Improving NGSS focused model-based learning curriculum through the examination of students' experiences and iterated models

Byung-Yeol Park, Todd Campbell, Miriah Kelly, Ron Gray, Chester Arnold, Cary Chadwick, Laura M. Cisneros, David Dickson, David M. Moss, Laura Rodriguez, John C. Volin & Michael R. Willig

Published online: 24 Sep 2021.

Submit your article to this journal ✏

View related articles 🔍

View Crossmark data

Routledge
Taylor & Francis Group

Check for updates

Improving NGSS focused model-based learning curriculum through the examination of students' experiences and iterated models

Byung-Yeol Park [a], Todd Campbell [a], Miriah Kelly [b], Ron Gray [c], Chester Arnold [d], Cary Chadwick[d], Laura M. Cisneros[e], David Dickson[d], David M. Moss[f], Laura Rodriguez [g], John C. Volin[e,h] and Michael R. Willig[i]

[a]Department of Curriculum & Instruction, Neag School of Education, University of Connecticut, Storrs, CT, USA; [b]Environment, Geography & Marine Sciences, Southern Connecticut State University, New Haven, CT, USA; [c]Center for Science Teaching and Learning, Northern Arizona University, Flagstaff, AZ, USA; [d]Center for Land Use Education & Research (CLEAR), University of Connecticut, Storrs, CT, USA; [e]Department of Natural Resources & the Environment and Institute of the Environment, University of Connecticut, Storrs, CT, USA; [f]Global Education, Neag School of Education, University of Connecticut, Storrs, CT, USA; [g]Department of Education, Eastern Connecticut State University, Willimantic, CT, USA; [h]School of Forest Resources, and Office of the Executive Vice President for Academic Affairs and Provost, University of Maine, Orono, ME, USA; [i]Department of Ecology & Evolutionary Biology and Institute of the Environment, University of Connecticut, Storrs, CT, USA

ABSTRACT

Background: It is crucial to support students in better understanding water and sustainability issues because water plays a vital role in maintaining global ecosystems, including human life. A wide range of curricular and instructional supports like those embodied in model-based learning (MBL) are necessary for teachers to engage students in the core epistemic commitments of the Next-Generation Science Standards (NGSS).

Purpose: The main purpose of this study is to inform theory about students' early attempts to engage in the complex kinds of sense-making experiences inherent in the Framework for K-12 Science Education and the NGSS.

Sample: Data for this study was collected from 74 10th grade students in a high school in the Northwest region of the New England state.

Design and Methods: An explanatory sequential mixed-method research design was used to examine students' learning outcomes, and to better understand these outcomes in connection to their experiences engaging in modeling in the MBL curriculum unit.

Results: The results indicated that students' model scores, the number of concepts in models, and the coherence and sophistication of models improved between their initial and final models. Additionally, the following patterns emerged related to ways in which students engaged in the practice of modeling: (1) students attempted to directly represent what they observed, (2) they

CONTACT Byung-Yeol Park ✉ byung-yeol.park@uconn.edu ▣ Department of Curriculum & Instruction, Neag School of Education, University of Connecticut, Storrs, CT, USA

struggled to pictorially express complex patterns or mechanisms, and (3) students experienced difficulties representing models from a diverse range of perspectives.

Conclusion: The patterns identified across student models, as well as their reports of experiences related to the MBL unit implementation, provided insight into student experiences with models, while also providing meaningful implications for the refinement of the MBL curriculum unit investigated in this research specifically, while informing approaches MBL curricular units aimed at supporting NGSS implementation efforts more generally.

1. Introduction

Science education in the United States is in the throes of a paradigm shift. This shift has been prompted by the development of the *Framework for K-12 Science Education*, subsequently referred to as the *Framework* (NRC 2012), and the *Next-Generation Science Standards* (NGSS) (NGSS Lead States 2013). These science education standards put forth a vision of science teaching and learning that is dramatically different than that which has been historically experienced by students in most classrooms (Reiser 2013; Wilson 2013). More specifically, the standards emphasize three-dimensional learning, whereby students engage in science and engineering practices to use disciplinary core ideas and cross-cutting concepts to explain phenomena or solve problems (National Research Council 2012). The most significant shift related to teaching and learning is the focus on engaging students in 'figuring out' how to explain real-world phenomena or solve problems, using important disciplinary concepts and through the epistemic practices of science, instead of 'learning about' isolated disciplinary concepts divorced of context (Krajcik 2015). Chief among the epistemic practices emphasized in the NGSS for 'figuring out' is modeling and the resultant products of modeling (i.e. models) that are revised over time to explain events that happen in the world through investigations, discourse, evidence, and argumentation (Ko and Krist 2019).

Developing and using models, as a science practice, focuses on 'the process by which scientists [and students] represent ideas about the natural world to each other, and then collaboratively make changes to these representations over time in response to new evidence and understandings' (Ambitious Science Teaching 2015, 1). A scientific model aims to describe, explain, and predict natural phenomena and communicate scientific ideas with others (Clement 2000; Oh and Oh 2011). Various strategies with geometric figures, lines, numbers, symbols, arrows, and even explanatory words can be used in developing a model, while the explanatory power of the model depends on how the model developer decides to represent information (Campbell et al. 2019; Osborne, Rafanelli, and Kind 2018). Because of the complexity of developing and using models, researchers have noted the benefits that can come from teachers being supported with curriculum and professional development as they learn to engage students in developing and using models (Schwarz et al. 2009; Zangori and Forbes 2016; Vo et al. 2015).

The writers of the *Framework* identified many key drivers that they believed would determine the success of NGSS implementation. Importantly, they pointed to the critical need for consideration of how implementation would proceed with respect to curriculum,

instruction, and assessment (National Research Council 2012). Within these identified areas of concern, curriculum, and instruction were foregrounded. Related to curriculum, others have noted the critical role of and need for curriculum resources supportive of the NGSS implementation. As an example, a recent convening sponsored by the National Academies of Science's Board on Science Education (BOSE) identified how only a limited quantity of high-quality curriculum resources exist. BOSE pointed to the need for high-quality instructional materials that are designed to attend to the essential epistemic elements prioritized in the NGSS (e.g. making sense of phenomena; three-dimensional learning) (Achieve Inc. 2017; Bybee and Chopyak 2017; NASEM 2018). However, at the time of this research, it was clear that high-quality curriculum was not yet available at scale to support the ambitious aims of implementing the NGSS (c.f. Banilower 2019; Smith 2020).

2. Theoretical framework and relevant literature

2.1. The framework and the NGSS

One of the main priorities from the *Framework* and the NGSS is to support students in understanding how science knowledge is developed and used to explain natural phenomena and solve problems (Next Generation Science Standards Lead States 2013; National Research Council 2012). The *Framework* and NGSS provide a vision whereby students engage in three-dimensional learning. Rather than focus on isolated decontextualized or generalized content and inquiry separately, three-dimensional learning can be understood as engaging students in science and engineering practices (SEPs; i.e. knowledge production practices) to use disciplinary ideas (DCIs) and crosscutting concepts (CCCs) to explain phenomena or solve problems of consequence (Krajcik 2015). In this type of three-dimensional learning that prioritizes a focus on practice, emphasis is placed on the use of practices (SEPs) to explore and refine ideas (DCIs & CCCs) in the context of meaningful pursuits (i.e. explaining natural phenomena or problem-solving) (Ford 2015). When this is considered in the context of high-quality materials that might support teachers and learners in classrooms, Ko and Krist (2019), referencing NSTA (2014), point to the following core epistemic features of learning outlined in the recent standards documents supported by high-quality curriculum: (1) an anchoring phenomenon as the problem space for instruction, (2) a model or explanatory account that is revised over time to explain the problem space, (3) the use of investigations to incrementally build and refine a model-based account, and (4) the use of discourse, evidence, and argumentation for generating claims supported by evidence to construct and refine the explanatory account. However, at the time of this research it was widely noted that high-quality curriculum is important but only sparsely available (cf. Bybee and Chopyak 2017; National Academies of Sciences, Engineering, and Medicine (NASEM) 2018).

