



Supporting Student System Modelling Practice Through Curriculum and Technology Design

Tom Bielik¹ · Lynn Stephens² · Cynthia McIntyre² · Daniel Damelin² · Joseph S. Krajcik³

Accepted: 6 October 2021 / Published online: 25 October 2021
© The Author(s) 2021

Abstract

Developing and using models to make sense of phenomena or to design solutions to problems is a key science and engineering practice. Classroom use of technology-based tools can promote the development of students' modelling practice, systems thinking, and causal reasoning by providing opportunities to develop and use models to explore phenomena. In previous work, we presented four aspects of system modelling that emerged during our development and initial testing of an online system modelling tool. In this study, we provide an in-depth examination and detailed evidence of 10th grade students engaging in those four aspects during a classroom enactment of a system modelling unit. We look at the choices students made when constructing their models, whether they described evidence and reasoning for those choices, and whether they described the behavior of their models in connection with model usefulness in explaining and making predictions about the phenomena of interest. We conclude with a set of recommendations for designing curricular materials that leverage digital tools to facilitate the iterative constructing, using, evaluating, and revising of models.

Keywords Modelling practice · Educational technology · Curriculum development

Introduction

Understanding the complexity of the natural world is empowering, providing us with a sense that we can explain phenomena and solve difficult problems. To engage in civic discourse about how to effect positive change requires an understanding of systems and how the components of a system can interact to produce emergent behaviour that is more than the sum of its parts. This understanding is the essence of system modelling. System models are powerful tools for investigating the world around us, allowing scientists and students to make sense of phenomena, represent complex causal relationships, solve problems, and share ideas. Therefore, modelling, systems thinking, and supports for associated cognitive skills such as causal reasoning should be significant components of every student's education.

In a previous publication (Bielik et al., 2019), we presented our framework of four aspects of system modelling. In this study, we expand this framework by focusing on the following research question: how are students engaging with the four aspects of system modelling, and what are the challenges they encounter in the process? To explore this question, we provide empirical evidence in support of student use of these aspects of system modelling practice and how they changed over the course of a 10th grade instructional unit in chemistry. By analyzing students' responses and produced artefacts, we deepen our understanding of these aspects while addressing the opportunities and challenges students face when engaging with curricular materials that include modelling environments. This is followed by a set of recommendations for educators and curricular materials developers who aim to support students as they learn system modelling.

✉ Tom Bielik
tom.bielik@fu-berlin.de

¹ Freie Universität, Berlin, Germany

² The Concord Consortium, Concord, MA, USA

³ Create for STEM Institute, Michigan State University, East Lansing, MI, USA

Literature Review

Scientific Modelling Practice

Models are generative epistemological tools, consisting of components and the relationships between them (Harrison & Treagust, 2000; Nicolaou & Constantinou, 2014; Schwarz

et al., 2009), and are essential to scientists and engineers for representing a system and for explaining phenomena and predicting possible outcomes (Harrison & Treagust, 2000; National Research Council, 2012). Developing and using models is one of the core scientific and engineering practices described in *A Framework for K-12 Science Education* (National Research Council, 2012). Schwarz et al. (2009) defined the modelling practice as the ability to construct, use, evaluate, and revise models of the natural world. These prior works influenced how we discuss the four aspects of system modelling practice. To build scientific models, students need appropriate modelling tools that include scaffolds to support their ability to build and use models as explanatory and predictive mechanistic tools of phenomena (Clement, 2000; Lehrer & Schauble, 2006), but most students have few opportunities to meaningfully engage with models (Louca & Zacharia, 2012; Nicolaou & Constantinou, 2014; Schwarz et al., 2009).

Systems Thinking

Systems thinking is required for investigating and learning about complex systems and typically includes ideas such as the ability to consider the system boundaries, the components of the system, the interactions between system components and between different subsystems, and emergent properties and behaviour of the system (Passmore et al., 2014; Russ et al., 2008; Wilensky & Resnick, 1999). The difficulties of understanding complex systems are well documented (Booth Sweeney & Sterman, 2000; Dörner, 1980; Hmelo-Silver & Pfeffer, 2004; Jacobson & Wilensky, 2006). We suggest that engaging in modelling through the use of a system modelling tool with appropriate scaffolds can support students in developing a systems thinking perspective.

System Dynamics Modelling Practices

System dynamics (SD) was created at MIT in the 1950s to help humans think about complex systems (Forrester, 1968). With the introduction of computer simulations based on stock and flow models (Richmond et al., 1987), SD became more accessible and eventually began to be used in educational settings with children and young adults (Gould-Kreutzer, 1993). Within the five stages of SD modelling (problem identification, system conceptualization, model formulation, testing, and using), Martinez-Moyano and Richardson (2013) compiled a list of 27 best practice statements regarded of highest importance by practitioners and experts in SD. These include identifying the components of the system and the relationships between them, iteratively developing the structure of the models by adding detail as needed, comparing simulated

with real behaviour, and iteratively testing and validating the model. Students face many challenges in thinking about complex dynamic systems (Stratford et al., 1998; Tadesse & Davidsen, 2020), and there has long been hope within the SD community for new tools to introduce these challenging ideas to an audience not as constrained by age and ability (Gould-Kreutzer, 1993).

Causal Reasoning

To understand a system, explain it, and troubleshoot system problems often requires understanding causal relationships between components in the system so that one can predict and compare with real-world system behavior (Jonassen & Ionas, 2008). Student difficulties with causal reasoning have been well studied (e.g., Schauble, 1996; Zimmerman, 2007). The causal structures in complex systems can be especially difficult to reason through and to teach about (Perkins & Grotzer, 2000; Yoon et al., 2017). According to Jonassen and Ionas (2008), system modelling tools are the only class of tools that can enable learners to model both covariational and mechanistic attributes of causal relationships necessary for full causal reasoning.

Given the theoretical background discussed above, there is a need for better understanding of how instructional practices and technology tools can support students in system modelling.

Four Aspects of System Modelling Practice

In Bielik et al. (2019), we presented a theoretical examination of four aspects of system modelling practice illustrated with several student exemplars. These aspects are as follows: (1) defining the boundaries of the system by including components in the model that are relevant to the phenomenon under investigation, (2) determining appropriate relationships between components in the model, (3) using evidence and reasoning to construct, use, evaluate, and revise models, and (4) interpreting the behavior of a model to determine its usefulness in explaining and making predictions about phenomena. The first two aspects primarily grew out of challenges we observed when students construct system models. The second two aspects are related to our initial design ideas for a tool that would support sensemaking with models through comparative data analysis and ease of model construction. In the “[Analysis](#)” section, we describe criteria to evaluate whether students engage with each aspect and the results of applying those criteria to student work.

