Designing an Assessment of Systems Thinking Skills Using the Context of COVID-19

Eric A. Kirk[1], Troy D. Sadler[1], Li Ke[1], Laura Zangori[2]

[1] School of Education, University of North Carolina at Chapel Hill

[2] College of Education and Human Development, University of Missouri

Paper presented at the Annual Meeting of NARST

Vancouver, BC, March 2022

Designing an Assessment of Systems Thinking Skills Using the Context of COVID-19

The COVID-19 pandemic is an archetypical socio-scientific issue (SSI): a societal dilemma that is entangled with science practices and scientific knowledge (Zeidler, 2014). As partisan debates about masking, business closures, and vaccinations embroil our population, it has become abundantly clear that there is a need to better support public understanding of the interactions between science and society and better develop our populations' ability to discern non-immediate consequences of actions (and inaction). Decision-making in socio-scientific issues is often a high-stakes affair. Conflicting interests between diverse stakeholders preclude straight-forward solutions based on simple, linear cause-effect reasoning. Individuals must recognize the inherent complexity of these issues and consider the complex interactions between the entangled components of the systems they operate within should they wish to predict behaviors, and design solutions that minimize unintended consequences (Sadler et al., 2007).

It is important for educators and researchers to be equipped with the tools that allow us to understand how students approach these types of issues. Without an understanding of the cognitive processes involved in navigating SSI, educators and researchers face an uphill battle if they wish to deliver instruction that is developmentally appropriate and targets specific skills we wish to develop. Researchers and educators alike need access to assessment tools that can help them uncover student knowledge to inform practice. This paper outlines the design decisions involved in designing an assessment that aims to fill this need.

Design Motivation

The complex and interdisciplinary nature of socio-scientific issues favors an approach that emphasizes the relationships and interconnectedness of the many facets of these issues. To better support this, Ke and colleagues (2021) advocate for the increased use of modeling in SSI-based curriculum. Using the term “socio-scientific models” to refer to models that account for both scientific and social factors, they posit that these models have the potential to be particularly useful in helping students negotiate complex societal issues; aiding students in drawing connections between scientific knowledge and relevant social dimensions while engaging in decision-making regarding possible solutions to these issues.

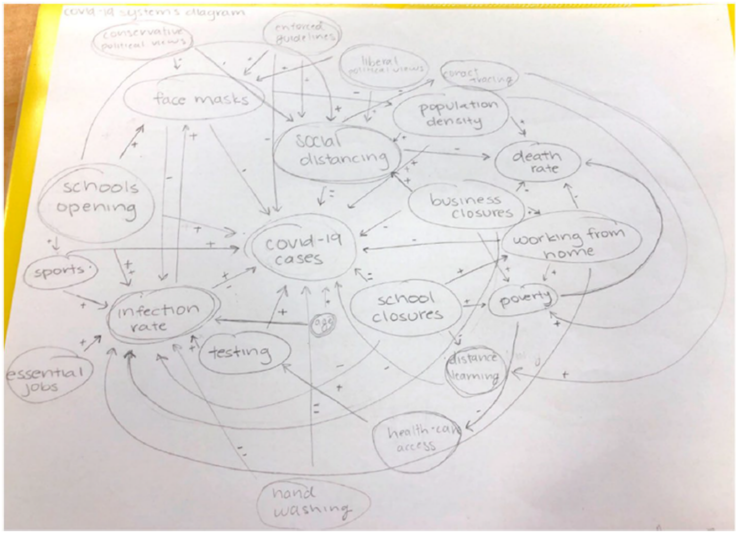
What sets these models apart from traditional scientific models is the inter-disciplinary nature of socio-scientific models. Whereas scientific models only seek to explain scientific phenomena and rely on scientific evidence, socio-scientific models incorporate knowledge from social domains such as the economic, historical, or political dimensions tied to the scientific phenomena. For example, a scientific model that represents a fishery collapse would focus on the unfolding ecosystem dynamics (e.g., predator/prey relations and water quality measures). A socio-scientific model may expand upon the scientific model by also incorporating the economic impact on the local fishing industry as well as relevant laws and regulations that dictate how many fish may be harvested, illustrating how these factors and the ecosystem dynamics shape one-another.

