Instructional Science (2022) 50:571–597
https://doi.org/10.1007/s11251-022-09586-5

ORIGINAL RESEARCH

Concept map as a tool to assess and enhance students' system thinking skills

Mojtaba Khajeloo[1] · **Marcelle A. Siegel**[1,2]

Received: 25 April 2021 / Accepted: 5 March 2022 / Published online: 12 April 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Concept map (CM) is introduced as a useful tool for studying students' system thinking (ST). However, it is more known to represent students' knowledge of system components and organization and less recognized as a tool to examine and enhance students' understanding about the underlying causal mechanisms in complex systems. In this study, through a mixed method approach, we investigated the potential of CM in demonstrating undergraduate students' (n = 173) ST. We also conducted a comparative analysis to examine the effects of different scaffolding on developing students' ST skills. Through a theoretical framework of causal patterns, we present a new perspective on what CM reveals about students' ST and what are its limitations in showing system complexities. The results indicated that CM can provide a platform for students to practice causal mechanisms such as domino, mutual, relational, and cyclic causalities, and accordingly, work as a tool for teachers to examine students' knowledge of such mechanisms. The results also showed that students improved in demonstrating ST by CM when they were scaffolded for showing causal mechanisms and building CM. Eventually, this study concludes that the CM is a highly relevant tool to increase and examine students' ST skills. To this end, we found it is important to explicitly teach students about causal patterns and guide them to construct CM with an emphasis on showing the interconnection among concepts.

Keywords Concept map · System thinking · Scaffolding system thinking · Complex system education · Causal mechanisms

Introduction

Every day, we encounter complex problems ranging from disease epidemics and environmental issues to novel challenges raised by social media. Addressing such problems requires system thinking (ST) skills. ST is the ability to conceptualize complex systems and problems and involves understanding the dynamic interrelationship between system

✉ Mojtaba Khajeloo
mk4rf@umsystem.edu

1 Department of Learning, Teaching, and Curriculum, University of Missouri, Columbia, MO 65211, USA

2 Department of Biochemistry, University of Missouri, Columbia, USA

components and the patterns and behaviors emerging from them (Hammond, 2017; Meadows, 2008).

The significance of ST has been well recognized among science education researchers over the past two decades (Yoon et al., 2018) and is accentuated in science education documents internationally (Boersma et al., 2010; National Research Council, 2010; NGSS Lead States, 2013). There are many benefits in embedding complex systems and ST practices in science education. It is found that engaging students in thinking about complex systems improves their science literacy (Assaraf & Orion, 2005; Ke et al., 2020; Sabelli, 2006). Jacobson et al. (2017) relate the value of learning about complex systems "both to the importance of these ideas in modern science as well as for the potential of complexity ideas to provide conceptual interconnections across different science subjects as a new perspective about scientific literacy'' (p. 1). Furthermore, the instructors who applied the system approach in their teaching observed that students gained a deeper understanding of content (Mathews et al., 2008; Verhoeff et al., 2008) and made more interdisciplinary connections between concepts (Fisher, 2011, 2018; Jacobson & Wilensky, 2006).

Science education researchers have proposed various methods to foster students' knowledge about complex systems. For example, explicit instruction on system characteristics is introduced as a proper approach to improve students' ST competencies and develop their system language (Hmelo-Silver et al., 2007; Jordan et al., 2013; Tripto et al., 2016, 2018). Considering ST as a metacognitive skill, Verhoeff et al. (2018) indicated that system characteristics should be used as metacognitive tools to help students practice ST throughout their learning trajectory. Gilissen et al. (2020) came up with four design guidelines in teaching to improve students' ST: (1) introducing the seven characteristics of systems in relation to system theories; (2) providing students opportunities to apply system characteristics in different contexts; (3) focusing on individual characteristics at a time; and (4) attending to system language and encouraging students to use them. Regardless of the teaching approaches, an appropriate assessment tool is needed to examine students' skills in ST. Assessing traditional learning is often challenging enough, while modern notions of science teaching, such as ST, require creative assessment designs (Abell & Siegel, 2011; Izci et al., 2020). Further, effective learning assessments should not just elicit and monitor learning but also foster and assist growth (Izci et al., 2020; Siegel et al., 2014; NRC, 2014; Pellegrino, 2013).

In these respects, the concept map (CM) has been suggested as a helpful tool (Sommer & Lücken, 2010). CM is a graphical tool depicting the relationship of concepts within a domain of knowledge and is used to organize and represent the structure of conceptual knowledge (Novak & Gowin, 1984; Novak & Cañas, 2008). Considering that the structure of knowledge is an indicator of the quality of understanding (Mintzes et al., 2005), many researchers have found CM a suitable tool to assess students' knowledge construction (Baxter et al., 1996; Edmondson, 2005). CM is indeed an external representation of one's mental model and entails concepts (nodes) connected to each other by labelled lines, in each case building a proposition (Yin et al., 2005). Propositions are regarded as the building blocks and meaningful units of CM that allow an individual to express their knowledge of a content area by them (Novak & Musonda, 1991; Yin et al., 2005). In other words, CM is a particular way of constructing and representing knowledge in which a concept finds its meaning in connection with other concepts, and the final product is a visual representation of students' cognitive structures (Nesbit & Adesope, 2006). CM building is often associated with a "constructivist" view of learning and its practice helps students integrate new information to their prior knowledge and visualize their understandings of a content area, as well as helps teachers become informed about the integrity of students' understandings in a domain of knowledge (Brandstädter et al., 2012; Conradty & Bogner, 2012; Schwendimann, 2015).

Johnson-Laird (2001) discussed the ways mental models can be informative about one's ability to do problem solving in complex systems. Finding the CM as an external representation of one's mental model, some researchers found CM an informative tool to study students' ST. For example, Tripto and her team used CM to externalize and analyze students' conceptual knowledge of the human body to draw conclusions about students' ST skills (Assaraf et al., 2013; Tripto et al., 2016, 2017, 2018). Furthermore, Buckley and Boulter (2000) found CM building a helpful approach for teaching and assessing students' understanding of multilevel structures such as complex and nonlinearly organized biological systems. It was also reported that CM provided insight into students' thinking and revealed the processes by which students constructed their cognitive structures (Hay et al., 2008; Ifenthaler, 2010; Shavelson et al., 2005).

Despite its benefits in ST studies, CM has been used more to represent a system's components and their organization and less to represent a system's underlying processes and mechanisms. For example, Tripto et al. (2013) examined the effectiveness of CM as a tool to assess students' ST skills. They found that students' CM emphasized a system's structural components more than its processes. Similar findings were evident in other studies as well; students tended to demonstrate the components of a system on their CM with little understanding about the interactions between system components (Hmelo-Silver & Azevedo, 2006; Hmelo-Silver & Pfeffer, 2004). Tripto et al. (2013) also noted that students had difficulty in demonstrating the underlying mechanisms and dynamic natures of a system by CM; they found CM to be a static representation of students' conceptual knowledge about systems and not a proper tool to display higher-order ST skills such as expression of multistep simultaneous processes.

Students need scaffolding to improve in ST (Hmelo-Silver & Azevedo, 2006). Scaffolding refers to various forms of assistance (e.g. social, linguistic, conceptual) that can support students' learning, reasoning, and participation (Sawyer, 2006). The literature indicates that scaffolding assists students to increase and integrate their higher order thinking skills to generate solutions to complex problems (Belland et al., 2017). Scaffolding students' CM building improves their learning and reasoning skills (Eggert et al., 2017). Yet, despite the benefits of scaffolding in students' learning, there are scarce studies that examine how scaffolding students' construction of CM improves their ST skills.

In this paper, we examine CM as a tool to study ST and how scaffolding students' CM building influence the demonstration of their ST competencies. Through a theoretical framework of causal patterns, we present a new perspective on what CM reveals about students' ST skills and what are its limitations in showing system complexities. Plus, through a comparative scaffolding strategy, we demonstrate that, accompanied with a proper scaffolding, CM is a powerful tool to examine and promote students' knowledge about the underlying causal mechanisms in complex systems. To this end, we sought to answer the following research questions:

1. What does CM reveal about students' knowledge of biological complexities and the underlying causal mechanisms in them?
2. How does scaffolding influence students' performance in showing the causal mechanisms in their CM?

Theoretical framework

Thinking is a process in which individuals coordinate their inferences from their knowledge to address their needs such as solving a problem or making a decision (Moshman & Tarricone, 2016). Based on this definition of thinking, we define ST as a process to coordinate one's knowledge of a system to explain the behavior or function of a system. In this sense, an individual's awareness about the components of a system is a matter of their knowledge about the system, and ST is how they interrelate and interconnect the system components to make meaningful patterns to explain the system's behaviors and functions. Emphasizing the thinking skills of students in addressing complex biological phenomena, in this study, we provided our participants with the components of the systems. We attempted to reduce the cognitive load of our assignments by providing students the required knowledge of the task (i.e., system components) to have them primarily focus on examining the interconnections among the system components to build meaningful patterns to explain biological phenomena. Accordingly, the more students could make various meaningful patterns on their CM, the higher ST skill we considered for them. To study those patterns, we drew on the patterns of underlying mechanisms in complex systems introduced by Grotzer (2012).

