Computers & Education 222 (2024) 105149

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Using multiple, dynamically linked representations to develop representational competency and conceptual understanding of the earthquake cycle

Christopher Lore [*], Hee-Sun Lee, Amy Pallant, Jie Chao

The Concord Consortium, United States

ARTICLE INFO

Keywords:
Teaching/learning strategies
Simulations
Pedagogical issues
Applications in subject areas

ABSTRACT

Using computational methods to produce and interpret multiple scientific representations is now a common practice in many science disciplines. Research has shown students have difficulty in moving across, connecting, and sensemaking from multiple representations. There is a need to develop task-specific representational competencies for students to reason and conduct scientific investigations using multiple representations. In this study, we focus on three representational competencies: 1) linking between representations, 2) disciplinary sensemaking from multiple representations, and 3) conceptualizing domain-relevant content derived from multiple representations. We developed a block code-based computational modeling environment with three different representations and embedded it within an online activity for students to carry out investigations around the earthquake cycle. The three representations include a procedural representation of block codes, a geometric representation of land deformation build-up, and a graphical representation of deformation build-up over time. We examined the extent of students' representational competencies and which competencies are most correlated with students' future performance in a computationally supported geoscience investigation. Results indicate that a majority of the 431 students showed at least some form of representational competence. However, a relatively small number of students showed sophisticated levels of linking, sensemaking, and conceptualizing from the representations. Five of seven representational competencies, the most prominent being code sensemaking ($\eta^2 = 0.053$, $p < 0.001$), were significantly correlated to student performance on a summative geoscience investigation.

1. Introduction

Computational methods have been increasingly incorporated into modern science practice (Henderson et al., 2007, pp. 195–196), changing the ways science research is carried out. There is a critical need to design student-appropriate learning experiences that leverage computational tools in ways that capture the essence of scientists' authentic practices (NGSS Lead States, 2013). The integration of science and computation is important because it allows educators to exploit the substantial overlap between computational practices and important skills in other subject areas (Waterman et al., 2020). To this end, many educational researchers have operationalized computational thinking for K-12 students (Grover & Pea, 2013; Wang et al., 2022; Weintrop et al., 2016). However, there

* Corresponding author. 25 Love Lane, Concord, MA, 01742, United States.
E-mail address: clore@concord.org (C. Lore).

https://doi.org/10.1016/j.compedu.2024.105149
Received 31 May 2024; Received in revised form 21 August 2024; Accepted 23 August 2024
Available online 28 August 2024

remains a need to identify effective methods for engaging students in scientific ways of investigating and reasoning that rely on computational methodologies while learning domain-specific content.

One area of research that has shown promise in the design of computationally integrated science activities is related to the study of multiple representations (Schnotz, 2014, pp. 72–103; Seufert, 2003). Several forms of representations (e.g., graphs, physical models, computational models, mathematical models, etc.) have been utilized for students to develop, clarify, and communicate knowledge about scientific phenomena (Evagorou, Erduran, & Mäntylä, 2015; Pauwels, 2006). However, there is considerable evidence to show that learners often fail to exploit the advantages of using multiple representations in science learning. Challenges include understanding the form of a representation, understanding the relationship between the representation and the domain, and understanding how to relate representations (Ainsworth, 2008). To help students effectively use multiple representations for science learning, researchers have begun to focus on building students' representational competencies, or the knowledge and skills that enable students to reason and solve domain-relevant tasks. Representational competencies have been shown to simultaneously support student learning of representations and learning of domain knowledge (Ainsworth, 2006; Council, 2006; de Jong & van Joolingen, 1998; Gilbert, 2005; McElhaney et al., 2015). By investigating the design of instructional activities, Rau (2017) found that effective learning activities support students in acquiring representational competencies and developing scientific understanding from the representations.

The field of geoscience, which focuses on understanding Earth's systems and processes, is increasingly leveraging computational techniques to research the planet's complex systems (Medunić et al., 2022). In geoscience, where large data sets and difficult-to-comprehend spatial and temporal scales pose challenges to student learning (Cheek et al., 2017), visualizations are often used to investigate Earth phenomena. For example, many seismic and geodesy organizations leverage networked observatories, such as GPS stations, that automatically collect large, high-quality seismic data around the globe to provide insight into Earth systems (Karpatne et al., 2019). If students are to engage in modern, authentic geoscience methods, they must be able to interpret multiple visualizations produced from models and data (Lore, Lee, & Pallant, 2024). Additionally, Earth science curricula in the US is still commonly organized into discrete bits of content, often with limited integrative connections (Whitmeyer et al., 2007). Geoscience is therefore an ideal context for developing students' representational competencies while simultaneously learning domain content through the representations.

In this study, we use multiple dynamically linked representations (MDLRs) to engage students in investigations about the earthquake cycle based on GPS data of land motion. The three representations used in the investigations are a *procedural representation* of block codes, a *geometric representation* of land deformation build-up, and a *graphical representation* of deformation build-up over time. We focus on three representational competencies necessary to carry out earthquake cycle investigations: *linking*, *sensemaking*, and *conceptualization*. Students' investigations are organized in competency-building steps in an activity sequence designed to engage students in computationally mediated investigations. The goal of the activity is to deepen students' understanding of the factors influencing the frequency and magnitude of earthquakes. Two research questions were addressed in this study.

(1) To what extent do students link, make sense of, and conceptualize procedural, geometric, and graphical representations in the context of the earthquake cycle?
(2) How does the extent of students' representation linking, sensemaking, and conceptualization in the activity correlate with their performances in a computationally mediated geoscience investigation?

2. Background

Visual representations have become increasingly important in the practice of science. Scientists use visual representations to reproduce or model nature, as well as to solve problems, fill gaps in knowledge, and facilitate the transfer of knowledge (Lynch, 2006). Visual representations include symbols, objects, maps, simulations, and graphs. They can be used as primary data (e.g., results from a simulation) to help in concept development, to find patterns and relationships, and to give form to abstract concepts (Evagorou, Erduran, & Mäntylä, 2015). Therefore, visual representations and the computational practices involved in creating and interpreting them play a critical role in developing, clarifying, and transmitting scientific knowledge (Evagorou, Erduran, & Mäntylä, 2015; Pauwels, 2006).

