Journal of Science Education and Technology (2025) 34:1039–1054
https://doi.org/10.1007/s10956-025-10242-z

Characterizing Integrated Learning of Disciplinary Core Ideas and Science Practices in a Computational Thinking (CT)–Integrated Biology Curriculum

Sugat Dabholkar[1,4] · **Golnaz Arastoopour Irgens**[2] · **Uri Wilensky**[3]

Accepted: 7 July 2025 / Published online: 23 August 2025
© The Author(s), under exclusive licence to Springer Nature B.V. 2025

Abstract

Calls for science education reforms emphasize integrated learning of science practices and disciplinary ideas established using those practices. In this paper, we present a novel approach to model and analyze student participation in Science and Engineering Practices (SEPs) and learning of Disciplinary Core Ideas (DCIs) in a Next Generation Science Standards (NGSS)–aligned curricular unit. We studied student participation and learning in a Computational Thinking (CT)–integrated biology unit about natural selection taught in an urban under-resourced high school. Students ($n = 88$) designed and conducted different experiments using an agent-based computational model and answered questions about their experimental investigations. Using Epistemic Network Analysis (ENA), we analyzed student responses ($n = 2026$) to model and investigate connections among DCIs and SEPs, which we call *epistemic connections*. An aggregate-level analysis of the centroids of networks of epistemic connections shows statistically significant clusters indicating differences in participation in practices and ideas in different lessons. The detailed analysis of the epistemic connections in networks of two students and their written responses showed how various kinds of participation in science practices supported students in making sense of disciplinary ideas, and their engagement in disciplinary ideas reciprocally supported the refinement of science practices. Ours demonstrates the usefulness of analyzing students' epistemic connections using ENA to investigate their integrated learning of practices and disciplinary ideas. We discuss the implications of such analysis for improving curricular designs and instructional strategies and studying student learning in NGSS-aligned curricular units.

Keywords Epistemic network analysis · Epistemic connections · Model-based inquiry · Science educationf · Agent-based models

Many calls for science education reforms emphasize that the goals of science education should not be limited to "knowing about science," rather they should include "learning to use science practices and tools to make sense of the world" (Abd-El-Khalick et al., 2004; Schwarz et al., 2017). This shift in goals entails supporting students in learning contemporary scientific inquiry practices in addition to disciplinary core ideas. Such learning would involve epistemologically meaningful engagement in science practice for sense-making rather than merely knowing about scientific inquiry (Berland et al., 2016; Lehrer & Schauble, 2012). In other words, student participation in science practices for constructing knowledge about the world in classrooms should be similar to how scientists participate in these practices. In contemporary scientific practices, the nature of scientific inquiry is rapidly changing due to the integration of computational tools for scientific investigations (Wilensky et al., 2014). It is important that learning environments are designed to engage students in these continuously advancing ways of science practice.

Because of the increased use of computational tools and methods, Computational Thinking (CT) has impacted

✉ Sugat Dabholkar
sugat.dabholkar@tufts.edu; sugat@iiserpune.ac.in

[1] Department of Education, Tufts University, Medford, MA, USA

[2] Department of Teaching and Learning, Vanderbilt University, Nashville, TN, USA

[3] School of Education and Social Policy, Northwestern University, Evanston, IL, USA

[4] Department of Science Education, Indian Institute of Science Education and Research Pune, Pune, India

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

how scientists engage in various science practices. CT has recently garnered a lot of attention in the field of Learning Sciences, in general and science education in particular (Grover & Pea, 2013; Lee et al., 2020; Weintrop et al., 2016; Wilensky et al., 2014). Researchers of education have defined CT in several different ways, from CT as thinking that Computer Scientists do to CT as a fundamental form of thinking that cuts across domains (Grover & Pea, 2013; Weintrop et al., 2016). In our work, we use a definition of CT in the context of a specific disciplinary domain such as science or mathematics (Weintrop et al., 2016). Weintrop and colleagues have defined a taxonomy of practices that scientists participate in when they use computational tools and methods. Integrating these CT practices in a science curriculum would foster student participation in contemporary science practices in curricular contexts (Arastoopour et al., 2020; Arık & Topçu, 2022; Swanson et al., 2019; Waterman et al., 2020).

The Next Generation Science Standards (NGSS) in the USA emphasize participation in a set of science and engineering practices (SEP) to engage in disciplinary core ideas (DCI) (NGSS Lead States, 2013). SEPs describe behaviors that scientists use as they investigate and make sense of phenomena in the natural world and the key set of engineering practices that engineers engage in as they design and build models and systems. DCIs are the key ideas that students should know in order to understand multiple science disciplines. NGSS recommends science learning that involves the integration of these dimensions. This means that these dimensions should not be used in isolation, nor should they be combined at a superficial level, instead they should be integrated while designing science curricula (Berland et al., 2016; Linn et al., 2016; Manz et al., 2020; Reiser et al., 2017).

In this paper, we study student learning in an NGSS-aligned CT-integrated curriculum designed to foster their participation in SEPs using computational models to learn about DCIs. We designed an agent-based computational model (Wilensky & Rand, 2015) to serve as an experimental system (Dabholkar & Wilensky, 2024). Using this model, we designed a CT-integrated high school biology curriculum that included activities for students to learn by participating in scientific investigations to construct knowledge about natural selection. Supporting science learning with such activities using computational tools raises new questions about student participation and learning, such as what counts as meaningful engagement in integrated learning of science practices and disciplinary ideas in a specific curricular unit; and how we can characterize student learning as they participate in such knowledge construction activities using a computational model. In this paper, we primarily focus on the second question about characterizing student learning by developing an analytical approach to study student participation in specific science practices as they engage with specific disciplinary ideas. We call these co-occurrences of specific practices and ideas *epistemic connections*. We use Epistemic Network Analysis (ENA; Shaffer et al., 2009) to visualize and analyze students' epistemic connections in two ways as they participated in the CT-integrated biology curriculum. First, the aggregate centroid analysis in ENA highlights patterns of students' science practices and disciplinary ideas that changed sequentially as they continued through the curriculum. Second, the detailed network analysis of epistemic connections and written responses of two students highlight individual differences in student participation in terms of how they engaged with the practices and ideas and how their participation in science practices in a disciplinary context supported learning of both.

Theoretical Framework

Epistemic Connections

Classroom learning environments need to be designed to support student participation in epistemic activities in disciplinary learning contexts (Barzilai & Chinn, 2018; Berland et al., 2016; Duschl, 2008; Wilkerson et al., 2018). An epistemic activity is a learning activity in which students participate in specific practices to construct knowledge. Such activities would involve construction, evaluation, and revision of knowledge of disciplinary ideas. Broadly speaking, student participation in epistemic activities in the science classroom would entail becoming familiar with the context of a phenomenon they are learning about, asking relevant questions that they can investigate the phenomenon using an experimental system, testing and verifying claims by designing and performing investigations, constructing and sharing explanations regarding the phenomenon based on their investigations, and reflecting on the process of the knowledge construction (Berland et al., 2016; Schwarz et al., 2017). A key to supporting such learning is to design an experimental system that would allow students to meaningfully investigate disciplinary ideas related to a phenomenon by participating in science practices. For this, we use an approach to design computational models called Emergent Systems Microworlds (ESMs) that combines two established theories of learning and design: agent-based modeling of complex systems (Wilensky, 2001, 2003; Wilensky & Jacobson, 2014) and constructionism (Papert, 1980). The core design features of an ESM, agent-based representations and abilities to manipulate and observe in a microworld, when combined, are known to facilitate learning of a complex phenomenon (Jacobson & Wilensky, 2006; Klopfer et al., 2005; Wilensky, 2003; Wilensky & Papert, 2010; Wilkerson-Jarde & Wilensky, 2010). In this paper, we study student