Materials and Methods

This study is part of a larger design-based research project aimed at understanding how students learn when given the opportunity to use a modelling tool that allows them

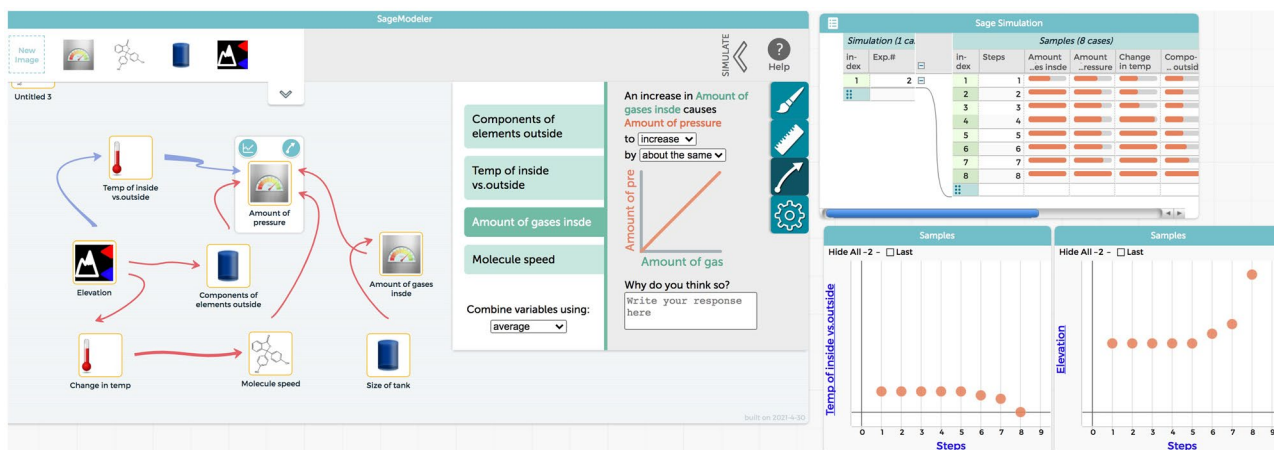


Fig. 1 Initial model of pair I (S13 and S20)

to create an instantiation of their conceptual understanding in the form of an external, runnable model. Their models can generate output that feeds back into the students' understanding of the phenomena and leads to iterations in both their conceptual understanding and their external, runnable models. The focus of the present analysis is a single instructional unit that provided opportunities for students to revise their models multiple times. In particular, we focus on evidence of students' engagement with the aspects as they revised their models.

The Modelling Tool

The modelling tool, a free, Web-based, open-source tool, was created by the authors in partnership with other project team members.¹ It is designed to scaffold student learning so that young students, beginning in middle school, can engage in systems thinking through constructing, using, evaluating, and revising models. The tool facilitates the modelling of a system and also makes it possible to calculate and visualize model output without requiring students to write equations or code (Damelin et al., 2017).

With our modelling tool, students begin by dragging images that represent components to a canvas and then linking the components with arrows to specify relationships (Fig. 1). For the system model to become a runnable model, each component is treated as a variable that can be calculated by the modelling engine. The next step for students is to define each relationship link in the model such that the impact of one variable can be calculated on all of the variables to which it is linked. Variable values are defined using a low-to-high scale. Students construct a verbal description

of how one variable affects another by using drop-down menus, e.g., "An increase in *temperature* causes *volume* to [increase] by [about the same]." The resulting relationship is also depicted by a graph showing a visual representation of this relationship (Fig. 1). Defining relationships with words helps students overcome the mathematical obstacles typically associated with creating system models and allows them to focus on a conceptual understanding of the relationships between variables (Damelin et al., 2017).

To "run" the model, a student uses a slider to move an independent variable through a range of values. To facilitate analysis of the impact of this variable on the rest of the model, our modelling tool integrates CODAP, the Common Online Data Analysis Platform, a tool designed to support student exploration of data, including comparisons of model-output and real-world data sets (Finzer & Damelin, 2016).

Context: High School Chemistry Unit About Emergent Properties of Gases

This study involves data collected from an enactment of a high school curricular unit conducted in the spring of 2017 in an honors chemistry class. The enactment included 11 lessons of about 70 min each. Prior to starting the unit on the emergent properties of gases, all students completed an Introduction to Modelling unit, which directly addressed the issue of appropriate variables for a system model. The chemistry unit focused on the emergent properties of gases and was co-designed by the authors, together with a high school chemistry teacher who enacted the unit in her classroom. The unit was designed to explore the anchoring phenomenon of an oil tanker that imploded after being steam cleaned, as shown in an online video. This builds towards the driving question of the unit: "How can something that can't be seen crush a 67,000 pound oil tanker made of half inch steel?" A full description of the

¹ The modelling tool can be freely accessed online at <https://sagemodeler.concord.org/>.

Table 1 Description of the gas laws instructional unit and expected target models

Investigation no	Phenomenon	Activities	Target model
1	Oil tanker implosion (video)	Reflection questions Two simulations and student-designed lab experiment	Initial model: <i>volume/pressure</i> : Model revision 1: revise volume/pressure relationship
2	Balloons in liquid (video)	<i>Volume/temperature</i> : article and student-designed lab experiment Reflection questions <i>Volume/temperature/kinetic energy</i> : simulation	Model revision 2: revise volume/temperature relationship Model revision 3: include kinetic energy
3	Egg in Erlenmeyer flask (demonstration)	<i>Pressure/temperature/kinetic energy</i> : article, video, student-designed lab experiment, simulation Reflection questions	Model revision 4: include pressure/temperature/kinetic energy relationships

curricular unit can be found in Appendix 1. Descriptions of phenomena, activities, and expected target models in the unit are provided in Table 1.

Participants

Data was collected from 20 tenth grade students (14 male, 6 female) in a public school in the northeastern USA. Students were from an average socioeconomic level (27.5% of the students were eligible for free or reduced lunch, 88.5% White, 5.5% Hispanic, 2.5% Asian, 1.5% African-American). The teacher was an experienced chemistry teacher. Researchers observed the enactment and supported the teacher.

Data Sources

Model Reflection Questions

Throughout the unit, students engaged in offline and online activities. The online component included illustrations, readings, labs, embedded assessments, opportunities to develop and revise models in the modelling tool, and text prompts with reflection questions on the model construction and revisions. Student reflections on their models were illustrative in exploring how and why they created their models, the reasons behind changes in their models, and insights into progress toward the aspects of system modelling practice considered here.

After the first model construction activity, students were asked the following question (among others): What are you still uncertain about in your model? Following each subsequent model revision, students were asked the following: (i) What did you change in your most recent model? (ii) What were your reasons for making these changes? (iii) What are you still uncertain about in your model? Students responded in pairs or individually (depending if their partner

was absent) to these questions, which were embedded in the online curriculum. However, due to student absences, the total number of responses for each revision is less than 20. A total of 9 pairs of student responses were taken for analysis, referred to below as pairs A–I.

Student Models

Student models were automatically collected via use of the Web-based modelling tool and analyzed by two members of the team. Students had four opportunities to modify their models, each following an activity designed to help them explore a feature of the phenomenon in a more focused way, typically through a lab experiment. Therefore, each succeeding model revision added content and relationships as students learned about them. Each revision was also preceded by peer review of the models and class discussion. Students worked in pairs on the models; nine of the ten student pairs submitted complete sets of models.

Student Interviews

In semi-structured interviews conducted by two of the authors during and after the unit, six students (five males, one female) who had worked in pairs were interviewed individually. They were asked to describe their models and to show how their models could explain the unit's driving question of an oil tanker implosion. The interviews were videotaped and transcribed.

Analysis

Model Reflection Questions

Using the list of aspects from Bielik et al. (2019), the second author examined a sample of student responses to

the model reflection questions and developed a tentative set of observables that gave evidence for engagement or lack of engagement for each aspect. Two team members who were experienced modelers then provided feedback on the validity of the observables. The resulting list was used as coding criteria, and applied to additional student answers and refined in an iterative process. Once the wording of the coding criteria had stabilized, the first two authors independently applied the codes to the entire set of student responses, and achieved Inter-Rater Reliability (IRR) ($\kappa = 0.73$). They then discussed differences to reach full consensus. Coding criteria are provided in Appendix 2.