Visual representations are now commonly used in the teaching and learning of science. Different types of representations are often combined as multiple representations, which can support the development of higher-quality conceptual models (Sell et al., 2006). Learning can be facilitated by working with multiple visual representations because they can: 1) complement each other in terms of the information they convey, 2) constrain the interpretation of each other, and 3) aid in constructing a deeper understanding of the domain under study (Ainsworth, 2006). Several studies have shown that using multiple representations that are conceptually linked to one another can deepen students' learning of discipline-specific knowledge (Ainsworth et al., 2002; Rau et al., 2017). As research on multiple visual representations has grown, the focus has shifted from whether learning with multiple representations is effective to what factors influence their effectiveness in teaching and learning (Ainsworth, 2006).

Several challenges exist in designing competency-building activities with multiple representations. Students first encountering a new representation often show difficulty moving across or connecting multiple representations (Rau et al., 2015; van der Meij & de Jong, 2006). At first, there is a high cognitive load associated with understanding how the representations represent the phenomena under study, how to translate between representations, and how representations are related to each other (Ainsworth, 2008; Spiro & Jehng, 1990). Students also need assistance mapping between visual features and concepts (Ainsworth et al., 2002; Rau et al., 2014). In particular, students with low prior knowledge often require support in moving beyond a focus on superficial surface features towards more conceptually relevant visual information (Bodemer & Faust, 2006; Stern et al., 2003).

To support students in building representational competency, instructional activities can be designed to make explanation-based processes explicit. Previous research has produced several principles for the instructional design of activities that support students' development of representational competencies. These include having students: 1) verbally explain the perceptual features of each representation and the complementary information between representations, 2) actively compare perceptual features across representations, and 3) receive scaffolded support in identifying relevant perceptual features (Rau et al., 2017). Our goal in this study was to use these design principles to engage students in an activity sequence that scaffolds students in developing representational competency through the use of MDLRs in computationally mediated scientific investigations in a geoscience context.

3. Theoretical framework

Several types of representational competencies are identified in the literature, each of which describes possible ways in which a learner may use a representation to access and reason about scientific concepts (Pande & Chandrasekharan, 2017). Each competency provides a unique way for classifying and analyzing students' uses of and reasoning about representations along with reasoning difficulties (Schönborn & Anderson, 2009). In this study, we engaged students in using MDLRs to investigate the phenomena of the earthquake cycle. We investigated students' representational competencies and their ability to reason as a result of using an activity sequence that scaffolds representational linking, sensemaking, and conceptualization.

Representational linking describes the ability to interpret and translate between different types of scientific representations. Students need to be able to establish relations between corresponding representations and identify overlapping information shown between them. In this study, the representations are dynamically linked, meaning learners act on one representation and see the results of those actions in another (Ainsworth, 2008). Dynamic linking supports students in making surface-level connections between the representations; however, the student must still connect the domain-relevant conceptual aspects of the representations.

Representational sensemaking involves learning how to identify domain-relevant concepts in perceptual features and how those features are connected across multiple visual representations (Ainsworth, 2006; DeLoache, 2000; Eilam and Ben-Peretz, 2012; Schnotz, 2014, pp. 72–103). Through sensemaking, students learn to sort out conceptual features from surface features by explaining their mappings to concepts (Ainsworth, 2006; Rau et al., 2014; Seufert, 2003) related to the underlying domain being represented.

Representational conceptualization is the ability to explain domain-relevant concepts using scientific representations (Rau, 2017). This competency involves determining which representations can be used to explain abstract scientific concepts. While sensemaking involves relating the representations to each other and the domain, conceptualization involves understanding how to best solve a specific conceptual task using or creating the appropriate type of representation (Edelsbrunner et al., 2023).

Implementing instruction that includes MDLRs requires careful consideration of the design of the representations and the scaffolded activity sequence that prompts students to make sense of the representations as they relate to each other and to the domain. Neither generalized representational competencies nor computational practices can be built through just one question or one activity.

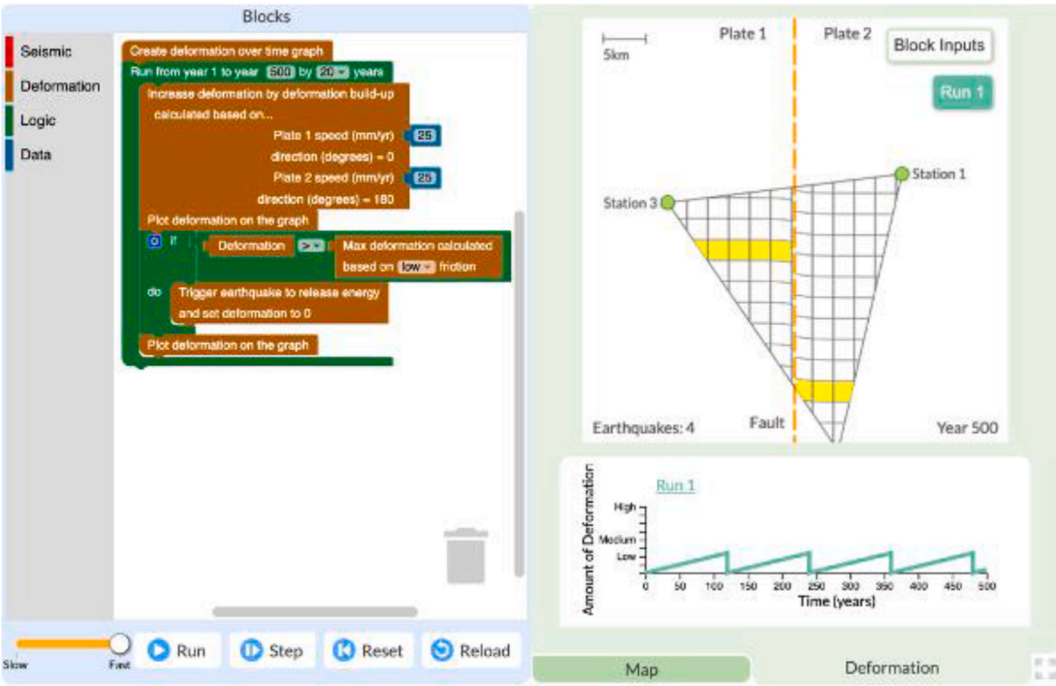

Fig. 1. The GeoCoder model shows block code (procedural representation, left), the Deformation Simulation (geometric representation, upper right), and the deformation over time graph (graphical representation, bottom right).

The goal of developing an activity sequence for building representational competency targets an important educational problem and addresses the lack of resources for authentically integrating computational investigations with geoscience content.

