2
Biology (4)	2.5	3	1	0	1	0	1	2
Biotechnology (3)	2	3	2	0	2	2	2	2
Chemistry and chemical engineering (6)	2	2	1	0	2	2	2.5	1
Environmental science (2)	2	3	2.5	1	2	2.5	2.5	2.5

In the qualitative assessment of responses, we noticed that most student teams described the *architecture* and *function* of the deep seabed mining system in general and non-discipline specific terms, e.g. choosing the subsea collector (the mining vehicle collecting the nodules), or the ship transporting the collector as the system of interest (SoI).

Discipline-specific aspects of the SoI were more explicitly present in responses related to the *outcome* (the problem that the system addressed, its benefits, detriment and related stakeholders). For instance, a team of biology students focused on the impact of the deep seabed mining system on a particular endangered species of squid, whereas a team of biotechnology students focused on improving the coating of the drill to minimize maintenance costs. In one cross-disciplinary feedback session, we observed the communication between these two teams, through which the students seemed to become aware of these differences as well. The students discussed their choices regarding system perspectives and whether these choices reflected aspects of, and perhaps embedded values within, their own discipline and other disciplines close to, but different, from their own. At the same time, the flexibility of the SAFO framework allowed for the students to “increase” or “decrease” complexity of the system by ‘zooming in and out’, realizing the multi-faceted interactions within or between problems, benefits and detriments of engineered systems as well as their interconnectedness with other socio-technical and natural systems. We intend to explore these potentials further in future studies.

With regards to the development of IST assessment rubric, we were able to refine the rubric further through five iterations (two iterations for adherence, three for correctness), adding more detail to the assessment guidelines to make the distinction between adherence and correctness clearer. Furthermore, we added another level to the scoring to make the assessment finer grained to provide better and more targeted feedback (example provided in Table 3 below).

While we did reach inter-rater agreement regarding both ‘adherence’ and ‘correctness’, we found that when scoring correctness, the concept of ‘precise’ was particularly challenging, as some student teams tended to ‘overdo’ their responses (e.g. providing multiple problems, benefits and detriments) at the expense of precision. Thus, teams who had engaged in discussions and identified multiple valid responses could potentially receive a lower score compared to less engaged teams who provided very brief (but precise) responses. However, this could potentially be addressed in future iterations by clarifying the instructions for students further, such as requesting full sentence responses and adjusting certain prompts, e.g. requesting "*one* key problem" instead of "*a* key problem", or "the group of people *most* affected by the problem" rather than "*a* group of people affected by the problem".

Table 3. Scoring guidelines for 'Problem' (Adherence)

	Original guidelines	Final iteration of guidelines*
Level 1 (weak)	Multiple problems/ needs/lacks/demands of stakeholders	Multiple non-distinct problems/needs/lacks/demands (phrased as such) not caused by the system's function
Level 2 (moderate)		One distinct problem/need/lack/demand (phrased as such) not related to any group of stakeholders AND not caused by the system's function OR One non-distinct problem/need/lack/demand (phrased as such) not caused by the system's function OR Multiple distinct problems/needs/lacks/demands (phrased as such) not caused by the system's function
Level 3 (strong)	One problem/need/ lack/demand of stakeholders	One distinct problem/need/lack/demand (phrased as such) related to at least one group of stakeholders mentioned AND not caused by the system's function.

* The full rubric in its final version is available in the Appendix.

Conclusions and future work

This research brief provides initial evidence-based insights into teaching introductory systems thinking in collaborative problem-solving in STEM using the SAFO framework, for which we were able to refine an assessment rubric and apply it in a collaborative CBL-setting. While student teams in the study generally described the 'architecture' (structure and behavior) of the system of interest and its 'function' (interactions with boundary systems) in similar ways regardless of discipline, they often approached the system 'outcome' (problem, benefits and detriments to stakeholders) from the perspective of their specific field of study, which facilitated interdisciplinary awareness in subsequent cross-disciplinary feedback sessions.

In future work, we aim to research this further by comparing findings from collaborative problem-solving in CBL and PBL, respectively, to explore the potential of SAFO for scoping and analyzing more open-ended and ill-defined problems within one discipline and using it to communicate and possibly connect problems across disciplines in a systems perspective. Furthermore, we will explore the application of the assessment rubric as a tool to foster peer-feedback and potentially AI-augmented feedback.

Finally, future research will further continue to explore the framework's applicability as a research tool to investigate the impact of variations in collaborative problem-solving (e.g. interdisciplinary project work) on students' systems thinking, as well as progression in systems thinking across educational levels, with the purpose of continuously and effectively preparing future STEM professionals and assessing their ability to tackle real-world and interconnected complex challenges.

References

[1] National Research Council, *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. The National Academies Press, 2012.

[2] M. J. Jacobson and U. Wilensky, "Complex systems in education: Scientific and educational importance and implications for the learning sciences," *The Journal of the Learning Sciences*, vol. 15, no. 1, pp. 11-34, 2006.