In her book *Learning Causality in a Complex World*, Grotzer (2012) introduces six causal patterns that underlie complex causalities: simple linear, domino, cyclic, spiraling, mutual, and relational causalities. All these patterns exist in biological systems, but they are not taught explicitly and with equal emphasis in biology education – including the studied course in this research. Below, except for spiral causality, we explain these patterns of causal mechanisms and their role in determining features of complex systems and illustrate their existence in the context of this study.

Simple linear causality

According to Grotzer (2012), stemming from the works of David Hume (1739–1740), simple linear causality is based on three basic principles: (1) a cause precedes an effect; (2) there is a cause and effect mechanism for any outcome; and (3) for something to happen, a cause must exist (determinism) (Morris & Brown, 2001). These principles lead to the assumptions of linear causality: (1) cause and effect work in one direction, in the order of cause preceding effect; (2) there is a direct link between cause and effect with a clear beginning and an end; and (3) there is usually one cause for one effect. In sum, linear causality implies that anything that happens in a system can be traced back to a cause in the same system, suggesting that a sequence of cause and effect explains behaviors of a system. For example, in our study, the reasoning pattern that natural selection causes the extinction of an organism is a form of simple linear causality.

Domino causality

Domino causality refers to a linked sequence of events that unfold over a period of time and typically has a beginning, middle, and end. In domino pattern of causality, causes induce effects, and effects turn into causes for subsequent effects. It is one directional and can have branches; therefore, one cause can bring about multiple effects and so forth. Domino causality explains some aspects of complex systems, in which a cause triggers a chain of reactions that subsequently leads to unpredicted outcomes. Climatic events that induce a cascade of challenges and mutations that occur at a molecular level and then bring about changes at an organism level illustrate this pattern of causality.

Domino causality is one of the central mechanisms by which complex systems display their behavior. It explains how within the complex systems a small change can generate huge impacts and how a cause can have both direct and indirect effects. Accordingly, this practice of thinking requires individuals to consider that the impact of a cause can travel over time and space and possibly be out of their sight. For instance, thinking about the effects of a species introduction into an ecosystem, students should consider that the effects of such measures happen over time and can be multiple and, to an extent, unpredictable.

Mutual causality

Mutual causality is a pattern of causal relation in which two events, acts, or processes mutually or bi-directionally affect each other. Therefore, the factors or agents involved in this relationship operate as both cause and effect on each other in a way that it is usually impossible to tell which factor precedes the other. In mutual causal relationship, involved factors can both benefit, or get harmed, or one side benefit and the other side get harmed from the interaction, and the process can be simultaneous or sequential. Symbiotic relationships between two species in which at least one species benefits from the relationship illustrates a pattern of mutual causality. Some of the main features of complex systems are the abilities of self-organization and self-transformation. Mutual causality is one of the mechanisms by which scientists explain how systems and their components engage in their own organization and transformation (Morgan, 2006). For instance, the mutual relationship between prey and predator population in an ecosystem explains how that ecosystem engages in balancing its resources.

Relational causality

In relational causality, the effect is caused by the relationship, one of balance or imbalance, between elements of a system. The variables, events, or processes involved in this causal relationship cannot be a cause by themselves. Therefore, in exploring relational causality one looks for more than one variable for a cause and consequently examines the relationship between variables or sets of variables to argue for the effect. The relationship between the variables can be in balance or imbalance and the change in the relationship can influence the outcome. On the other hand, the proportional equal change in all variables in a relational causal relationship does not change the effect or outcome.

The natural world and consequently the complex systems exhibit many of their features and behaviors through the mechanism of relational causality between their components. For instance, the fitness of an organism in an ecosystem is determined out of the relationship between the phenotype of the organism and its environment; neither the environment nor the phenotype solely determine the fitness.

Cyclic causality

Cyclic causality refers to a pattern of relationship in which variables in an action, event, or situation are connected in a circle, and typically there is no clear beginning and end to them. In cyclic causality, an engaged component can be both cause and effect simultaneously. Also, there is a form of feedback loop or reciprocity in cyclic causal patterns that makes the cycle continue. The feedback loop can feed to maintain the status quo or amplify or play down the effect over time. Cyclic causality is one of the main mechanisms that demonstrates how complex system components are interlocked to each other; thus, all components in a system are of high significance and failure of one component can disrupt the cyclic pattern and continuation. It also explains the way complex systems are sustainable and continue to maintain themselves. The natural world is full of cyclic patterns that operate through cyclic causal mechanism. As an example, the relationship among evolution, environment, and organisms' phenotype and fitness explains how ecosystems maintain and develop biodiversity within themselves. Through the mechanisms of natural selection, evolution occurs over environments. The change, or lack thereof, in the environment affects the phenotype and fitness of the organisms in that environment. And again, through natural selection, organisms evolve that possibly change the environment. This cyclic process can either maintain the existing biodiversity within an ecosystem or develop new ones.

Methods

In this study, we used convergent mixed method design (Creswell & Plano Clark, 2011). This approach draws on both quantitative and qualitative data to compare and contrast their results to provide a complete understanding of a phenomenon and compare multiple levels within a system (Creswell & Plano Clark, 2011). Drawing on both quantitative and qualitative methods, this study provides a detailed insight into students' approach toward CM building to demonstrate the interrelationship among the components of biological systems. We also present the patterns of causal mechanisms existing in students' CM and the effect of scaffolding on their performance in showing those patterns.

Research design and participants

For this study, the first author developed a master CM (Appendix A) out of 16 big concepts taught during an introductory biology course. The concepts addressed different levels of biological organization from sub-organismic entities (e.g., DNA, RNA, protein) to supra-organismic entities like population. Then three variations of a CM assignment were designed. Providing different levels of scaffolding, the assignments asked students to demonstrate the effect of introducing a new predator species into an environment while showing the relationship between different biological concepts. The scaffolding was practiced as follows: assignment A provided students with the concepts to build CM; assignment B gave students the skeleton of the master CM to fill out the boxes by the given concepts and label the links; and assignment C was the same as assignment A, with the only difference that students were provided with scaffolding for the patterns of causal mechanisms at the beginning of the assignment (Appendix B). The CM building task was followed by three questions: (1) how students evaluated their CM in regard of expressing their thoughts on the scale of 1–5, with 5 being very satisfied; (2) what aspects of their thoughts they could express satisfactorily; and (3) what thoughts they could not express by their CM. The follow-up questions served two purposes. First, they provided us with students' general satisfaction in presenting their thoughts by each type of the CM assignment. This assessment helped us to know students' general expectations from a CM task. Second, they prompted students to reflect on their CM regarding how it allowed them to express their thoughts—or not. Accordingly, students' answers to the follow-up questions provided us a rich data on how CM, as an assessment platform, favored expression of some thoughts and disfavored some others. The assignment was given to students in Week 13 of the semester by which they were taught about all the concepts. The assignment was a take-home task, and students had two weeks to do it. Of note is that students were taught about CM building in a 50-min session in the first half of the semester, but the causal patterns were not explicitly taught throughout the semester.

Participants of this study were 173 students from an introductory biology course at a public university in the Midwest region of the USA. They were randomly put into three groups, and each group was assigned to a different variation of a CM assignment. We also interviewed 14 students (five students from each group of A and C, and four students from group B). They were recruited by sending out an email to the whole class and a $5 gift card was offered as an incentive for interview participation. The interviews were conducted in a semi-structured format and the students were asked to elaborate on the overall structure of their CM as well as the patterns that existed in them. Students were also prompted about the different (causal) relationships that could possibly exist among the given concepts and were given time to demonstrate them on their CM. The interviews lasted around 25 min and were audio recorded and later transcribed for detailed analysis.

The interview data was insightful in different ways. First, it revealed students' views toward the CM task and their previous experiences with it. This data was significant as it showed students' general approach toward CM building that favored expression of some thoughts over the others. Next, it allowed students to express their thoughts about the interaction between the given concepts and whether CM enabled them to demonstrate those thoughts. Last, it revealed students' knowledge about the causal patterns, providing insight into whether students who demonstrated little knowledge was due to the lack of knowledge about the patterns or lack of skill in constructing interactive CM.

Data analysis

Students' views expressed by CM: qualitative analysis

Qualitative analysis sought to examine students' overall view toward the relationship among biological components, and how they could express them by a CM. To this end, the first author conducted a qualitative inductive analysis (Thomas, 2006) that entailed students' CM, their explanations regarding the expression of their thoughts by the CM, and their interviews. He first carried out multiple readings and interpretations of raw data. Then, based on interpretations, he developed categories of thoughts and ideas students were capable and incapable to demonstrate by the CM. Finally, through the lens of our theoretical framework, the first author analyzed the data for the patterns and reasons that enabled or impeded students' expression of thoughts by the CM.