2.2. Modeling and model-based learning (MBL)

Modeling. Researchers in science education and the most recent standards documents (i.e. Next Generation Science Standards Lead States 2013; National Research Council 2012) have called for positioning students to use modeling as an epistemic practice for

organizing their day-to-day sensemaking work across the arc of learning experiences to incrementally make progress in explaining events that happen in the world (Campbell and Fazio 2018; Ford 2008; Ko and Krist 2019). In this, our definition of modeling aligns with Passmore, Gouvea, and Giere (2014) in how it takes into account the forms that models take on in the representation of targets (e.g. phenomena) and their function in support of reasoning. Here, we specifically focus on modeling in classrooms as the process in which students make representations of a phenomenon or events by drawing pictures, using diagrams, and using text (Gilbert and Boulter 2000) (see Figure 1). However, most important is the way in which models function in the classroom. Functionally, this act of representing serves as a sense-making tool as students collaboratively propose, argue, and decide among ideas, especially related to how they think ideas might be connected for explaining or predicting natural phenomena or events. Additionally, the process of modeling does not end after students construct an initial model, as students recursively engage in modeling overtime across a unit in connection to learning and reflection activities (e.g. investigations that generate evidence).

Modeling (Figure 1) can be envisioned as the interplay between a phenomenon or an event that happens in the world, a question developed by an agent (e.g. student or scientist) that is aimed at better understanding something uncertain about the
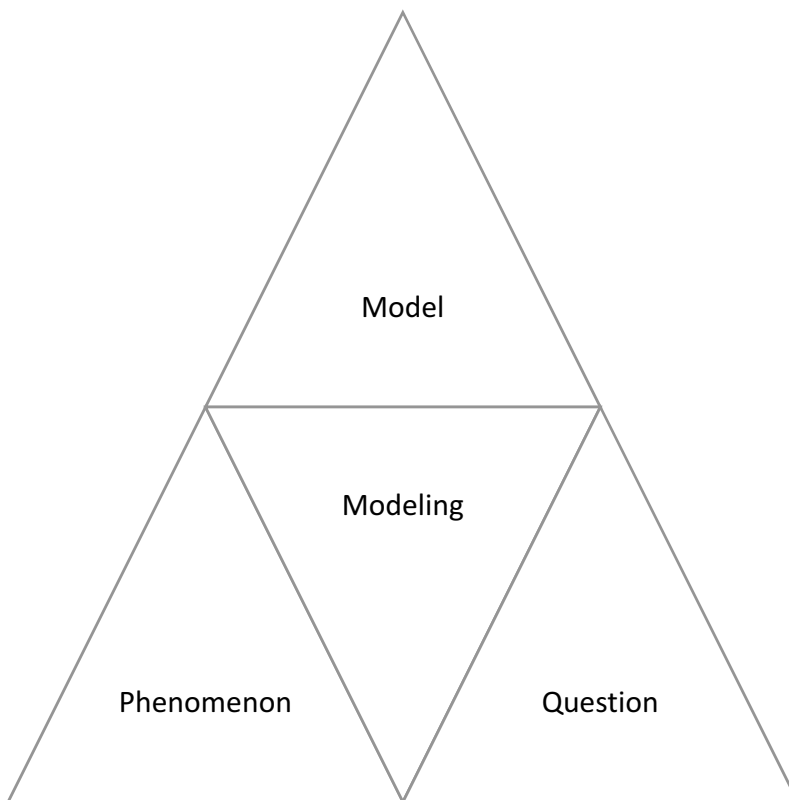

Figure 1. The Modeling Triangle depicts important interacting features that contribute to the process of modelling (modified from Model Based Biology 2017).

phenomenon, and the model that is constructed by the agent for the purpose of answering the question (Model Based Biology 2017). Here, it can be seen that the practice of modeling is inextricably connected to the epistemic aim of the agent (e.g. students) (Gouvea and Passmore 2017).

Further still, students' modeling experiences are grounded in common practices found in scientific inquiry, such as explanation, conjecture, testability, revision, and generativity, so that they are able to understand scientific practices, as well as the content of scientific disciplines as they engage in the practice of modeling (Passmore and Svoboda 2012). As students engage in investigations (broadly construed) (Ko and Krist 2019) or are introduced to disciplinary ideas and given space to reason with them as part of interactive direct instruction (Windschitl, Thompson, and Braaten 2008), they are able to refine their initial intuitive ideas and use these experiences in subsequent modeling opportunities (Clement 2000). Cumulatively, through a series of recursive modeling opportunities, students can enhance the explanatory, predictive, and communicative power of their models to support their experiences engaging in collaborative sense-making with others (Baumfalk et al. 2019; Böttcher and Meisert 2011; Kenyon, Davis, and Hug 2011). In the end, through positioning students to engage in modeling, they will better understand the role of SEPs in the knowledge production process that is central to the scientific enterprise (Gouvea and Passmore 2017; Krajcik 2015).

2.2.1. Model-based learning

In this research, MBL was defined, in alignment with others, as teaching and learning that prioritizes modeling as an epistemic practice that can serve as an anchor around which students necessarily engage in other SEPs in their sense-making pursuits (Passmore, Stewart, and Cartier 2009; Svoboda and Passmore 2011). This is important, since Ko and Krist (2019) identified how the recursive revisions of a model or an explanatory account to explain a specified problem space is an essential epistemic element of NGSS-aligned curriculum materials. MBL curriculum recursively engages students in developing, testing, and revising models as a central support for student learning (Schwarz et al. 2009; Windschitl, Thompson, and Braaten 2008). In fact, research indicates that MBL helps students to develop facility in explaining phenomena in ways that both build on and challenge their existing ideas. Examples of such research across a range of disciplinary foci include engaging students in reasoning about phenomena in relation to the following: buoyancy (Campbell and Fazio 2018), plants (Zangori and Forbes 2016), properties of and changes in materials (Acher, Arcà, and Sanmartí 2007), water cycle (Vo et al. 2015), and molecular structures (Khan 2007). In addition, many model-based learning focused studies involved developing instructional and conceptual frameworks (Passmore, Stewart, and Cartier 2009; Schwarz and Gwekwerere 2007). Schwarz et al. (2009) developed a learning progression for model-based inquiry to support curricular development. While many studies have been undertaken to help teachers support students' model-based learning, concerns have surfaced in connection to how model-based curriculum and teachers perceive and frame modeling for students. As an example, researchers have identified how learning in science classrooms with modeling remains focused on students' acquisition of canonical science knowledge (Guy-Gaytán et al. 2019; Russ and Berland 2019; Thompson et al. 2016). In fact, researchers have noted how in some model-based learning classrooms, more emphasis was placed on students having correct

answers, rather than being supported to iteratively build scientific explanations from their initial ideas (Russ and Berland 2019). Gouvea and Passmore (2017) identified this tension in classrooms and framed it as a distinction between the use of 'models of' scientific concepts that mainly focus on students learning about canonical scientific knowledge compared to 'models for' 'figuring out' how to apply knowledge to explain events that happen in the world. However, while a limited number of studies have been undertaken to consider modeling as a critical epistemic tool supportive of student learning, far less research is available that illustrates how teachers' implementation of the NGSS, anchored in a focus on MBL curriculum, can support students model-based explanations.

2.3. *Water and sustainability*

Water plays a vital role in maintaining global ecosystem as well as all living things, including humans. The conservation and protection of natural resources like water depend on the sustainability of interactions between air, soil, plants, animals and water (Dickerson et al. 2007). As an example, a stream, river, pond, and lake not only provide nutrition to living things but, through circulation and filtering, help clean and refresh water. However, population growth over the past half century has led to increased demands for clean and safe water (UNWWAP, 2014). According to the 2012 report by United States Census Bureau, the world's population, that was around 2.5 billion in 1950, is predicted to reach around 9 billion by 2050. This population increase correlates positively with the consumption of more food and the need for more homes, among other impacts, that increase water consumption use (UNWWAP, 2015). Globally, water in many regions of the world it is not accessible. More specifically, each person uses approximately between 20 and 50 liters of clean water daily for drinking, cooking, and hygiene, yet more than one in six people still suffer from lack of clean and safe water, especially in the developing world such as in the Middle East and north Africa (United Nations World Water Assessment Programme 2014). Relatedly, the changing landscape to meet the demands for food, homes, and facilities also influences the lack of access to water resources. The 2015 and 2016 reports by the United Nations Food and Agriculture Organization (2015, 2016) estimated that the Earth's forest area fell by 129 million hectares (i.e. 3.1%) in the period 1990–2015 as a direct result of human activities.