Student Models

In the course of a single unit, it was not practical to evaluate improvement in the modelling practice in terms of the models alone. To a large extent, changes in the models were due to students adding new variables to represent new features of the phenomenon as they learned about them. In addition, as students learned more about the four aspects, they sometimes disassembled their models and began rebuilding them due to new ways of thinking about the system. Therefore, the emphasis was not on producing a perfect final model, but on the process of modelling. However, evidence from the models can be triangulated with other results, particularly with respect to Aspects 1 and 2. To this end, the variables and connections in the models were analyzed independently from the written explanations. The coding criteria are presented in the Analysis section. For Aspect 1, defining the boundaries of the system by including components in the model that are relevant to the phenomenon under investigation, the models were examined jointly by one of the authors with a researcher/observer who was familiar with how the boundaries were defined during the classroom implementation of the curricular unit. For Aspect 2, determining appropriate relationships between components in the model, two team members independently coded the models for the presence of indirect or redundant connections and had 100% agreement.

Student Interviews

The interview transcripts were reviewed as a potential source of additional information about student thinking. However, only limited use is made of these data in the present analysis. Two of the models discussed in this paper were from a pair of students we interviewed, and we use quotations from one of their final interviews to suggest a possible explanation for some surprising features of their models and written explanations.

Results and Discussion

Examples of student responses are used to illustrate how the four aspects of modelling practice can be used as a lens for analyzing student progress in engaging in modelling, systems thinking, and causal reasoning. Student quotations are drawn from the written responses to the reflection questions and in one instance from a student interview transcript. Screenshots of models from two student pairs are used as exemplars. Together these data provide a window into how students exhibited aspects of system modelling practice and allow us to characterize changes that occurred during the unit.

Aspect 1: Defining the Boundaries of the System by Including Components in the Model That Are Relevant to the Phenomenon Under Investigation

This aspect is characterized by the two following features:

A Distinguishing between objects and variables.

Directing students to define the components in their models as measurable variables and not as objects in order to run a simulation of their model is not a simple task. It requires explicit focus of the teacher and repeated experiences. In the enacted unit, the teacher emphasized this issue when supporting the students in developing their models. Although most students produced initial models that included only variables that could be defined on a low-to-high scale, there were some exceptions. For example, in Fig. 1 of pair I initial model, the components *Change in temp* and *Amount of pressure* are defined appropriately as variables, but the component named *Components of elements outside* describes an object rather than a variable with specific characteristics, and cannot be defined on a low-to-high scale in any practical way. These students might have intended to describe the ratio of different elements in the air and to imply that the ratio might affect the air pressure. In any case, these students removed this variable in their first model revision (Fig. 2).

B Choosing relevant variables through consideration of appropriate size and scope.

To exemplify this aspect, the model in Fig. 1, *Change in temp* is a variable of appropriate size and scope for the tanker phenomenon; temperature changes have an important effect on the outcome. *Elevation*, however, is an example of a variable that is not of appropriate size or scope. In their first revision (Fig. 2), these students associated elevation with *Density of gas inside vs outside* the tanker. Although the ratio of densities outside and inside is crucial to the phenomenon, the density outside does not change appreciably

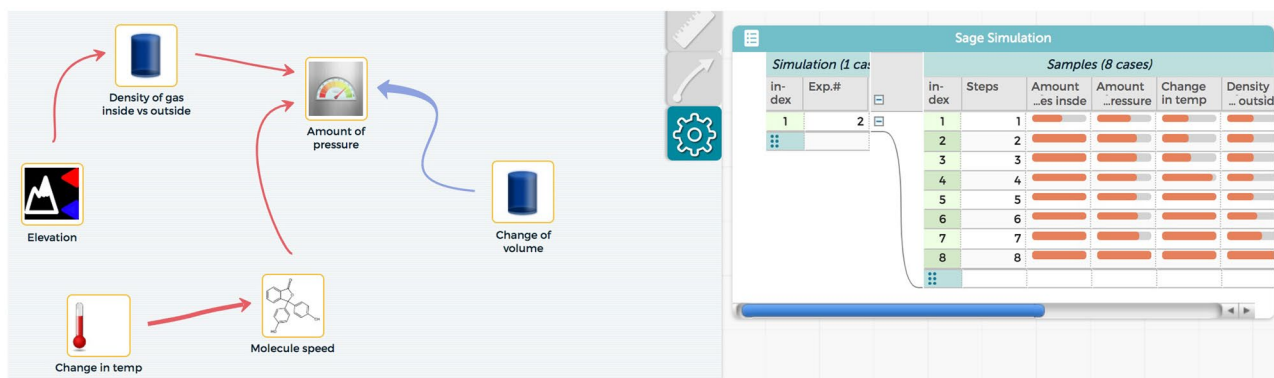


Fig. 2 First model revision of pair I

in the scenario and so does not help to explain the phenomenon. After a class discussion in which models containing elevation were discussed, these students (pair I) removed both *Elevation* and *Density of gas inside vs outside* as they were not needed within the scope of these students' explanatory model. After their second model revision (Fig. 3) they wrote, 'We did not think the elevation and density of the molecules in the tanker was crucial to the model.'

In their responses to the model reflection questions, students commonly mentioned that what they changed in their models was related to Aspect 1, in particular, variables (over 50% of responses after each model revision; over 80% after the second and third revisions; see Table 2). Students' responses as to why they made these changes fell into one or more of the following categories: to have them better explain and describe the phenomenon under investigation, to make the models more correct, to have the models make more sense or be more logical, to include important or missing

variables, to have more specific variables, or to remove variables that are not significant.

Aspect 1 Discussion

Although the Introduction to Modelling unit had directly addressed the issue of appropriate variables for a system model, in the early models of the chemistry unit (initial model through Revision 2), seven of nine student pairs included variables that were outside the boundaries of the tanker system. Of these seven, only two pairs failed to make progress; in over half the final models all variables were relevant (see Table 3). Moreover, students' written responses show active engagement in thinking about which variables should be included to best represent that system, considering both size and scope, and ensuring that each variable represented a measurable quantity (see Table 2).