Socio-scientific System Models

Although socio-scientific models can take many forms, one form that may be particularly helpful is a system model (see Figure 1). System models are readily used to understand phenomenon encountered in science education (e.g., biogeochemical cycles and food webs). These models represent components that make up a system and the relationships that exist between these components, allowing the behaviors of a system to be better understood. Important system components are represented within a labeled circle, and relationships between factors are conveyed using arrows running between two or more interrelated factors. Socio-scientific system models differ from other system models traditionally encountered in science classes, however, in that they focus not only on the material conditions of a phenomenon, but also the interplay of these material conditions with society. Socio-scientific system models can explicitly represent the relationship between the scientific and social factors of the issue in question. Students often struggle to identify complex causal relationships such as domino causality, feedback effects, non-obvious causes, and distributed causality (Grotzer, 2012), all of which can be immensely consequential to the behavior of a complex socio-scientific issue. By explicitly identifying and representing the causal relationships of a complex system, students may be in a better position to navigate the complexity of these issues as they work to understand system behaviors, predict system changes, or design interventions to achieve a desired outcome.

Figure 1

Example of a Student Developed COVID-19 Pandemic Socio-Scientific Model



Note: This figure originally appeared in Ke et al. (2021)

Systems Thinking

For students to use socio-scientific system models to navigate SSI such as the COVID-19 pandemic, students must possess systems thinking (ST) skills: skills that support an ability to understand and interpret complex systems (Evagorou et al., 2009). These ST skills are applicable across different contexts, functioning as scaffolds that support student thinking about the specific context being investigated (Yoon, 2018). ST skills allow students to consider solutions to complex problems in ways that may not be possible when relying on simple, linear causal reasoning. This can help minimize the likelihood of unpredicted or unwanted outcomes by considering the problem holistically (Mehren et al., 2018). The science education community has recognized the value of these skills, underscoring their importance by including "systems and system models" as a crosscutting concept in the United States' *Next Generation Science Standards* (NGSS Lead States, 2013).

Systems thinking skills can be classified as domain-general or domain-specific depending on their transferability across contexts; both are important. Although researchers have presented collections of domain-general skills that vary both in numbers and specifics (c.f., Ben-Zvi Assaraf & Orion, 2010; Mehren et al., 2018), these skills find common ground around the ability to identify components and processes that constitute a system, to understand dynamic relationships among the components within the system, and to organize these components into a usable framework to explain and predict behavior (Yoon, 2018). Even though complex systems are regularly found in scientific and social settings that we frequently navigate, these skills have been shown to be incredibly difficult to develop, requiring significant changes to cognitive structures like personal epistemologies and ontologies (Jacobson & Wilensky, 2006; Wilensky & Jacobson, 2014), as well as schema used to understand causation (Grotzer, 2012; Jacobson & Wilensky, 2006; Wilensky & Jacobson, 2014).

Despite the cross-cutting nature of domain-general skills, these skills are not applied independently of domain-specific knowledge, they must be placed in conversation with system specifics. For example, domain specific knowledge in ecology such as predator-prey relationships impacts a students' ability to make use of domain-general skills like identifying organizational features and predicting system behavior of a food web (Mambrey et al., 2020). For COVID-19, the domain-specific skill of understanding how the SARS-CoV-2 virus is transmitted between individuals is necessary to accurately predict system behaviors or design solutions that minimize transmission, both being domain-general skills.