In order for students develop environment literacy, especially about water resources, the *Framework for K-12 Science Education* (National Research Council 2012) elevates a focus on water in the standards through explicit and situated contexts in the following earth and space science disciplinary core ideas: 'the roles of water in Earth's surface processes', 'natural resources', and 'human impacts on Earth systems'. Importantly, these disciplinary core ideas orient learners to the mechanism and role that water plays in all the planet's dynamic systems and provides learners with insight into the importance of the 'balance between economic, social, environmental, and geopolitical costs and risks, as well as benefits' (191). Moreover, the *Framework* (National Research Council 2012) provides the example of 'the need for supplies of clean water' as one of the major global challenges humanity faces. Because the *Framework* and the NGSS make up a set of compendium documents that collectively serve as the standards science education in the U.S., these same disciplinary core ideas highlighted in the *Framework* are also

foundation to the performance expectations identified in the NGSS. This further points to the importance and need to support learners environmental literacy, especially connected to local water resources which is the focus of the MBL curriculum examined in this research.

3. Research questions

Based on the noted importance of curriculum and instruction (e.g. National Academies of Sciences, Engineering, and Medicine (NASEM) 2018), and the small number of studies about student learning in science education related to NGSS-designed curriculum, we took advantage of the implementation of a USDA-funded science teacher professional development (PD) project that elevated the epistemic practice of modeling and the incremental development of student models over time to better understanding the experiences of students as they engaged in MBL, which represents and instantiation of NGSS-designed curriculum. In addition to informing improvement to our particular MBL curriculum unit, we sought to inform theory about MBL units design more generally, as well as to better understand students early attempts to engage in the complex kinds of sensemaking experiences inherent in the *Framework* and NGSS (i.e. modeling and the refinement of models over time), especially in relation to the emergent student products (i.e. pre- and post-models) and the students' reported experiences as they thought about their respective work across an NGSS-designed unit.

Given this aim, the following questions guided our research:

(1) To what extent and in what ways does an MBL curricular unit support secondary students' model-based explanations of water and sustainability in local contexts across a unit?
(2) What patterns were observed in student models that can inform focused improvements to MBL units and student learning with models?

4. Methods

4.1. Context

The project included a three-day PD workshop for secondary science teachers, designed and facilitated by multidisciplinary experts in water resources, land use, climate science, science education, and geospatial technology, with the aim of engaging teacher participants in exploring local water resources issues. The MBL curricular unit as designed by the project leaders (Freidenfelds et al. 2020), was grounded in the four epistemic commitments of NGSS, as outlined earlier by Ko and Krist (2019) (e.g. anchoring phenomenon to set problem space for instruction, model incrementally revised over time to explain the problem space). Additionally, the structure of the PD design was such that teacher participants engaged in 'learner hat' and 'teacher hat' experiences throughout the workshop, so that they could be supported in learning more about the NGSS, while simultaneously preparing to enact the curricular unit in their classrooms (Reiser 2013).

The MBL curriculum unit was developed with a unit planning template aligned to Stroupe and Windschitl (2015) framework for Ambitious Science Teaching which focuses on '1) planning a unit around a "fundamental science idea", 2) eliciting and activating students' prior knowledge about a puzzling phenomenon (for the purpose of adapting instruction), 3) helping students make sense of scientific ideas and activities, and 4) pressing students to construct evidence-based explanations'. The unit was focused on a water theme that facilitates connections between larger-scale issues (global and local natural resources) to environmental problems like water quality and water contaminants associated with land-use. The unit began as students were introduced to the phenomenon (e.g. three different stream sampling sites surrounded by differing land use or land cover). The first component of the phenomenon presentation consists of aerial photos of three different areas surrounding the study sites. When presented with the aerial photos, students were asked, 'What do you predict is the state of the local water resources at the three sites, and why?' This was initial part of the unit is focused on eliciting students' initial ideas as students were supported to construct their initial group models of the phenomenon. In their models, students tried to explain the differences that might be found at each of the sites due to the surrounding land cover (e.g. agricultural, forest, urban) and other factors they thought were important. After eliciting initial ideas, students were engaged in a series of activity in the middle of the MBL unit that were intended to 'put ideas on the table' so that students could engage with these ideas before or as they considered how these ideas could help them explain the state of the local water resources at the three sites that served as the anchoring phenomena for the unit. Activities in the middle of the unit, 'to put ideas on the table', included data collection trips to the local water resources to collect water quality and biodiversity data for comparison across sites, using an interactive online watershed runoff simulation tool, Wikiwatershed – https://wikiwatershed.org, to determine the relationships between land cover, hydrologic soil groups, runoff, infiltration, and evapotranspiration during a storm, an exploration of a storymap of the state where the students lived to examine how human activities have changed the landscape over the last several decades, and a reading set focused on ecological services. After students considered the evidence and ideas they were introduced to in the middle of the unit, they return to their initial models at the end of the unit to revisit the driving question and revise their initial models based on what they have learned to help explain the phenomenon.

After the workshop, using purposive sampling (Patton 2002), we selected one science teacher willing to participate in our study that agreed to implement the unit with her students, collected pre- and post-student models as well as journal entries. The unit focused on water and sustainability as students were oriented to predicting and explaining the water quality of two different local bodies of water through the use of models. The workshop took place in August 2017, and students engaged in the MBL unit during the 2017–2018 school year.

4.2 Research design

We employed an explanatory sequential mixed-method research design (Creswell and Plano Clark 2011). Within this design, researchers collect quantitative and qualitative data sequentially in two phases – first collecting quantitative data and then collecting

qualitative data to help interpret or elaborate on the quantitative results (Creswell 2012). More specifically, we first quantitatively analyzed students' learning outcomes (i.e. pre- and post-student models) to address the first research question. To address the second research question, we looked for trends in the ways in which students engaged in modeling, before we sought to better understand these trends by qualitatively analyzing student reflections about their models and modeling experiences as captured in their journal entries. Furthermore, the evidence-based patterns in models and modeling identified through the analysis and interpretation of the second research question helped us to better understand and illustrate what features of student learning with models would benefit from more attention in future iterations of the our NGSS designed unit implementation or that of others.

4.3. Participants and data collection

Purposive sampling involves identifying and selecting participants who are available and willing to participate, communicate, and reflect on their experiences (Etikan, Musa, and Alkassim 2016; Patton 2002). Among 21 science teachers who participated in the workshop, 19 teachers expressed that they intend to implement the developed unit for their classroom during the 2017–2018 school year. At the beginning of the school year, three teachers had a specific implementation plan. We selected one teacher from these three teachers by using purposive sampling, based on how much her teaching plan was aligned to what we proposed during the workshop and the number of the classrooms she led that could provide us with students' modeling examples. 74 10th grade students between the ages of 14 and 16 years old participated in this research. All students were from one teacher's four classrooms (i.e. 19 students each from three classrooms and 17 students from one classroom) as she implemented the MBL unit. The teacher held a master's degree in education and was a veteran teacher with 17 years of teaching experience in Biology. At the time of this study, she was a teacher of Biology for 7 years in a high school in the Northwest region of the New England state in which we collected the data. From the student participants organized into 23 groups of 3 or 4 students each, 23 (initial-final) paired group models were collected. Additionally, we collected student journal reflections about their experience learning across the MBL unit after their pre- and post-modeling experiences. To do so, we used questions such as: *In what ways do you think developing the model may have helped your group develop a more accurate explanation of the water quality of the sites?*

4.4. Data analysis

To assess students' modeling outcomes in their pre- and post- group models and examine observable trends or patterns in students' models (research question two), we developed a rubric for assessing group models. Our water sustainability context-specific modeling rubric was based on the framework created by Schwarz et al. (2009) and in collaboration with our multidisciplinary project team. Schwarz et al.'s (2009) rubric as part of a learning progression project aimed at understanding models as tools for students to predict and explain phenomena or events that happen in the world. Within their rubric, four levels were described with two combined dimensions of metaknowledge (e.g. students'

understanding of the nature and purpose of models, and understanding of modeling performances that makes their activity meaningful and guides their practices in modeling) and elements of practice (e.g. constructing, using, evaluating, and revising models). Specifically, in Level 1, 'students see models as a means of showing others what the phenomenon looks like' (Schwarz et al. 2009, 640). They explained that students at this level use models simply to describe or depict the surface features of the phenomenon to others, rather than explain why or how it occurs. In Level 2, students use models to explain how a phenomenon occurs. To describe the process, they draw on scientific ideas in coordination with observational and experimental evidence. Level 2 involved phenomenon- or context-specific mechanistic scientific ideas and evidences useful for explaining how the phenomenon occurred, the unit's sustainability-specific science ideas and evidences were included within the three different explanatory sub-categories (i.e. observational or experimental evidence; authoritative evidence; and non-visible evidence to explain a phenomenon). Students in Level 3 not only describe the process of how the phenomenon occurs, but also generate predictions about how new phenomena or related phenomena might occur in relation to the phenomenon they have already learned. Finally, in Level 4 students are able to use models in combination with other models or to develop further questions to explore in efforts to strengthen the explanatory power of their models as tools for explaining events that happen in the world. Aligned with the structure of Schwarz et al.'s framework, our rubric was constructed with four levels; from Level 1 (lower) to Level 4 (higher) (see Table 1).