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

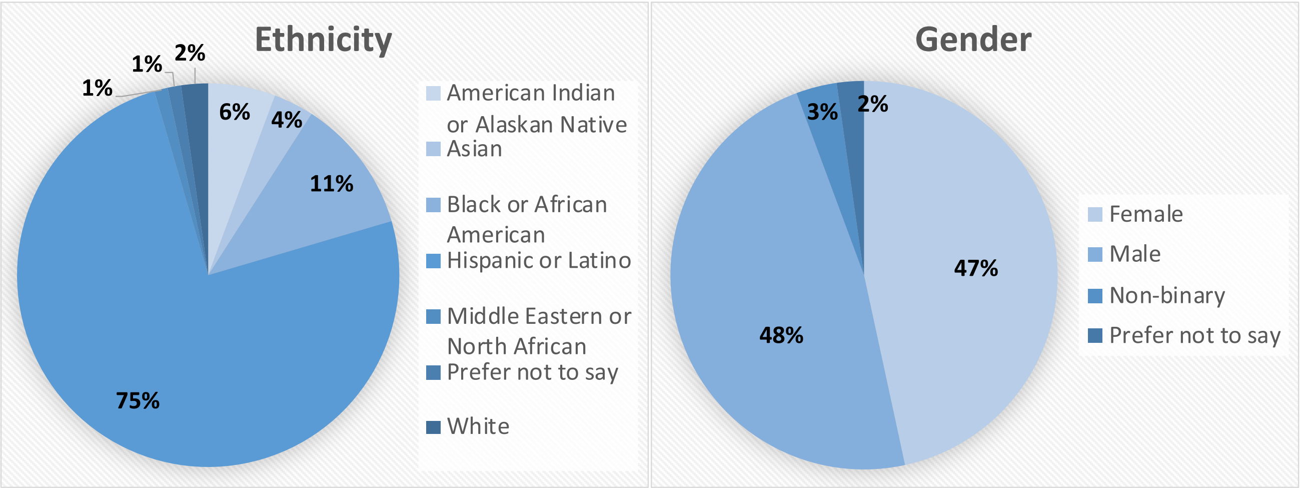


Fig. 1 Demographic distribution of participating students (total students = 88)

participation in epistemic activities using a computational model designed as an ESM (Dabholkar et al., 2018a). For example, in a curricular unit about natural selection, such participation in epistemic activities would involve asking questions, analyzing data, and constructing explanations using specific ideas such as heritability, environment, and survival.

In our analysis, we focus on the co-occurrence of disciplinary ideas and practices in student participation in epistemic activities, which we call *epistemic connections*. Epistemic connections are connections between Practices ↔ Practices, Ideas ↔ Ideas, and Practices ↔ Ideas. We analyze changes in students' epistemic connections to characterize the learning of science practices and disciplinary ideas. We hypothesize that using ENA to visualize and measure students' epistemic connections would help in understanding the connections between practices and ideas and specifically, how a disciplinary context provided opportunities to learn science practices and how engagement in science practices help students in learning about disciplinary ideas. In this paper, we investigated the following research question:

What do students' *epistemic connections* reveal about their learning of science practices and disciplinary core ideas?

Methods

Research Context

Evolution of Populations is a 10-day CT-integrated biology unit designed by the lead author in consultation with high school biology teachers (Dabholkar et al., 2018b). The unit is freely available for editing and using in classrooms (https://ct-stem.northwestern.edu/curriculum/preview/681/). The computational models in the unit are built using NetLogo, an agent-based modeling platform that is intentionally designed to foreground emergent systems modeling for educational and research purposes (Wilensky, 1999). The unit focused on predator–prey dynamics, competition among individuals, and natural selection. The unit was taught by a biology teacher, Ms. Lydia (pseudonym), in a large Midwestern city's public high school. Ms. Lydia taught the unit in four 9th grade high school biology classes to students of ages 14–15 years over 2 weeks. All the classes were conducted in person. One hundred percent of the participating students were on free lunch (an indicator of belonging to an underserved population). Students were predominantly of Hispanic or Latino ethnicity with a roughly equal number of male and female identifying and a few non-binary students (Fig. 1).

The unit used an anchoring phenomenon (Reiser, 2014) about a population of rock pocket mice in New Mexico's Valley of Fires reported in a Howard Hughes Medical Institute (HHMI)[1] *BioInteractive* classroom resource. The unit consisted of six lessons, three of which were about natural selection. In our analysis, we included only student participation in the lessons that were about natural selection. The Natural Selection lessons progress from introducing the anchoring phenomenon to students performing scientific investigations using the ESM. There were two main tools embedded in the curriculum: (a) a video developed by HHMI to introduce students to the anchoring phenomenon and (b) a computational model designed as an ESM

---

[1] https://www.biointeractive.org/classroom-resources/making-fittest-natural-selection-and-adaptation.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

(Dabholkar & Wilensky, 2024) that simulated natural selection and adaptation in a rock pocket mice population. The pages of the lessons typically had embedded tools, informative text, and question prompts. Student responses to the prompts were recorded on the platform for teachers as well as for researchers.

The curricular activities were delivered through an online portal and were split into lessons. Each lesson consisted of three to four pages. Typically, on each page, students read a prompt with a description of a computational model and suggestions for exploration. Then, students answered two to five embedded questions on the same page. Figure 2 shows a page in lesson 2 in which students explored the ESM (using the drop-down menu and sliders to change parameters and run simulations with different conditions) and wrote answers to embedded questions in the lessons. Each student had an account on the curriculum portal, which they accessed on their laptop computers to use the embedded model and write responses to the question prompts. The data analyzed in this paper includes written responses of 88 consented students to the embedded questions. In the next section, we discuss how we analyzed the student responses to infer their participation in science practices and engagement with different aspects of disciplinary ideas.

The computational model (Fig. 2) is designed so the students can manipulate several parameters about the system such as the composition of the initial population of rock pocket mice, the background colors—dark, light, or mixed, and so on. Such manipulations are useful for students to investigate how system-level changes emerge because of agent-level interactions (Dabholkar & Wilensky, 2024). The predation probability in the model depends on how well a mouse camouflages with the color of its surroundings. The button “Add a mutant" adds a heterozygous mutant at a random location. The model allows temporal visualizations of population changes under specified conditions. This unit incorporates specific disciplinary ideas in the NGSS (Table 1) related to the evolution of populations for high school students (NGSS Lead States, 2013).

Agent-based representations were used in the following manner for students to investigate specific aspects of the phenomenon related to genetic inheritance, survival, environmental factors (Table 2).

The task structures in the curriculum used constructionist design principles such as facilitating personally meaningful self-driven explorations and expressing and testing ideas using a constructionist microworld (Papert, 1980). Students worked in groups of three or four to discuss their answers but wrote answers individually on their laptops. The teacher projected pages from each lesson on a projector screen and discussed curricular activities. The teacher then asked students to participate in the activities as groups. As students participated in curricular activities, the teacher walked around the class, checked on the progress of each student group, and answered their questions. The teacher occasionally held class-level discussions to facilitate sharing among student groups.

Identification of Epistemic Connections

We characterize student learning of practices and ideas together by studying changes in their epistemic connections using Epistemic Network Analysis (ENA) (Shaffer et al., 2009, 2016). ENA involves creating network models based on when and how often learners connect domain-relevant elements. ENA has been demonstrated to be an effective way to visually and statistically compare networks; it allows researchers to reflect the weighted structure of connections and quantitatively compare the networks in a variety of domains (e.g., Arastoopour et al., 2014; Bagley & Shaffer, 2015; Siebert-Evenstone & Shaffer, 2019). These network visualizations and statistical comparisons furthermore allow researchers to characterize and assess student learning as they express their ideas (Arastoopour Irgens et al., 2015).

We build on our work (Arastoopour Irgens et al., 2020) to present novel methodological insights into the operationalization of the newly defined analytical construct—*epistemic connections*. In this paper, we analyzed student responses to embedded questions in three different lessons of the unit to study their engagement in science practices and disciplinary ideas. We treat student responses as a part of the discourse between students and the curriculum.