Students' CM performance: quantitative analysis

To conduct quantitative analysis, we developed a scoring scheme to examine the presence of causal patterns in students' CM. On this, we developed a list from combination of propositions that accurately demonstrated the patterns of domino, mutual, relational, and cyclic causalities. Figure 1 shows examples of proposition combinations used for scoring each

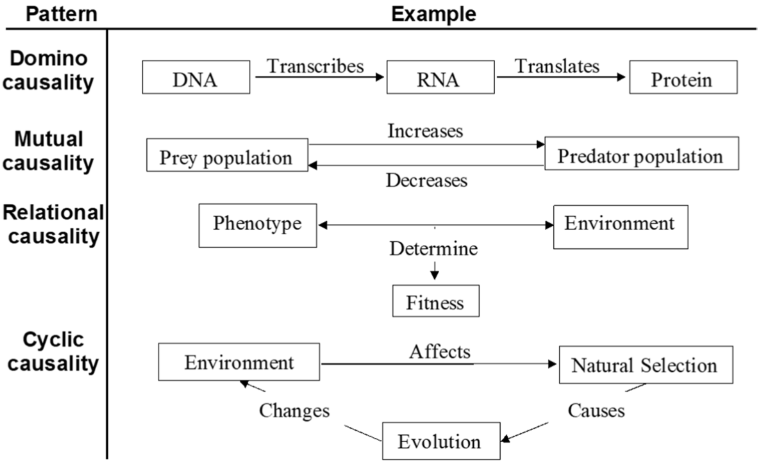


Fig. 1 Examples of proposition combinations making causal patterns

causal pattern. To make the list, first, we compiled all possibly accurate combinations of propositions based on our own knowledge (both authors have master's degrees in biological sciences) and the repeated patterns in students' CM. Then, we had the list edited and approved by two biology experts. We used this list as a reference to score the presence of different causal patterns in the CM. The causal patterns were scored as correct (2 points) if the concepts and describing labels were correct, partially correct (1 point) if the connected concepts were correct but without or with wrong labels, and incorrect (0 point).

This pattern-based scoring method enabled us to examine students' CM at a higher organizational level than the conventional method of scoring CM based on single propositions. Indeed, it framed students' CM in bigger chunks of meaningful propositions, and accordingly, reduced the disorderliness of the students' CM when it came to scoring. As a result, this approach made it easier for us to score the CM and have consensus over the results.

We sought the presence of patterns; therefore, several repetition of one pattern on one CM did not add to the score of that pattern. Moreover, considering that any proposition in a CM represents a pattern of simple linear causality, we did not analyze this pattern separately as a causal mechanism in this study.

Finally, for quantitative analysis, we calculated the percentage of scores for each pattern to determine their prevalence in students' CM across the groups. Then, using Kruskal Wallis H Test, we compared students' scores for each pattern between the three groups to determine how different scaffolding approaches affected students' performance in demonstrating causal patterns in their CM.

Validation of findings

To validate findings, we provided a third researcher, who had a science education background, with 20 percent of the entire data set, the scoring scheme, and the findings of the study. Some discrepancies were brought up by the researcher that were discussed and addressed in a meeting with the researcher, and adjustments were made accordingly to our data analysis and the results.

Results

This study, in general, investigated CM as an assessment tool to represent students' views toward biology as a complex system and in specific as a tool to examine and promote students' knowledge about underlying mechanisms in complex systems. In this section, we first provide a qualitative analysis about what students were able and unable to demonstrate by CM regarding the interconnection and complexity of biological systems. Then, through merging quantitative and qualitative results, we describe the patterns of causal mechanisms present in students' CM and how scaffolding affected their performance in showing those patterns.

Aspects of complexities in biological systems revealed in students' CM

We delineated four major categories in the data featuring biological complexities. These categories often were not mutually exclusive in students' CM.

Cluster of concepts; interconnections within organizational levels in biology

Students expressed part of their views about biology by creating cluster of concepts. Appendix C represents an example of a cluster of concepts CM constructed by students. The concepts of a cluster were often organized based on their association to a particular concept or their affiliation to an organizational level in biology. For example, Grace (all names are pseudonyms) from group A explained that she was satisfied with her CM because the way she "organized the main thoughts and then broke the topics into subunits of each other show[ed] how they all are connected to an ecosystem's biodiversity." In her CM, she had built four clusters of concepts connecting them to the central concept of the ecosystem biodiversity. Likewise, Skylar (group C) explained in her interview that she "initially put evolution in the center […] trying to make the connections in between every individual thing." She was not satisfied with her CM (scoring her CM 2 out of 5) but believed that she could "explain [the] relationship between individual parts of whole." Similar approach was adopted by group B students to fill out their pre-designed CM. Bella, in her interview, explained that "I try to put the concepts that I thought were more related … closer to each other; for example, DNA and RNA [are] very closely related as are fitness and natural selection closely related." In fact, we observed that creating a cluster of concepts was an approach of CM building in which students focused more on organizing and categorizing the concepts than trying to show the meaningful interconnections among them.

Sequence of concepts; hierarchy and flow among biological concepts

A group of students expressed their views by ordering the given concepts sequentially. Appendix D represents an example of a sequence of concepts CM constructed by students. They believed that their CM showed how each concept related to the other one. For example, Molly (group A) explained, "I think I accurately and satisfactorily expressed all of the terms on my concept map by branching them off of each other. Starting with larger concepts and breaking them down to smaller parts." Molly's CM was one directional with some branches indicating that there was an order in how biological concepts were related. Similarly, Clara (group B) said that she liked "how everything flows from DNA to overall interacting evolution." Kaela elaborated in her interview that "I was satisfied with that I put some kind of flow into it and the fact that I started at a more like molecular level and worked my way up to broader concepts and [...] put those altogether." We noted that organizing concepts sequentially was helpful for students in expressing some of their views about biological complexities. However, it also acted counterintuitive in showing various interconnections in biological systems as it focused on depicting the relationship among biological concepts through chain of connections.

Biological processes; protein synthesis and evolution

Biological processes were also demonstrated in students' CM. The process of protein synthesis was one of them; the processes of DNA translation to RNA, and RNA transcription to protein were shown and mentioned by many students. Evolution and natural selection were the other processes students mentioned depicting in their CM. Sadie (group A) explained that her CM showed "the evolution process and how it correlates to natural selection [...] and how it changes over time due to [...] adaptations and mutations." We noted that to explain the processes, students used domino pattern to express their thoughts. This approach was more helpful in demonstrating the process of protein synthesis than evolution. We observed consistency and accuracy among students in presenting how DNA and RNA played a role in proteins synthesis and how proteins attributed to an organism phenotype. However, there was inconsistency and often inaccuracy in students' patterns showing how, for example, the prey and predator population and the concepts of adaptation, extinction, and fitness played a role in the process of evolution. In sum, students could express some biological processes such as protein synthesis in their CM, but struggled to demonstrate others such as evolution.

System dynamics; matters of change and effect

A few students mentioned that their CM demonstrated the changes and effects induced by the events of mutation and the introduction of a new species to an ecosystem. For instance, Katy (group C) said that her CM clearly showed "how the ripple effect of a mutation change the original species." In her CM, she demonstrated a mutation bringing about domino of changes from DNA to protein and eventually to phenotype, and she also exhibited that phenotype affects fitness and consequently influences evolution. Similarly, Emily (group B) believed that her CM showed "how a mutation can lead to an adaptation and have a better fitness for [an] animal." Likewise, some students mentioned that their CM showed how the introduction of a new species could lead to extinction or adoption of other organisms in an ecosystem. Overall, students were able to show how an event could bring about changes and effects into a system. They did so by showing how an event unfolded sequentially to particular results.

Indeed, CM as a tool allowed students to express their thoughts about both interconnections among biological components and dynamics in biological systems. To this end, students applied two methods; they either organized the concepts according to their biological level association or ordered them in a particular sequence of relationships. We found that these methods were helpful in presenting some aspects of biological complexities but, as will be discussed in the following section, were not suitable for showing some other complex features of biological systems.

Missing aspects of biological complexities in students' CM

We found three categories in respect to thoughts students could not present in their CM.

System dynamics; matters of probability, magnitude, simultaneous interactions and ongoing processes

Some dynamic aspects of biological systems were challenging for students to demonstrate by CM. They were often related to the matters that required: simultaneous exhibition of multiple factors involved in a cause to bring about different effects, presentation of the causes and effects magnitude, and demonstration of ongoing processes. Showing the role of mutation in biological processes was mentioned by many students in their explanations and in interviews. For example, by a unidirectional arrow, Miles (group C) had shown that mutation "alter"[ed] fitness but did not believe such presentation expressed his view on "how a mutation can increase fitness and lead to evolution." Clara (group B) referred to a similar dissatisfaction by saying that she had difficulty showing "mutations being a mechanism of evolution." In the interview, Adeline said that she wanted to show how mutations could change organisms and elaborated that "you change something on a smaller level, it can affect all the larger tiers to like till biodiversity" but she did not know how to show a small change could bring about big effects on a CM.

The impact of a new predator on an ecosystem was another challenging issue for students to display by CM. The challenge was related to the interaction of the predator with a prey population and its overall effect on the ecosystem biodiversity and evolution. One student, who had shown on her CM that a new predator population can lead to extinction of a prey population, was dissatisfied with her presentation of the matter because she could not show "how a predator population can have a positive impact." There were students who showed the bidirectional relationship between the prey and predator population, but some of them also expressed dissatisfaction about how they could not demonstrate the role of other factors involved in such relationship and how they could bring about different outcomes. However, it is worth noting that many students did not attempt to show the impact of a new species to an ecosystem by CM, which likely was due to the design of the assignment. It turned out that the primary focus of our participants was on connecting the concepts and creating CM, and many of them missed showing the possible effects of a species introduction to an ecosystem on their CM.