4. Materials and methods

4.1. The curriculum

This study was situated within an online curriculum called "Assessing Seismic Hazards and Risk with Code," the seismic module hereafter, developed for middle and high school students to carry out computationally supported scientific investigations using MDLRs. The earthquake cycle, the focus of the activity, is the phenomenon in which earthquakes repeatedly occur on the same fault as the result of continual accumulation of deformation and periodic deformation release. The seismic module is delivered to students as web-based interactive pages that include background information, instructions, the GeoCoder model (described below), and open-response and multiple-choice questions. The activity described in this study consisted of eight pages. In this paper, we focus on four pages that most directly relate to students exhibiting representational competency. The activity scaffolds student use of and interpretation of three dynamically linked representations within the GeoCoder model: block code, a geometric simulation called the Deformation Simulation (DS), and a graph. Each page is intended to engage students in investigations and questions that prompt them to relate the representations to each other and to earthquake cycle content.

4.2. The GeoCoder

To engage students in computer programming related to an investigation of factors related to the earthquake cycle, we developed the GeoCoder, a block-based programming workspace combined with a simulation of ground deformation and a graph depicting the amount of deformation (Fig. 1).

4.2.1. Procedural representation: Block code

The programming workspace is similar to other block-based programming environments such as Scratch (Resnick et al., 2009) and Google's Blockly software, from which the GeoCoder blocks were derived and adapted, where students build code from a selection of pre-made blocks. GeoCoder blocks were designed to produce unique visual outputs as the code is run. Blocks can result in the execution of the DS or data plotted in the graph. Using block code allows an easy-to-understand syntax that is not available with other coding languages, such as JavaScript. In addition, we contextualize the wording of the blocks to provide more information to students about the process of the earthquake cycle.

We refer to the block code as a "procedural representation" because the order and syntax of the blocks describe the actions and events that take place as the model runs. By combining the syntax of the blocks and the order in which they are executed, the code describes which events occur in the other representations. Each block is highlighted as it is executed, making it easier for students to make connections between how the execution of the code affects the outputs in the other representations.

4.2.2. Geometric representation: Deformation Simulation

The Deformation Simulation, shown on the right side of the GeoCoder (Fig. 1), represents the deformation of land along a fault due to the differential movement of two tectonic plates. Several features of the DS have been designed to contextualize the visualization around earthquakes. The green dots represent GPS stations (used by scientists to measure plate movement) and the orange line represents a fault. The grid lines help students see how land deforms as tectonic plates move. When the code is run, the two plates in the DS slide past each other based on the speeds set in the code. The movement of the plates causes the land to deform as represented by the grid lines that bend and stretch along the fault. The deformation is a product of the two sides of the fault being locked together due to the frictional forces between them while the rock that makes up the land bends and stretches. Once the deformation becomes too great and reaches a breaking point, an earthquake occurs. To represent this, the grid lines snap and become straight again with the lines offset from their original position.

The background equation governing the functioning of the DS is based on the Savage and Burford (1973) model of deformation along a linear fault experiencing perfect strike-slip motion. The model contains several parameters associated with deformation along a fault (time, locking depth, distance from the fault, and horizontal displacement). To simplify the equation for the GeoCoder, many parameters were set to constant values. We chose two parameters for students to vary–speed of differential plate movement and friction level between the plates–because of their importance to the earthquake cycle.

We refer to the DS as a "geometric representation" because it is a mathematical calculation shown as movement in a grid. The goal of this representation is to give students a mental model of land deformation. The DS is a bird's eye view of land around a fault. As one plate moves north and the other moves south, students can visualize how deformation builds up over time and eventually releases as an earthquake.

4.2.3. Graphical representation: deformation over time graph

The third GeoCoder representation is the graph of deformation over time shown on the bottom right side of the GeoCoder. The x-axis is time and the y-axis is the amount of deformation with three levels marked: low, medium, and high. As the model runs, the graph automatically updates based on the amount of deformation in the DS. As the model runs, the line in the graph increases until the deformation exceeds the maximum deformation level set in the code. When an earthquake occurs, the line drops to zero, indicating the complete dissipation of deformation. The slope of the line in the graph indicates the speed of deformation build-up on the land surface, and the level it reaches before an earthquake indicates the friction level that students set in the code. Depending on the plate speeds and friction level students define in the code, multiple earthquakes are possible throughout a single run, and the graph will have a saw-toothed pattern.

The graph serves as an abstract representation of deformation build-up over time. Its value comes from recording the differential speed, level of maximum deformation, and the number of earthquakes that occur throughout a simulation run.

These three representations are dynamically linked because actions performed on one representation (the block code) are automatically shown in both other representations (the DS and graph) (van der Meij & de Jong, 2006). The blocks were specifically designed so that each block, when highlighted during a run, produces a visual output in the DS and the graph. The three representations together show overlapping information, such as the build-up of deformation over time in the DS and graph. Yet each representation also shows unique information, such as the block code syntax indicating why earthquakes repeatedly occur. The additive value of all three dynamically linked together gives a more rounded and robust model of the earthquake cycle than any of the representations individually.

4.3. The activity sequence

The four focal pages of this activity scaffold two parallel learning goals: 1) building representational competency and 2) learning about the earthquake cycle. The pages interweave questions pertaining to these learning goals in order to develop students' abilities to conduct and explain investigations around the earthquake cycle. For example, in order to learn about the controlling factors of the earthquake cycle, the curriculum slowly walks students through the control of variables method over the course of three pages to investigate the independent effects of plate speed and friction level on the earthquake cycle. While doing this, students are prompted to link the outputs of multiple representations to describe the effects of the variables on the occurrence of earthquakes. The activity starts simply, with only a few blocks available to build code and only the DS visible. As the activity progresses, the deformation over time graph is made visible, and additional blocks are used, which produce additional features in the DS and graph. By going back and forth between questions regarding representational competencies and the earthquake cycle, we aim to slowly build students' abilities to investigate and make sense of the earthquake cycle phenomena. The four pages of interest are described in more detail below (Fig. 2).

On Page 1, the curriculum prompts students to identify links between the block code and the DS. Using one function block and two number blocks, students input data of specific GPS sites from around California into the program and observe the grid lines of the DS bend and stretch based on those data. On this page earthquakes are disabled and the grid lines deform for the entirety of the run and never break. Students are prompted to run multiple iterations with plate speeds from different GPS stations to explore the relationship between differential plate speed and the amount of land deformation. As this is the first time students use block code to run the DS, the curriculum focuses on linking the procedural code blocks and the geometric DS representation by asking the following question (Q7): "What happens every time the 'Increase deformation … ' block is executed?" In answering this question, students exhibit their ability to link the function of the block code to the output visualization in the DS.