The importance of systems thinking and socio-scientific issues-based education have been well recognized. Unfortunately, educators and researchers lack easy to administer, evidence-based tools to assess student thinking as they make sense of socio-scientific systems. The purpose of this paper is to present the design process of an instrument that assesses three key skills for student systems thinking in the context of COVID-19. What follows is an overview of the design process. Next, we discuss the literature that we used to shape major design choices embodied by this assessment. We then go on to discuss the design goals, constraints, and key design decisions that emerged throughout the design process that were included in the final product. This paper concludes with a discussion of the limitations and significance of this assessment.

Systems Thinking Framework

Although there have been several instruments developed in recent years to assess students' ability to understand complex systems (e.g., Grotzer et al., 2016; Mehren et al., 2018), the assessment developed by Mambrey and colleagues (2020) most closely aligned with our intended design goals. Before addressing specific design challenges and our approach to addressing these challenges, we discuss the assessment framework we adopted to guide our design as well as features of the assessment that served as a model for our assessment. This background information provides context pertinent to our design rationale in the following section.

Mehren and colleagues (2018) outline a framework for assessing ST ability that identifies three core skills necessary for analyzing systems. First, students must be able to identify the components of the system in question, as well as understand how those components are organized in relation to one-another (system organization, SysOrg). Second, students must understand how systems behave when a system component is modified (system behavior, SysBeh). Finally, students must be able to manipulate a system to achieve a desired outcome (system-adequate intention to act, referred to as system modeling in this paper, SysMod). These skills closely resemble those identified by Ben-Zvi Assaraf and Orion (2005; 2010) in their hierarchical conception of systems thinking skills, providing further empirical support for the selection of this framework. The assessment developed from this framework by Mehren and colleagues assessed systems thinking using qualitative and quantitative items in the context of geography, featuring systems that include both social and scientific factors.

Competence in each of these skills is determined by the ability to correctly respond to items that vary both in the complexity of causal relationships, but also the structural complexity of the system itself. In Stage 1, students answer questions about simple, direct relationships (X influences Y) in systems defined by linearly connected factors that are not cross-linked. For Stage 2, students evaluate more complex systems that present linear but indirect relationships (X influences Y, and Y influences Z, therefore X influences Z). Finally, students encounter the most complex systems in Stage 3; they analyze complex, indirect relationships. These relationships differ from those featured in Stage 2 in that there are multiple pathways between two factors that must be considered rather than one direct path (e.g., W influences X and Y, X influences Y and

Z, Y influences Z). Features of systems that determine structural complexity include the number of system components, the number of connections between system components, and ways in which system components are connected. A structure index calculation is used to calculate a system's structural complexity (Mehren et al., 2015).

Mambrey and colleagues (2020) used this framework in the design of an assessment to evaluate ecology systems thinking skills in German 5th and 6th grade biology students using single-select, multiple-choice items. Although the test developed by Mambrey and colleagues is designed for a different population and utilizes a different anchoring phenomenon, we felt that this assessment served as a helpful model for our COVID-19 systems assessment because it satisfied many of the pragmatic and theoretical design constraints for this project.

Ultimately, the assessment framework detailed in Mambrey and colleagues' paper provided us with a matrix of question possibilities that we drew upon as we designed the Covid systems thinking test. Our goal was to create six different system models spread across three stages of difficulty. Each stage features two distinct system models. Although each system model was unique, we feel this framework ensured that this test was structured such that student systems thinking ability can be assessed in a clear, reliable, systematic way.

Design Process

After discussing the intended application and constraints our assessment should consider, we began a review of the literature on systems thinking assessments. Based on the available literature and the constraints we previously identified, the design team made the decision to model our assessment after the assessment developed by Mambrey and colleagues (2020) as it seemed best suited to our goals. Next, we explored areas of alignment and difference between the features and behaviors of the socio-scientific system we wished to use as context, and the systems used by Mambrey and colleagues. Prototype items and socio-scientific system models were then developed and critiqued by team members, leading to a refined set of items that were used to guide the development of the remaining items.