Top-down Coding of SEPs

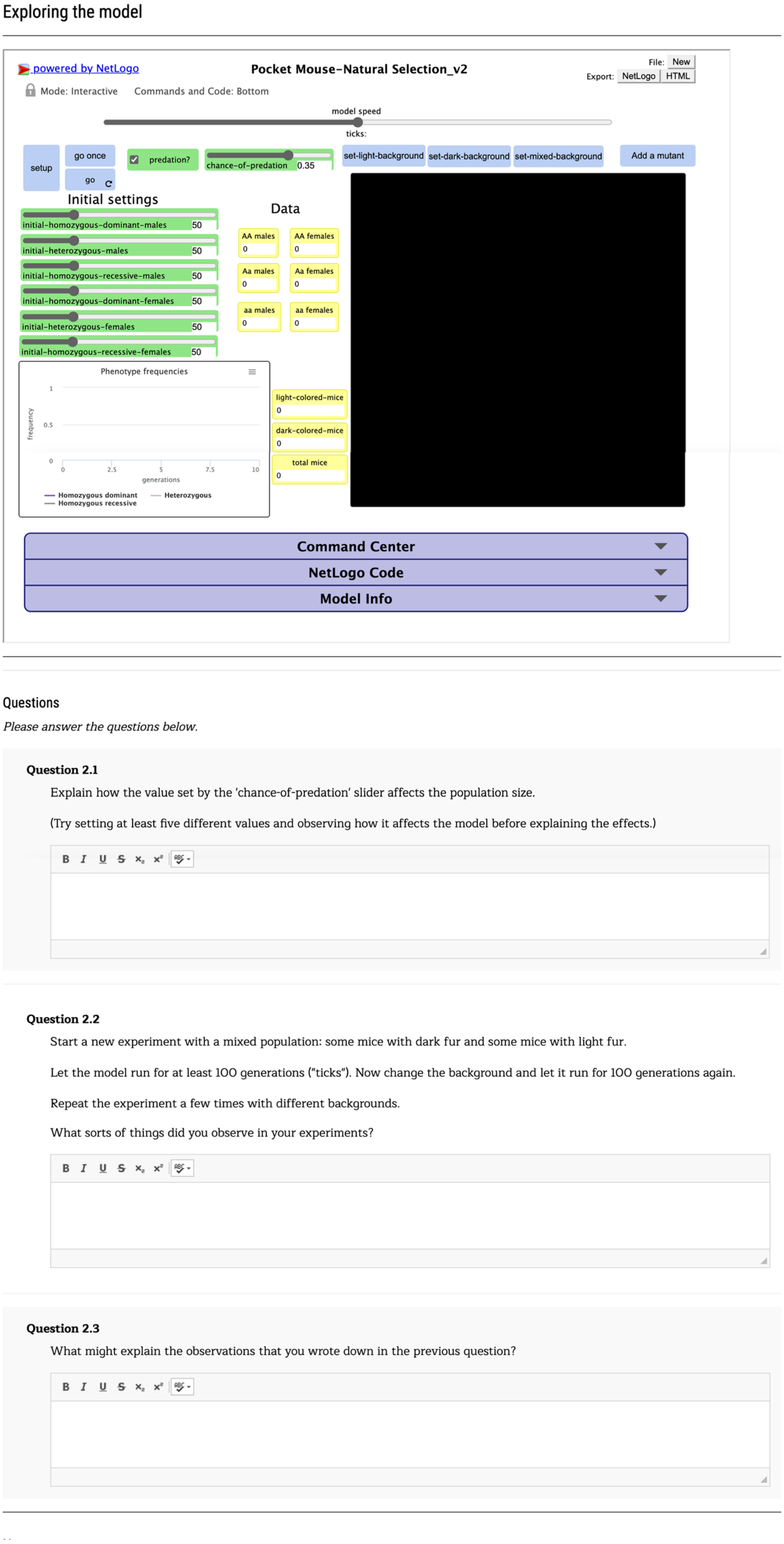
We coded for students’ explicit engagement in science practices such as, using a model or analyzing data, as well as explicit mentions of aspects of disciplinary core ideas (DCIs) such as adaptation or inheritance. We used both a top-down and a bottom-up approach to develop codes (Miles et al., 2019). To identify student responses that showed student engagement in science practices (SEPs[2]), we used top-down codes derived from NGSS science practices (NGSS Lead States, 2013) and Weintrop et al.’s (2016) taxonomy of computational thinking practices: Asking Questions, Developing and Using Models, Planning and Carrying Out Investigations, Analyzing and Interpreting Data, and Constructing

---

[2] Though we primarily focused on science practices, we use the term SEP and not SP because SEP is a more common term among researchers and practitioners who use the NGSS framework.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Fig. 2 A page from lesson 2 in which students explored a NetLogo model of rock pocket mice (Dabholkar & Wilensky, 2020) about the predator–prey relationship and natural selection



Exploring the model

☑ predation?

chance-of-predation 0.35

initial-homozygous-dominant-males 50

initial-heterozygous-males 50

initial-homozygous-recessive-males 50

initial-homozygous-dominant-females 50

initial-heterozygous-females 50

initial-homozygous-recessive-females 50

AA males 0

AA females 0

Aa males 0

Aa females 0

aa males 0

aa females 0

light-colored-mice 0

dark-colored-mice 0

total mice 0

Questions

Please answer the questions below.

Question 2.1

Explain how the value set by the 'chance-of-predation' slider affects the population size.

(Try setting at least five different values and observing how it affects the model before explaining the effects.)

Question 2.2

Start a new experiment with a mixed population: some mice with dark fur and some mice with light fur.

Let the model run for at least 100 generations ("ticks"). Now change the background and let it run for 100 generations again.

Repeat the experiment a few times with different backgrounds.

What sorts of things did you observe in your experiments?

Question 2.3

What might explain the observations that you wrote down in the previous question?

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Explanations. Please see Appendix for the definitions and examples of the codes.

Bottom-up Coding of DCIs

We used a bottom-up coding approach to devise codes for characterizing student knowledge of DCIs. In 10% of written responses, we identified the parts about disciplinary ideas related to natural selection and used an inductive coding method (Miles et al., 2019) to label the contextual engagement of different aspects of disciplinary ideas in students written discourse (Shaffer & Ruis, 2021). Based on iterative refinement of coding categories using constant comparison (Miles et al., 2019), we arrived at the following coding categories for disciplinary ideas: Populations/Individuals/Agents, Phenotypic Properties or Characteristics (Phenotype), Genotypic Properties or Characteristics (Genotype), Environments, Heritability, Survival, Adaptation Mechanism, and Change/Mutation/Variation (see the codebook for explanations of codes and examples in the Supplementary Materials). Because the data (of 88 students) contained a large number ($n = 2026$) of responses, we developed an automated coding algorithm that used keywords and regular expressions (see Arastoopour Irgens et al., 2020 for a similar methodological approach), refined the coding scheme, and conducted final pairwise inter-rater reliability tests among two human raters and the algorithm using Cohen's Kappa and Shaffer's Rho (Shaffer, 2017) (see Table 3).

The automated coding algorithm used keywords and regular expressions to code for student participation in SEPs and their engagement with DCIs. Regular expressions are sequences of characters and special symbols that delineate a

Table 1 NGSS Disciplinary Core Ideas (DCIs) incorporated in the unit

HS-LS4-2	Construct an explanation based on evidence that the process of evolution primarily results from four factors: (1) the potential for a species to increase in number, (2) the heritable genetic variation of individuals in a species due to mutation and sexual reproduction, (3) competition for limited resources, and (4) the proliferation of those organisms that are better able to survive and reproduce in the environment
HS-LS4-4	Construct an explanation based on evidence for how natural selection leads to adaptation of populations

Table 2 Modeled agent-based representations and emergent outcomes in the agent-based model

Agents	Mice
Agent properties	Genotype—homozygous/heterozygous; phenotype—dark/light fur color
Agent behaviors	Movement, reproduction, inheritance, change in genotype and phenotype because of mutation, death because of predation
Agents' environment	Color of the surroundings (to camouflage), predators
Emergent outcomes	Changes in the genotype and phenotype frequencies in the population (adaptation and survival)

Table 3 Code categories and inter-rater reliability values for each code (*Shaffer's Rho < .05)

Code category	Code	Cohen's Kappa between rater 1 and rater 2, rater 1 and automation, and rater 2 and automation
Scientific Inquiry Practices	Asking Questions and Defining Problems	1.0*, 1.0*, 1.0*
	Developing and Using Models	.92*, .91*, .83
	Planning and Carrying Out Investigations	.91*, .73*, .77*
	Analyzing and Interpreting Data	.85*, .91*, .77
	Constructing Explanations	.86*, .65, .81*
Disciplinary Ideas	Populations and Individuals	1.0*, .92*, .92*
	Phenotypic Properties	1.0*, .78*, .78*
	Genotypic Properties	1.0*, .82*, .82*
	Environments	1.0*, .79*, .79*
	Heritability	.98*, .75*, .77*
	Survival	.92*, .83*, .91*
	Adaptation Mechanism	.94*, .81*, .88*
	Variation and Mutation	.92*, .76*, .83*

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

search pattern for text strings. For example, for SEP.asking.questions one regular expression that was developed was ^/bwhy/b. In this regular expression, the ^ symbol indicates to only search at the start of the text string, the /b indicates a word boundary, and why indicates searching for the word "why." Another regular expression for SEP.asking.questions was ^\bwhat(?!I had)\b. This regular expression searches for the word "what" but not if "I had" follows. This expression was used to identify student's questions that started with "what," but not to identify when students made statements about their work, such as "What I had was 152 mice." (The keywords and regular expressions for all the SEPs and DCIs are listed in the Appendix.)

Two human coders and the regular expressions algorithm coded 50 student responses for all the SEP and DCI categories. We then used nCoder, an online software for developing, validating, and implementing automated classifiers. The tool allows for testing inter-rater reliability among two human coders and the automated algorithm (Eagan et al., 2017; Shaffer et al., 2015). When we had a disagreement between the algorithm and a human for any coding category, we refined the keywords and regular expressions until we reached an acceptable agreement and rho values using an unused set of 50 student responses. After reaching acceptable agreements for all the code categories, we applied the algorithm to the full dataset.