On the second page, students are prompted to link deformation shown in the DS to the deformation over time graph. To do this, students are instructed to add two function blocks to the code they previously built that will: 1) create a graph of deformation over time and 2) plot data from the DS on that graph. In this case, as the code is executed, it shows not only the progressive bending and stretching of the grid lines in the DS as on page 1, but also the line on the graph linearly increasing over time, reinforcing the idea that deformation is increasing. To help students link the visualizations of these representations, Q12 asks, "In this snapshot, which part of

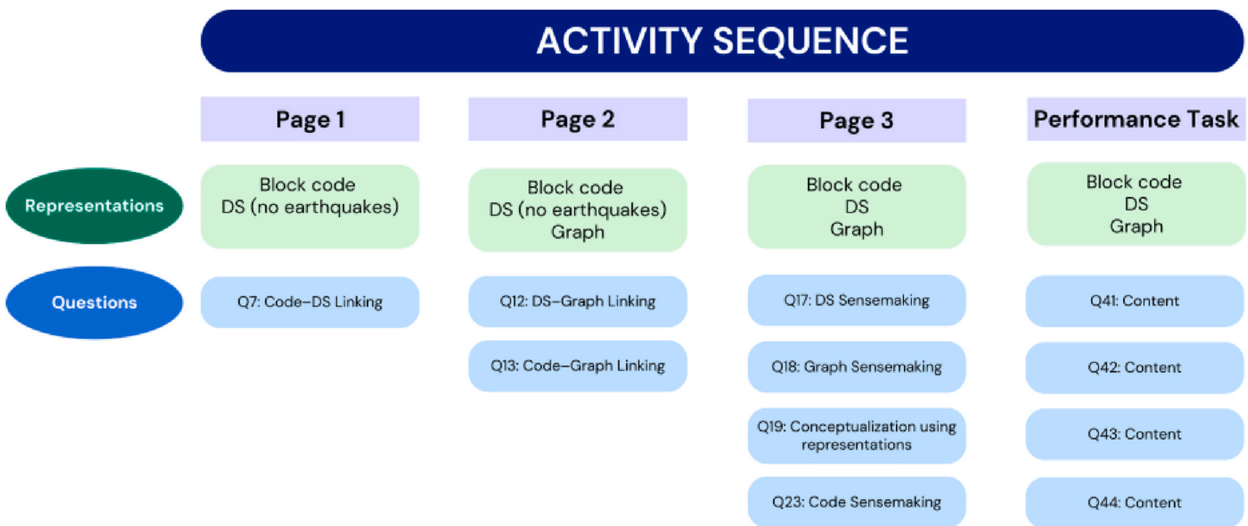

Fig. 2. The activity sequence. Two parallel and interacting scaffolded sequences were designed into the focal activity.

the graph corresponds to the amount of deformation shown in the grid?" Students are then prompted to link the block code with the deformation over time graph in Q13, which asks, "Describe what happens in the graph during each iteration (when you run through each 20-year time segment)." By the end of this page, students have interacted with all three representations and attempted to link each representation to the others.

On the third page, students add several blocks to the code to produce earthquakes in the DS. These blocks include a conditional statement and a function block, which allows them to set the friction level between the two plates. When the code is run with these new blocks, the grid lines bend and stretch up to a friction limit, which is defined in the conditional, and then break and snap back to straight. This breaking and snapping represents the phenomena in which earthquakes on the fault are the result of periodic deformation release and how the land disconnects on each side of the fault at that moment. Three sensemaking questions (Q17, Q18, Q23) prompt students to relate each of the representations to earthquake cycle concepts. The representational conceptualization question (Q19), "Why do earthquakes repeatedly occur?" asks students to use any of the representations in the GeoCoder to explain the earthquake cycle phenomenon.

The fourth focal page in the activity, hereafter referred to as the "performance task", serves as an embedded assessment of students' knowledge of the earthquake cycle after completing the previous pages of the seismic activity. This task engages students in an investigation where they use the GeoCoder to explain the earthquake cycle along a fault at three different locations in California (Table 2). In each location, we provided a pair of GPS stations with their actual velocities and the estimated friction level along the fault between the stations. Students are prompted to enter the speeds and friction levels for the three pairs of stations into the code to simulate the earthquake cycle in each location. The four questions on this page assess students' ability to conduct the investigation using the GeoCoder, interpret the representations, and apply their understanding to earthquake cycle concepts. These questions engage students in using the representational competencies from earlier in the activity.

4.4. Participants

During the spring semester of the 2023 school year, the seismic module was implemented by four teachers at a suburban high school in Colorado, two teachers at an urban middle school in Illinois, and one teacher at a suburban middle school in Illinois. Of the 431 students who completed the pre-test, 25% were 6th graders, 40% were 8th graders, 32% were 9th graders, and 1% were 10th graders. Forty-eight percent of students identify as female, 44% as male, 3% as non-binary, and 1% preferred not to answer. Eighty-seven percent of students spoke English as a first language. Many of the students (79%) had used computers to learn science prior to the module; 55% had prior coding experience.

4.5. Analyses

4.5.1. Scoring of representational competency items

Students' responses to multiple-choice (MC) and open-response (EX) questions in the seismic module were collected automatically. Throughout the entire activity, students were asked 44 MC, EX, and snapshot questions. For this study, we focus on 11 of the questions as they highlight the representational linking, sensemaking, and conceptualization tasks that students completed in the activity. Questions 7, 12, and 13 prompted representational linking. Questions 17, 18, and 23 prompted representational sensemaking. Question 19 prompted representational conceptualization (Table 1and 2).

The MC items were scored automatically and given a 1 for correct and 0 for incorrect. EX items were scored as follows: 0 (no information), 1 (non-normative), 2 (normative but simple), 3 (normative with simple linking, sensemaking, or conceptualization), and 4 (normative with complex linking, sensemaking, or conceptualization). For example, Score 2 was assigned to students' explanations that mentioned one of the representations or concepts but did not make any link between them. Score 3 was assigned to students' explanations that recognized the connection between representations or concepts but did not fully elaborate on the explicit connection between them. Score 4 was assigned to students' answers that showed strong competency in linking or sensemaking. The progression within this rubric can be viewed as higher sophistication in students' representational competency, leading to higher levels of understanding of both the representations themselves and the earthquake cycle.