Our analysis of students' CM aligned with their expression of discontent about not showing some system dynamics. Students struggled to show how multiple factors simultaneously affected each other and how the interactions among them could attribute to various outcomes. They also did not show aspects of complexity that required manifestation of nonlinear and probabilistic effects over a period of time.

CM structure; not representative of the whole view

The overall structure of CM was a matter of discontent for many of our participants as it did not fully reflect their views on biology. They believed that biological systems were more interrelated than what they had shown in their CM. For example, Renee (group A) had three clusters of concepts in her CM, and she felt that the "sub-sections worked well but not the whole thing." She further elaborated that "I couldn't find a way to fit everything together as a comprehensive thought." Students' interviews revealed part of the reason why their CM did not represent their whole view. Sadie, for instance, said that she knew that "everything is interrelated" but she intentionally avoided drawing all the connections; otherwise, it would have been a "jumbled mess." A similar approach was adopted by some other interviewees as well. Tiana decided "to keep it simple," and Addison avoided showing some interactions because she did not want to have "a lot of information on [her] concept map."

In sum, students seldom said that the overall structure of their CM represented their view on how the components of biological systems were interconnected. They believed that there could be more interconnections and interactions among the given concepts, but they did not show them for two reasons; either they did not know how to do so, or they avoided them so that their CM could clearly express their other points.

Some concepts did not fit together

Students referred to the relationship between some concepts as what they could not express well in their CM. We noted that those relationships were of two types. One type was similar to what students mentioned in the cluster of concepts above; they were associated with a particular topic or an organizational level in biology, but this time students did not know about them. For example, a few students said that they lacked the knowledge about the evolution and so they could not express well the relationship among the concepts related to it. The other type of concepts students mentioned struggling to put together required integration of biological levels. Rylee (group A) had difficulty with showing "how microbiology relates to the study of the environment and ecosystem." Similarly, Emily (Group B) could not see "how genes influence biodiversity," and Derek (group C) was dissatisfied about the way he had tied "the lower half" of his CM, which entailed evolutionary concepts, to "the upper half," which were about molecular biology. Overall, we observed that students found it challenging to put the concepts into a bigger picture and often the issue exacerbated as the two concepts were far from each other in biological levels.

In sum, contrary to what students had depicted in their CM, they believed that the given concepts should have been more interconnected to each other to present a more realistic image of their view. This issue, to an extent, overlapped with the difficulties to demonstrate dynamic aspects of biological systems and the interconnection between organizational levels in biology. Accordingly, students were more confident and satisfied about parts of their CM, not its whole.

Patterns of underlying causal mechanisms present in students' CM

In this section, we present the prevalence of various patterns in students' CM according to our theoretical framework. To this end, first, through a descriptive analysis, we provide a statistical summary from all the data and how it complements our qualitative analysis. Then, through a comparative analysis, we reveal the results about the impact of scaffolding on students' CM building performance in respect to showing each causal pattern.

Causal patterns in students' CM: descriptive analysis

The first author scored the existence of domino, mutual, relational, and cyclic patterns in each student's CM. Table 1 presents the number and percent of students demonstrating

Table 1 Percentage of students showing each pattern of causality in their CM

Pattern	Group A (N = 62)			Group B (N = 57)			Group C (N = 54)		
	Correct (%)	Semi-correct (%)	Absent (%)	Correct (%)	Semi-correct (%)	Absent (%)	Correct (%)	Semi-correct (%)	Absent (%)
Domino	42	27	31	32	14	54	50	19	31
Mutual	5	6	89	14	39	47	19	20	61
Relational	6	0	94	7	21	72	7	13	80
Cyclic	6	3	91	0	18	82	4	7	89

each pattern. It shows that the domino pattern was the main causal mechanism that a majority of students (69 percent—correct and semi-correct scores combined) in groups A and C demonstrated in their CM. Domino causality was the second prevalent pattern (46 percent) in the CM of group B students after mutual causality (53 percent). The table also reveals that the relational and cyclic patterns were notably absent in students' CM; on average, over 80 percent of our participants did not show those patterns.

The qualitative analysis of data confirmed that the domino causality was the major causal mechanism employed by our participants. Group A and C students usually had a linear CM with a beginning and end with some branches, and group B students presented the relationship between most of the concepts through spotting a pattern of domino connection among them. Also, in the interviews, students used a sequential narration to explain how the concepts related to each other. For example, elaborating on her CM, Anne brought up the mice example in which a mutation "changed the DNA, RNA, proteins, genes, and then eventually changed the [mice] phenotype" and then how this process ended in creating two species. We found students more familiar with and adept in presenting the relationship between biological concepts in linear unidirectional pattern.

Our findings here align with the thoughts students believed they could express well by CM. In fact, it was largely through a domino pattern that students connected the concepts in their organizational levels, showed the flow and hierarchy between the concepts, talked about some biological processes, and showed change and effect in their CM. On the other hand, the lack of the other patterns explains part of the students' difficulty in showing some of their thoughts. For example, students struggled to show simultaneous and ongoing processes in their CM. This problem could be eased if students were skilled in demonstrating mutual and cyclic causalities. Similarly, students wrestled to show the interrelationship among evolutionary concepts that required understanding of relational causality. Moreover, whereas students were satisfied with parts of their CM, they were discontented with its overall structure. This finding indicated that they were aware of the multiple and various interconnections that existed in biological systems but could not show them on their CM. This issue, to a considerable extent, was associated with students' lack of skills in showing the various causal relationships among biological concepts.

Causal patterns in students' CM: comparative analysis

We assumed that by providing scaffolds for various causal relationships, students would be able to demonstrate more complex relationships in their CM. To examine this assumption, we did a comparative analysis on demonstration of each causal pattern in students' CM. First, a Shapiro–Wilk test of normality was performed, and the result was significant ($p < 0.05$) indicating that the data had a non-normal distribution. Then, we conducted a Kruskal–Wallis test to determine if there was any significant difference between groups in demonstrating each causal pattern.

Regarding domino causality, Table 2 shows a statistically significant difference [$H(2) = 6.822$, $p = 0.033$] between the score of students in different groups. Post hoc comparison results revealed that the significant difference ($p = 0.046$) was between groups B and C. This result indicated that students without the pre-designed CM but with prompts for causal mechanisms (i.e., group C with the mean ranks of 95.44) had a better performance to demonstrate the pattern of domino causality than the students with the pre-designed CM (i.e., group B with the mean rank of 74.00).

Table 2 The results of the Kruskal–Wallis test and post hoc comparisons regarding domino causality

Domino causality							
Groups	N	Mean rank	Chi-Square	df	p	Pairwise comparisons of group	p[a]
A	62	91.60	6.822	2	0.033	Group A—Group B	0.118
B	57	74.00				Group A—Group C	1.000
C	54	95.44				Group B—Group C	0.046

[a]Significance values have been adjusted by the Bonferroni correction for multiple tests

Results further confirm that students had the knowledge and skills to demonstrate the relationship among the biological concepts in a domino order. Therefore, given the pre-designed CM with limited domino patterns, group B students found it difficult to demonstrate some of their thoughts. Many of the Group B students "wished" that they were not given a pre-designed CM, when talking about their dissatisfaction about their CM. Emily mentioned in her interview that she could have had a better performance on her CM if she had built it herself from scratch. Accordingly, we observed that students in group B evaluated their CM performance (M = 2.55, SD = 0.748) lower than the other two groups (group A: M = 3.01, SD = 0.905; group B: M = 3.08, SD = 0.908). Students tended to show the relationship among the concepts in a sequential pattern, which explains why group A students (mean rank = 91.60) did better, though not significantly, than group B (mean rank = 74.00) with pre-designed CM and almost as good as group C (mean rank = 95.44) with scaffolding for domino causality. Because in any way group A students were going to show the relationship among the concepts through the sequential order.

Regarding mutual causality, Table 3 reveals a statistically significant difference [H(2) = 21.382, p = 0.000] between the score of students in different groups. Post hoc comparison results revealed significant difference between groups A and B (p = 0.000) and groups A and C (p = 0.004). This result indicated that scaffolds of mutual causality positively affected students' performance in demonstrating such a pattern in their CM. In fact, when students were scaffolded for mutual causality either through getting tips about the concept of mutual causality (group C) or having the pattern in a pre-designed CM (group B), they performed better in finding such a relationship between some concepts. Students in group A tended to show a unidirectional relationship between prey and predator populations, and more students in groups B and C demonstrated that relationship as mutual.

In the interviews, more students could demonstrate mutual causality when they received different levels of scaffolding. Anna (group A), for example, talked about how a new predator and environment affected each other bidirectionally when she was prompted about mutual causality. Tara (group B) had not been able to find a meaningful mutual relationship

Table 3 The results of the Kruskal–Wallis test and post hoc comparisons regarding mutual causality

Mutual causality							
Groups	N	Mean rank	Chi-square	df	p	Pairwise comparisons of group	p[a]
A	62	67.98	21.382	2	0.000	Group A—Group B	0.000
B	57	102.07				Group A—Group C	0.004
C	54	92.93				Group B—Group C	0.746

[a]Significance values have been adjusted by the Bonferroni correction for multiple tests

between any two concepts; however, during the interview, when Tara received more explanations about mutual relationship, she found the relationship between prey and predator "mutual, obviously."