Once initial iterations of all 19 items had been developed, the test was reviewed and critiqued by our design team. Changes and recommendations were incorporated into the assessment. The test was then reviewed by two educators who were not familiar with the assessment. These educators' feedback was then incorporated into the next iteration. The resulting assessment was then administered to a small sample of high-school students in a school located in the Midwest United States (n=34) via Qualtrics to identify any glaring issues such as problematic items or difficulties in administering the assessment. These students were selected out of convenience; they had already given consent to participate in the larger study this assessment was developed for.

Final Design

The resulting assessment asks students to analyze six systems across five different levels of complexity. Our assessment comprises 19 multiple-choice, single-select items. Each item is designed to assess one of three specific systems thinking skills. Items are to be scored dichotomously; therefore, the maximum score is 19. This assessment features systems with varying levels of structural complexity, ranging from simple, linear systems composed of five factors up to highly interconnected systems featuring up to eight factors. Three items accompany each system such that all three skills were tested for each system. The final, most complex system, features a fourth item resulting in a total of 19 items.

Design Considerations

Throughout the design process, several features of our assessment emerged as critical to its success. There were constraints imposed by the context that shaped the forms this test could take. These constraints provided non-negotiable boundaries to which the design had to conform. Within these boundaries, we also made decisions that helped achieve our ultimate goal as researchers: better understanding students' systems thinking skills in the context of a socio-scientific issue. The following section presents a discussion of significant challenges, as well as the resulting design decisions made to address them. Furthermore, this section illustrates why the choice of the assessment framework detailed above was best suited to our needs. A brief summary of these challenges and features can be found in Table 1.

Table 1

Summary of Design Challenges and Features

Challenge	Design Feature
Large Scale Implementation with Partner Teachers	Fewer than 20 multiple-choice items
Adaptable for future pandemics.	A focus on domain-general skills
Understanding systems thinking	Systems thinking framework
Accounting for complex system relationships	Identifying causal relationships Focusing on readability Explicitly identifying relationship directionality
Accounting for variations in prior knowledge	Embedded content supports

Large Scale Implementation with Partner Teachers

Our instrument was developed to be used for a large-scale, multi-year study on the use and integration of multiple models, including systems models, in SSI based units. We intend this assessment to be administered to large sample sizes during instructional time. Because of this, it was important to design the assessment in such a way that it could be administered easily by teachers, and graded rapidly. Additionally, it was important to protect teachers' instructional time; something teachers often express concerns about given the amount of time it takes to implement an SSI unit (Ekborg et al., 2013; Tidemand & Nielsen, 2017).

Design Decision: 20 Multiple Choice Items

To address these challenges, we made the decision to use only multiple choice, single-select questions for our assessment. We also imposed a constraint that limited test length to 20 items or less. These decisions address the aforementioned concerns in several ways. First, multiple choice items are easily adapted to online formats such as Qualtrics, allowing the test to be rapidly disseminated and collected by researchers without unnecessary time investment by teachers, thus protecting our partner teachers' instructional time. Imposing a limit of 20 items also supported the goal of protecting instructional time, ensuring that the amount of class time taken to administer the test is kept to a minimum, again protecting our teachers' instructional time. Second, by delivering multiple choice items through Qualtrics, researchers and teachers alike are saved the process of manually scoring items. This provides teachers with rapid feedback on their students' performance while also expanding our capacity to work with larger sample sizes. The test developed by Mambrey and colleagues (2020) demonstrates that it is indeed possible to assess systems thinking skills through a multiple-choice assessment using the framework described above.

Adaptable to Future Pandemics

The project for which this assessment was developed focuses on supporting student learning about global pandemics caused by respiratory viruses such as COVID-19 through the design of curricular materials (Sadler et al., 2021) and research. Although COVID-19 is the specific anchoring phenomenon, the project this assessment was developed for initially entailed the development of materials that could be adapted for use during future viral pandemics. Thus, we felt the design of the assessment should be flexible enough to be adapted to future contexts with minimal effort and minimal threat to item integrity. Should another pandemic arise, practitioners and researchers will have access to curriculum and instruments that support the teaching and learning of these topics, allowing teachers to focus on addressing concerns that arise from challenges presented by the unfolding situation.