To illustrate scoring, we use Question 7, which asks students to describe the changes in the DS as they run through the code one block at a time. This question targeted students' ability to link changes in the DS to the execution of the code blocks as they step through the code over and over again. The rubric developed to score student explanations to this question sorted student descriptions into five levels. Score 0 was given to students who did not provide any information, such as blank or off-topic ideas. Score 1 was given

Table 1
Questions in the performance task on the last page of the activity. EX = open response question. MC = multiple choice question.

No.	Type	Question
41	EX	How many earthquakes occurred at the fault between each pair of stations over 500 years?
42	EX	Compare the graphs and explain why the area along the fault between P513 and P544 GPS stations has less frequent earthquakes than the other two locations.
43	EX	The Salton Sea, which is between stations P491 and CACT, has experienced many small earthquakes over the years. Using evidence from your experimentation, explain why this is true.
44	MC	Between which set of stations would you expect the largest earthquake to occur?

Table 2
Percentage of responses in each score by question and the type of representational competency assessed with each question.

No.	Type	Connection	Score				
			0	1	2	3	4
7	EX	Code–DS Linking	2	9	27	46	17
12	MC	DS–Graph Linking	64	36	N/A	N/A	N/A
13	EX	Code–Graph Linking	1	13	46	21	19
17	EX	DS Sensemaking	0	25	44	29	2
18	EX	Graph Sensemaking	2	21	43	31	3
19	EX	Conceptualization	1	31	52	16	N/A
23	EX	Code Sensemaking	0	4	45	18	32
41	EX	Content	0	48	22	30	N/A
42	EX	Content	5	55	27	13	N/A
43	EX	Content	7	45	27	20	N/A
44	MC	Content	38	62	N/A	N/A	N/A

to students whose responses included non-normative ideas, such as "It changes the code." Score 2 was given to the response "The ground in the simulation moves" because it describes only the outcome of the DS and not any effect of running the code on the DS. Score 3 was given to the response "The simulations starts moving" because it does mention a simple connection about how the code affects the outcome in the DS. Score 4 was given to the response "Each time you click the step button, station 3 moves up, while station 1 and 2 move down" because it explicitly mentions how the code affects the DS outcome and makes a reference to the execution of the code.

4.5.2. Scoring of performance task items

We combined the scores of the four questions in the performance task to gain an overall score of students' ability to conduct, interpret, and explain computationally mediated investigations of the earthquake cycle. We scored the MC item 1 for correct and 0 for incorrect. The EX items were scored as follows: 0 (no response), 1 (non-normative response), 2 (normative with simple causal or functional response), and 3 (normative with complex causal or functional response).

4.5.3. Mixed effects general linear modeling

We used a mixed-effects general linear model to analyze the student performance score on the performance task. We created the performance score as a sum of the scores each student received on the four performance task items. We entered five student variables as fixed effects, including race (White vs. Non-White), English (first vs. second language), gender (male vs. female/non-binary), prior experience with using computers for science learning (yes vs. no), and prior experience in computer programming (yes vs. no). We chose to include female and non-binary students together because they are two genders who are underrepresented in STEM disciplines. We want our results to speak to the representational competency ability of these underrepresented genders, rather than understanding each gender individually. We entered the teacher variable as a random effect to account for variations that occur across teaching contexts. Fixed effects entered into the mixed effects general linear model were student scores related to three representation linking items (i.e., code to DS, DS to graph, and code to graph), three representation sensemaking items, and one representation conceptualization item. We used the partial eta squared (η^2) values along with p values to determine the significance and magnitude of each variable. We interpret the size of the effects using $\eta^2 = 0.01$ as a small effect, $\eta^2 = 0.06$ as a medium effect, and $\eta^2 = 0.14$ as a large effect.

Table 3
Mean performance task scores by student variables.

Variables	n	M	SD	t	p
Gender					
Female and other	233	5.82	2.24	0.59	0.55
Male	198	5.70	2.03		
Race					
Non-White	151	5.58	2.15	−1.34	0.18
White	294	5.86	2.13		
English					
Second language	43	5.77	2.17	0.01	0.98
First language	388	5.76	2.14		
Prior uses of computers for science learning					
Yes	81	5.49	2.13	−1.25	0.21
No	350	5.83	2.14		
Prior computer programming experience					
Yes	188	5.58	2.10	−1.57	0.12
No	243	5.91	2.17		

5. Results

Our first research question was to identify the extent to which students link, make sense of, and conceptualize procedural, geometric, and graphical representations in the context of the earthquake cycle. Table 3 shows the scoring results for each of the 11 questions in this study, including the type of question and the targeted representational competency. The percentage of responses in each score is shown.

In seven questions addressing representational competencies, we found that a majority of students were able to demonstrate some form of representational competence. However, few students showed strong competency in linking, sensemaking, and conceptualizing from the representations. In only one of the seven questions (Q7, code–DS linking) did over half of students make at least a simple connection (Score 3 or 4).

5.1. Descriptive statistics on representational competencies

5.1.1. Representational linking

The representational linking questions (Q7, Q12, and Q13) centered around students' ability to link the correlated visual output of two representations. Sixty-three percent of students made a simple connection between code and DS in Q7 and 40% did between code and graph in Q13. On Q7, 46% of students made a simple connection (Score 3) and 17% made a complex connection (Score 4). For example, in response to the question "What happens [in the DS] every time the 'Increase deformation … ' block is executed?" a sample Score 4 response is "It moves the plate 25 mm every year for 500 years." This response shows a sophistication in linking the elements of the code (eg. the data block specifying speed at 25 mm/year), the DS (eg. the plate motion), and the dynamic effect the changing code has on the output of the DS. Many of the responses earning a score of 4 included the values of plate speed in their answer to show their understanding of how those values in the code manifested in the speed at which the plates shift past each other.

5.1.2. Representational sensemaking

The representational sensemaking questions (Q17, Q18, and Q23) prompted students to relate the visual output of each representation to earthquake cycle concepts. Question 17 prompted students to sensemake using the DS, Q18 prompted students to sensmake using the deformation over time graph, and Q23 prompted students to sensemake using the code. Of these three questions, students did best relating the code to content (Q23), with 32% of responses including a sophisticated connection. This question, which asked students "What is the requirement for triggering an earthquake?" mirrored the language in the block code. An answer of "Deformation to the point of breaking" was scored as a 3 because the response describes the general relationship, but with no specific evidence from the representations. The response, "The requirement for triggering an earthquake is that the land deformation passes the max deformation determined by the amount of friction in the plates" scored 4 because the student is directly relating the procedural structure of the code to explain the phenomenon. It is clear this student has comprehended the meaning of the syntax in the code blocks enough to apply it in answering this question.