We applied Epistemic Network Analysis (ENA) (Shaffer, 2017; Shaffer et al., 2016) to the coded data of student responses and operationalized epistemic connections in terms of co-occurrences among the codes. The accumulation of the co-occurrences of codes for each student was represented as a weighted node-link network. The nodes in the networks represent scientific practices and disciplinary ideas and the links represent how often a student linked particular science inquiry practices and core disciplinary ideas. In addition to a network representation, we used ENA to visualize the centroid of each student's network and plotted the centroids in a fixed *x*–*y* axis space determined by the ENA algorithms. In this representation, the centroid of each student's network is projected into a two-dimensional space so that all student networks can be viewed together at one time. Since the position of the centroid is dependent on the prominent nodes in a student network, this representation allows us to identify the nodes that were prominent in students' epistemic connections as they progressed through the unit.

Additionally, we performed a detailed, micro-level analysis of individual students' network representations and their responses to curricular questions. We present a purposeful sample (Palinkas et al., 2015) of two student cases: Alejandro, a Hispanic male, and Jane, an Asian female (all pseudonyms). We identified the two student networks based on the changes in their epistemic connections—one with prominent changes in connections with a Science and Engineering Practice (sep.asking.questions) and another with prominent changes in the connections with a Disciplinary Core Idea (dci.heritability). Our analytical choice of purposeful sampling is justified because we use these cases to illustrate how analysis of network representations with a focus on changes in epistemic connections of individual students helps in characterizing student learning of practices and disciplinary ideas and discuss possibilities for the general validity of the analytical approach and inferences about student learning (Hammer et al., 2018). We identified written responses of the two students that corresponded to their networks in different parts of the unit. We used these responses to interpret students' epistemic connections in terms of how they engaged with specific disciplinary ideas by participating in specific practices and what changes in the connections meant in terms of their learning.

Findings

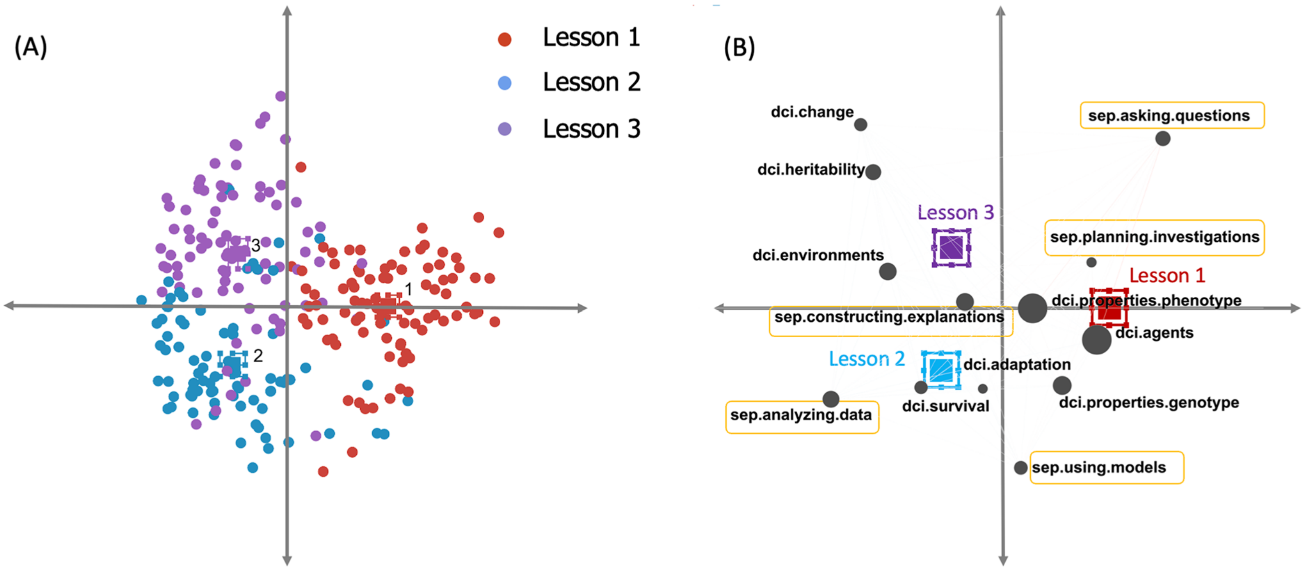
In this section, we first present a broad, macro-level epistemic network analysis of student progression using the centroid representation. We then present a micro-level discourse analysis of student responses in which we focus on two students.

Aggregate Analysis of Student Progression through ESM Lessons

For each lesson, students had statistically distinct clusters of epistemic connections between disciplinary core ideas (DCIs) and science practices (SEPs), as represented by the means (Fig. 3A, Table 4). Thus, students' engagement in SEPs and DCIs in their written responses were significantly different for each lesson (Fig. 3B).

The centroids of student networks for lesson 1 (red dots) are in the same cartesian space where the nodes about the *phenotypic and genotypic properties* of the mice agents and practices of *asking questions* and *planning investigations* are situated. This indicates that in lesson 1, students' epistemic connections were primarily between properties of the agents and practices of asking questions and planning investigations. This aligns with the intended learning goals of lesson 1 about asking research questions about the phenomenon that they can answer using the model. The centroids of student networks in lesson 2 are in the space with practices nodes—*using models*, *planning investigations*, and *analyzing data* and nodes of disciplinary ideas about *survival*, *environments*, and *heritability* (Fig. 3B). This aligns with the learning goals of lesson 2 of using the model to investigate disciplinary ideas related to natural selection by analyzing data. The lesson 3 centroids (purple dots) are in the space where the disciplinary ideas nodes for *change*, *heritability*, and *environment* and practices node

Content courtesy of Springer Nature, terms of use apply. Rights reserved.



Each dot represents centroid of an epistemic network of a student.

Fig. 3 (**A**) Centroids of networks for all students for lesson 1 (red), 2 (blue), and 3 (purple). Average is represented as a square with confidence intervals. Axes represent the first two dimensions of the multidimensional scaling in ENA to maximize variance in the data. (**B**) Mapping of DCI and SEP nodes on the ENA space. The first quadrant (positive *x*, positive *y*) represents asking questions and planning investigations; the second quadrant (negative *x*, positive *y*) represents core natural selection ideas such as heritability, mutation, and environment; the third quadrant (negative *x*, negative *y*) represents analyzing data and other key ideas such as survival and adaptation; and the fourth quadrant (positive *x*, negative *y*) represents using models and knowledge about agents and their properties. The SEP codes are marked with orange squares

Table 4 Comparison between means of centroids of students' epistemic networks for each lesson

	X-axis				*Y*-axis			
	Mean	SD	*P*-value (two-sample *t*-test)	Cohen's *d* (effect size)	Mean	SD	*P*-value (two-sample *t*-test)	Cohen's *d* (effect size)
Lesson 1 (*N* = 87)	0.85	0.44	0.00* (Lesson 2) 0.00* (Lesson 3)	2.82 (Lesson 2) 3.05 (Lesson 3)	0.01	0.43	0.00* (Lesson 2) 0.00* (Lesson 3)	1.12 (Lesson 2) 0.85 (Lesson 3)
Lesson 2 (*N* = 84)	−0.46	0.49	0.00* (Lesson 1) 0.51 (Lesson 3)	2.82 (Lesson 1) 0.10 (Lesson 3)	−0.48	0..45	0.00* (Lesson 1) 0.00* (Lesson 3)	1.12 (Lesson 1) 1.77 (Lesson 3)
Lesson 3 (*N* = 86)	−0.41	0.39	0.00* (Lesson 1) 0.51 (Lesson 2)	3.05 (Lesson 1) 0.10 (Lesson 2)	0.45	0.59	0.00* (Lesson 1) 0.00* (Lesson 2)	0.85 (Lesson 1) 1.77 (Lesson 2)

for *constructing explanations* are situated (Fig. 3B). The co-occurrences in lesson 3 indicate that students constructed explanations about natural selection by primarily discussing change, heritability, and environment, which are fundamental ideas to understand how natural selection works.

Qualitative Analysis of Student Epistemic Connections

We present a qualitative analysis of written responses and epistemic networks of two students to investigate student participation and learning science practices and disciplinary ideas.