Design Decision: A Focus on Domain General Skills

Focusing on domain-general skills such as identifying structural features of a system or predicting system behaviors increases the versatility of an assessment, allowing these assessments to be adapted to future contexts through the modification of domain-specific details without requiring extensive modification of the deep structure of the assessment. The work done by Mambrey and colleagues (2020), and Mehren and colleagues (2018) shows that the domain-general skills identified in the underlying framework used for this assessment are transferrable to multiple contexts, including socio-scientific systems. Because this framework has been applied successfully to multiple systems, it stands to reason that the constructs identified in this framework are robust enough to withstand the minor modifications we anticipate being needed when adapting this assessment to a future pandemic or other emergent, socio-scientific crises.

Additionally, Mambrey and colleagues effectively designed two, parallel tests that were administered at the same time. One test was based on an aquatic habitat, whereas the other was based on a terrestrial habitat. Mambrey and colleagues changed the organisms portrayed in their food webs; however, the structures of the systems remained unchanged between habitats suggesting that it is possible to modify the factors portrayed on our system models without making significant changes to the system's structure. The parallel nature of their test items is relevant to our design rationale as it demonstrates the potential for this test to be adapted across multiple systems that share a similar underlying structure without threatening the assessment's validity.

It stands to reason that future viral pandemics will share many features with the currently unfolding pandemic. For example, although there are many factors that predict COVID-19 vaccine uptake in the United States; political affiliation is one of the most significant (Milligan et al., 2021). Should political affiliation not predict vaccine uptake in a future pandemic, this system factor could be replaced with a predictive factor that is more applicable without modifying the underlying structure of the system or the nature of the questions asked of students. Mambrey and colleagues' (2020) adaptation of items across aquatic and terrestrial habitats models this type of adaptability nicely.

Understanding Systems Thinking

We sought to design the assessment with varying degrees of system complexity, task difficulty, and task type. We felt it was important that these variations occur in deliberate ways, following specific rules, so that we can better explore differences that emerge in domain-general skills across students as they engage with socio-scientific systems. To ensure that questions are designed in a systematic manor, we sought an organizing framework to guide item development.

Design Decision: Systems Thinking Framework

The skills that comprise the framework developed by Merhen and colleagues (2018) (i.e., SysOrg, SysBeh, SysMod) are domain general skills that have been used to guide a number of assessments and have been applied across varying contexts. This framework provided a consistent structure to help us design our questions. By specifying these three skills, we were able to standardize the language used to assess each of these skills across assessment items. By standardizing language, we can more confidently attribute variations in student responses across items to variations in ability, rather than interpretation of the items. Likewise, by naming these three skills, we were able to design questions specifically to test one skill at a time, thus enabling us to see differences in individual skills that support systems thinking.

Accounting for Complex System Relationships

Although many of the features of Mambrey and colleagues' test were suitable for our needs, significant changes needed to be made to individual items. Unsurprisingly, the most significant changes were made to the systems themselves, although there were noteworthy changes made to the questions as well. Notably, there is no uniform "predator" and "prey" heuristic that guides the design and interpretation of the COVID-19 system. Likewise, unlike the predator/prey relationships and the flow of energy depicted in a typical food web, the relationships in the COVID-19 pandemic system are not as predictably directional.

Design Decision: Identifying Causal Relationships

Mambrey and colleagues assessed a student's ability to analyze SysOrg by asking them to identify predator-prey relationships. These relationships are non-existent in our system of interest: therefore, we needed to reconsider how these items should be designed. Rather than focusing on predator-prey relationships, we focused on causal relationships more broadly. Changes to upstream items (causes) drive changes to items downstream (effects). For these items we did not ask students to identify what changes would occur, simply which factors would impact one another. Causality was represented in systems using arrows, with arrows leaving downstream causes and pointing towards upstream effects.