The DS sensemaking (Q17) and graph sensemaking (18) proved more difficult for students. Only two percent of students were able to make a sophisticated connection (Score 4) between the DS and geoscience content and 3% of students were able to make a sophisticated connection between the graph and content. In these two questions, students struggled to ground the representations in the domain-relevant language, despite the multiple visual and textual aids. For example, students' language often described only the surface features of the DS such as "one half goes up and one goes down" rather than the domain-relevant conceptual features that the DS was representing, making no mention of how the representation connects to land deforming due to plate movement.

Table 4
Mixed effects general linear model results.

Question	Source	df	F	p	Partial Eta Squared
–	Intercept	1	70.77	***	0.154
–	Gender[1]	1	0.13	0.72	0.000
–	Race[1]	1	0.38	0.85	0.000
–	English second language[1]	1	0.75	0.39	0.001
–	Prior uses of computers for science learning[1]	1	0.47	0.50	0.001
–	Prior computer programming experience[1]	1	0.001	0.97	0.000
Q7	Code-DS linking[1]	1	2.58	*	0.025
Q12	DS-graph linking[1]	1	4.28	*	0.011
Q13	Code-graph linking[1]	1	0.59	0.67	0.006
Q17	DS sensemaking[1]	1	2.46	*	0.024
Q18	Graph sensemaking[1]	1	0.77	0.54	0.008
Q23	Code sensemaking[1]	1	5.60	***	0.054
Q19	Conceptualization using representations[1]	1	3.49	*	0.026
–	Teaching context[2]	6	1.18	0.31	0.018

Note: [1]Fixed effect, [2]Random effect; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

5.1.3. Representational conceptualization

The sole question assessing students' ability to relate earthquake cycle concepts to the representations overall is Q19. This question asks, "Why do earthquakes repeatedly occur?" Students could draw their explanations from any, or multiple, of the representations in the GeoCoder. For example, students could describe the loop function in the code, which causes deformation to build up, even after an earthquake occurs. Or they could explain how the lines in the DS begin to bend and stretch directly after an earthquake. Or they could use the graph to describe how deformation immediately builds up again after an earthquake. In this question, 16% of students earned a top score of 3 by using evidence from any one of the representations to effectively answer this question. Fifty-two percent of students earned a score of 2 on this question because they were able to identify evidence from the representations that could answer this question, but were unable to use the evidence to explain the phenomenon.

5.2. General linear model analysis results

To investigate our second research question (How does the extent of students' representation linking, sensemaking, and conceptualization in the activity correlate with their performances in a computationally mediated geoscience investigation?), we conducted mixed effects general modeling. The results of this analysis (Table 4) showed that students' performance task scores were not affected by student demographic variables such as gender, race, English language learner status, prior uses of computers for science learning, and prior computer programming experience. Nor did students' performance task scores depend on the teaching context variable. The only significant correlations were with five representational competency items answered by students. The most prominent effect was code sensemaking ($\eta^2 = 0.054$, near medium effect), indicating the importance of students understanding how the code procedure relates to the earthquake cycle. The next significant effect was student performances on DS sensemaking ($\eta^2 = 0.024$, between small and medium effect). Other significant effects were found in code to DS linking ($\eta^2 = 0.025$, medium effect), conceptualization using representations ($\eta^2 = 0.026$, medium effect), and DS to graph linking ($\eta^2 = 0.011$, near small effect).

6. Discussion

This paper reports on the extent of the connections students make between multiple dynamically linked representations and on the representational competency scores that correlate with their performance in conducting and explaining a computationally mediated geoscience investigation. The goal of this study was to engage students with MLDRs so that they could develop necessary representational competencies in the context of investigating the earthquake cycle. In discussing the results of this study, we focus on students' expression of representational competencies, which competencies significantly correlated with students' performance task scores, and how representations in a scaffolded activity sequence enabled students' use of a computational tool for geoscience investigations.

6.1. Students' representational competencies

Several studies on multiple representations have reported success in building students' understanding of domain knowledge (Ainsworth, 2008; de Jong & van Joolingen, 1998; McElhaney et al., 2015). Presenting information in multiple modalities (e.g., simulations, images, graphs, and maps) is advantageous to learners because it allows them to actively map and process complementary information from each representation (Ainsworth, 2006). By engaging learners in integrating different representations, they actively construct a conceptual image that illustrates relationships between different kinds of information from diverse representations into a coherent structure (Schnotz & Rasch, 2005). However, before learning from the representations, students must fully grasp the intricacies and nuances associated with linking between the representations. In this study, many students were able to describe the dynamic links between each of the representations. Students were able to point out the automatic effect of the code on the geometric and graphical representations and articulate which code blocks caused specific changes in the outputs during the model run. They were also able to link the changing outputs of the geometric and graphical representations together, identifying how those representations provided complementary information of increasing deformation over time due to plate movement. Students who did not notice the visual cues between representations were the ones who had trouble linking between them, a result in line with Anzai (2001) and Schoenfeld et al. (1993).

Another essential competency defined in the literature necessary to build domain knowledge using multiple representations is sensemaking. Sensemaking involves using the perceptual features of representations to develop a conceptual understanding of domain-relevant material (Rau et al., 2017). Students who exhibited sophisticated sensemaking were able to see the conceptual aspects of the representations rather than only the surface-level aspects. For example, when explaining how the Deformation Simulation connects to the earthquake cycle, students who showed strong sensemaking used contextualized words such as "plates," "north and south," "land," and "transform boundary." These students also correctly identified events in the representations, such as how the snapping of the gridlines in the DS represents an earthquake. Students who showed weak sensemaking used descriptive words such as "two halves," "gridlines," and "up and down." Research suggests that being able to reason about how each representation relates to domain-relevant concepts that are not explicitly shown displays high sensemaking competency, and it is a significant result that many students in this study were able to exhibit this ability.

The third representational competency studied in this paper is conceptualization. Conceptualization competencies detail the knowledge and skills students use to explain real-world phenomena using representations and why a particular visual representation could be chosen for a task. In previous research (Mayer, 2005; Schnotz, 2014, pp. 72–103), conceptual competencies are involved when students select information for further processing by identifying meaningful visual features and mapping these to conceptual knowledge. This study found that many students were able to identify evidence from one or more representations in the GeoCoder to answer conceptualization questions, yet few students were able to use the evidence to effectively explain the phenomena. In other words, students identified the relevant aspects of the representations to be used in the explanation yet could not extend their interpretation to write an effective explanation. This result is unsurprising, as we believe the conceptualization competency is the most difficult for students to achieve. It requires that students understand what the question is asking, what evidence can answer that question, and which representations can be used to gather that evidence.