Regarding relational causality, Table 4 shows a statistically significant difference [H(2) = 8.427, p = 0.015] among different groups. Post hoc comparison results indicated that the significant difference (p = 0.013) was between groups A and B. On a pre-designed CM, significantly more students could make a meaningful relational relationship among the given concepts. Results also revealed that when students got scaffolded for the concept of relational causality (group C), they performed better in showing such relationship in their CM (mean rank = 89.05), but this better performance was not statistically significant from group A (mean rank = 77.69).

Although the scaffolding improved students' demonstration of relational causality, the majority of students in all groups found it difficult to show such relationships in their CM. This was visible in students' interviews as well. When prompted about relational causality, few students could talk and show such a relationship among some concepts. Emily, when prompted about the relational causality, pointed out that the relationship between prey and predator affects their adaptation. Referring to the tip in the assignment, Kelsey mentioned that she struggled to show relational causality in her CM because "I kind of don't really see where two things can equally affect C (third factor)." However, after getting more explanation about the relational causality, Kelsey was able to give the example of relationship between genotype and environment determining phenotype. However, the majority of interviewees could not give an example of relational causality after getting prompts about it. They struggled to distinguish between the case when the interaction between two factors determined the third one (e.g., genotype and environment determine the phenotype) and the case when two factors were involved in a process to determine the third one (e.g., DNA and protein determine the phenotype).

Quantitative analysis of cyclic patterns showed no significant difference between our groups, which indicated that scaffolding did not influence students' ability to exhibit cyclical patterns. Also, as mentioned previously, students did not do well in demonstrating cyclic patterns according to our scoring approach. However, this finding did not align with our qualitative examination of students' CM. In general, we observed that some students had attempted to show that all concepts were interrelated by making a form of loop or cycle in their CM, which indicated that they were familiar with the pattern of cyclic causality but lacked the knowledge and skill to show such relationship scientifically accurate because it was a multistep process. Students needed to make at least three accurate or semi-accurate proportions in a cyclic form to get a score for the pattern.

On the other hand, interviews showed students did not find it challenging to depict a cyclic relationship in their CM. Kelsey had not made a cyclic pattern in her CM but illustrated that she knew the concept by giving the following example: "Well for example, but this could be

Table 4 The results of the Kruskal–Wallis test and post hoc comparisons regarding relational causality

Relational causality							
Groups	N	Mean rank	Chi-square	df	p	Pairwise comparisons of group	p[a]
A	62	77.69	8.427	2	0.015	Group A—Group B	0.013
B	57	95.18				Group A—Group C	0.204
C	54	89.05				Group B—Group C	1.000

[a]Significance values have been adjusted by the Bonferroni correction for multiple tests

wrong, but natural selection can lead to extinction and extinction of species can lead to change in the ecosystem biodiversity and depending on that how like how that affects the ecosystem in terms of like resources or predation or anything like that, that can also affect natural selection again and starts total." In fact, we found that cyclic pattern, though much less frequent than domino causality, was one of the main approaches adopted by students to explain the interconnectivity of the biological concepts. Therefore, we assume that the low score of correct cyclic patterns in students' CM was less due to their unfamiliarity to the concept of cyclic causality and more related to the extent the task was demanding and therefore required coherent knowledge of multiple steps in cyclic relationship in a biological system.

In summary, for our first research question, we found students at different levels of knowledge and skill in showing patterns of causal mechanisms to demonstrate the complex interactions in biological systems. They largely gravitated to organize concepts in a sequential order to make meaningful connections among concepts. Students were also somewhat familiar with the concept of mutual causality, but they needed scaffolding to display it on their CM. Relational and cyclic causality were notably absent in students' CM. We found relational causality the most abstract causal relationship for students to grasp; they struggled to discern the time the relationship between two factors determined the third factor from the time two factors were involved in the process of determining the third factors. Regarding cyclic causality, students knew about it, but they could not show it accurately on their CM probably because the task required relatively high cognitive load and wide knowledge of concepts and their relationships.

Regarding our second research question, based on our quantitative data analysis, we found that scaffolding students with pre-designed CM (group B) and prompts about causal mechanisms (group C) improved their performance in demonstrating different causal patterns. In specific, group B participants did better—statistically significant—than group A students in showing mutual and relational relationships in their CM, and group C students did better—statistically significant—than group A and B students in showing mutual and domino causalities, respectively. Overall, these results reveal the benefits of scaffolding in improving students' ST skills; though, they do not necessarily indicate what type of scaffolding treatment (B or C) is better.

Drawing on our different data sources, we infer that prompting students with causal patterns would be a better scaffolding approach than providing students with pre-designed CM to improve students' ST skills. Such inference is based on the data that group B students had the lowest satisfaction level in expressing their thoughts through CM. This finding was also supported by our qualitative data analysis wherein we found many group B students believed that they could have had a better performance, provided they were not given a pre-designed CM.

We conclude that scaffolding students' CM building either by a pre-designed CM or prompts about causal mechanisms can improve their demonstration of different causal pattern in complex systems. However, the latter approach would be more beneficial as it would allow students to express their thoughts more freely.

Discussion

In this study, we adopted causal patterns as our theoretical framework and drew on both qualitative and quantitative data to investigate the CM as a tool to assess and enhance undergraduates' ST skills. Specifically, we sought to examine how students demonstrate their understanding using CM and how their ST skills are affected by scaffolding them to demonstrate causal patterns. To this end, we were able to tease out which aspects of system complexities students could show by CM. Results indicated that, through CM, students could demonstrate some aspects of their views about biology as a complex system. To different degrees, they showed structural aspects of biological systems like organizational levels and how concepts were associated with each other both within and across the organizational levels. Students also displayed some processes and dynamics of biological systems on their CM. These findings corroborate with extant literature on how CM is a helpful tool to examine students' understandings about a system components, hierarchy, and behavior (Assaraf & Orion, 2005; Buckley & Boulter, 2000; Tripto et al., 2013).

This study also examined the ways students undertook to depict their views on CM. According to our theoretical framework, students largely used the pattern of domino causality to organize their knowledge about biological concepts. In this way, they showed how each concept is related to its predecessor and built a structure of concepts that depicted biological hierarchy and processes.

Several reasons caused the prevalence of domino pattern in students' CM. First, the linear and sequential expression of relationship among the concepts was probably easier for students, as it resembled the way they learned about the concepts from the textbooks or course materials. Second, the probability of making meaningful domino patterns was relatively higher than other patterns by the concepts in our study. Therefore, if students ordered some concepts by chance due to their lack of knowledge about them, they had higher chance of making meaningful patterns from them. Third, organizing concepts through a chain of connections and putting them in hierarchical order were the conventional approaches to constructing CM. For example, aligned with a constructivist view toward CM building, Reinagel and Speth (2016) helped students to make integrative gene-to-phenotype CM by first teaching them isolated relationships between pairs of molecular genetic structure and then having students construct their CM upon those pairs of relationship blocks. This approach corresponds with hierarchical CM building approach proposed by Novak and Gowin (1984) to examine students' prior knowledge and build new knowledge upon it. Therefore, considering the mentioned reasons, we found the prevalence of domino pattern unsurprising in students' CM, as it largely aligned with both their knowledge of the concepts and their skills of CM construction.

Our findings also revealed that students had difficulties showing some aspects of system complexities through CM, due in part to their lack of knowledge about the concepts, specifically evolutionary concepts. In this regard, CM was a useful tool to diagnose students' lack of knowledge and their possible misconceptions about biological complexities. Of this, CM as an assessment and learning tool has been supported by research studies (Bergan-Roller et al., 2020; Brandstädter et al., 2012). Out of students' CM, Bergan-Roller et al. (2020) found that students had simple understanding about cellular respiration and had a better understanding about the process of glycolysis than fermentation. On the other hand, results indicated that students could not show certain aspects of system complexities by CM despite their knowledge about them. Two factors attributed to this issue: students' skills of CM construction and the nature of CM itself.

Students, as a typical approach to CM construction, organized the concepts. Largely, they tried to put the concepts in hierarchical orders, show a flow of relationship among them and/or categorize them in cluster of concepts associated with a topic. This approach is helpful in comprehension and retention of information by structuring and organizing concepts (Novak & Cañas, 2008). Yet, according to our findings, the overall structure of the CM built by this approach was not a true representation of students' view of the way biological concepts were interconnected; students believed in more interactions among the concepts but avoided them as they had more emphasis on organizing and categorizing the concepts rather than showing their dynamics and interconnections. In fact, we found it counterintuitive to demonstrate system complexities by CM through its conventional approaches because the focus on organizing and categorizing concepts discouraged students to demonstrate all the interconnections they knew about biological systems.

Moreover, we observed that some dynamic aspects of complex systems were considerably absent in our participants' CM. Tripto et al. (2013) had similar observations, finding that students could hardly show homeostasis mechanisms by CM. Drawing on literature as well as their own results, Tripto et al. attributed the issue to the fact that "understanding homeostasis requires several cognitive abilities, such as discerning that multiple phenomena occur simultaneously and comprehending that every process is comprised of several stages" (p. 251). Likewise, we noted that students struggled to show multistage processes that presented ongoing phenomena and sustainability in systems. This challenge particularly manifested itself in students' poor performance in making meaningful cyclic causal patterns as the mechanism embodied the features of ongoing multistage processes.