Design Decision: Focusing on Readability

Whereas energy always flows from prey to predator as it ascends the trophic pyramid, relationships in the complex socio-scientific system we aimed to assess do not always follow a common heuristic. To account for this we abandoned the strict, hierarchical organization featured in Mambrey and colleagues' test. Our system does not feature an intuitive directionality, thus, we abandoned the "bottom to top" organizational structure in favor of structures that presented the most user-friendly visual representations of our systems, ensuring that connecting arrows did not overlap or intersect as this may cause unnecessary confusion. By prioritizing readability during system design, we hope to improve the reliability of our assessment, decreasing the likelihood of errors caused by misinterpretation of system features and behavior caused by designs that prioritized other organizational features.

Design Decision: Explicit Identification of Relationship Directionality

The relationships in our system of interest are not always correlated in easily predictable ways. Whereas interactions commonly portrayed between trophic levels in food webs follow predictable rules (e.g., an increase in energy availability in lower trophic levels can support larger predator populations) the relationships exhibited in socio-scientific systems such as the COVID-19 pandemic do not align with a uniform set of governing rules. Thus, we found it important to include this complexity in our system models. This complexity was included through the explicit labeling of arrows with "+" or "—" to indicate whether a relationship is positively (if X increases so does Y) or negatively correlated (if X increases Y decreases). This feature also helped address another prominent challenge unique to this assessment: the unfolding, controversial nature of the relationships within the system and the large amount of misinformation and disinformation that can impact student content knowledge.

Accounting for Variations in Prior Knowledge

Because of our specific interest in students' abilities to interpret and use system models, we found it necessary to design this test to minimize the impact content knowledge could have on our results. As mentioned previously, our goal was to better understand domain-general skills, skills that can be applied across system contexts. Because these skills have been shown to be influenced by system specific knowledge (Mambrey et al., 2020), varying levels of content knowledge and exposure to misinformation could obfuscate the patterns in the application of domain-general skills we are ultimately interested in uncovering.

Design Decision: Embedded Content Supports.

In an effort to minimize these confounding factors, we provided content supports for students directly within the assessment. To encourage students to rely on the information we provide, directions and items explicitly instructed students to base their answers using only the information presented in the model being displayed. We felt this to be an acceptable approach because models are inherently over-simplifications of a phenomenon, and our aim was to assess how students interpret and use models presented to them – not their conceptual understanding of the COVID-19 pandemic.

Second, the models were designed to make explicit not only the connections between factors, but also whether these connections were positively or negatively correlated. We accomplished this by labeling arrows to represent positively and negatively correlated relationships. These are the same conventions used by the curriculum materials designed as a part of this project (Sadler et al., 2021), ensuring alignment between resources should students be exposed to the curriculum prior to taking this assessment. We also created a splash page with a diagrammatic and text-based description of these conventions that students must click through to begin the test. This page communicates these conventions to students who are not already familiar with them, or may have forgotten them.

We acknowledge the impact content knowledge and misconceptions can have on student systems thinking performance (Mambrey et al., 2020, 2022). The measures described above are designed to act as content supports for students with varying levels of exposure to these ideas. For example, whether a student rightly believes that masking is an effective way to manage infection rate, the model they are using explicitly specifies this relationship. Students correctly interpreting the model and following the directions provided should answer based on this information, not their prior knowledge. Despite this support, it is possible that students may not follow this assumption. This represents a possible direction for future research

Limitations

As with any test, there are limitations. Although a multiple-choice assessment is limited in its ability to capture a full range of student thinking and modeling practices; we found this limitation to be acceptable for two reasons. First, this test aims to understand how students interpret and understand system models, not construct them. Second, this limitation is directly addressed by our intent to collect observation data of student work. Supplementing these quantitative findings with qualitative data allows us to make inferences not possible with either methodology alone. Ideally, the data generated by this test and through observational work will be placed in conversation with one another, providing us with a richer understanding of the interactions between the practices involved in the creation of models and the skills needed to interpret models.