6.2. Representational competency correlation with performance

This study found that all three representational competencies are significantly correlated with future performance. Students who showed strong ability in linking, sensemaking, and conceptualizing the representations in the three pages of the focal activity scored higher on the questions in the performance task. Students who did not score well on the representational competency questions did not do well in the performance task. This finding adds to the growing body of evidence that developing proficiency in representational competencies is crucial for students to effectively learn scientific concepts across various disciplines.

This study found students' levels of responses to all but two competency questions are significantly correlated with performance task scores. Interestingly, these two questions involved linking graphical representation to code (Q13) and sensemaking of graph (Q18). This indicates that successfully linking and sensemaking with the graph was important but not significant in determining student performances on the performance task. It is well known that students struggle to read and interpret graphs, and this finding could speak to the affordances of particular representations to help build a mental model of scientific phenomena. The graph is the most abstract representation in GeoCoder. Both the DS and the syntax of the code convey more information about the processes of the earthquake cycle than the graph does. It is possible that information from the graph was not useful for students in building their mental model of the earthquake cycle. This finding is in line with previous studies which show that although curriculum designers may perceive equivalent and useful information in multiple representations, students may not (Ainsworth, 1999; Ainsworth et al., 2002). Additionally, Kuo et al. (2017) demonstrated that students often prefer to use only one or two representations in tasks despite being asked to use multiple. As code and DS other representations were possibly more useful to students, they relied on those to build their understanding.

6.3. Design considerations

Cory and Garofalo (2011) highlighted two essential design considerations when building a curriculum with technology-based representations: (1) the careful curation required to select and sequence the representations and (2) the instructional design required to scaffold and guide students' use of the technology. In this study, we deliberately sequenced the representations and representational competency questions to help students link, sensemake, and conceptualize the representations and the earthquake cycle content. We introduced the representations and the visual cues that connect them in sequenced pages so as not to overwhelm students when first using them. For example, the graph was not shown on the first page of the activity so that students could focus only on the block code and DS. Additionally, when the graph was added on the second page, earthquakes were still not included in the function of the model to help students focus on making connections between the rise of the deformation as represented on the Y-axis of the graph and the bend of the lines in the DS. The full functionality of the model, with earthquakes, was added on the third page of the sequence.

To go along with the step-by-step introduction of model features, the curriculum slowly scaffolded students in linking, sensemaking, and conceptualization using the representations. As Ainsworth (2006) points out, in order to utilize a representation, learners must first understand how the information is encoded in the representation, how it relates to its context or domain, and how to choose the best representation according to the problem. Only after students understand these elements of the representations can they effectively engage in scientific inquiry. Other researchers have studied how the order of introduction of representational competencies affects learning (Rau et al., 2017). In this study, we show an example of one way to sequence competencies. Linking is first because in making connections across representations, learners develop mental models that may lead to a deeper understanding of the domain (Ainsworth & van Labeke, 2004). Linking is followed by sensemaking and conceptualization because those competencies require students to operationalize their knowledge of the representations. This study cannot speak to the efficacy of this sequence compared to a different sequence, but we believe learning *about* the representations is required before learning *from* the representations.

7. Limitations

There are several limitations in this study. First, although we found that students were successful in exhibiting representational competencies, there is considerable room for improvement. To some extent, the low scores may reflect the noted struggle that students have when encountering new representations, especially when investigating complex systems such as the earthquake cycle. Also, there is only one question in the study assessing students' conceptualization competency, and we have made conclusions about the correlation between this competency and students' future performance based only on this one question. It would be advantageous in the future to have more questions prompting students to exhibit their conceptualization competency. Additionally, there are no experimental variations, and therefore we cannot say that there is a causal correlation between the three individual competencies and the learning outcomes on the performance task. Lastly, this study reports only a snapshot of students' performance at one time and there are no pre-post tests or general abilities that would allow controlling for confounding variables. Future work could show whether students are able to build representational competence throughout an authentic computational investigation such as this using pre- and post-assessments to measure their learning gains.

8. Conclusion

Multiple representations can play a significant role in engaging students in simultaneous learning of both computational practices and domain content. However, teachers are struggling to find ways to effectively integrate computational practices with disciplinary core ideas in science. The seismic module used in this study leveraged multiple dynamically linked representations to build students' representational competencies while carrying out computationally-mediated geoscience investigations. In this study, the use of the three representations was scaffolded within the instructional activities so that students could slowly build representational competencies to conduct investigations within the computational learning environment. Research is warranted to further explore representational competencies essential to building a deeper understanding of domain content knowledge. This is especially true when working in domains such as geoscience, where representations such as such as simulations, maps, and graphs allow students to study systems that occur across a wide range of spatial and temporal scales. This research may have a broad impact on teaching in many science domains as representational competencies are significantly correlated with future performance in science investigation tasks and using computational methods to produce and interpret multiple scientific representations is now a common practice in many science disciplines.

CRediT authorship contribution statement

Christopher Lore: Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis, Data curation, Conceptualization. **Hee-Sun Lee:** Writing – original draft, Validation, Methodology, Formal analysis, Conceptualization. **Amy Pallant:** Writing – review & editing, Supervision, Funding acquisition. **Jie Chao:** Methodology, Conceptualization.

Declaration of competing interest

none.

Data availability

The authors do not have permission to share data.

Acknowledgements

This material is based on work supported by the National Science Foundation under grant nos. DRL-1841928 and DRL-2241021. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compedu.2024.105149.

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*(2–3), 131–152. https://doi.org/10.1016/S0360-1315(99)00029-9.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*(3), 183–198. https://doi.org/10.1016/j.learninstruc.2006.03.001

Ainsworth, S. (2008). The Educational Value of Multiple-representations when learning complex scientific concepts. *Visualization: Theory and Practice in Science Education*, 191–208.

Ainsworth, S., Bibby, P., & Wood, D. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *The Journal of the Learning Sciences, 11*(1), 25–61.

Bodemer, D., & Faust, U. (2006). External and mental referencing of multiple representations. *Computers in Human Behavior, 22*(1), 27–42. https://doi.org/10.1016/j.chb.2005.01.005

Cheek, K. A., Ladue, N. D., & Shipley, T. F. (2017). Learning about spatial and temporal scale: Current research, psychological processes, and classroom implications. *Journal of Geoscience Education, 65*(4), 455–472. https://doi.org/10.5408/16-213.1

Council, N. R. (2006). *Learning to think spatially*. National Academies Press. https://doi.org/10.17226/11019

de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*(2), 179–201.