Based on our findings, we attribute the students' poor performance in demonstrating certain system dynamics to their lack of knowledge and skills in depicting the different causal mechanisms. For example, through domino causality, many students could accurately show how DNA, RNA, and Protein were related to each other but the same approach did not work for the evolution process. Many evolutionary concepts make sense in light of understanding mutual and relational causal mechanisms; for example, fitness emerges as a result of interaction between an organism's phenotype and its environment, or adaptation and extinction emerge as a result of interaction among multiple factors such as the environment and prey-predator population. The linear presentation of these concepts did not portray an accurate image of the evolutionary process.

Results indicated that when students received scaffolding for the patterns of causal mechanisms, they pictured more dynamic aspects of complexity on their CM. For example, getting prompts about the pattern of mutual causality, students could show the concepts of equilibrium and dynamic relationship between prey and predator population. In the interviews, we observed that after being prompted about the cyclic pattern of connections, some students were able to demonstrate cyclic relationship among the concepts. These results supported our stand that students' inadequacy in showing certain system dynamics by CM was due to their lack of knowledge about causal mechanisms. Consequently, informing students about the patterns of causal mechanisms can improve their ST skills through CM building practices.

Research has shown that scaffolding students' CM building improves their reasoning skills (Eggert et al., 2017). Similarly, we observed a better performance in students' ST skills demonstration by scaffolding their CM building. By providing students with a pre-designed CM, we could help students to find more diverse patterns of interactions among biological concepts. This approach, however, had some downsides. First, it limited students to express some of their thoughts, which was particularly evident by students' level of satisfaction about their CM. Next, a pre-designed CM provided a possibility for students to demonstrate different causal patterns without truly recognizing them on their CM. Though this possible limitation of a pre-designed CM was not specifically explored in our study, we found it plausible. While intentionally filling out a pre-designed CM to demonstrate the relationship between some patterns, students inevitably showed connections between some concepts—as it pre-existed—that they did not necessarily intend to show. Therefore, in our study, we found the introduction and illustration of causal patterns the most effective scaffolding approach to improve students' ST skills by CM. Having been introduced to the patterns and concepts of causal mechanisms, participants in group C demonstrated more diverse and meaningful interconnections and relationships among biological concepts than students in group A with no scaffolding, and accordingly they depicted more dynamic and complex view toward biological systems. These results suggest that students should be explicitly instructed about the concepts of different causal mechanisms alongside scaffolding them to show such patterns in CM.

Eventually, we found CM was not a suitable tool for some features of complex systems. For example, it was impractical to show non-linear relationships with probabilistic outcomes and various intensities on a CM. We observed that in order to demonstrate change because of a mutation, students were bound to provide one or two scenarios about the possible outcomes with deterministic results. Likewise, when designing this research study, the authors also found it implausible to examine and scaffold spiral causality by CM as it required demonstration of interactions with outcomes of different intensities over a period of time. These limitations of CM in showing time dependent processes have been noted by other researchers as well (Tripto et al., 2013).

Conclusion

Identifying and assessing students' ST patterns is essential to "develop and facilitate a pedagogical scaffolding that allows students to engage in counterintuitive modes of thought and overcome the variety of cognitive barriers that can prevent them from fully understanding the system's complexity" (Tripto et al., 2018, p. 673). Also, knowing about students' thinking patterns, educators can present biological systems in a way that develop students' metacognition and improve their skills in applying and adapting their mental models in new contexts (Dauer et al., 2013). Our study provided insight into students' ST patterns related to underlying causal mechanisms in complex systems. It also revealed that students needed to know more about various causal mechanisms to be able to explain broader aspects of biological complexities and behaviors. Thus, taking our definition of ST into account, we think learning about the patterns of causal mechanisms will enable students to better coordinate their knowledge of a system to reason about its behaviors and functions and accordingly improve their ST skills.

Furthermore, this study showed that CM can be a powerful tool in assessing and improving students' ST skills. We found that CM could reveal students' understandings about concepts in association with other concepts and expose their knowledge about a system's structure and behavior. To show the relationship among the concepts, students largely applied domino pattern of causality, and they did so with the purpose of organizing and categorizing the concepts. Despite its benefits, this approach provided a narrow view to students' ST skills and limited their presentation of dynamic relationships on CM. Therefore, to make CM a more powerful tool to assess and develop students' ST skills, we believe that students should be instructed on how to create CM with the emphasis on showing the interactions and interconnections among concepts and about the underlying causal mechanisms in complex systems so as to be able to implement them in their CM.

In summary, it is crucial that students understand complex systems and develop their ST skills to better navigate our contemporary world (Boersma et al., 2010; Ison & Straw, 2020; National Research Council, 2010; NGSS Lead States, 2013). Not only do they need to learn about the components of systems, but also they must know how those components through various interconnections and interactions create systems. Students should know about the different types of interactions within systems and practice them. This study found CM a suitable tool as it forced students to think about the relationship between the concepts and revealed their knowledge about the concepts and their interactions. This study also adds to the literature how scaffolding students' CM building can enhance their ST skills. Accordingly, it proposes two instructional steps for having CM building practices improve students ST competencies. First, students should receive explicit instructions and scaffolding about the various causal mechanisms; without the familiarity with the different interactions in complex systems, students will not know what skills to practice on and demonstrate for ST. Second, students should be trained for CM building with an emphasis on demonstrating the interconnection among the components of a system, rather than organizing and categorizing them. Despite the benefits of CM for assessing and enhancing students' ST skills, it demonstrated limitations on how to show some dynamic aspects of complexities in systems. Therefore, we recommend that teachers use CM in concert with other tools to gain broader perspective on students' ST skills.

Limitations

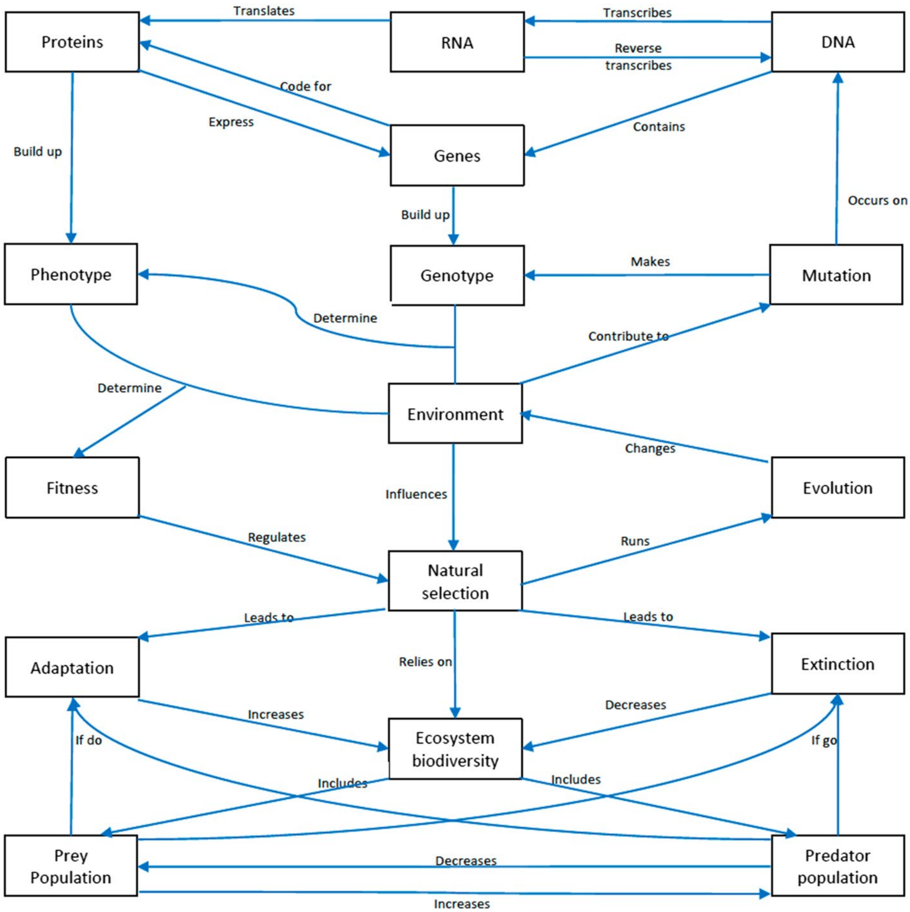
This study concluded that CM is a powerful tool for both improving and examining students' ST skills. Questions remain about the relationship between students' CM building skills and their demonstration of ST skills by CM, which leads us to three limitations. First, despite having a 50-min session of CM building, students had different approaches and skills for building a CM. Second, the CM assignment was a take-home task and the amount of time and effort students put on it varied among them. Last, students had different background knowledge about the given concepts for the CM task. These interrelated factors affected the patterns and interconnections students showed on their CM. Therefore, though our study provides insight into how students demonstrated causal patterns on their CM, it does not show the extent students' CM building skills were decisive on such demonstrations. Moreover, our study drew on multiple approaches to demonstrate the potentials of CM in improving students' ST skills. Though this is a strongpoint for a research study, it is not a plausible approach for classroom teaching. Therefore, further studies are required to transfer the potentials of CM practices into actual classrooms for improving students' ST skills.