Another notable limitation of this assessment stems from our focus. By focusing our efforts on the three systems thinking skills, we inherently constrained the types of student knowledge we could understand through this instrument. The static presentation of systems in this test limits our ability to understand how students conceive of many features of complex systems and causal relationships that have temporal components like steady-states and simultaneous causality (Grotzer, 2012). Also, our test deliberately omitted probability, and magnitude from the relationships depicted in the models. Although complex systems rarely operate in a sequential, deterministic, “all or nothing” fashion, we felt these understandings would be best assessed in a more naturalistic setting. Because this test is to be paired with student observations and interviews, and the unit this test is situated in contains a computational modeling component that could support student exploration of these ideas, we were willing to accept this limitation for our purposes. Likewise, incorporating these features would require students to engage in mathematical calculations, introducing mathematical ability as a confounding variable and dramatically increasing the amount of time necessary to administer the test – a burden we were not willing to impose on our partner teachers.

Finally, the extent to which our embedded content supports are effective remains undetermined. We did not embed items that specified relationships that run counter to knowledge accepted by the scientific community to see whether students would rely on the information provided in these items or default to their background knowledge. We acknowledge the position of ourselves and our partner teachers as knowledgeable authorities on the pandemic and feel that it is our obligation to ensure that our research does not inadvertently contribute to misconceptions or misinformation that pose threats to the health of students and their communities.

Conclusion

Although SSI-based instruction makes a meaningful step in developing students' ability to think about the social dimensions of scientific issues, more targeted interventions are needed to support student reasoning as they grapple with highly complex issues such as the COVID-19 pandemic. The use of socio-scientific modeling in the form of system models may be a productive tool in supporting student decision making in SSI. Unfortunately, however, research on the use of systems assessments in SSI-based instruction is limited, as are instruments that can be used to assess the efficacy of this kind of instruction. This paper presents the rationale that drove the design of an instrument with the explicit intent of helping fill this gap in the literature. Data generated by this assessment may be used to support research and curriculum development, potentially yielding new insights into the efficacy of instructional interventions, ultimately leading to higher quality instructional materials.

Should the science education community wish prepare our students to confront the grand challenges of their lifetime, we must strive to develop our students' ability to apply science to problems in ways that account for the inherent complexity of our world. It is becoming increasingly obvious that these problems demand a systemic approach should we wish to solve them. Ethical applications of scientific knowledge should consider the broader social impacts of their actions, as science does not operate within a vacuum. As we have seen throughout this pandemic, a failure to look beyond the initial obvious impacts of an action has the potential for loss of livelihoods and life.

References

Assaraf, O. B.-Z., & Orion, N. (2010). System thinking skills at the elementary school level. *Journal of Research in Science Teaching*, *47*(5), 540–563. https://doi.org/10.1002/tea.20351

Ekborg, M., Ottander, C., Silfver, E., & Simon, S. (2013). Teachers' Experience of Working with Socio-scientific Issues: A Large Scale and in Depth Study. *Research in Science Education*, *43*(2), 599–617. https://doi.org/10.1007/s11165-011-9279-5

Evagorou, M., Korfiatis, K., Nicolaou, C., & Constantinou, C. (2009). An Investigation of the Potential of Interactive Simulations for Developing System Thinking Skills in Elementary School: A case study with fifth-graders and sixth-graders. *International Journal of Science Education*, *31*(5), 655–674. https://doi.org/10.1080/09500690701749313

Grotzer, T. A. (2012). *Understandings of Consequence: Learning Causality in a Complex World*. Rowman & Littlefield Education.