Refining a Science Practice in a Disciplinary Context

The unit asked students three times to propose and refine the research questions about natural selection that they wanted to investigate. In lesson 1, after watching the introductory video, Alejandro mentioned two questions that he wanted to investigate: "*i would investigatesf [investigate] how many small pocket mouse [mice] are there in that area. also [,] another question is how long did it take the lave [lava] to cool down.*" Here, Alejandro proposed a question to investigate the number of mice in the New Mexico desert and another to ask about the time it took for the lava to cool down. These open-ended questions indicate Alejandro's curiosity about the contents of the video. His question

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

considered mice population and the environment separately, but not in the context of the unit's focus on population change due to natural selection.

Later in the same lesson, he refined this question: "*Want to investigate how many black mice there would be in a[n] area that had white, and have the background of the floor white on one side and black on the other side.*" Alejandro's second question was about investigating the survival of mice in different environments (sandy vs rocky) based on their fur color. Between question (A) and question (B), Alejandro performed a series of structured but self-driven learning activities about using the ESM model. He answered questions that required playing around with the model parameters to deduce how mice genotype and phenotype are modeled in the ESM. Alejandro's new question included several specific aspects of disciplinary ideas such as environment (color of the background) and mouse phenotypes (color of mice fur). Compared to the first question, Alejandro's second question was more specific, more feasible, and more aligned to the context of ESM.

Alejandro's network in lesson 1 (Fig. 4) shows the prominent co-occurrences in his responses to all the questions in this lesson. The connections between sep.asking.questions and other nodes are highlighted. Three of these connections are evident in his refined question: dci.environments ("the background of the floor white on one side and black on the other side"), dci.phenotype ("black mice"), and sep.planning.investigation ("want to investigate"). However, he did not make a connection in lesson 1 with dci.heritability, which is an important aspect of understanding natural selection as a process that causes changes in populations after several generations.

After conducting computational mini-experiments about the mechanism of inheritance and predation in lesson 2, Alejandro chose to investigate the following question in lesson 3: "*If we have a few [mice] with light fur coat in a population of homozygous dominant (dark fur coat) mice that are moving in a mixed background environment, how will the population change after 200 generation(s).*" This question is about investigating change in the population when light-colored mice are introduced in a population of dark-colored mice. Here, Alejandro framed a question for which he could design an experiment to study two important aspects of natural selection, heritability and change in the environment, using the model. His question considered several aspects that are important to understand natural selection: *change* in the population, *inheritance* of fur coat color, *genotype* (homozygous dominant), and *phenotype* (dark fur coat color) of mice.

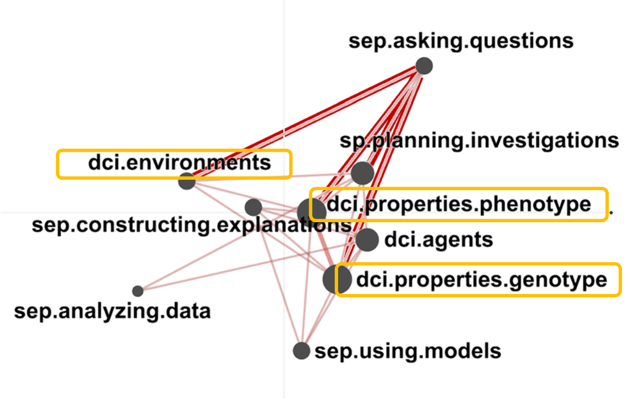


Fig. 4 Alejandro's Network (Lesson 1). Alejandro's SEP node of sep.asking.questions in connected to dci.environment, dci.properties.phenotype and dci.properties.genotype, which shows that his question in lesson 1 was about these three aspects of DCIs

In lesson 2, Alejandro connected an important aspect of disciplinary ideas, namely, the *heritability of a trait* to *asking questions* (Fig. 5A). Alejandro's responses in lesson 2 illustrate how he made these connections. After he conducted an experiment to see how a population of mice with a specific initial genotype combination changed, he noted "*after generations past there was only white mice, totals mice was 489 white mice 242 aa male and 247 aa female.*" This shows how Alejandro described the change in the mouse population concerning the inheritance of the fur coat color in his response after performing a computational experiment using the ESM.

Understanding natural selection as a process that happens over generations because of changes in the environment and mutations in inheritable traits requires an understanding of both heritability *and* environment. In lesson 1, Alejandro made connections with dci.environment, and in lesson 2, he made connections with dci.heritability. Alejandro considered these two aspects together in lesson 3 (Fig. 5B), in which his question "*If we have a few [mice] with light fur coat in a population of homozygous dominant (dark fur coat) mice that are moving in a mixed background environment, how will the population change after 200 generation(s)*" asks about population changes in a mixed environment (dci.environment) after several generations (dci.heritability). The lessons in the unit were designed to provide opportunities for students to revise their research questions. Analysis of Alejando's epistemic connections shows that making connections with central aspects of disciplinary ideas (the idea of heritability in the case of natural selection) in different parts of the unit is likely to have helped him in refining his research question. His refined question was more specific and included the core aspects of the phenomenon related to natural selection.

The states of ESMs corresponding to Alejandro's research question (Fig. 6) indicate how an experimental investigation using the ESM designed to address Alejandro's question can potentially result in identifying changes in a population because of natural selection. Alejandro's written responses to other questions in the unit show that he performed such an

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

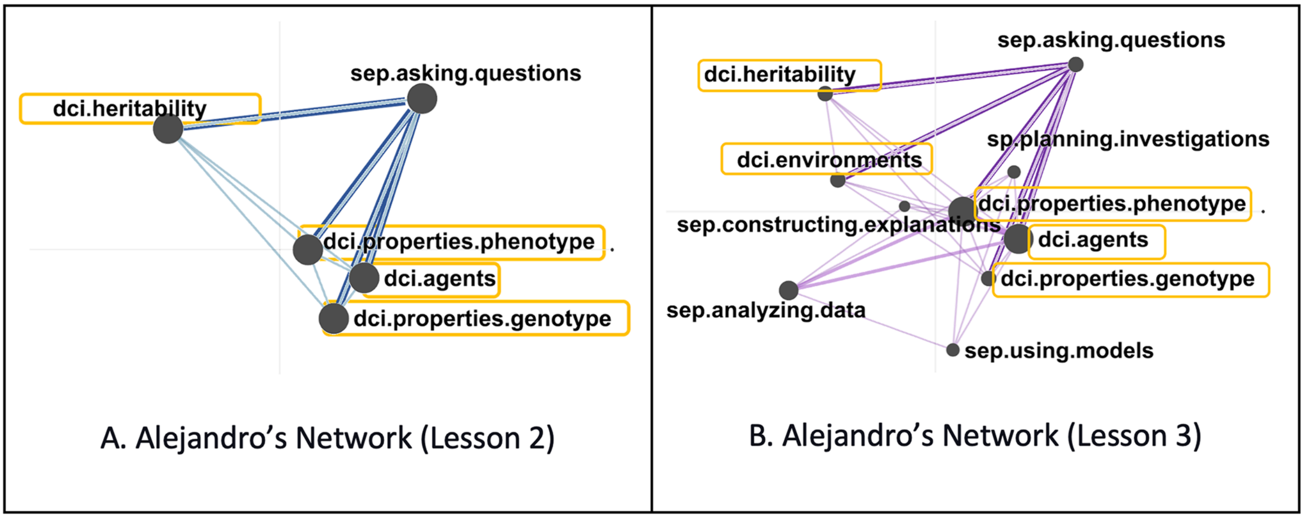


Fig. 5 Alejandro's epistemic networks in lesson 2 and lesson 3. (**A**) Alejando's epistemic network in lesson 2 has four DCI nodes connected to his SEP node, sep.asking.questions, but it is missing a node dci.environment, which was there in his lesson 1 network (Fig. 4). (**B**) In Alejandro's network in lesson 3, all the five DCI nodes (from his lesson 1 and lesson 2 networks) are connected to sep.asking.questions node, indicating that he has incorporated all those dci aspects while asking a question in lesson 3

investigation in lesson 2 using the ESM. This analysis shows how Alejandro refined the practice of asking questions as he participated in many other practices in connection with disciplinary ideas and did not learn *asking questions* as an isolated research skill.