Fig. 3 Second model revision of pair I

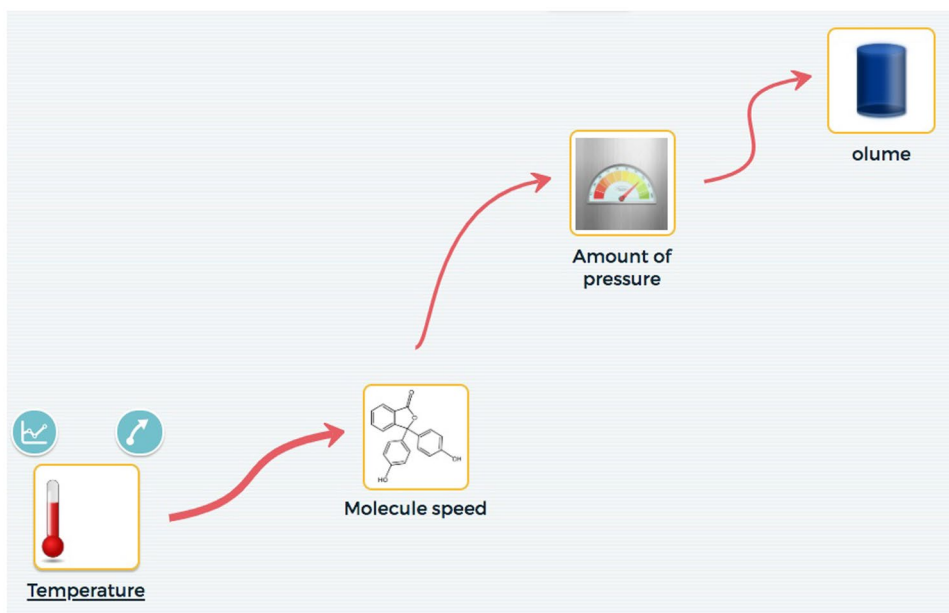


Table 2 Evidence for Aspects 1–4 in student written explanations^a

	Aspect 1	Aspect 2	Aspect 3	Aspect 4
Initial model ($n = 19$)	11%	53%	0%	11%
Revision 1 ($n = 19$)	58%	74%	21%	26%
Revision 2 ($n = 16$)	81%	75%	13%	19%
Revision 3 ($n = 19$)	84%	100%	0%	11%
Revision 4 ($n = 19$)	68%	100%	0%	11%

^aStudents' written explanations were coded for evidence that the student had addressed each of four aspects of system modelling competence. Each row represents the class result at the end of a different modelling cycle. The numbers reflect the number of students present. Out of 20 students, there was no day when everyone was present

Aspect 2: Determining Appropriate Relationships Between Components in the Model

This aspect is characterized by the two following features:

A Defining scientifically accurate relationships to represent interactions between variables.

There are several ways a link between two variables could be incorrect.

1. There may be no relationship between variable A and variable B.

Table 3 Number of variables that were outside the size or scope in student models^a

Student pair	Initial	Revision 1	Revision 2	Revision 3	Revision 4
A (S1 and S16)	2	2	2	2	1
B (S2 and S6)	0	0	0	0	0
C (S3 and S5)	2	1	1	1	1
D (S4 and S21)	2	1	0	0	0
E (S7 and S17)	1	2	2	2	2
F (S9 and S15)	0	0	1	1	1
G (S10 and S12)	0	0	0	0	0
H (S11 and S14)	0	1	0	0	0
I (S13 and S20)	2	1	0	0	0
Class total	9	8	6	6	5

^aEach cell represents an individual model created by a pair of students. The table shows the number of variables in each model that were out of bounds of the system or not relevant to the driving question

2. There is a relationship, but the way the relationship is defined does not match the real-world behavior of the interaction between variable A and variable B.
3. The direction of causality is reversed.

An example of this feature can be found in the initial model of pair I in Fig. 1; the sequence of two connections that link *Size of tank* to *Amount of pressure* do not match the real-world behavior of the system. The students changed this during their first model revision. As S13 wrote, "I got rid of this part of the model because an increase in volume does not lead to an increase in pressure. Actually, an increase in volume causes a decrease in pressure" (Fig. 4, third revision of pair I model). There was an additional change in their next revision (Fig. 5) that was also related to feature A, a change in the direction of causality between *Amount of pressure* and *[V]olume*. In the tanker scenario, a change in volume was not what caused the decrease in pressure; rather the decrease in pressure caused the volume to change—suddenly. These students continued to change relationships throughout each model revision, which was typical for all students.

In their responses to the model reflection questions about what they changed in their models, most students mentioned the relationships between the variables (74 and 75% after the first and second revisions, 100% after the third and fourth revisions; see Table 2). In their written reasons for making changes to these relationships, students mentioned: having the model make better sense and be more logical, aligning the model with new findings from experiments, making the model better fit what was learned in the lessons, having the model better explain the phenomenon under study, and including the necessary and relevant relationships.

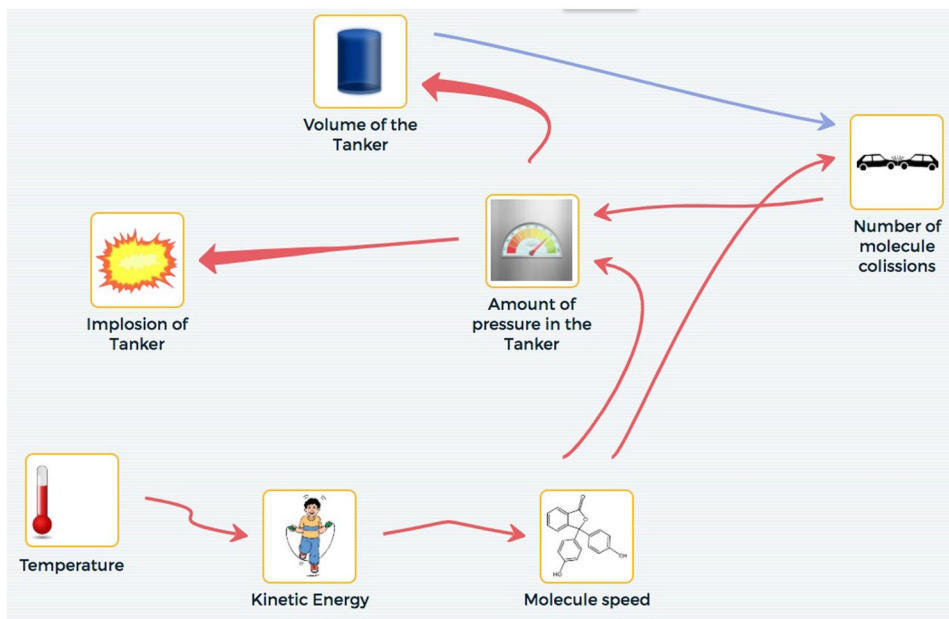
B Defining direct relationships between variables.

There are two ways problems with the directness of relationships manifest themselves:

1. There may be large gaps in the causal chain.
2. There may be indirect relationships included between variables.

Gaps in causal chains sometimes resulted when students leapfrogged one or more steps in a causal chain of variables. Students were encouraged to add microscopic causal mechanisms, helping them to explain why one variable had an effect on another. Pair I students explicitly mentioned adding *Kinetic Energy* to fill a gap they perceived in their causal chain. After their third revision (Fig. 5), they wrote, "We added kinetic energy as a variable in between temperature and molecule speed in our model and had kinetic energy increase the same amount as temperature. These changes help respond to the unit's driving question because they

Fig. 4 Third model revision of pair I



show how as temperature increases, so does kinetic energy and molecule speed about the same.” This is a situation where the model was used to explicate covariational and mechanistic attributes of the relationship between temperature and pressure. In a gas, temperature is defined in terms of average kinetic energy of the molecules. Explicitly including kinetic energy as a variable in their model allowed a direct link to molecule speed, which, in turn, allowed a direct link between speed and number of collisions, which provided a direct link to gas pressure (Fig. 4).

In this same model, the students added a causal mechanism, *Number of molecule collisions*, between *Molecule*

speed and *Amount of pressure in the tanker*. The single link that also connects *Molecule speed* and *Amount of pressure* represents an indirect relationship. It complicates the model and adds no new information, and it was later removed.

Five of the nine student pairs who submitted complete model sets had indirect links in all of their model revisions (Table 4). However, after an initial increase in indirect links there was a drop in such links in their final models. In the third revision, there were 10 total indirect links in class models while in the final revision, there were 6 indirect links.