DeLoache, J. S. (2000). Dual representation and young children's use of scale models. *Child Development, 71*(2), 329–338. https://doi.org/10.1111/1467-8624.00148

Edelsbrunner, P. A., Malone, S., Hofer, S. I., Küchemann, S., Kuhn, J., Schmid, R., Altmeyer, K., Brünken, R., & Lichtenberger, A. (2023). The relation of representational competence and conceptual knowledge in female and male undergraduates. *International Journal of STEM Education, 10*(1), 44. https://doi.org/10.1186/s40594-023-00435-6

Eilam, B., & Ben-Peretz. (2012). *Teaching, learning, and visual literacy: The dual role of visual representation*. Cambridge University Press. https://doi.org/10.1017/CBO9781139026611

Evagorou, M., Erduran, S., & Mäntylä, T. (2015). The role of visual representations in scientific practices: From conceptual understanding and knowledge generation to 'seeing' how science works. *International Journal of STEM Education, 2*(1), 11. https://doi.org/10.1186/s40594-015-0024-x.

Gilbert, J. K. (2005). Visualization: A metacognitive skill in science and science education. In J. K. Gilbert (Ed.), *Visualization in science education* (pp. 9–27). Netherlands: Springer. https://doi.org/10.1007/1-4020-3613-2_2.

Henderson, P. B., Cortina, T. J., Hazzan, O., & Wing, J. M. (2007). Computational thinking. *Proceedings of the 38th ACM SIGCSE technical symposium on computer science education (SIGCSE '07)*. https://doi.org/10.1145/1227310.1227378

Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H. A., & Kumar, V. (2019). Machine learning for the geosciences: Challenges and opportunities. *IEEE Transactions on Knowledge and Data Engineering, 31*(8), 1544.

Lore, C., Lee, HS., Pallant, A., et al. (2024). Integrating computational thinking into geoscientific inquiry about volcanic eruption hazards and risks. *International Journal of Science and Mathematics Education, 22*, 1173–1195. https://doi.org/10.1007/s10763-023-10426-2.

Lynch, M. (2006). The production of scientific images: Vision and re-vision in the history, philosophy, and sociology of science. In *Visual cultures of science: Rethinking representational practices in knowledge building and science communication* (pp. 26–40). Dartmouth College Press.

Mayer, R. E. (2005). *Multimedia learning*. Cambridge University Press.

McElhaney, K. W., Chang, H.-Y., Chiu, J. L., & Linn, M. C. (2015). Evidence for effective uses of dynamic visualisations in science curriculum materials. *Studies in Science Education, 51*(1), 49–85. https://doi.org/10.1080/03057267.2014.984506

Medunić, G., Chakravarty, S., & Kundu, R. (2022). *Computational skills in geosciences higher education system for the 21st century* (Vol. 8).

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. National Academies Press.

Pande, P., & Chandrasekharan, S. (2017). Representational competence: Towards a distributed and embodied cognition account. *Studies in Science Education, 53*(1), 1–43. https://doi.org/10.1080/03057267.2017.1248627

Pauwels, L. (2006). *Visual cultures of science: Rethinking representational practices in knowledge building and science communication*. UPNE.

Rau, M. A. (2017). A framework for educational technologies that support representational competencies. *IEEE Transactions on Learning Technologies, 10*(3), 290–305. https://doi.org/10.1109/TLT.2016.2623303

Rau, M. A., Aleven, V., & Rummel, N. (2017). Making connections among multiple graphical representations of fractions: Sense-making competencies enhance perceptual fluency, but not vice versa. *Instructional Science, 45*(3), 331–357. https://doi.org/10.1007/s11251-017-9403-7

Rau, M. A., Aleven, V., Rummel, N., & Pardos, Z. (2014). How should intelligent tutoring systems sequence multiple graphical representations of fractions? A multi-methods study. *International Journal of Artificial Intelligence in Education, 24*(2), 125–161. https://doi.org/10.1007/s40593-013-0011-7

Rau, M. A., Michaelis, J. E., & Fay, N. (2015). Connection making between multiple graphical representations: A multi-methods approach for domain-specific grounding of an intelligent tutoring system for chemistry. *Computers & Education, 82*, 460–485. https://doi.org/10.1016/j.compedu.2014.12.009

Resnick, M., Maloney, J., Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. (2009). Scratch: Programming for everyone. *Communications of the ACM, 52*, 60–67.

Savage, J. C., & Burford, R. O. (1973). Geodetic determination of relative plate motion in central California. *Journal of Geophysical Research, 78*(5), 832–845. https://doi.org/10.1029/JB078i005p00832

Schnotz, W. (2014). Integrated model of text and picture comprehension. *The cambridge handbook of multimedia learning*. https://doi.org/10.1017/CBO9781139547369.006

Schnotz, W., & Rasch, T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology Research and Development, 53*(3), 47–58. https://doi.org/10.1007/BF02504797.

Schönborn, K. J., & Anderson, T. R. (2009). A model of factors determining students' ability to interpret external representations in biochemistry. *International Journal of Science Education, 31*(2), 193–232. https://doi.org/10.1080/09500690701670535

Sell, K. S., Herbert, B. E., Stuessy, C. L., & Schielack, J. (2006). Supporting student conceptual model development of complex Earth systems through the use of multiple representations and inquiry. *Journal of Geoscience Education, 54*(3), 396–407. https://doi.org/10.5408/1089-9995-54.3.396

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction, 13*(2), 227–237.

Spiro, R. J., & Jehng, J.-C. (1990). In D. Nix, & R. J. Spiro (Eds.), *Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter* (pp. 163–205). Lawrence Erlbaum Associates.

Stern, E., Aprea, C., & Ebner, H. G. (2003). Improving cross-content transfer in text processing by means of active graphical representation. *Learning and Instruction, 13*(2), 191–203. https://doi.org/10.1016/S0959-4752(02)00020-8

van der Meij, J., & de Jong, T. (2006). Supporting students' learning with multiple representations in a dynamic simulation-based learning environment. *Learning and Instruction, 16*(3), 199–212.

Whitmeyer, S. J., Fichter, L. S., & Pyle, E. J. (2007). New directions in wilson cycle concepts: Supercontinent and tectonic rock cycles. *3*(6), 511–526.