Refining Disciplinary Ideas by Engaging in Science Practices

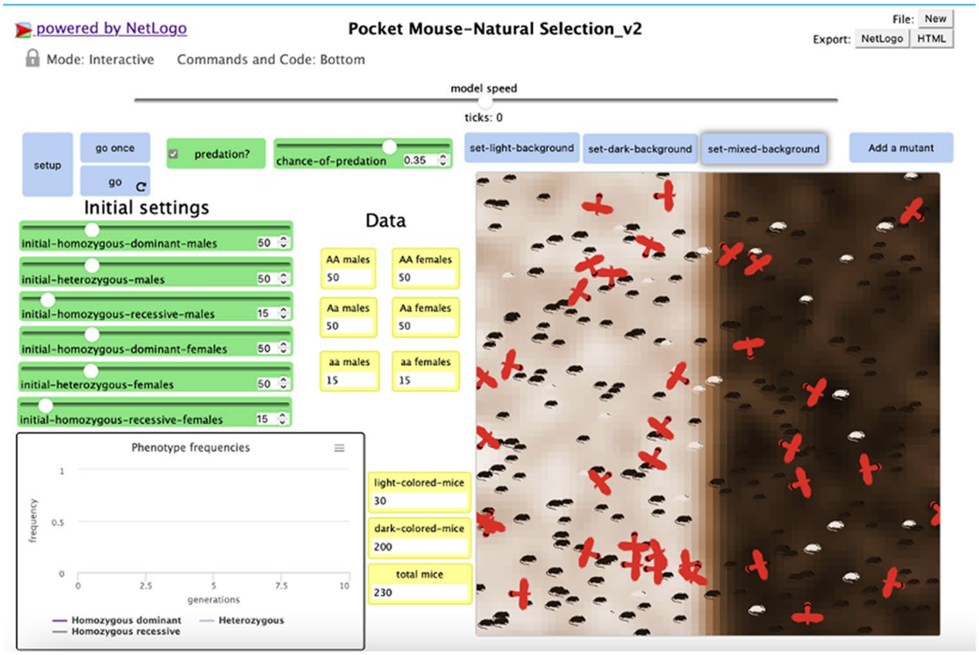
A question in the unit asked Jane to predict the inheritance in a population of homozygous recessive (light-fur-colored) mice. Jane's prediction was that she would have ***mostly*** light-colored mice. However, after performing a mini-experiment using the ESM, she refined her idea because she noticed something different. She responded: "*My population is **all** light pocket mice when I run an experiment after 47 generations.*" Jane's response shows that she ran the experiment for several generations to establish that "all" the population remained light colored, which allowed her to learn how light fur color inheritance works in this ESM.

This observation possibly helped Jane in designing, conducting, and interpreting another mini-experiment to conclude an important insight regarding how Mendelian inheritance worked in the model. She wrote, "*When both parents are homozygous dominant, ALL of the babies are all dark. When both parents are heterozygous, the babies could be dark or light.*" Jane mentioned two initial experimental setups with different initial population genotypes and highlighted the outcome differences. She used all-caps for "ALL" to highlight an aspect that she found crucial in her results: Homozygous dominant initial genotypes will result in all dark fur offspring, while heterozygous genotypes may result in dark or light offspring.

Network representation of Jane's epistemic connections (Fig. 7) provides insights into how the connections she made helped her overall learning of natural selection and her participation in the science practices. In lesson 2, Jane made a connection between sep.using.models and dci.heritability. Both these nodes were absent in Jane's epistemic network in lesson 1. in lesson 3, the node dci.heritability is connected to dci.change, dci.environment, dci.properties.phenotype, dci.agents, and sep.analyzing.data in Jane's network. Her experiment design in lesson 3 reflects what she learned about dci.heritability and dci.properties.genotype. She wrote: "*I will make the homozygous [dominant] males/females to 0, heterozygous males/females to 0, homozygous recessive males to 150 and homozygous recessive females to 100. I would use a light background with 0.1 chance of predation. I would have 3 mutant[s] who are dark mice. For every trial, I would add 2 mutants and add 0.05 chance of predation per trial.*" Jane's answer included a clear mention of the genotype of her initial population. In addition, she mentioned in her response to the next question that she ran all the trials for 270 generations, which indicates that she tracked the inheritance of the fur color over many generations.

This analysis of Jane's participation and co-occurrences of SEPs and DCIs illustrates how Jane corrected her understanding of how the inheritance of fur coat color worked for rock pocket mice. Thinking about how heritability worked across many generations while considering phenotypes of mice agents in the model is likely to have helped Jane

Content courtesy of Springer Nature, terms of use apply. Rights reserved.



ticks: 0

☑ predation?

chance-of-predation 0.35

initial-homozygous-dominant-males 50
initial-heterozygous-males 50
initial-homozygous-recessive-males 15
initial-homozygous-dominant-females 50
initial-heterozygous-females 50
initial-homozygous-recessive-females 15

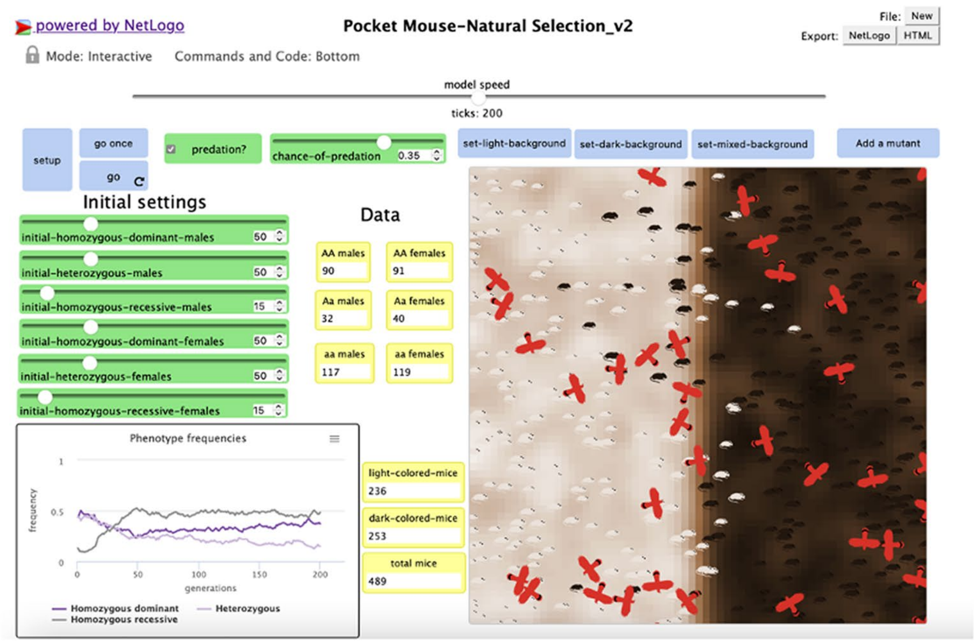
AA males	AA females
50	50

Aa males	Aa females
50	50

aa males	aa females
15	15

light-colored-mice 30
dark-colored-mice 200
total mice 230

A. The ESM setup based in the initial conditions mentioned in Alejandro's research question



ticks: 200

☑ predation?

chance-of-predation 0.35

initial-homozygous-dominant-males 50
initial-heterozygous-males 50
initial-homozygous-recessive-males 15
initial-homozygous-dominant-females 50
initial-heterozygous-females 50
initial-homozygous-recessive-females 15

AA males	AA females
90	91

Aa males	Aa females
32	40

aa males	aa females
117	119

light-colored-mice 236
dark-colored-mice 253
total mice 489

B. An ESM state after running the model for 200 generations

Fig. 6 Initial and final states of the ESM corresponding to Alejandro's research question. (**A**) The mice with dark and light fur colors are uniformly distributed irrespective of the background color. (**B**) After 200 generations, the mice with dark fur color are predominantly present in the dark background areas and the same is true for the mice with light fur color

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

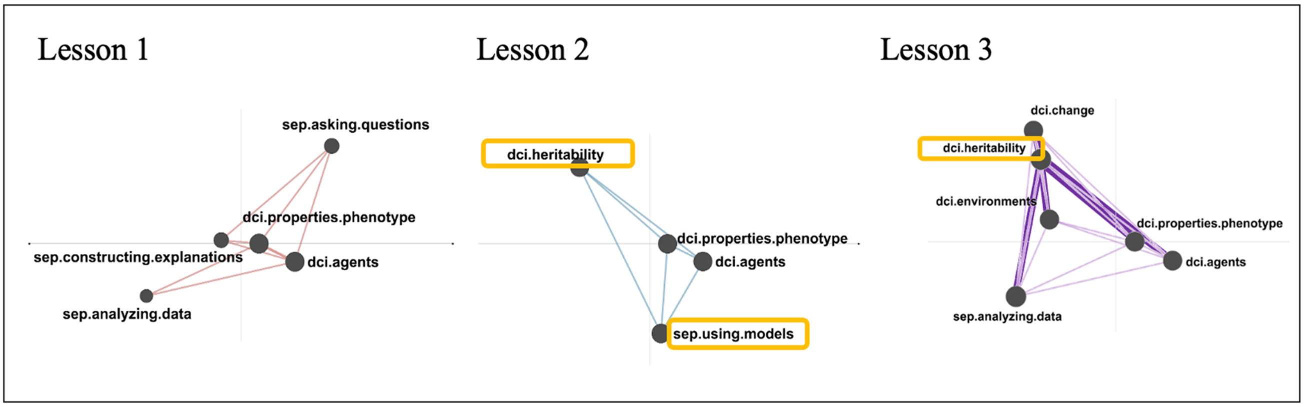


Fig. 7 Jane's epistemic networks in three lessons. In lesson 2, Jane made a connection between sep.using.models and dci.heritability, the nodes that were absent in her network in lesson 1. In lesson 3, the node dci.heritability is connected to dci.change, dci.environment, dci.properties.phenotype, dci.agents, and sep.analyzing.data which shows how new DCI and SEP nodes got incorporated in Jane's network

develop a nuanced understanding of the dominant expression of the dark fur trait.