Within these four levels, we identified 28 possible distinguished items that could be included in students' models. Further, when considering and rating whether and how each item was used in the model, the following scale and ratings were used: 0 – not used, 1 – included but without any function, 2 – included with some explanation how it works, and 3 – included with some explanation how and why. Finally, the lowest level (i.e. Level 1) was assessed with only a Yes (1) or No (0) to indicate whether or not the model moved beyond a literal representation of the phenomenon. Because of our method of data collection (i.e. student models and journal entries), it was difficult to detect the extent to which students performed at Level 3 (i.e. generating predictions with models) and Level 4 (i.e. combining models to form new questions), so most of the comparisons in this research reveal only differences in Level 1 (i.e. showing what a phenomenon looks like) and Level 2 (i.e. explaining a phenomenon) of the framework developed by Schwarz et al. (2009). Even still, this rubric is important because a shift toward using models to explain phenomena (i.e. Level 2) is a relatively new learning performance for most students in U.S. classrooms (see Schwarz and White 2005; Windschitl, Thompson, and Braaten 2008).

The score of each model was calculated as the sum of scores for each of the 28 distinguished items in Table 1. As an example, the model of one group of students contained two 'water condition' relative concepts (i.e. item B-a-iv. in Level 2, Table 1) and was rated 2 (i.e. with an explanation of how the phenomenon occurs) and 3 (i.e. with an explanation of how and why the phenomenon occurs) for each. In this case, the cumulative score of this item on the model was 5. Beyond this rating strategy, a second score was recorded for each model that was comprised only of the number of distinguished items or concepts included in each model. Here, each distinguished item was measured only once.

Table 1. Pre- and Post- Model Rubric

Level	Students Performances	Counts	Scale*
4	A. Consider how the world could behave according to various models	A.	0 — 1 — 2 — 3
	B. Construct and use models to generate new questions about the behavior of phenomena	B.	0 — 1 — 2 — 3
	a. Develop questions that can be tested against evidence from the phenomena		
	C. Consider changes in models to enhance the explanatory power prior to obtaining evidence supporting these changes	C.	0 — 1 — 2 — 3
	D. Consider combining aspects of models that can enhance the explanatory and predictive power	D	0 — 1 — 2 — 3
3	A. Make prediction about phenomena using mechanism and process	A.	0 — 1 — 2 — 3
	B. Explain a cluster of related phenomena (multiple aspects of the phenomena)	B.	0 — 1 — 2 — 3
	C. Consider alternatives in constructing model	C.	0 — 1 — 2 — 3
	D. Evaluate the advantages and weaknesses of the model	D	0 — 1 — 2 — 3
2	A. Consider observational or experimental evidence to explain how/why a phenomenon occurs	A. a.	0 — 1 — 2 — 3
	a. Consider what they observed thus far in specific area	A. b.	0 — 1 — 2 — 3
	b. Include process or mechanism (with arrows or explanations)		
	c. Use own prior knowledge and relate this to phenomenon	A. c.	0 — 1 — 2 — 3
	d. Use the knowledge from peers, family, or others, and relate this to phenomenon	A. d.	0 — 1 — 2 — 3
	B. Consider authoritative evidence to explain how/why a phenomenon occurs	B. a. i.	0 — 1 — 2 — 3
	a. Gotta Have List	B. a. ii.	0 — 1 — 2 — 3
	i. Components Interaction in Water		
	ii. Connection Between Water and Earth System	B. a. iii.	0 — 1 — 2 — 3
	iii. Human Activity		
	iv. Water Condition	B. a. iv.	0 — 1 — 2 — 3
	v. Improvement and Sustainability	B. a. v.	0 — 1 — 2 — 3
	b. Evidence from teacher, textbook etc.	B. b. i.	0 — 1 — 2 — 3
	i. Agricultural areas use fertilizers and pesticides which can run off into local waters	B. b. ii.	0 — 1 — 2 — 3
	ii. Forested areas provide absorption and natural pollutant processing for rainfall and surface waters	B. b. iii.	0 — 1 — 2 — 3
	iii. People use forest and agricultural lands to build houses and commercial areas		
	C. Consider non-visible evidence to explain how/why a phenomenon occurs	C. a. i.	0 — 1 — 2 — 3
	a. Include non-visible structural component (e.g., microscopic particles)	C. a. ii.	0 — 1 — 2 — 3
	i. N, O, P, N_2, O_2, NH_3, NO_3^-, (pH)	C. b. i.	0 — 1 — 2 — 3
	ii. Water Vapor, Sunlight, Energy, Microorganism, Temperature	C. b. ii.	0 — 1 — 2 — 3
	b. Include non-visible process or mechanism	C. b. iii.	0 — 1 — 2 — 3
	i. Agricultural areas high N & P high algae	C. b. iv.	0 — 1 — 2 — 3
	ii. Forested areas low N & P low algae		
	iii. Evapotranspiration	C. b. v.	0 — 1 — 2 — 3
	iv. Sedimentary		
	v. Decomposition		
1	A. Students only show literal illustrations of the phenomena		
	B. No connections and explanations showing mechanisms or processes		Yes (1) / No (0)
	a. No explanatory constructs, just a picture of watershed		

Scale* 0=Not used, 1=Included but without any function, 2=Included with some explanation how it works, 3=Included with some explanation how and why.

To maintain consistency in evaluation of models, one researcher (i.e. first author) rated all models. However, to ensure reliability of the first author's ratings of the models, a second rater, an undergraduate science major, was trained with the rubric over a three-month period through the course of six meetings. After the training period, the second rater scored 13% (6 models) of the total 46 models included in this current research. Inter-Rater Reliability (IRR) was assessed using a consistency, two-way mixed, average-measures Intraclass Correlation Coefficient (ICC) (McGraw and Wong 1996). In the end, an ICC value of 0.829 was obtained in the excellent range (Cicchetti 1994). This indicated that raters had a high degree of agreement (Hallgren 2012). For research question one, to evaluate the extent to which participation in the learning exercise affected attributes of group modeling, we conducted a suite of multivariate and univariate statistical tests, all involving repeated measures.

For the validity of the rubric, we developed a 'Gotta Have List' based on the disciplinary core ideas from the NGSS (Next Generation Science Standards Lead States 2013) that we identified as necessary for both teachers and students to fully explain the phenomena. This list involved five key concepts and mechanisms that are 1) how the components interact with each other in water cycle, 2) what connections exist between the water cycle and Earth systems, 3) how human activities (e.g. land use, agriculture, water use, etc.) impact the water cycle, 4) how conditions of the watershed are expected to change over time, 5) how humans can sustain or improve water quality. The student performances evaluated with the rubric were judged in comparison to the target model-based explanation developed by the research group. Further, student performances were divided into four levels based on their level of sophistication outlined in accordance with Schwarz et al.'s (2009) rubric specifications discussed earlier.