[3] T. Bielik, I. Delen, M. Krell, and O. B. Z. Assaraf, "Characterising the literature on the teaching and learning of system thinking and complexity in STEM education: A bibliometric analysis and research synthesis," *Journal for STEM Education Research*, vol. 6, no. 2, 2023.

[4] K. E. Dugan, E. A. Mosyjowski, S. R. Daly, and L. R. Lattuca, "Systems thinking assessments in engineering: A systematic literature review," *Systems Research and Behavioral Science*, vol. 39, no. 4, pp. 840-866, 2022.

[5] 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," *Thinking Skills and Creativity*, vol. 28, pp. 110-130, 2018.

[6] S. York, R. Lavi, Y. J. Dori, and M. Orgill, "Applications of Systems Thinking in STEM Education," *Journal of Chemical Education*, vol. 96, no. 12, pp. 2742-2751, 2019.

[7] R. Lavi, and L. B. Bertel, "The System Architecture-Function-Outcome framework for fostering and assessing systems thinking in first-year STEM education and its potential applications in case-based learning," *Education Sciences,* vol. 14, no. 7, 2024.

[8] W. Hung, D. H. Jonassen, and R. Liu, "Problem-based learning," in *Handbook of Research on Educational Communications and Technology*, 3rd ed., Routledge, 2008, pp. 485-506.

[9] R. Lavi, L. O. R. I. Breslow, M. M. Salek, and E. F. Crawley, "Fostering and assessing the systems thinking of first-year engineering students using the system architecture-function-purpose framework." *International Journal of Engineering Education,* vol. 39, 2023.

[10] S. Nagarajan and T. Overton, "Promoting systems thinking using project-and problem-based learning," *Journal of Chemical Education*, vol. 96, no. 12, pp. 2901-2909, 2019.

[11] W. Hung, "Enhancing systems-thinking skills with modelling," *British Journal of Educational Technology*, vol. 39, no. 6, pp. 1099-1120, 2008.

[12] Y. J. Dori, and R. Lavi, "Teaching and assessing thinking skills and applying educational technologies in higher education." *Journal of Science Education and Technology,* vol. 32, no. 6, pp. 773-777, 2023.

[13] L. B. Bertel, M. Winther, H. W. Routhe, and A. Kolmos, "Framing and facilitating complex problem-solving competences in interdisciplinary megaprojects: An institutional strategy to educate for sustainable development," *International Journal of Sustainability in Higher Education,* vol. 23 nr. 5, pp. 1173-1191, 2022.

Appendix: Assessment rubric for SAFO

Terms and Descriptions

- Behavior: ‘Interaction’ must be a causal relation, not a structural one. Interactions can involve matter, energy, or information. Behavioral interactions are time-dependent and dynamic.
- Function: a boundary system shares one or more parts with the target system (explicitly or implicitly). If it does not, then it is not a boundary system.

Instruction prompts

System architecture

- structure: name five key parts of the system.
- behavior: name four causal interactions between the parts you mentioned in the previous question. An interaction should be between two or more parts.

System function

- input: describe a system that exists on the boundary of our system of interest and which provides our system with input. What is this 'input system'? What is its input into our system of interest?
- output: describe a system that exists on the boundary of our system of interest and which receives output from our system. What is this 'output system'? What is the output it receives?

System outcome (provide responses as full sentences)

- key stakeholders: what group of people is most affected by the problem which the system function solves or improves?
- problem: what is one key problem the system is designed to solve for its key stakeholders? If multiple choose just one.
- key benefit: Describe one key positive intended outcome of the system when it functions as intended. The outcome should affect the key stakeholders of the system.
- key detriment: Describe a key negative expected outcome of the system when it functions as intended. The outcome should affect key stakeholders of the system, either the group already mentioned or another key stakeholder.

Score key

- 3 - Strong: No changes required
- 2 - Moderate: Few and/or minor changes required
- 1 - Weak: Many and/or major changes required
- 0 - Unacceptable: Cannot be improved, need to redo

Adherence x Correctness = Introductory Systems Thinking (Potential total scores: 0, 1, 2, 3, 4, 6, 9)

Adherence of system description to assignment instructions: Knowledge of the framework

The adherence of a response should be based on the instructions provided to students as well as on inferences that can reasonably be made based on the totality of students' responses and on the content of those sources.