Discussion

As new curricular design approaches are being advocated for and developed to support students' integrated learning of science practices and disciplinary ideas (Berland et al., 2016; Linn et al., 2016; Manz et al., 2020; NGSS Lead States, 2013; Reiser et al., 2017; Weintrop et al., 2016), it is important to develop analytical strategies to study student participation and learning. Integrating Computational Thinking (CT) through model-based inquiry curricula is one such approach that supports students in participating in contemporary ways of doing science (Weintrop et al., 2016). In this paper, we studied student learning in a biology unit about natural selection that was integrated with CT learning activities using computational models designed as Emergent Systems Microworlds (ESMs) (Dabholkar & Wilnesky, 2024).

We developed an ENA-based analytical approach to visualize and measure students' participation in science practices in a model-based inquiry unit designed to support students' integrated learning of science practices and disciplinary ideas. We conceptualize "epistemic connections" as connections among knowledge-building practices (SEPs in a science curriculum) and disciplinary core ideas (DCIs) that are being used to make sense of a phenomenon. We argue that student participation and learning in curricula designed to knowledge building activities (Barzilai & Chinn, 2018; Berland et al., 2016) can be studied by modeling and measuring students' epistemic connections. Our research question investigates student learning of SEPs and DCIs by analyzing their epistemic connections.

We use ENA to model and analyze integrated learning of SEPs and DCIs by characterizing epistemic connections in a CT-integrated biology curriculum to investigate student learning. The findings show how ENA can be a useful methodological approach to operationalizing epistemic connections for analytical purposes. With ENA, one can visualize the student engagement with SEPs and DCIs at an aggregate level and analyze which SEPs and DCIs that most students participated in during different parts of the unit. An individual-level analysis of students' epistemic connections using ENA, along with qualitative analysis of their written responses, can be useful to investigate how participation in specific SEPs can support learning of specific DCIs and vice versa.

The aggregate-level ENA analysis of centroids of individual student networks in the Natural Selection unit demonstrated statistically different engagement in sets of practices and disciplinary ideas during different lessons. These changes in the networks at the aggregate level revealed temporal shifts in participation in practices as students moved from asking questions to constructing explanations about system-level emergent patterns, such as changes in a population because of the introduction of a mutant phenotype. This analysis also showed that students made epistemic connections with more core aspects of natural selection as they progressed through the curricular unit, moving from genotypic and phenotypic properties of individuals to heritability, the impact of environmental factors on survival, and changes in a population. As core ideas and new practices became prominent nodes in student networks, they also contained and were connected to earlier nodes, which shows that as students integrated new ideas and practices into their works, they continued using the earlier ideas and practices. To make sense of changes in a population because of natural selection, connections of multiple ideas such as genotypic and

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

phenotypic variation, heritability of a trait, and environmental factors influencing the chances of survival and reproduction are important (e.g., Nehm & Schonfeld, 2008). Analysis of how students integrate these ideas as they participate in science practices can be helpful in understanding their sense-making of the phenomena as well as specific aspects of disciplinary ideas (Gouvea, 2023). Our work shows how ENA-based analysis of epistemic connections is useful for studying how students engage in different practices to investigate specific aspects of the phenomenon.

We further used ENA to conduct in-depth qualitative analyses of the changes in two students' epistemic connections in the unit to investigate what the temporal shifts entailed in terms of learning DCIs and SEPs. This analysis of students' epistemic connections revealed how different kinds of participation in SEPs and investigations of various aspects of DCIs supported learning of SEPs and DCIs in a reciprocal manner. Alejandro's participation in the science practice of asking questions increased in complexity as he made more epistemic connections with different disciplinary ideas. Jane's engagement with a disciplinary idea, heritability, increased as she made epistemic connections with specific science practices. Such analysis of epistemic connections is helpful in understanding whether and how students participated in integrated learning of practices and ideas (Berland et al., 2016; Linn et al., 2016; Reiser et al., 2017). This analysis also demonstrates how a CT-integrated model-based inquiry unit supported student learning of practices and ideas in a sequential yet integrated manner through iterative refinement. Over time, students participated in more refined and sophisticated aspects of practices and used core aspects of the disciplinary ideas to construct.

Conclusions and Implications

In this paper, we presented an ENA-based novel analytical approach to investigate students' *epistemic connections* in an NGSS-aligned CT-integrated biology curriculum about natural selection. Our characterization of students' epistemic connections using ENA at aggregate and individual levels showed how ENA could be useful for investigating the changes in students' epistemic connections in terms of their refined participation in practices and understanding nuanced aspects of disciplinary ideas. Such analysis of epistemic connections can help study specific aspects of student participation and learning in NGSS-aligned science curricular units as we demonstrated in this paper. For example, an aspect of student participation we studied in this paper is the refinement of practices and ideas by creating multiple opportunities for students during the unit and checking if and how iterative participation helps in the refinement of ideas in a model-based curriculum.

Additionally, the analysis of aggregate networks (Fig. 3) and individual student networks (Figs. 4 and 5) can be helpful from both the perspective of analyzing the usefulness of a curriculum and classroom instruction. Understanding which practices and ideas students engage with and learn about in different lessons is useful in modifying the lessons to enhance participation in the ideas and practices that are not present as expected. For example, in lesson 1, connections with dci.heritability are not as prominent as connections with dci.genotype or dci.phenotype. The low engagement with dci.heritability in lesson 1 indicates that lesson 1 did not engage students in thinking about the transmission of genes and traits across generations. Access to aggregate students' networks of practices and ideas in a teacher dashboard can help a teacher highlight those ideas in class discussions. Access to individual student networks can help teachers support students in participating in and learning about practices and ideas that are less prominent in their networks. For example, while Alejandro refined his asking questions practice, his network for lesson 3 does not have a connection between sep.using.models and dci.heritability. This lack of epistemic connection indicates that he may not have considered how the idea of heritability is implemented in the model. A teacher could ask Alejandro a question to verify if he has understood the idea of heritability and its connection to natural selection.

Limitations and Future Directions

This work also provides insights into how ENA as a methodology and studying student learning by characterizing their epistemic connections in NGSS-aligned model-based inquiry units can be improved further. Regarding the use of ENA in science education research, findings in this paper suggest that using ENA to analyze students' epistemic connections can be a powerful way to characterize students' integrated learning in a science classroom. However, there were limitations in using ENA to operationalize epistemic connections. In this study, we demonstrated how it is possible to infer student participation in science practices based on their written responses in CT-integrated model-based learning environments. However, such inferences may not be feasible in all kinds of learning environments, and we might have missed out on modeling student participation in some practices they did not write about. This limitation can be overcome by leveraging multimodal ENA approaches that have been demonstrated to be effective in modeling participation in specific practices, such as surgical residents doing simulated hernia repair (Lund et al., 2017; Ruis et al., 2018). Analysis of logs of student actions using a NetLogo model (Levy & Wilensky, 2005; McBride et al., 2016) can be helpful in future work to refine this analytical approach further.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

In addition, ENA networks are mathematically constructed by counting co-occurrences within a given temporal segmentation. However, after the construction is complete, the interpretation of the links is left up to the researcher to go back to the discourse data and close the interpretative loop (Arastoopour Irgens & Eagan, 2022). For example, a connection between sep.using.models and dci.environment could mean that a student used the model to investigate the background color of the environment or they investigated the impact of predators in the environment. The interpretation of the relationship between SEPs and DCIs is critical to further conceptualizing epistemic connections. Thus, leveraging computational assistance in interpreting the links themselves and helping the researcher close the interpretive loop would be helpful. For example, ENA could integrate natural language processing tools to help the researcher identify the context within which co-occurrences of each pair of codes occur and draw the researcher's attention to particular configurations of co-occurrences, which would aid in the interpretation of the different forms of epistemic connections that students make.