To evaluate the overall changes in modeling, we performed a Repeated-Measures (pre- and post-scores) Multivariate Analysis of Variance (MAOVA) on ratings for each of the 28 performance attributes (Table 1). Similarly, we conducted a separate paired t-test (pre- and post-scores) for total score and for total number of concepts per model. To assess the extent to which overall changes in performance were associated with each of the Levels of Learning, we separately performed a Repeated-Measures (pre- and post-scores) MANOVA on ratings for each of the performance attributes within each Learning Level (i.e. separate analysis for four performance attributes of Level 4, four performance attributes of Level 3, and 19 performance attributes of Level 2). To control for inflated Experiment-Wise Error Rates, we performed Bonferroni Sequential adjustments before declaring significance for any particular MANOVA. Finally, to determine the extent to which each of the particular performance attributes of models changed in response to the learning exercise, we separately conducted a paired t-test (pre- and post-scores) for each of the 28 performance attributes. Again, to control for inflated Experiment-Wise Error Rates, we performed Bonferroni Sequential adjustments before declaring significance for any particular paired t-test.

To answer the second research question, we sought to identify more contextual evidence to highlight and better understand trends or patterns in student group models. Consequently, student journals were qualitatively analyzed using thematic analysis (Miles, Huberman, and Saldaña 2013). More specifically, student journal

entries were analyzed to code data and to identify themes. Here, any statement made by the student in their journal identified as capable of lending insight into their experiences in the MBL unit, or their modeling practice was identified as a unit of meaning, before units of meaning were grouped into themes to answer the research question. In the end, this resulted in student-reported descriptions of their experience being analyzed to help explain the patterns we identified in their models or in connection to the MBL unit more broadly.

5. Results

5.1. To what extent and in what ways does an MBL curricular unit support secondary students' model-based explanations of water and sustainability in local contexts across a unit?

According to the repeated measures MANOVA with two dependent variables (i.e. the means of the numbers of concepts and score), significant differences occurred between pre- and post-models (F(2, 21) = 47.22, p < 0.001). Subsequent univariate test indicated significant mean differences (F(1, 22) = 68.68, p < 0.001), between numbers of concepts for pre- (M = 12.1, SD = 3.83) and post-models (M = 26.7, SD = 8.32). Similarly, a univariate test indicated a significant mean difference (F(1, 22) = 98.53, p < 0.001) between pre- (M = 14.52, SD = 4.122) and post-scores (M = 40.87, SD = 12.571).

A repeated-measures MANOVA with the numbers of concepts for each level and sub-level of student performances indicated that the difference between the means for Level 4, Level 3, and Level 2 B-b, Level 2 C-b were not different between pre- and post-models. However, significant differences between pre- and post-models characterized Level 2 A (F (1, 45) = 18.62, p < 0.001), Level 2 B-a, (F(1, 45) = 8.48, p = 0.006), and Level 2 C-a (F(1, 45) = 35.57.62, p < 0.001). Because our data were limited to student models and journal entries, these statistical differences mainly reflect changes in Level 2 aspects of student models. Importantly, a paired t-test for the numbers of concepts for each item and model rubric provided insight into student performance in relation to Level 2 expectations (i.e. student's ability to explain phenomena with models).

Moreover, important nuances were revealed to help to better understand the three different clusters of performance items (see Table 2). First, students were arriving to the instruction with resources they were already drawing on, like the capability to depict what was observed in an area that thought were relevant to the water quality of an area before engaging in the watershed unit (i.e. L2-A-a). More specifically, student models containing information about the location of cars and trees, the direction of the road, the shape or color of things, and the kinds of the trash that they observed in the area. Even though these features in pre-models were relatively simple compared to those included in post-models, the ways that students developed pre- and post-, models included their previous observations that could begin to help explain water quality. This was evidenced in the following mean comparisons between pre- (M = 2.2, SD = 2.70) and post-(M = 2.3, SD = 2.92) models, and lack of significant changes (paired-t = 0.890, df = 22; p = 0.383) overtime with respect to item 2-A-a (i.e. consider what they observed thus far in a specific area). This suggested that students were already facile at depicting some relevant observations in an area before engaging in the watershed unit.

Second, students included resources or ideas for the first time in their post-models, like the following: 'include process or mechanism with arrows or explanations' (L2-A-b); 'improvement and sustainability' (L2-B-a- ⅴ); and 'decomposition' (L2-C-b- ⅴ). In these cases, even when some of these were not significantly increased (i.e. improvement and sustainability, decomposition), it was evident that they began to appear for the first time in post-models. In cases like these, we inferred that those items that were only found in the post-models were influenced by the unit implementation, especially related to influencing students' understanding of processes or mechanisms that could help explain the target phenomenon (i.e. water quality of the watershed), or of how arrows and explanations were important for describing the phenomenon. It is encouraging that statistical differences (paired-t = 5.609; df = 22; $p < 0.001$) were found related to the use of arrows and explanations in describing phenomenon: pre- (M = 0.0, SD = 0.00) and post- (M = 1.4, SD = 1.15) model outcomes. Consequently, we concluded that the focus on students' understanding of processes or mechanisms across the unit affected students' facility with constructing explanations of phenomena.

Third, some resources or ideas were included in models as a focus in the unit, but did not emerge as important features for explaining the phenomenon in students' post-models. In particular, the following are example items, among others, that were not incorporated in student models.: 'agricultural areas use fertilizers and pesticides that can run off into local waters' (L2-B-b- ⅰ) and 'forested areas provide absorption and natural pollutant processing for rainfall and surface waters' (L2-B-b- ⅱ). The relations between agriculture runoff and forested area land-use and conservation services were important parts of the unit that were experienced by the teachers as part of PD (i.e. 'learner hat' experiences) and were emphasized by the inclusion of particular tasks. The fact that these items were not included in post-models points to how these approaches or tools were not as effective in supporting learners to think about these aspects of watershed sustainability as we originally anticipated. Consequently, this pointed to the need to revisit these particular tasks within the unit, both related to the nature of the tasks as an area for future improvement.

Table 2. Comparison of model items between post and pre-models.

Items	MD	SD	t(22)	p-value
L2-A-a	0.174	0.937	0.890	0.383
L2-A-b	1.348	1.152	5.609	< 0.001***
L2-B-a- ⅰ	0.565	1.472	1.842	0.079
L2-B-a- ⅱ	1.478	2.998	2.365	0.027*
L2-B-a-ⅲ	2.304	3.007	3.676	0.001**
L2-B-a-ⅳ	2.304	1.845	5.991	< 0.001***
L2-B-a- ⅴ	0.087	0.417	1.000	0.328
L2-C-a- ⅰ	4.739	4.169	5.451	< 0.001***
L2-C-a- ⅱ	1.478	1.620	4.376	< 0.001***
L2-C-b- ⅴ	0.043	0.209	1.000	0.328

*0.01 < p < 0.05 ** 0.001 < p < 0.01 *** p < 0.001

5.2. What patterns were observed in students' models that can inform focused improvements to the MBL unit and student learning with models?

Consistent with what others have noted (e.g. Lesh and Lehrer 2003) we found that as students engaged in modeling, sometimes for the first time, to consider possible explanations of real-world phenomena, large variation characterized student understanding. This variation among student models, offered a chance to identify noteworthy patterns that could be used for further refinement of the curriculum, or to better understand how student learning with models unfold, especially in students' early attempts to model. When examining student models, three patterns emerged. First, the students tended to try to represent directly what they observed – most of the models developed by students appeared as if they were photographs (See Figures 2 and 3).

One explanation for this pattern may be connected to students' beliefs that they needed to draw everything that they observed in their models to explain every interaction between each component, rather than highlighting particular features within their models that may have been more important than others when considering watershed sustainability. Student journal entries like the following provided additional insight into these student beliefs: 'the most difficult was having to put all of our ideas based on the ponds health in one picture' or 'trying to add everything that could affect the pond's environment'. Consequently, because a majority of students believed that they needed to represent most of all things they observed, they repeatedly and unnecessarily drew the same objects (e.g. tree, car, and fish in Figures 2 and 3), instead of representing one object and focusing on mechanistic accounts or causal relationships that elevated the importance of representing any single object because of its importance in explaining the phenomenon. In the end, we believe this may have detracted from focusing on the explanatory purpose of engaging in modeling.

Second, students generally had difficulties pictorially representing complex patterns or mechanisms. More specifically, they added text descriptions to the model when expressing complex relationships (see Figure 4).