Fig. 5 Fourth model revision of pair I

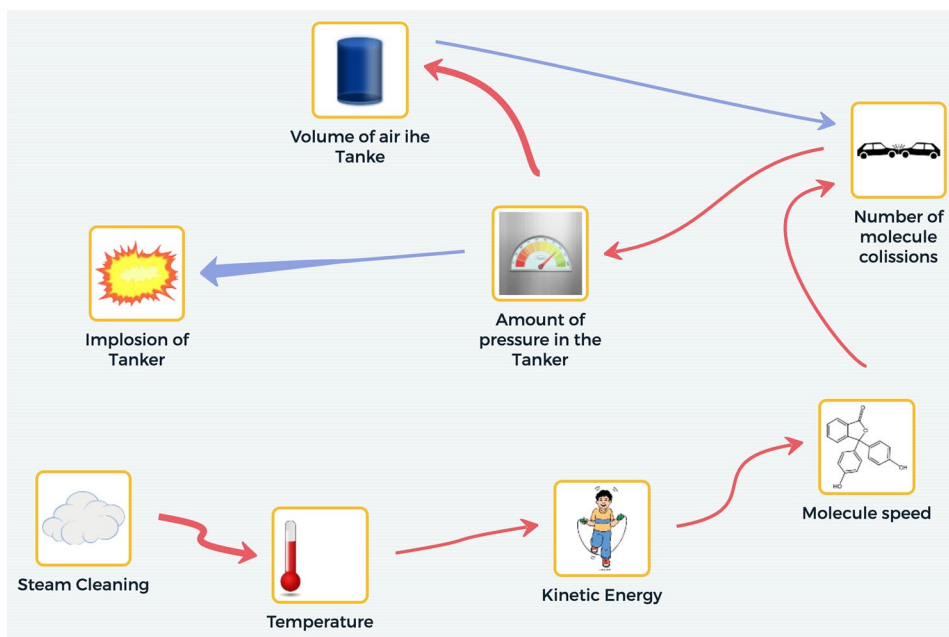


Table 4 Number of indirect links in student models^a

Student pair	Initial	Revision 1	Revision 2	Revision 3	Revision 4
A	2	2	1	2	1
B	0	3	3	2	1
C	0	1	0	0	0
D	0	2	3	2	1
E	0	0	0	0	0
F	0	1	2	2	2
G	0	0	0	0	0
H	0	1	1	1	1
I	0	0	0	1	0
Class total	2	10	10	10	6

^aEach cell represents an individual model created by a pair of students. The table indicates the number of places in each model where a single link connects two variables already connected by a causal chain, usually resulting in a redundant or indirect link

Aspect 2 Discussion

Defining appropriate relationships between variables in the system under study is crucial for producing a model that is useful for understanding and making predictions about a phenomenon. Examples of student work illustrate the two features related to this aspect: defining scientifically accurate relationships between variables and defining direct relationships between variables. Students made steady improvements in defining the relationships between variables. These improvements took the form of relationships that showed appropriate directionality in cause and effect, correct definitions in how one variable affects the other, and more direct linkage between variables. Nevertheless, indirect links were widespread within student models, remaining pervasive in their third revision, where over half of the models had one or two such links. In their final models, there were fewer indirect links. A possible explanation is that students were learning new mechanisms and adding new links during each revision, and in their final models noticed that these new links made some of their earlier links unnecessary.

Aspect 3: Using Evidence and Reasoning to Construct, Use, Evaluate, and Revise Models

After designing and carrying out experiments, including using a molecular dynamics simulation of gas behavior, students revised their models. Some students cited evidence from the experiments and/or simulations as inspiration for changes in their models. Explicit references to evidence from experiments or other class activities were mentioned after revisions 1 and 2 (21% and 13% of the responses, respectively), but not mentioned after the third and fourth revisions (Table 2). This could have been due

to the structure and order of the questions asked of the students, or perhaps because students tended to focus more on having the appropriate variables and relationships in their later model revisions. Observation notes suggest that more could have been done to support students in making links between their experimental results and the predictions their models generated.

An example where a student cited evidence from experimental results was S21's written comments after the first model revision: "The biggest reason as to why I made the changes was because of the recent experiment that we conducted in which we learned about the relationship between pressure and volume. Also, based on this new knowledge, I removed things I thought weren't applicable to the model." Sometimes the link between experiment and model revision was more tenuous. For instance, to explain why she and her partner added a link to their model, S11 wrote, "We connected amount of air to pressure.... We did this because we learned that the amount of air affects pressure." She did not mention where she had learned this relationship, although she correctly answered questions about an experiment she had conducted that explored it.

Aspect 3 Discussion

The instructional unit included explicit cycles of evidence gathering through exploration of phenomena and activities and model building to promote development of this aspect of modelling practice (see Table 1). Model revisions suggest that students were incorporating new evidence into their models as they encountered it, although they seldom cited empirical evidence or their reasoning about it as justification for those revisions. One reason may be that the teacher and curricular materials did not ask students to draw explicit connections between the sources of evidence and the changes they made to their models. Due to the critical nature of using evidence to justify model design, a greater emphasis on this could have been incorporated into class discussions and model reflection questions.

Aspect 4: Interpreting the Behaviour of a Model to Determine Its Usefulness in Explaining and Making Predictions About Phenomena

S13 and S20 exhibited this aspect of system modelling practice. They created a model that was testable, compared it with expected real-world behavior, and revised it accordingly. They also revised it to more clearly address the unit's driving question, drawing a line of cause and effect all the way from *Steam Cleaning to Implosion of Tanker* (Fig. 6), although aspects of that chain were still problematic. They ran their model multiple times, and attempted several