SAFO aspect		3 - Strong	2 - Moderate	1 - Weak
Architecture	Structure	[number in instructions] different parts AND majority are technological	[number in instructions] different parts AND majority are not technological OR [+ or - 1 number in instructions] different parts AND majority are technological	[+ or - 2+ number in instructions] different parts where at least one is technological
	Behavior	Explicit interactions (cause-and-effect relationships) that together cover every part mentioned under 'Structure'	Explicit interactions that together cover at least two parts but not every part mentioned under 'Structure' OR Implicit interactions that together cover every part mentioned under 'Structure'	Implicit interactions that together cover at least two parts but not every part mentioned under 'Structure'
Function	Input*	One/Two Boundary System/s AND one interaction between Target System and Boundary System, for each system. The direction of the interaction is from Boundary System to Target System. Input and output systems can be the same. Target system or its parts can be implied.	One/Two Boundary System/s AND multiple interactions between Target System and Boundary System, for each system. The direction of all the interactions is from Boundary System to Target System. OR One/Two Boundary System/s AND one interaction between Target System and Boundary System for one of the systems AND no interaction between Target System and Boundary system for the other system. The direction of all the interactions is from Boundary System to Target System. OR One/Two non-Boundary System AND one interaction between Target System and non-Boundary System, for each system. The direction of the interaction is from the non-Boundary System to the Target System.	Any number of boundary System/s without interactions
	Output*	One/Two Boundary System/s AND one interaction (cause-and-effect relationship) between Target System and Boundary System, for each system. The direction of the interaction is from Target system to Boundary System. Input and output systems can be the same. Target system or its parts can be implied.	One/Two Boundary System/s AND multiple interactions (cause-and-effect relationships) between Target System and Boundary System, for each system. The direction of all the interactions is from Target system to Boundary System OR One/Two non-Boundary System AND one interaction (cause-and-effect relationship) between Target System and non-Boundary System, for each system. The direction of the interaction is from Target System to non-Boundary System. OR	Any number of boundary System/s without interactions

			One/Two non-Boundary System AND one interaction between Target System and non-Boundary System, for each system. The direction of the interaction is from the Target System to the non-Boundary System.	
Outcome	Stake-holders	One distinct group	One non-distinct group OR Multiple distinct groups	Multiple non-distinct groups
	Problem	One distinct problem/need/lack/demand (phrased as such) related to at least one group of stakeholders mentioned AND not caused by the system's function**	One distinct problem/need/lack/demand (phrased as such) not related to any group of stakeholders AND not caused by the system's function** OR One non-distinct problem/need/lack/demand (phrased as such) not caused by the system's function** OR Multiple distinct problems/needs/lacks/demands (phrased as such) not caused by the system's function**	Multiple non-distinct problems/needs/lacks/demands (phrased as such) not caused by the system's function**
	Benefit	One direct positive outcome related to at least one group of the stakeholders mentioned	One direct positive outcome not related to any group of stakeholders OR Multiple direct positive outcomes OR One indirect positive outcome	Any number of indirect positive outcome on non-distinct stakeholders
	Detriment	One direct negative outcome on distinct stakeholders. No causal explanation required	Multiple direct negative outcomes on distinct stakeholders. No causal explanation required OR One indirect negative outcome on distinct stakeholders. No causal explanation required OR One direct negative outcome on non-distinct stakeholders. No causal explanation required	Any number of indirect negative outcome on non-distinct stakeholders

* People, natural phenomena, etc., can also be considered as systems.
** Based on students' responses to the "architecture" and "function" elements.

Correctness of system description - domain knowledge

Each aspect should be scored independently of all other aspects. Do not score responses that received '0' for adherence.

The correctness of a response should be based on the sources of truth provided to students and that students reported as having used, as well as on inferences that can reasonably be made based on the totality of students' responses and on the content of those sources.

SAFO aspect		3 - Strong	2 - Moderate	1 - Weak
Architecture	Structure	All parts are correct and precise	Half or more of parts are correct and precise, but not all	At least one part but less than half are correct and precise
	Behavior[1]	All behavioral interactions are correct and precise	Half or more of behavioral interactions are correct and precise, but not all	At least one behavioral interaction but less than half are correct and precise
Function	Input	Boundary system/s (if mentioned) are correct and precise and Interaction/s between boundary system/s and SoI (if mentioned) are correct and precise	If multiple boundary systems and/or interactions are mentioned: half or more of are correct and precise If one boundary system and/or interaction is mentioned: correct but imprecise	If multiple boundary systems and/or interactions are mentioned: at least one boundary system but less than half are correct and precise
	Output	Boundary system/s (if mentioned) are correct and precise and Interaction/s between boundary system/s and SoI (if mentioned) are correct and precise	If multiple boundary systems and/or interactions are mentioned: half or more of are correct and precise If one boundary system and/or interaction is mentioned: correct but imprecise	If multiple boundary systems and/or interactions are mentioned: at least one boundary system but less than half are correct and precise
Outcome	Stake-holders	Every response is correct and precise	If multiple responses: half or more are correct and precise, but not all If one response: correct but imprecise	If multiple responses: at least one response but less than half are correct and precise
	Problem	Every response is correct and precise	If multiple responses: half or more are correct and precise, but not all If one response: correct but imprecise	If multiple responses: at least one response but less than half are correct and precise
	Benefit	Every response is correct and precise	If multiple responses: half or more are correct and precise, but not all If one response: correct but imprecise	If multiple responses: at least one response but less than half are correct and precise
	Detriment	Every response is correct and precise	If multiple responses: half or more are correct and precise, but not all If one response: correct but imprecise	If multiple responses: at least one response but less than half are correct and precise

[1]Ignore non-behavioral relations, if any are mentioned