Within the model (Figure 4), one of the text descriptions is '[o]ur pond has average of 7–6 for pH supporting fish life. If it were to be above 9, the fish would have a difficult time excreting ammonia from their bodies'. While we recognize the accuracy of this textual claim is problematic, we highlight it to illustrate how, like other students, when it came to representing complex mechanistic explanations within models, students used text, instead of or without pairing their textual mechanistic explanations with pictorial representations. We believe that this pattern in student models may have arisen from, among other factors, students' difficulties visualizing those features of their models that were not visible to the naked eye or when students perceived that what they wanted to represent didn't have a physical form. This pattern became evident when students were asked what difficulties they had developing models when they reported in their journals, as they stated things like, 'knowing exactly what could be in the pond and figuring out how to put them in the model [was difficult]'.

Figure 2. Student model example 1.

Figure 3. Student model example 2.

Lastly, students had difficulties reflecting various perspectives in their models. Most of students' models had a top-down or birds-eye viewpoint. Specifically, among the 23 sets of pre-post models, 20 sets were based on a top-down viewpoint, whereas only 3 sets of pre-post models had a side viewpoint (See Figure 5).

Here, most models looked like a map (Figures 2, 3, and 4). Although there is nothing inherently wrong about students using the top-down viewpoint, recognizing that most models relied on one particular viewpoint (i.e. the top-down viewpoint), and not others (e.g. a side viewpoint) seemed noteworthy, since we believe that the collective class orientation to the phenomenon might have limited discussion of certain aspects of the phenomenon (e.g. puzzling aspects of the watershed that might only be considered when considering the side view of the watershed).

6. Discussion

6.1. To what extent and in what ways does an MBL curricular unit support secondary students' model-based explanations of water and sustainability in local contexts across a unit?

Scores for student models and the number of science concepts used in the models significantly increased at the end of the unit. This reflected the improved explanatory power of models, and the increased incorporation of science ideas or concepts as important mechanistic features of their models. Moreover, when students' performance at Level 2 (i.e. use models to explain how a phenomenon occurs) were considered, some important features of their explanations improved significantly. At the end of unit implementation, students were able to consider particular concepts that we identified

Figure 4. Student model example 3.

Figure 5. Student model example 4.

as important in constructing their own explanations about water quality and sustainability, especially as it related to local watersheds. In some cases, students incorporated important concepts only in post-models. This pointed to the effectiveness of MBL curriculum in supporting development of more complex and evidence-based organization of their understanding over time as a result of their modeling experiences. These results indicate, as other researchers have revealed, that student learning with models and modeling is improved as high-quality curriculum coupled with supportive teacher PD can be effective in supporting students to engage in sense-making about real-world events or natural phenomena (Acher, Arcà, and Sanmartí 2007; Baumfalk et al. 2019; Campbell and Fazio 2018; Forbes, Zangori, and Schwarz 2015; Schwarz and White 2005; Vo et al. 2015; Zangori et al. 2017).

Our results also revealed that some features (i.e. science ideas or concepts) of models that we identified as important were already present in initial models. The results also revealed how some features (i.e. Level 2-B: consider the use of agriculture areas or the interaction between forested areas and watershed; Level 2-C: non-visible process or mechanism) were not in pre- or post-group models, even though PD and curricular unit were intentionally designed to foreground and support students in learning taking up and using these features in their explanatory accounts.

The presence of important concepts or science ideas in pre-models demonstrates that students were already facile at depicting things that they could observe in an area that they thought might influence water quality before engaging in the watershed unit. We were able to understand this finding better as we considered similar examples discussed by prior researchers. More specifically, researchers previously identified how students can more easily identify and describe certain elements of phenomena that they can more easily observe or experience, when compared to other elements of phenomena (e.g. Baumfalk et al. 2019; Zangori et al. 2017). As an example, Baumfalk et al. (2019) pointed

out how it might be more important to focus student learning opportunities on those features of phenomena that are less-easily observable, especially in comparison to those features of phenomena that are more-easily observed and already more likely to be considered as students use models to construct explanations. Given this, as we iterate the MBL curricular unit, and ways we can support student learning with models, we can focus more on those less-easily observed features to support students in recognizing their importance in explaining and subsequently predicting watershed sustainability.

6.2. What patterns were observed in student models that can inform focused improvements to the MBL unit and student learning with models?

We identified a number of salient patterns in student models. In relation to the first pattern (i.e. students represented what they observed directly), Clement (2008) pointed out how this occurs when students focus on representing the easily observable features of a phenomenon rather than focusing on representing patterns or explaining underlying causal mechanisms that explain the real-world event or condition(s) (Cheng and Brown 2015; Schwarz et al. 2009). Constructing and evaluating scientific knowledge claims as part of students' model-based learning not only requires students to understand the 'what' (e.g. representation of a phenomenon), it requires that students attend to the 'how' and 'why' (e.g. patterns and causal mechanism) that underlie the phenomenon. However, in this first pattern, we found students using the same concepts repeatedly with a focus on describing everything, without focusing on the patterns or causal mechanisms of how and why the phenomenon happened. Given this, our findings pointed to early foci that students are likely to attend to in modeling, and make more explicit as part of our MBL unit a more explicit focus on the epistemic functional aspects of the explanatory features of modeling, while working to move away from representational aspects of modeling (Gouvea and Passmore 2017; Passmore, Gouvea, and Giere 2014).

Other researchers have revealed instances where students rely on written descriptions without drawing on pictorial representations to explain complex patterns or causal mechanisms in models. For example, Baumfalk et al. (2019) concluded that elementary students struggle to represent non-visible and complex ideas (e.g. water vapor or groundwater) in models about the hydrosphere, but did not evidence such problems when characterizing the representation of visible components (e.g. sun, rain, or lakes) at smaller or larger scales. Further, Baumfalk et al. (2019) noted how students used written descriptions and explanations for models to describe mechanistic features (e.g. sun's impact on the hydrological systems) because of the relative difficulty in representing these features pictorially. Given this, we concluded that more can and perhaps should be done in future iterations of our MBL curriculum afford students resources in continually enriching their pictorial–spatial representations to expand their use of semiotic modes in their models (Ramadas 2009). As an example, it might be possible to consider whether students could be offered pictorial representational choices related to the non-visible complex ideas that might add an important and needed explanatory feature to their models (e.g. water vapor or groundwater).

As well, students generally had difficulties reflecting various perspectives in their models. More specifically, Van Der Valk, Van Driel, and De Vos (2007) described accessibility as one criterion that can be used to think about models as epistemic tools for

developing explanations. Here, accessibility can be understood as those factors that enhance the ways that students orient to the phenomenon as they develop models. We believe that accessibility was influenced by the technical ways in which student groups chose to orient to the phenomena that involved the choice of direction (e.g. top-down view, side view), among other features. We do not believe that any one view is inherently wrong. However, we suspect that having limited ways to orient to a phenomenon might reduce the overall explanatory power of models. As an example, by considering only the top-down view of the watershed, more attention might be paid to runoff, as an example, whereas little to no attention might be paid to ground-water in considering the sustainability practices related to watersheds.

7. Conclusion, implication, and limitation

7.1. Conclusion and implication

Since the emergence of the *Framework* (National Research Council 2012) and the *Next-Generation Science Standards* (NGSS) (Next Generation Science Standards Lead States 2013), examples of effective models of NGSS implementation strategies have emerged (e.g. Campbell et al. 2019; Knight-Bardsley and McNeill 2016; Penuel, Harris, and DeBarger 2015). However, at the time of this research, a need still existed for more research that provided practical support for teachers in the form of NGSS-designed curriculum (Penuel and Reiser 2018; Reiser 2013). Consequently, we took advantage of the opportunity to conduct this research as part of an NGSS-designed unit implementation. The analysis of this curricular unit implementation of model-based learning provided initial and promising evidence of how modeling benefited student learning (i.e. statistical improvements in students modeling outcomes). Beyond this, our research uncovered insightful patterns in how students use models that can serve to support our own and others research and implementation of model-based learning curriculum units. What we found in current study provided meaningful implications for teacher PD. Specifically, researchers focused on teacher professional learning in science education have pointed to the need to focus on how interactions happen in classrooms (Reiser 2013), especially with a focus on supporting teachers with sufficient opportunities to translate the instructional strategies learned in PD into their classroom practice (Garet et al. 2001; Putnam and Borko 2000). Further, Reiser (2013) pointed to the importance of teachers experiencing the same type of sensemaking experiences that are planned for students. Relatedly, research by Knight-Bardsley and McNeill (2016) revealed how instruction changed to be more aligned with what was modelled during PD when teachers had more opportunities to try out and reflect on instructional resources from the PD as they moved into their classrooms. Connected to this, our research provided insight to promising positive outcomes of implementation of an MBL curricular unit, while also surfacing challenges that others can consider and we can use as we seek to refine our approaches to MBL unit design. In the end, we believe that our work represents a grain-sized focus on model-based learning that is seldom approached. More specifically, this research offers meaningful insights into the implementation of NGSS-designed materials in connection to student learning

outcomes and experiences using models. We expect to build on what we learned here in the future as we continue to engage in research and iterative refinement of our own MBL unit design.