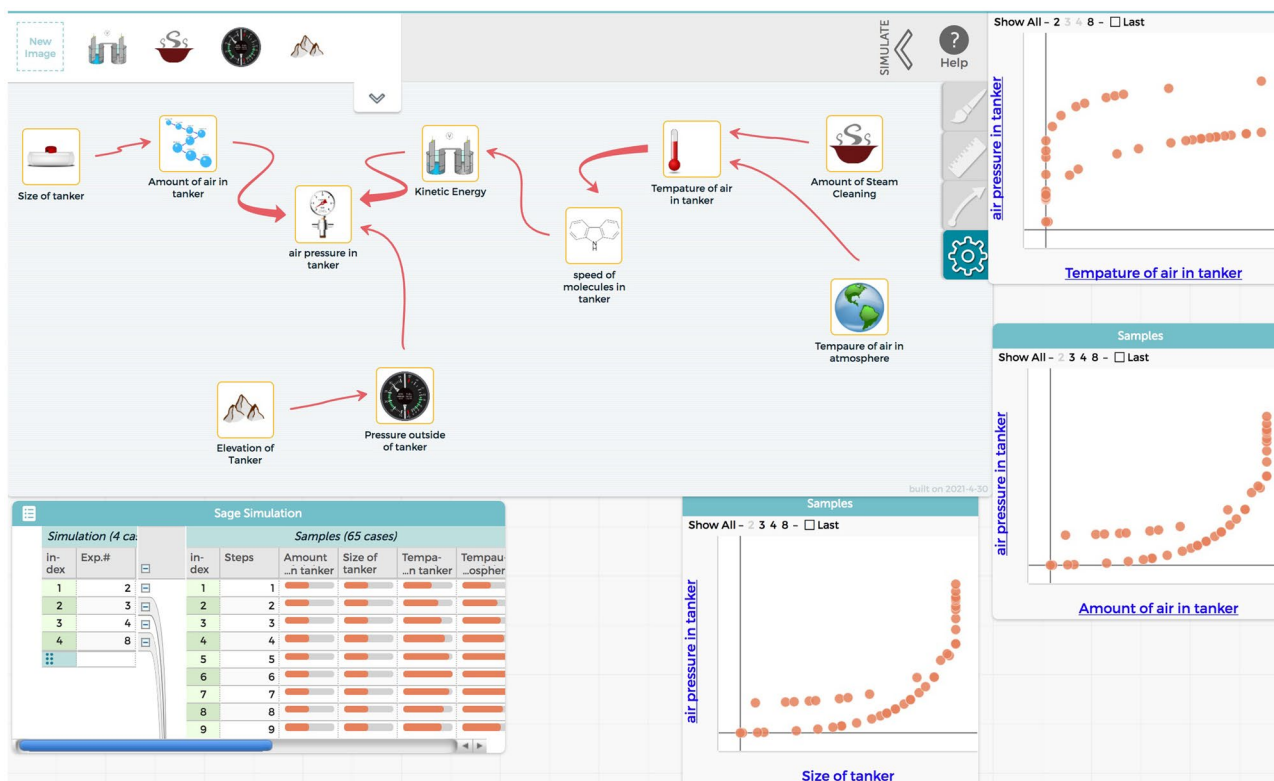


Fig. 6 Third model revision of pair E (S7 and S17)

visualizations of the output, but appeared to have trouble deciding which relationships would be useful to explore. Nevertheless, they exhibited a clear ability to interpret the behavior of the model in terms of the visual representations of relationships, writing after their final revision, ‘We added steam cleaning because it increased the temperature, which increased the kinetic energy, which increased the molecule speed and number of collisions, which increased the pressure of the tank.’ This is an impressive five-link chain that correctly described most of their model.

However, many of the students struggled with this aspect. This struggle was evident in how pair E students (S7 and S17) had *air pressure in tanker* as the outcome variable in all of their models (the last two of which are shown in Figs. 6 and 7), but never explicitly connected this to the driving question in any of their model reflection responses. Nonetheless, it was evident from their interviews that both students understood the connection between *air pressure in tanker* and the driving question about the tanker implosion. To interpret the behavior of their model, they ran it and constructed graphs of the relationships between three different variables and *air pressure in tanker*. Two of their graphs showed expected relationships; one did not. They could have used this unexpected behavior to evaluate the appropriateness of their model to address the driving

question and to troubleshoot relationships in their model, but they did not do so. An excerpt from the final interview with S17 offers one possible explanation for why they were unable to do this on their own. After the student explained their model (Fig. 7) and exhibited some understanding of the phenomena involved, the interviewer asked him to trace lines of cause and effect between *Size of tanker* and *air pressure in tanker*. S17 responded: ‘... this is saying that the bigger the size of the tanker, the larger the amount of air in the tanker. So if there’s more air in the tanker, I’m guessing that it would have a higher air pressure.’ In this short, two-link causal chain, S17’s reasoning about each separate link would be correct only if the tanker could change size for the first relationship (bigger tanker means more air), but remain a constant volume for the second relationship (more air means higher pressure in tanker). Further consideration might have shown him that this reflected an inconsistency in the logic of his model and perhaps in his conceptual understanding of the cumulative effects of relationships.

In written responses to the model reflection questions, nine students (from five of the pairs) did mention changing their model to make it better explain why the tanker imploded or to connect the behavior of their model to real-world behavior, but most did so only once (Table 2).

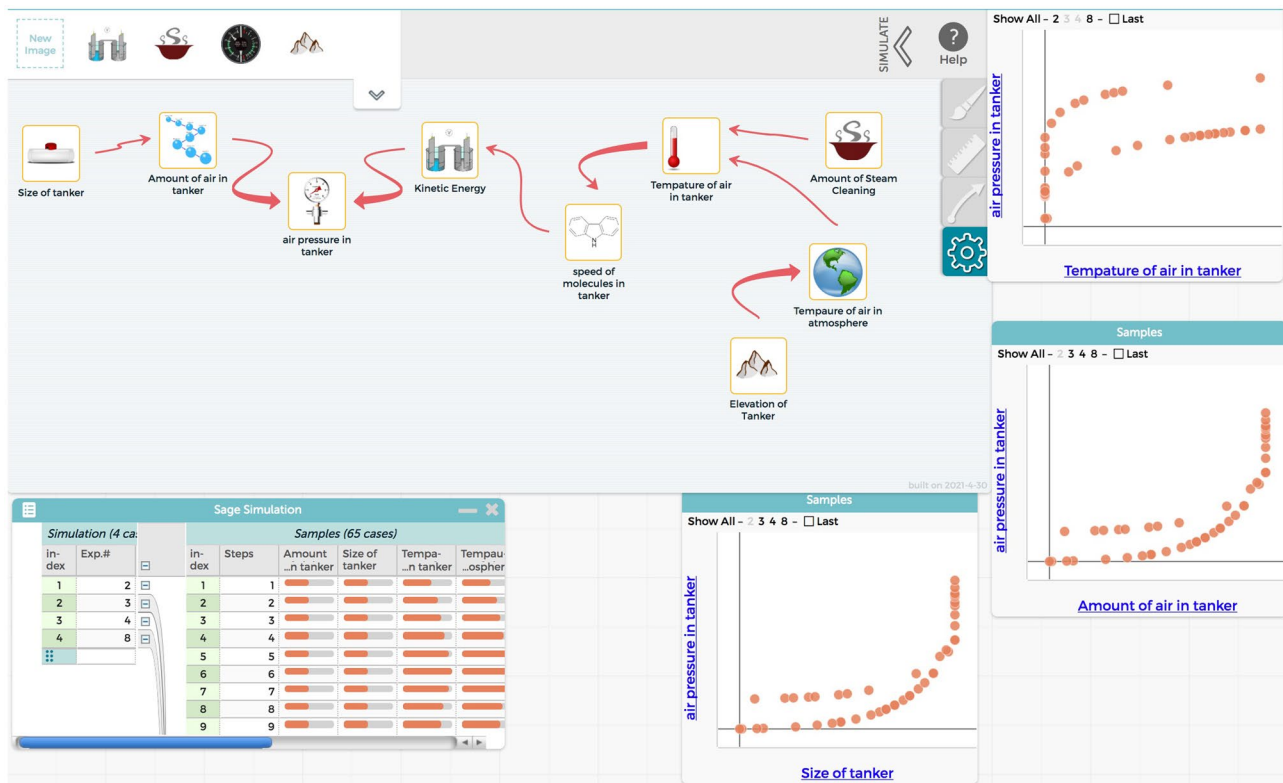


Fig. 7 Fourth model revision of pair E

Aspect 4 Discussion

Students varied in how explicitly their models were designed around the driving question versus a more generic model of the emergent properties of gases. Regardless of the model focus, some students appeared to evaluate the behavior of their model one link at a time but did not evaluate the entire chain of cause and effect extending from input variables to outcome for the driving question. These results indicate a limitation in the extent to which these students interpreted the behavior of their models in relation to their usefulness in explaining real-world phenomena. This suggests that additional supports to bring the focus of the modelling activity back to answering the driving question may be needed to help students achieve this aspect of system modelling practice. We believe this should be explicitly addressed in teachers' professional learning programs, in the curriculum, and in classroom instruction.