Grotzer, T. A., Tutwiler, M. S., Kamarainen, A. M., Derbiszewska, K. M., Metcalf, J., & Dede, C. J. (2016). Students' Reasoning Tendencies about the Causal Dynamics of Ecosystems and the Impacts of MUVE vs. Non-MUVE Instructional Contexts. *In The Next Phase of Research in Complex Systems in Science Education*, 29. https://clic.gse.harvard.edu/selected-publications

Jacobson, M. J., & Wilensky, U. (2006). Complex Systems in Education: Scientific and Educational Importance and Implications for the Learning Sciences. *Journal of the Learning Sciences*, *15*(1), 11–34. https://doi.org/10.1207/s15327809jls1501_4

Ke, L., Sadler, T. D., Zangori, L., & Friedrichsen, P. J. (2021). Developing and Using Multiple Models to Promote Scientific Literacy in the Context of Socio-Scientific Issues. *Science & Education*. https://doi.org/10.1007/s11191-021-00206-1

Mambrey, S., Schreiber, N., & Schmiemann, P. (2022). Young Students' Reasoning About Ecosystems: The Role of Systems Thinking, Knowledge, Conceptions, and Representation. *Research in Science Education*, *52*(1), 79–98. https://doi.org/10.1007/s11165-020-09917-x

Mambrey, S., Timm, J., Landskron, J. J., & Schmiemann, P. (2020). The impact of system specifics on systems thinking. *Journal of Research in Science Teaching*, *57*(10), 1632–1651. https://doi.org/10.1002/tea.21649

Mehren, R., Rempfler, A., Buchholz, J., Hartig, J., & Ulrich-Riedhammer, E. M. (2018). System competence modelling: Theoretical foundation and empirical validation of a model involving natural, social and human-environment systems. *Journal of Research in Science Teaching*, *55*(5), 685–711. https://doi.org/10.1002/tea.21436

Mehren, R., Rempfler, A., Ulrich-Riedhammer, E. M., Buchholz, J., & Hartig, J. (2015). Wie lässt sich Systemdenken messen? Darstellung eines empirisch validierten Kompetenzmodells zur Erfassung geographischer Systemkompetenz / How to measure systems thinking? Presentation of an empirically validated competence model for assessing geographic system competence. *Geographie aktuell & Schule*, *37*(215), 4–16.

Milligan, M. A., Hoyt, D. L., Gold, A. K., Hiserodt, M., & Otto, M. W. (2021). COVID-19 vaccine acceptance: Influential roles of political party and religiosity. *Psychology, Health & Medicine*, *0*(0), 1–11. https://doi.org/10.1080/13548506.2021.1969026

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States* (p. 18290). National Academies Press. https://doi.org/10.17226/18290

Sadler, T. D., Barab, S. A., & Scott, B. (2007). What Do Students Gain by Engaging in Socioscientific Inquiry? *Research in Science Education*, *37*(4), 371–391. https://doi.org/10.1007/s11165-006-9030-9

Sadler, T. D., Rawson, R., Kirk, E. A., Elsner, J., Ke, L., Apple, S., Elmey, C., Kinslow, A. T., McKee, R., Miller, D., Plato, J., Rockett, J., Wagner, B., Friedrichsen, P. J., & Zangori, L.

(2021). *COVID-19: A model-oriented issues-based science unit*. University of North Carolina at Chapel Hill School of Education. https://epiclearning.web.unc.edu/covid/

Tidemand, S., & Nielsen, J. A. (2017). The role of socioscientific issues in biology teaching: From the perspective of teachers. *International Journal of Science Education*, *39*(1), 44–61. https://doi.org/10.1080/09500693.2016.1264644

Wilensky, U., & Jacobson, M. J. (2014). Complex Systems and the Learning Sciences. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (2nd ed., pp. 319–338). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.020

Yoon, S. A. (2018). Complex Systems and the Learning Sciences. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International Handbook of the Learning Sciences* (1st ed., pp. 157–166). Routledge.

Zeidler, D. (2014). *Socioscientific Issues as a Curriculum Emphasis: Theory, Research and Practice.* (pp. 697–726).