7.2. Limitation

The qualitative nature of our analysis focused on a small sample of students' performance could be considered as a limitation of this study because of the challenges of generalizing what was learned in this research to broader populations. However, our smaller sample afforded us a more in-depth understanding than would have been possible with a larger sample, especially in relation to illuminating how NGSS-designed curriculum can be found supporting the implementation of the NGSS in connection to student learning and patterns in the ways that students engaged with models.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Byung-Yeol Park http://orcid.org/0000-0002-7474-9693
Todd Campbell http://orcid.org/0000-0001-6844-5303
Miriah Kelly http://orcid.org/0000-0002-9744-2897
Ron Gray http://orcid.org/0000-0002-0512-063X
Chester Arnold http://orcid.org/0000-0001-8533-5931
Laura Rodriguez http://orcid.org/0000-0002-7029-9701

References

Acher, A., M. Arcà, and N. Sanmartí. 2007. "Modeling as A Teaching Learning Process for Understanding Materials: A Case Study in Primary Education." *Science Education* 91 (3): 398–418. doi:10.1002/sce.20196.

Achieve Inc. 2017. *Primary Evaluation of Essential Criteria (PEEC) for Next Generation Science Standards Instructional Materials Design*. Washington, DC: Achieve.

Achieve Inc. and National Science Teachers Association (NSTA). 2014. "EQuIP Rubric for Lessons and Units: Science." https://www.nextgenscience.org/resources/equip-rubric-lessons-units-science

Ambitious Science Teaching. 2015. "Models and Modeling: An Introduction." http://ambitiousscienceteaching.org/wp-content/uploads/2014/09/Models-and-Modeling-An-Introduction1.pdf

Banilower, E. R. 2019. "Understanding the Big Picture for Science Teacher Education: The 2018 NSSME." *Journal of Science Teacher Education* 30 (3): 201–208. doi:10.1080/1046560X.2019.1591920.

Baumfalk, B., D. Bhattacharya, T. Vo, C. Forbes, L. Zangori, and C. Schwarz. 2019. "Impact of Model-based Science Curriculum and Instruction on Elementary Students' Explanations for the Hydrosphere." *Journal of Research in Science Teaching* 56 (5): 570–597. doi:10.1002/tea.21514.

Böttcher, F., and A. Meisert. 2011. "Argumentation in Science Education: A Model-Based Framework." *Science & Education* 20 (2): 103–140. doi:10.1007/s11191-010-9304-5.

Bybee, R., and C. Chopyak. 2017. *Instructional Materials and Implementation of Next Generation Science Standards: Demand, Supply, and Strategic Opportunities*. New York, NY: Carnegie Corporation.

Campbell, T., and X. Fazio. 2018. Epistemic frames as an analytical framework for understanding the representation of scientific activity in a modeling-based learning unit. Research in Science Education. https://doi.org/10.1007/s11165-018-9779-7

Campbell, T., T. J. McKenna, X. Fazio, A. Hetherington-Coy, and P. Pierce. 2019. Negotiating Coherent Science Teacher Professional Learning Experiences Across a University and Partner School Settings. *Journal of Science Teacher Education* 30(2): 179–199.

Cheng, M. F., and D. E. Brown. 2015. "The Role of Scientific Modeling Criteria in Advancing Students' Explanatory Ideas of Magnetism." *Journal of Research in Science Teaching* 52 (8): 1053–1081. doi:10.1002/tea.21234.

Cicchetti, D. V. 1994. "Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology." *Psychological Assessment* 6 (4): 284–290. doi:10.1037/1040-3590.6.4.284.

Clement, J. 2000. "Model Based Learning as a Key Research Area for Science Education." *International Journal of Science Education* 22 (9): 1041–1053. doi:10.1080/095006900416901.

Clement, J. 2008. *Creative Model Construction in Scientists and Students: The Role of Analogy, Imagery, and Mental Simulation*. Amherst, MA: Springer.

Creswell, J. W. 2012. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Boston, MA: Pearson Education.

Creswell, J. W., and V. L. Plano Clark. 2011. *Designing and Conducting Mixed Methods Research*. LA: Sage.

Dickerson, D. L., J. E. Penick, K. R. Dawkins, and M. Van Sickle. 2007. "Groundwater in Science Education." *Journal of Science Teacher Education* 18 (1): 45–61. doi:10.1007/s10972-006-9019-2.

Etikan, I., S. Musa, and R. Alkassim. 2016. "Comparison of Convenience Sampling and Purposive Sampling." *American Journal of Theoretical and Applied Statistics* 5 (1): 1–4. doi:10.11648/j.ajtas.20160501.11.

Forbes, C. T., L. Zangori, and C. V. Schwarz. 2015. "Empirical Validation of Integrated Learning Performances for Hydrologic Phenomena: 3rd-Grade Students' Model-Driven Explanation-Construction." *Journal of Research in Science Teaching* 52 (7): 895–921. doi:10.1002/tea.21226.

Ford, M. 2008. "'Grasp of Practice' as a Reasoning Resource for Inquiry and Nature of Science Understanding". *Science & Education* 17 (2–3): 147–177. doi:10.1007/s11191-006-9045-7.

Ford, M. 2015. "Educational Implications of Choosing "Practice" to Describe Science in the Next Generation Science Standards." *Science Education* 99 (6): 1041–1048. doi:10.1002/sce.21188.

Freidenfelds, N., L. Cisneros, L. Rodriguez, B-Y. Park, T. Campbell, C. Arnold, C. Chadwick, D. Dickson, M. Dietz, D. Moss, J. Volin, and M. Willig. 2020. Investigating Human Impact on Local Water Resources and Exploring Solutions. *The American Biology Teacher* 82(9): 619–623.

Garet, M. S., A. C. Porter, L. Desimone, B. F. Birman, and K. S. Yoon. 2001. "What Makes Professional Development Effective? Results from a National Sample of Teachers." *American Educational Research Journal* 38: 915–945. doi:10.3102/00028312038004915.

Gilbert, J. K., and C. J. Boulter. 2000. *Developing Models in Science Education*. New York: Springer.

Gouvea, J., and C. Passmore. 2017. "'Models Of' versus 'Models For': Toward an Agent-Based Conception of Modeling in the Science Classroom." *Science & Education* 26 (1–2): 49–63. doi:10.1007/s11191-017-9884-4.

Guy-Gaytán, C., J. S. Gouvea, C. Griesemer, and C. Passmore. 2019. "Tensions between Learning Models and Engaging in Modeling." *Science & Education* 28: 843–864. doi:10.1007/s11191-019-00064-y.

Hallgren, K. A. 2012. "Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial." *Tutorials in Quantitative Methods for Psychology* 8 (1): 23–34. doi:10.20982/tqmp.08.1.p023.

Kenyon, L., E. A. Davis, and B. Hug. 2011. "Design Approaches to Support Preservice Teachers in Scientific Modeling." *Journal of Science Teacher Education* 22 (1): 1–21. doi:10.1007/s10972-010-9225-9.

Khan, S. 2007. "Model-Based Inquiries in Chemistry." *Science Education* 91 (6): 877–905. doi:10.1002/sce.20226.

Knight-Bardsley, A. M., and K. L. McNeill. 2016. "Teachers' Pedagogical Design Capacity for Scientific Argumentation." *Science Education* 100 (4): 645–672. doi:10.1002/sce.21222.