Concluding Discussion

In line with the goals of scientific modelling practice (National Research Council, 2012; Schwarz et al., 2009), we found evidence of students engaging with system modelling,

especially Aspects 1 and 2, as they constructed, used, evaluated, and revised their models to explain the phenomenon under investigation. This engagement was demonstrated by determining, testing, and revising the boundaries of the system and the relationships between the variables in the models, as suggested by Harrison and Treagust (2000). This was not an easy and straightforward process in many cases. Nonetheless, most students were able to make progress toward achieving Aspects 1 and 2 of system modelling practice.

In Table 2, we can see that the students focus on Aspect 1, considering which variables to include and how to label them, decreased during the last model revision. This makes sense because at this point in the unit, all aspects of the phenomenon to be introduced had been introduced, and the focus was on refining the relationships, not on adding new variables. In fact, we do see that attention to Aspect 2, considering the relationships and how to revise them, increased to 100% for the last two model revisions.

However, focus on Aspects 3 and 4, using evidence and relating the model behaviour to the real world, also waned. A focus on these aspects may have peaked during the first revision, after students had watched videos and completed an experiment. One take-away is that the connection to real world aspects (Aspects 3 and 4) may need to be strengthened

at points in the unit when students are not conducting an experiment, especially near the end of the unit when this information could help them evaluate and revise their models. In addition, students may need a different kind of scaffolding in how to use real world information to help them improve their models.

The four aspects of system modelling practice (Bielik et al., 2019) were used here as a way to evaluate student engagement in the process of understanding phenomena through constructing, using, evaluating, and revising models, as well as changes in that engagement as a curriculum unit progressed. We suggest that these aspects can also provide a framework for curricular designers who want to promote young students' systems thinking, causal reasoning, and modelling practice. The four aspects can also be useful as an epistemic framework for teachers and students when reflecting on how to construct, use, evaluate, and revise models in the classroom.

Several challenges with causal reasoning were experienced by the students in this study. These challenges relate to Aspects 3 and 4 and included providing evidence and reasoning for their chosen variables and relationships in the models and explaining how their models address the driving question of the unit. These challenges align with those described by Schauble (1996) and Koslowski and Masnick (2002). However, these students improved their model-based explanation in each model revision, which suggests that the technology-rich environment and curricular materials supported students' causal reasoning abilities. For instance, the students used the models as a common referent when discussing cause and effect in the system.

One limitation of this study is that, as mentioned in the context section, it was performed by honor students with high academic achievements, and future studies should test this conclusion in other classrooms.

Challenges students face when engaging in systems thinking and understanding complex models were evident in our results. It may not be realistic to expect students to achieve a high level of proficiency in all four aspects of system modelling practice following a brief Introduction to Modelling unit and a single instructional unit. While our results indicate that the modelling tool and curricular supports provided students with a strong foundation to develop their systems thinking, students will likely require repeated experiences in multiple learning environments to achieve mastery of the modelling practice. It is important to mention that the unit was revised following the enactment described in this study to focus on engaging students in comparing their model outputs to their collected experimental data from their self-generated laboratory experiments.

In conclusion, students progressed in Aspects 1 and 2 with their ability to choose appropriate variables, determine relevant relationships, and clarify causal mechanisms

to make their relationships more direct. Less success was observed in Aspects 3 and 4 with respect to using evidence to support model design and explicitly linking the overall behaviour of the model to the driving question about the phenomenon under investigation. Our classroom implementation suggests that identifying appropriate curricular and teacher supports may be key.

Recommendations

Based on the in-depth analysis of the results of the implementation described here, we provide the following recommendations for teachers wishing to provide an introductory system modelling experience for students. These can be implemented in activities that engage secondary students in scientific systems modelling, especially those that take advantage of technology-rich modelling environments. In addition to the specific recommendations below, we suggest that the four aspects of system modelling can be used as an epistemic framework to support teachers and students when engaging in the modelling practice and when constructing, using, evaluating, and revising models.

1. Focus on using evidence to support evaluation of model components and the relationships defined between them. Running simulations to evaluate the outcome of a model in comparison with real-life data is an important advantage when using a modelling tool. This is a crucial checkpoint in each step of constructing, using, evaluating, and revising the model. In the results of the unit enactment, we found that students did not often evaluate their models in comparison with real-life data without explicit support. This is in line with findings from Chinn and Brewer (1993), who note that when a conflict between real-world data and model output occurs, students do not necessarily consider revising their models or theory, but sometimes ignore or disregard conflicting results. Although it is beyond the scope of this study, we suggest that teachers explicitly address these cases and be able to support students in resolving such conflicts. This recommendation should be further examined in future research. We also recommend using activities in the curricula and learning materials, such as text boxes in the software, to direct students to explicitly state the goal of their model and the evidence they have to support their claims.
2. Evaluate models in whole-class and small-group discussions. Interviews suggest that getting students to talk through their models and graphs can be helpful in prompting them to recognize inconsistencies and problematic model behavior. Student-centered discussions are powerful tools for promoting the sharing of ideas related

to the phenomenon being modelled and for engaging in activities that support growth in system modelling practice. These discussions can include peer review and whole-class discourse around student models.

3. Frequently revisit the overarching phenomenon and the driving question the models are intended to address. Students can easily lose the big picture of what they are modelling. In the enacted unit, the teacher often directed the students to consider the driving question and results indicate that students focused on making sense of the anchoring phenomenon (the imploding tanker) in their models. We suggest that teachers frequently refer back to the driving question and collect student questions and comments related to it. We also suggest that the driving question should be consistently visible for the students while they develop and revise their model. This could be done by adding a text box in the modelling tool that includes the driving question.

Appendix 1. Description of Curricular Unit

The phenomenon at the core of the unit was described in (Bielik et al., 2019) and involved an oil tanker that imploded after steam cleaning. This compelling phenomenon was used to support students in reasoning about the causes of air pressure and the relationships between air pressure, volume, and temperature of a gas.

One possible application of the three-dimensional learning in *A Framework for K-12 Science Education* (National Research Council, 2012) is to orient student work around a central question related to a particular phenomenon. The driving question for this unit was, “How can something that can’t be seen crush a 67,000 pound oil tanker made of half inch steel?”.

To develop the unit we focused on a performance learning goal from the Next Generation Science Standards (NGSS Lead States, 2013): HS-PS3-2—*Develop and use models to illustrate that energy at the macroscopic scale can be accounted for as a combination of energy associated with the motion of particles (objects) and energy associated with the relative positions of particles (objects)*. The imploding gas tanker is a phenomenon where the gas particles and their energy of motion can be accounted for both on the macroscopic scale (the emergent observable behaviors) and the atomic scale (in relation to the motion and energy of the gas particles themselves).

An initial classroom brainstorm around the driving question was used to set the context of the unit and provide information about students’ prior knowledge related to pressure, volume, temperature, and motion of gas particles. Students engaged in several online and offline activities to learn about appropriate variables and to create relationships

between variables in the modelling tool. Next, using their prior knowledge about particles in different states of matter and phase changes, together with information gathered from a video of the phenomenon and small-group discussions, students developed an initial model to explain the tanker phenomenon, using the modelling tool to demonstrate the cause and effect relationships between potential variables that could affect a system of gas molecules. The idea was for students to use what they had learned from the gas laws to construct a causal model of why the implosion occurred, not to model the implosion itself (which would have required a threshold-type relationship between pressure and volume).