The scaffolding intervention results were also affected by students' lack of knowledge about causal mechanisms and how they could be demonstrated on CM. This was limiting in respect to knowing the extent students' demonstration of causal patterns were meaningful to them, particularly for group B participants who had the possibility to randomly fill out their pre-designed CM. Moreover, considering that the three groups in this study participated in different scaffolding treatments, the authors could not do blind coding; therefore, their biases might have crept into their analysis despite their attempts at objectivity. In conclusion, this study indicates that scaffolding students for causal patterns and CM building can improve their ST skills. It also acknowledges that there is a need for further exploration to know the extent such improvement was based on students' understandings of causal mechanisms or the interference of other factors dismissed in the study.

Appendix A

The master concept map.

The master concept map developed in the preliminary stage of the study. This concept map without the concepts and the labels were given to students in group B in order to be completed



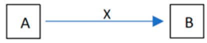
Appendix B

Assignment C scaffolding.

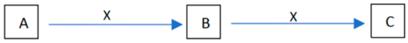
Different factors are related to each other in various ways. Below are five ways that factors can be related to each other. Try to apply these patterns in your concept map so as to make it more comprehensive. (Note: below "X" is any word or phrase that can explain the cause and effect relationship between the factors.)

The various causal relationships introduced to students in Group C

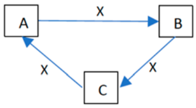
1. Simple linear relationship.
   One directional and direct effect relationship.



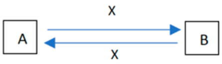
2. Domino relationship.
   One effect makes another effect, the latter effect makes another effect and so on.



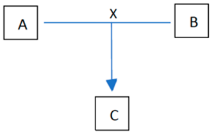
3. Cyclic relationship.
   There is no clear beginning or ending.
   Factors in a cyclic way affect each other.



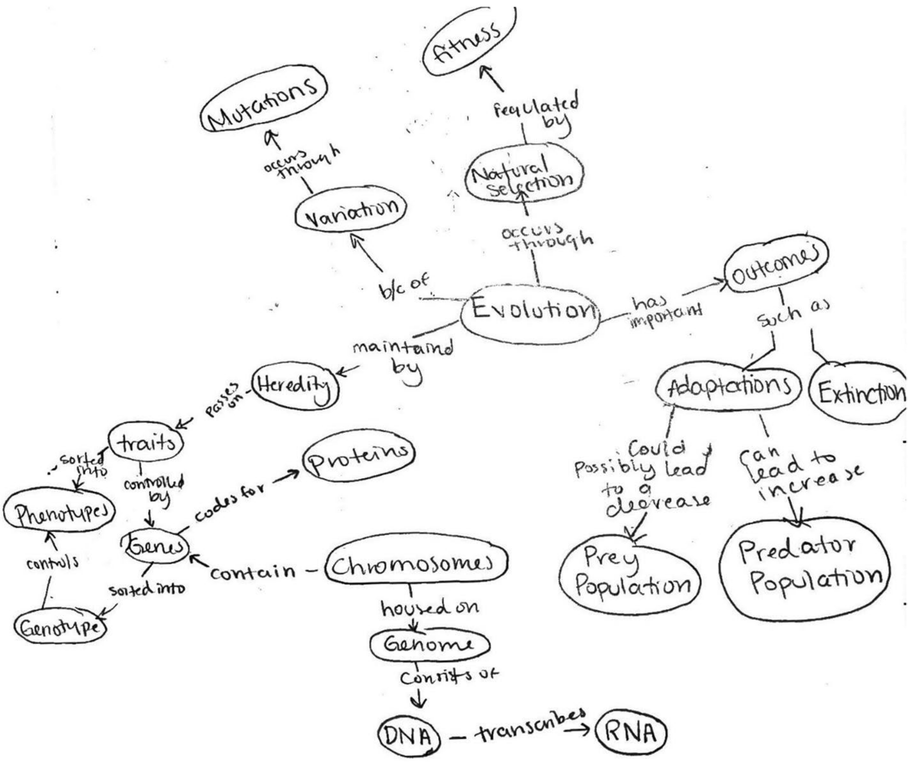
4. Mutual relationship.
   Two factors affect each other simultaneously.



5. Relational relationship.
   The relationship between two factors determine the effect on the third factor.



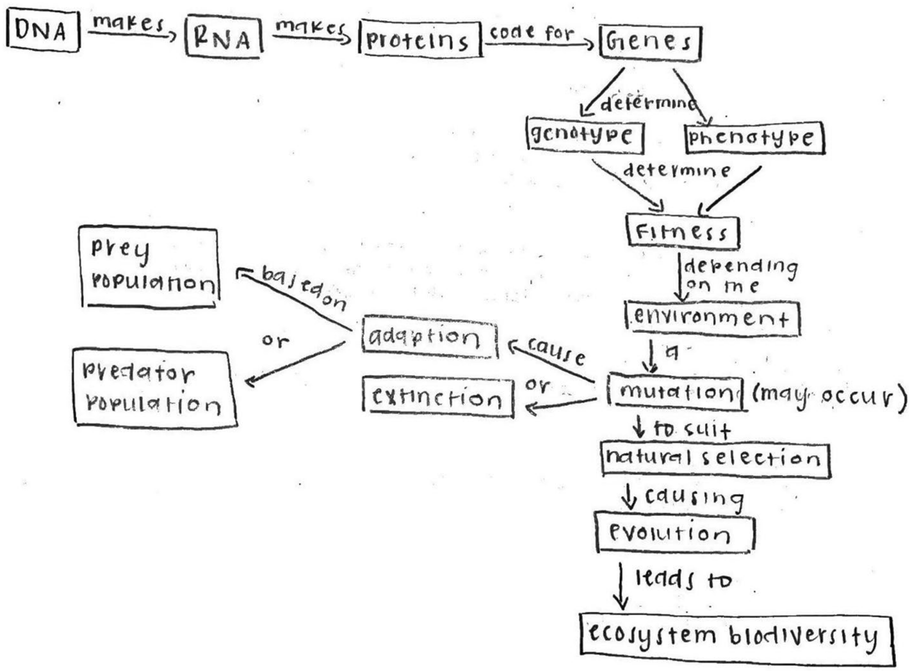
Appendix C



Concept map with cluster of concepts.
An example of a concept map with cluster of concepts representing one of the students' general approaches to concept map building

Appendix D

Concept map with sequence of concepts.
An example of a concept map with sequence of concepts representing one of the students' general approaches to concept map building



References

Abell, S. K., & Siegel, M. A. (2011). Assessment literacy: What science teachers need to know and be able to do. *The professional knowledge base of science teaching* (pp. 205–221). Springer.

Assaraf, O.B.-Z., Dodick, J., & Tripto, J. (2013). High school students' understanding of the human body system. *Research in Science Education, 43*(1), 33–56. https://doi.org/10.1007/s11165-011-9245-2

Assaraf, O.B.-Z., & Orion, N. (2005). Development of system thinking skills in the context of earth system education. *Journal of Research in Science Teaching, 42*(5), 518–560. https://doi.org/10.1002/tea.20061

Baxter, G. P., Elder, A. D., & Glaser, R. (1996). Knowledge-based cognition and performance assessment in the science classroom. *Educational Psychologist, 31*(2), 133–140. https://doi.org/10.1207/s15326985ep3102_5

Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research, 87*(2), 309–344. https://doi.org/10.3102/0034654316670999

Bergan-Roller, H. E., Galt, N. J., Helikar, T., & Dauer, J. T. (2020). Using concept maps to characterise cellular respiration knowledge in undergraduate students. *Journal of Biological Education, 54*(1), 33–46. https://doi.org/10.1080/00219266.2018.1541001

Boersma, K. T., Kamp, M. J. A., Oever, L. van den, & Schalk, H. (2010). Naar actueel, relevant en samenhangend biologieonderwijs. Retrieved from, https://repository.ubn.ru.nl/handle/2066/115258

Brandstädter, K., Harms, U., & Großschedl, J. (2012). Assessing system thinking through different concept-mapping practices. *International Journal of Science Education, 34*(14), 2147–2170. https://doi.org/10.1080/09500693.2012.716549

Buckley, B. C., & Boulter, C. J. (2000). Investigating the role of representations and expressed models in building mental models. In J. K. Gilbert & C. J. Boulter (Eds.), *Developing models in science education* (pp. 119–135). Springer.

Conradty, C., & Bogner, F. X. (2012). Knowledge presented in concept maps: Correlations with conventional cognitive knowledge tests. *Educational Studies, 38*(3), 341–354. https://doi.org/10.1080/03055698.2011.643100

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). SAGE Publications.

Dauer, J. T., Momsen, J. L., Speth, E. B., Makohon-Moore, S. C., & Long, T. M. (2013). Analyzing change in students' gene-to-evolution models in college-level introductory biology. *Journal of Research in Science Teaching, 50*(6), 639–659. https://doi.org/10.1002/tea.21094

Edmondson, K. M. (2005). Chapter 2—Assessing science understanding through concept maps. In J. J. Mintzes, J. H. Wandersee, & J. D. Novak (Eds.), *Assessing science understanding* (pp. 15–40). Academic Press.