Ko, M. L. M., and C. Krist. 2019. "Opening up Curricula to Redistribute Epistemic Agency: A Framework for Supporting Science Teaching." *Science Education* 1–32. doi:10.1002/sce.21511.

Krajcik, J. 2015. "Three-Dimensional Instruction: Using a New Type of Teaching in the Science Classroom." *Science and Children* 53 (3): 6–8. doi:10.2505/4/sc15_053_03_6.

Lesh, R., and R. Lehrer. 2003. "Models and Modeling Perspectives on the Development of Students and Teachers." *Mathematical Thinking and Learning* 5 (2–3): 109–129. doi:10.1080/10986065.2003.9679996.

McGraw, K. O., and S. P. Wong. 1996. "Forming Inferences about Some Intraclass Correlation Coefficients." *Psychological Methods* 1 (1): 30–46. doi:10.1037/1082-989X.1.1.30.

Miles, M. B., A. M. Huberman, and J. Saldaña. 2013. *Qualitative Data Analysis: A Methods Sourcebook*. Thousand Oaks, CA: Sage.

Model Based Biology. 2017. "Model Based Biology: Students Making Sense of Biological Phenomenon." https://www.modelbasedbiology.com

National Academies of Sciences, Engineering, and Medicine (NASEM). 2018. *Design, Selection, and Implementation of Instructional Materials for the Next Generation Science Standards (NGSS): Proceedings of a Workshop*. Washington, DC: National Academies Press. doi:10.17226/25001.

National Research Council. 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: National Academies Press.

Next Generation Science Standards Lead States. 2013. *Next Generation Science Standards: For States, by States*. Washington, DC: National Academies Press.

Oh, P. S., and S. J. Oh. 2011. "What Teachers of Science Need to Know about Models: An Overview." *International Journal of Science Education* 33 (8): 1109–1130. doi:10.1080/09500693.2010.502191.

Osborne, J., S. Rafanelli, and P. Kind. 2018. "Toward a More Coherent Model for Science Education than the Crosscutting Concepts of the Next Generation Science Standards: The Affordances of Styles of Reasoning." *Journal of Research in Science Teaching* 55 (7): 962–981. doi:10.1002/tea.21460.

Passmore, C., J. Stewart, and J. Cartier. 2009. "Model-Based Inquiry and School Science: Creating Connections." *School Science and Mathematics* 109 (7): 394–402. doi:10.1111/j.1949-8594.2009.tb17870.x.

Passmore, C. M., J. S. Gouvea, and R. Giere. 2014. "Models in Science and in Learning Science: Focusing Scientific Practice on Sense-Making." In *International Handbook of Research in History, Philosophy and Science Teaching*, edited by M. R. Matthews, 1171–1202. Dordrecht, Netherlands: Springer.

Passmore, C. M., and J. Svoboda. 2012. "Exploring Opportunities for Argumentation in Modelling Classrooms." *International Journal of Science Education* 34 (10): 1535–1554. doi:10.1080/09500693.2011.577842.

Patton, M. Q. 2002. *Qualitative Research & Evaluation Methods*. 3rd ed. Saint Paul, MN: Sage.

Penuel, W. R., and B. J. Reiser. 2018. "Designing NGSS-Aligned Curriculum Materials." *Committee to Revise America's Lab Report*. Washington, DC: National Academies of Science, Engineering, and Medicine.

Penuel, W. R., C. J. Harris, and A. H. DeBarger. 2015. "Implementing the Next Generation Science Standards." *Phi Delta Kappan* 96 (6): 45–49. doi:10.1177/0031721715575299.

Putnam, R., and H. Borko. 2000. "What Do New Views of Knowledge and Thinking Have to Say about Research on Teacher Learning?" *Educational Researcher* 29 (1): 4–15. doi:10.3102/0013189X029001004.

Ramadas, J. 2009. "Visual and Spatial Modes in Science Learning." *International Journal of Science Education* 31 (3): 301–318. doi:10.1080/09500690802595763.

Reiser, B. J. 2013. "What Professional Development Strategies are Needed for Successful Implementation of the Next Generation Science Standards." In Paper written for the Invitational Research Symposium on Science Assessment, Washington DC.

Russ, R. S., and L. K. Berland. 2019. "Invented Science: A Framework for Discussing A Persistent Problem of Practice." *Journal of the Learning Sciences* 28 (3): 279–301. doi:10.1080/10508406.2018.1517354.

Schwarz, C. V., B. J. Reiser, E. A. Davis, L. Kenyon, A. Achér, D. Fortus, Y. Shwartz, B. Hug, and J. Krajcik. 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. doi:10.1002/tea.20311.

Schwarz, C. V., and B. Y. White. 2005. "Metamodeling Knowledge: Developing Students' Understanding of Scientific Modeling." *Cognition and Instruction* 23 (2): 165–205. doi:10.1207/s1532690xci2302.

Schwarz, C. V., and Y. N. Gwekwerere. 2007. "Using a Guided Inquiry and Modeling Instructional Framework (EIMA) to Support Preservice K-8 Science Teaching." *Science Education* 91 (1): 158–186. doi:10.1002/sce.20177.

Smith, S. 2020. "What Does a National Survey Tell Us about Progress toward the Vision of the NGSS?" *Journal of Science Teacher Education* 31 (6): 601–609. doi:10.1080/1046560X.2020.1786261.

Stroupe, D., and M. Windschitl. 2015. "Supporting Ambitious Instruction by Beginning Teachers with Specialized Tools and Practices." In *Newly Hired Teachers of Science: A Better Beginning*, edited by J. Luft and S. Dubois, 181–196. The Netherlands: Sense Publishers.

Svoboda, J., and C. M. Passmore. 2011. "Models as Epistemic Anchors: A Framework for Model-Based Instruction." Annual Meeting of the American Educational Research Association, New Orleans, LA.

Thompson, J., S. Hagenah, H. Kang, D. Stroupe, M. Braaten, C. Colley, and M. Windschitl. 2016. "Rigor and Responsiveness in Classroom Activity." *Teachers College Record* 118 (5): 1–58.

United Nations Food and Agriculture Organization. 2015. "Global Forest Resources Assessment 2015." Rome. http://www.fao.org/forest-resources-assessment/en

United Nations Food and Agriculture Organization. 2016. "State of the World's Forests 2016. Forests and Agriculture: Land-use Challenges and Opportunities." Rome. http://www.fao.org/publications/sofo/en

United Nations World Water Assessment Programme. (2014). *The United Nations World Water Development Report 2014: Water and Energy*. Paris, UNES.

United Nations World Water Assessment Programme. 2015. *The United Nations World Water Development Report 2015: Water for a Sustainable World*. Paris: UNES.

Van Der Valk, T., J. H. Van Driel, and W. De Vos. 2007. "Common Characteristics of Models in Present-day Scientific Practice." *Research in Science Education* 37 (4): 469–488. doi:10.1007/s11165-006-9036-3.

Vo, T., C. T. Forbes, L. Zangori, and C. V. Schwarz. 2015. "Fostering Third-Grade Students' Use of Scientific Models with the Water Cycle: Elementary Teachers' Conceptions and Practices." *International Journal of Science Education* 37 (15): 2411–2432. doi:10.1080/09500693.2015.1080880.

Wilson, S. M. 2013. "Professional Development for Science Teachers." *Science* 340 (6130): 310–313. doi:10.1126/science.1230725.

Windschitl, M., J. Thompson, and M. Braaten. 2008. "Beyond the Scientific Method: Model-Based Inquiry as a New Paradigm of Preference for School Science Investigations." *Science Education* 92 (5): 941–967. doi:10.1002/sce.20259.

Zangori, L., and C. T. Forbes. 2016. "Development of an Empirically Based Learning Performances Framework for Third-Grade Students' Model-Based Explanations about Plant Processes." *Science Education* 100 (6): 961–982. doi:10.1002/sce.21238.

Zangori, L., T. Vo, C. T. Forbes, and C. Schwarz. 2017. "Supporting 3rd-Grade Students' Model-Based Explanations about Groundwater: A Quasi-Experimental Study of A Curricular Intervention". *International Journal of Science Education* 39 (11): 1421–1442. doi:10.1080/09500693.2017.1336683.