Students worked in pairs using sensors and other equipment to conduct student-designed investigations into the relationships in each of the three gas laws. After each investigation, students compared the results of their experiments to data generated by their models and were instructed to re-evaluate and revise their models. For example, to explore the relationship between volume and temperature, a common experiment designed by the students was to measure the change in volume of a balloon placed over a flask that was inserted into water baths of different temperatures. They then compared the results of these experiments with the parts of their models that addressed temperature and volume.

As the unit progressed, students began to include covariational and mechanistic attributes of the causal relationship between changes in temperature and pressure. The unit included opportunities for peer evaluation of the models and class discussions. The teacher prompted the class to respectfully review each model and predict the effect if one or more independent variables were changed. After revising their models, students usually worked in pairs to answer follow-up questions online, although they did not have to type the same response.

Appendix 2. Description of Criteria Used to Evaluate the Four Aspects

Aspect 1: defining the boundaries of the system by including components in the model that are relevant to the phenomenon under investigation.

Feature A—distinguishing objects and variables.

- Mentions that variables need to be measurable
- Mentions that they changed a variable to be measurable on low-to-high scale

Feature B—appropriate size and scope.

- Mentions deleting a variable because they decided it was not needed or not relevant or did not change appreciably, or added a variable because they thought it was needed

to help explain what happened or because they thought it had an important effect

- Includes question about whether a variable is appropriate for model

Aspect 2: determining appropriate relationships between components in the model.

Feature A—appropriate relationships: logically correct.

- Mentions they added or subtracted a relationship because otherwise the model did not show what really happened or did not make sense
- Mentions they changed a relationship from positive to negative or vice versa because otherwise that relationship did not show what really happened or did not make sense
- Mentions they changed the direction of an arrow because they thought it made more sense
- Question about appropriate relationships

Feature B—appropriate relationships: direct vs indirect.

- Mentions directness as a factor in their decision about a link
- Question about directness of relationships
 - Aspect 3: using evidence and reasoning.
- Mentions evidence from experiment or from some other class activity as a reason for making a change.
 - Aspect 4: interpreting behavior of model to determine usefulness.
- Mentions changing the model to make it better able to explain why the tanker imploded
- Mentions graphical relationship in connection w real world behavior
- Mentions behavior of part of model (multiple connections) or model as a whole in connection with real world behavior
- Includes question about agreement between model behavior and real world

Acknowledgements This material is based upon work supported by the National Science Foundation under Grant Nos. DRL-1842035, and DRL-1842037. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Author Contribution All authors contributed to the study conception and design, material preparation, data collection and analysis. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

We declare full compliance with the required ethical standards in this study. All participants of this study provided informed written parental

consent and personal verbal consent under the research institute IRB regulations.

Ethical Statement The data set was collected with IRB approval from the fifth's author organization.

Consent Statement Consent forms were collected from all the participants of this study.

Conflict of Interest The authors declare no competing interests

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bielik T., Stephens L., Damelin D., & Krajcik J. (2019). Designing Technology Rich Environments to Support Student Modeling Practice. In Upmeir Zu B., Kruger D., & Van Driel J. (Eds.), *Towards a Competence-based View on Models and Modeling in Science Education*. Springer (pp. 275-290). Springer International Publishing.
- Booth Sweeney, L., & Sterman, J. D. (2000). Bathtub dynamics: Initial results of a systems thinking inventory. *System Dynamics Review*, 16(4), 249–286.
- Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, 63(1), 1–49.
- Clement, J. (2000). Model based learning as a key research area for science education. *International Journal of Science Education*, 22(9), 1041–1053.
- Damelin, D., Krajcik, J., McIntyre, C., & Bielik, T. (2017). Students Making Systems Models: An Accessible Approach. *Science Scope*, 40(5), 78–82.
- Dörner, D. (1980). On the difficulties people have in dealing with complexity. *Simulation & Games*, 11(1), 87–106.
- Finzer, W., & Damelin, D. (2016). Design perspective on the Common Online Data Analysis Platform. In *Proceedings of the 2016 Annual Meeting of the American Educational Research Association (AERA)*, Washington, DC, USA.
- Forrester, J. (1968). *Principles of systems* (2nd ed.). Pegasus Communications.
- Gould-Kreutzer, J. (1993). Forward: System dynamics in education. *System Dynamics Review*, 9(2), 101–112.
- Harrison, A. G., & Treagust, D. F. (2000). A typology of school science models. *International Journal of Science Education*, 22(9), 1011–1026.
- 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.

- Jacobson, M. J., & Wilensky, U. (2006). Complex systems in education: Scientific and educational importance and implications for the learning sciences. *The Journal of the Learning Sciences*, 15(1), 11–34.
- Jonassen, D. H., & Ionas, I. G. (2008). Designing effective supports for causal reasoning. *Educational Technology Research and Development*, 56(3), 287–308.
- Koslowski, B., & Masnick, A. (2002). The development of causal reasoning. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 257–281). Blackwell Publishing Ltd.
- Lehrer, R., & Schauble, L. (2006). Cultivating model-based reasoning in science education. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 371–387). Cambridge University Press.
- Louca, L. T., & Zacharia, Z. C. (2012). Modeling-based learning in science education: Cognitive, metacognitive, social, material and epistemological contributions. *Educational Review*, 64(4), 471–492.
- Martinez-Moyano, I., & Richardson, G. (2013). Best practices in system dynamics modelling. *System Dynamics Review*, 29(2), 102–123.
- National Research Council (NRC). (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. The National Academies Press.
- NGSS Lead States. (2013). *Next Generation Science Standards: For states, by states*. The National Academies Press.
- Nicolaou, C. T., & Constantinou, C. P. (2014). Assessment of the modeling competence: A systematic review and synthesis of empirical research. *Educational Research Review*, 13, 52–73.
- Passmore, C., Gouvea, J. S., & Giere, R. (2014). Models in science and in learning science: Focusing scientific practice on sense-making. In M. R. Matthews (Ed.), *International handbook of research in history, philosophy and science teaching* (pp. 1171–1202). Springer.
- Perkins, D., & Grotzer, T. (2000). Models and moves: Focusing on dimensions of causal complexity to achieve deeper scientific understanding. Paper presented at the *American Educational Research Association Annual Conference*, New Orleans, LA.
- Richmond, B., Peterson, S., & Vescuso, P. (1987). *An academic user's guide to STELLA*. High Performance Systems, Inc.
- Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education*, 92(3), 499–525.
- Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, 32(1), 102–119.
- Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46(6), 632–654.
- Stratford, S. J., Krajcik, J., & Soloway, E. (1998). Secondary students' dynamic modeling processes: Analyzing, reasoning about, synthesizing, and testing models of stream ecosystems. *Journal of Science Education and Technology*, 7(3), 215–234.
- Tadesse, A., & Davidsen, P. (2020). Framework to support personalized learning in complex systems. *Journal of Applied Research in Higher Education*, 12(1), 57–85.
- Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, 8(1), 3–19.
- Yoon, S. A., Anderson, E., Koehler-Yom, J., Evans, C., Park, M., Sheldon, J., & Klopfer, E. (2017). Teaching about complex systems is no simple matter: Building effective professional development for computer-supported complex systems instruction. *Instructional Science*, 45(1), 99–121.
- Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, 27(2), 172–223.

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