Eggert, S., Nitsch, A., Boone, W. J., Nückles, M., & Bögeholz, S. (2017). Supporting students' learning and socioscientific reasoning about climate change—The effect of computer-based concept mapping scaffolds. *Research in Science Education, 47*(1), 137–159. https://doi.org/10.1007/s11165-015-9493-7

Fisher, D. M. (2011). "Everybody thinking differently": K–12 is a leverage point. *System Dynamics Review, 27*(4), 394–411. https://doi.org/10.1002/sdr.473

Fisher, D. M. (2018). Reflections on teaching system dynamics modeling to secondary school students for over 20 years. *Systems, 6*(2), 12. https://doi.org/10.3390/systems6020012

Gilissen, M. G. R., Knippels, M.-C.P.J., & van Joolingen, W. R. (2020). Bringing systems thinking into the classroom. *International Journal of Science Education, 42*(8), 1253–1280. https://doi.org/10.1080/09500693.2020.1755741

Grotzer, T. (2012). *Learning causality in a complex world: Understandings of consequence*. Rowman & Littlefield Education.

Hammond, D. (2017). Philosophical foundations of systems research. In M. C. Edson, P. Buckle Henning, & S. Sankaran (Eds.), *A guide to systems research: Philosophy, processes and practice* (pp. 1–19). Springer.

Hay, D., Kinchin, I., & Lygo-Baker, S. (2008). Making learning visible: The role of concept mapping in higher education. *Studies in Higher Education, 33*(3), 295–311. https://doi.org/10.1080/03075070802049251

Hmelo-Silver, C. E., & Azevedo, R. (2006). Understanding complex systems: Some core challenges. *Journal of the Learning Sciences, 15*(1), 53–61. https://doi.org/10.1207/s15327809jls1501_7

Hmelo-Silver, C. E., Marathe, S., & Liu, L. (2007). Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. *Journal of the Learning Sciences, 16*(3), 307–331. https://doi.org/10.1080/10508400701413401

Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science, 28*(1), 127–138. https://doi.org/10.1207/s15516709cog2801_7

Ifenthaler, D. (2010). Relational, structural, and semantic analysis of graphical representations and concept maps. *Educational Technology Research and Development, 58*(1), 81–97. https://doi.org/10.1007/s11423-008-9087-4

Ison, R., & Straw, E. (2020). *The hidden power of systems thinking: Governance in a climate emergency*. Routledge.

Izci, K., Muslu, N., Burcks, S. M., & Siegel, M. A. (2020). Exploring effectiveness of classroom assessments for students' learning in high school chemistry. *Research in Science Education, 50*(5), 1885–1916.

Jacobson, M. J., Markauskaite, L., Portolese, A., Kapur, M., Lai, P. K., & Roberts, G. (2017). Designs for learning about climate change as a complex system. *Learning and Instruction, 52*, 1–14. https://doi.org/10.1016/j.learninstruc.2017.03.007

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
Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Sciences, 5*(10), 434–442. https://doi.org/10.1016/S1364-6613(00)01751-4
Jordan, R. C., Hmelo-Silver, C., Liu, L., & Gray, S. A. (2013). Fostering reasoning about complex systems: Using the aquarium to teach systems thinking. *Applied Environmental Education & Communication, 12*(1), 55–64. https://doi.org/10.1080/1533015X.2013.797860
Ke, L., Sadler, T. D., Zangori, L., & Friedrichsen, P. J. (2020). Students' perceptions of socio-scientific issue-based learning and their appropriation of epistemic tools for systems thinking. *International Journal of Science Education, 42*(8), 1339–1361. https://doi.org/10.1080/09500693.2020.1759843
Mathews, L. G. J., Szostak, R., & Repko, A. (2008). Using systems thinking to improve interdisciplinary learning outcomes: Reflections on a pilot study in land economics. Issues in Interdisciplinary Studies. Retrieved from, https://our.oakland.edu/handle/10323/4442
Meadows, D. H. (2008). *Thinking in systems: A primer*. Chelsea Green Publishing.
Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (2005). *Teaching science for understanding: A human constructivist view*. Academic Press.
Morgan, G. (2006). *Images of organization* (Updated). Sage Publications.
Morris, W. E., & Brown, C. R. (2001). David hume. Retrieved from, https://plato.stanford.edu/entries/hume/?fbclid=IwAR2RNvkYTwX3G5oQUdalb8rKcVrDm7wTt55aWyauFXptJWEbxAXRQVY6_-M
Moshman, D., & Tarricone, P. (2016). Logical and causal reasoning. *Handbook of epistemic cognition* (pp. 54–67). Routledge.
National Research Council. (2010). *Standards for K–12 engineering education*. National Academies Press.
National Research Council. (2014). *Developing assessments for the next generation science standards*. The National Academies Press.
Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*(3), 413–448. https://doi.org/10.3102/00346543076003413
NGSS Lead States. (2013). *Next generation science standards: For states, by states*. The National Academies Press.
Novak, J. D., & Cañas, A. J. (2008). *The theory underlying concept maps and how to construct and use them*. Florida Institute for Human and Machine Cognition.
Novak, J. D., Gowin, D. B., & Bob, G. D. (1984). *Learning how to learn*. Cambridge University Press.
Novak, J. D., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal, 28*(1), 117–153. https://doi.org/10.3102/00028312028001117
Pellegrino, J. W. (2013). Proficiency in science: Assessment challenges and opportunities. *Science, 340*, 320–340. https://doi.org/10.1126/science.1232065
Reinagel, A., & Speth, E. B. (2016). Beyond the central dogma: Model-based learning of how genes determine phenotypes. *CBE-Life Sciences Education, 15*(1), 454. https://doi.org/10.1187/cbe.15-04-0105
Sabelli, N. H. (2006). Complexity, technology, science, and education. *Journal of the Learning Sciences, 15*(1), 5–9. https://doi.org/10.1207/s15327809jls1501_3
Sawyer, K. R. (2006). The new science of learning. In K. R. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 1–16). Cambridge University Press.
Schwendimann, B. A. (2015). Concept maps as versatile tools to integrate complex ideas: From kindergarten to higher and professional education. *Knowledge Management & E-Learning, 7*(1), 73–99.
Shavelson, R. J., Ruiz-Primo, M. A., & Wiley, E. W. (2005). Windows into the mind. *Higher Education, 49*(4), 413–430. https://doi.org/10.1007/s10734-004-9448-9
Siegel, M. A., Menon, D., Sinha, S., Promyod, N., Wissehr, C., & Halverson, K. L. (2014). Equitable written assessments for English language learners: How scaffolding helps. *Journal of Science Teacher Education, 25*(6), 681–708.
Sommer, C., & Lücken, M. (2010). System competence—Are elementary students able to deal with a biological system? *Nordic Studies in Science Education, 6*(2), 125–143.
Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation, 27*(2), 237–246. https://doi.org/10.1177/1098214005283748
Tripto, J., Assaraf, O.B.-Z., & Amit, M. (2013). Mapping what they know: Concept maps as an effective tool for assessing students' systems thinking. *American Journal of Operations Research (AJOR), 3*, 245–258. https://doi.org/10.4236/ajor.2013.31A022

Tripto, J., Assaraf, O. B. Z., & Amit, M. (2018). Recurring patterns in the development of high school biology students' system thinking over time. *Instructional Science, 46*(5), 639–680. https://doi.org/10.1007/s11251-018-9447-3

Tripto, J., Assaraf, O.B.-Z., Snapir, Z., & Amit, M. (2016). The 'What is a system' reflection interview as a knowledge integration activity for high school students' understanding of complex systems in human biology. *International Journal of Science Education, 38*(4), 564–595. https://doi.org/10.1080/09500693.2016.1150620

Tripto, J., Assaraf, O. B. Z., Snapir, Z., & Amit, M. (2017). How is the body's systemic nature manifested amongst high school biology students? *Instructional Science, 45*(1), 73–98. https://doi.org/10.1007/s11251-016-9390-0

Verhoeff, R. P., Knippels, M.-C.P.J., Gilissen, M. G. R., & Boersma, K. T. (2018). The theoretical nature of systems thinking. Perspectives on systems thinking in biology education. *Frontiers in Education*. https://doi.org/10.3389/feduc.2018.00040

Verhoeff, R. P., Waarlo, A. J., & Boersma, K. T. (2008). Systems modelling and the development of coherent understanding of cell biology. *International Journal of Science Education, 30*(4), 543–568. https://doi.org/10.1080/09500690701237780

Yin, Y., Vanides, J., Ruiz-Primo, M. A., Ayala, C. C., & Shavelson, R. J. (2005). Comparison of two concept-mapping techniques: Implications for scoring, interpretation, and use. *Journal of Research in Science Teaching, 42*(2), 166–184. https://doi.org/10.1002/tea.20049

Yoon, S. A., Goh, S.-E., & Park, M. (2018). Teaching and learning about complex systems in K–12 science education: A review of empirical studies 1995–2015. *Review of Educational Research, 88*(2), 285–325. https://doi.org/10.3102/0034654317746090

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.