The Next Generation Science Standards (NGSS) advocates three-dimensional learning, which involves learning Disciplinary Core Ideas (DCIs), Science and Engineering Practices (SEPs), and Cross Cutting Concepts (CCCs). Since we designed the curricular unit to integrate Computational Thinking through SEPs, our analysis focused on SEPs and DCIs. The future work will extend this approach to model NGSS-aligned three-dimensional learning to incorporate CCCs related to Natural Selection, such as Scale, Proportion, and Quantity; Systems and Systems Models; and Stability and Change (NGSS Lead States, 2013). This design and analytical shift would require changing the curricular activities and question prompts to draw students' attention to these CCCs and coding student responses for these CCCs in addition to SEPs and DCIs to study three-dimensional learning.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10956-025-10242-z.

Acknowledgements We are grateful to Mike Horn, Hillary Swanson, Teresa Granito, Kevin Hall, Connor Bain for their contributions to the design of the unit. We also express our sincere gratitude towards the teacher and students who participated in this research work.

Author Contribution Sugat Dabholkar conceived the study, designed curricular materials, conducted analysis, and wrote the paper. Golnaz Arastoopour Irgens conducted analysis and contributed to writing the paper. Uri Wilensky conceived the study and contributed to the design of curricular materials and writing the paper. We are grateful to Mike Horn, Hillary Swanson, Teresa Granito, Kevin Hall, Connor Bain for their contributions to the design of the unit. We also express our sincere gratitude towards the teacher and students who participated in this research work.

Funding This work is part of a larger CT-STEM Project which is made possible through generous support from the National Science Foundation (grants CNS-1138461, CNS-1441041, DRL-1020101, DRL-1640201, and DRL-1842374) and the Spencer Foundation (Award #201600069). The work presented in the paper was funded through National Science Foundation grant (DRL-1842374).

Data Availability The curricular materials developed for the project are publicly available.

Declarations

Ethics Approval The ethical approval to conduct this study was obtained from Northwestern University's Institutional Review Board. We obtained informed consents form the parents of the students and assents from the students to participate in the study and use the audio, video, and written responses data of their participation for published research.

Competing Interests All the authors declare no conflicts of interest.

References

Abd-El-Khalick, F., BouJaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., Niaz, M., Treagust, D., & Tuan, H. (2004). Inquiry in science education: International perspectives. *Science Education,88*(3), 397–419. https://doi.org/10.1002/sce.10118

Arastoopour, G., Chesler, N. C., & Shaffer, D. W. (2014). Epistemic persistence: A simulation-based approach to increasing participation of women in engineering. *Journal of Women and Minorities in Science and Engineering,20*(3), 211–234. https://doi.org/10.1615/JWomenMinorScienEng.2014007317

Arastoopour Irgens, G., Shaffer, D. W., Swiecki, Z., Ruis, A. R., & Chesler, N. C. (2015). Teaching and assessing engineering design thinking with virtual internships and epistemic network analysis. *International Journal of Engineering Education,32*(3), 1492–1501.

Arastoopour Irgens, G., Dabholkar, S., Bain, C., Woods, P., Hall, K., Swanson, H., Horn, M., & Wilensky, U. (2020). Modeling and measuring high school students' computational thinking practices in science. *Journal of Science Education and Technology, 29*(1), 137–161. https://doi.org/10.1007/s10956-020-09811-1

Arastoopour Irgens, G., & Eagan, B. (2022). The foundations and fundamentals of quantitative ethnography. *International Conference on Quantitative Ethnography*, 3–16.

Arık, M., & Topçu, M. S. (2022). Computational thinking integration into science classrooms: Example of digestive system. *Journal of Science Education and Technology, 31*(1), 99–115. https://doi.org/10.1007/s10956-021-09934-z

Bagley, E. A., & Shaffer, D. W. (2015). Comparing virtual and face-to-face mentoring in an epistemic game. *Journal of Computer Assisted Learning,31*(6), 606. https://doi.org/10.1111/jcal.12092

Barzilai, S., & Chinn, C. A. (2018). On the goals of epistemic education: Promoting apt epistemic performance. *Journal of the Learning Sciences, 27*(3), 353–389.

Berland, L. K., Schwarz, C. V., Krist, C., Kenyon, L., Lo, A. S., & Reiser, B. J. (2016). Epistemologies in practice: Making scientific practices meaningful for students. *Journal of Research in Science Teaching, 53*(7), 1082–1112. https://doi.org/10.1002/tea.21257

Dabholkar, S., Anton, G., & Wilensky, U. (2018a). GenEvo - An emergent systems microworld for model-based scientific inquiry in the

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

context of genetics and evolution. In Kay, J. & Luckin, R. (Eds.), Rethinking Learning in the Digital Age: Making the Learning Sciences Count, 13th International Conference of the Learning Sciences (ICLS) 2018, Volume 3. International Society of the Learning Sciences.

Dabholkar, S., Woods, P., Bain C., Hall, K., Horn M., & Wilensky U. (2018b). Evolution of populations. [curricular unit]. Center for Connected Learning and Computer-Based Modeling, Northwestern University. https://ct-stem.northwestern.edu/curriculum/preview/681/

Dabholkar, S., & Wilensky, U. (2020). NetLogo natural selection - Camouflage model. Center for Connected Learning and Computer-Based Modeling, Northwestern University. https://ccl.northwestern.edu/netlogo/models/NaturalSelection-Camouflage

Dabholkar, S., & Wilensky, U. (2024). Designing computational models as emergent systems microworlds to support epistemically agentive learning of emergent biological phenomena. *International Journal of Science Education, 47*(8), 1028–1050. https://doi.org/10.1080/09500693.2024.2356228

Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education, 32*(1), 268–291. https://doi.org/10.3102/0091732X07309371

Eagan, B. R., Rogers, B., Serlin, R., Ruis, A. R., Arastoopour Irgens, G., & Shaffer, D. W. (2017). Can we rely on IRR? Testing the assumptions of inter-rater reliability. *International Conference on Computer Supported Collaborative Learning*.

Gouvea, J. (2023). Processing misconceptions: Dynamic systems perspectives on thinking and learning. *Frontiers in Education,8*, Article 1215361.

Grover, S., & Pea, R. (2013). Computational thinking in K-12: A review of the state of the field. *Educational Researcher*. https://doi.org/10.3102/0013189X12463051

Hammer, D., Gouvea, J., & Watkins, J. (2018). Idiosyncratic cases and hopes for general validity: What education research might learn from ecology. *Journal for the Study of Education and Development, 41*(4), 625–673.

Jacobson, M. J., & Wilensky, U. (2006). Complex systems in education: Scientific and educational importance and implications for the learning sciences. *Journal of the Learning Sciences*. https://doi.org/10.1207/s15327809jls1501_4

Klopfer, E., Yoon, S., & Perry, J. (2005). Using palm technology in participatory simulations of complex systems: A new take on ubiquitous and accessible mobile computing. *Journal of Science Education and Technology, 14*(3), 285–297. https://doi.org/10.1007/s10956-005-7194-0

Lee, I., Grover, S., Martin, F., Pillai, S., & Malyn-Smith, J. (2020). Computational thinking from a disciplinary perspective: Integrating computational thinking in K-12 science, technology, engineering, and mathematics education. *Journal of Science Education and Technology, 29*(1), 1–8.

Lehrer, R., & Schauble, L. (2012). Cultivating model-based reasoning in science education. *In the Cambridge Handbook of the Learning Sciences*. https://doi.org/10.1017/cbo9780511816833.023

Levy, S. T., & Wilensky, U. (2005). An analysis of student patterns of exploration with NetLogo models embedded in the connected chemistry environment. *Proceedings of the Annual Meeting of the American Educational Research Association*.

Linn, M. C., Gerard, L., Matuk, C., & McElhaney, K. W. (2016). Science education: From separation to integration. *Review of Research in Education, 40*(1), 529–587.

Lund, K., Quignard, M., & Shaffer, D. W. (2017). Gaining insight by transforming between temporal representations of human interaction. *Journal of Learning Analytics, 4*(3), 102–122.

Manz, E., Lehrer, R., & Schauble, L. (2020). Rethinking the classroom science investigation. *Journal of Research in Science Teaching, 57*(7), 1148–1174.

McBride, E., Vitale, J. M., Gogel, H., Martinez, M. M., Pardos, Z., & Linn, M. C. (2016). Predicting student learning using log data from interactive simulations on climate change. *Proceedings of the Third (2016) ACM Conference on Learning@ Scale*, 185–188.

Miles, M., Huberman, A., & Saldana, J. (2019). Qualitative data analysis: A methods sourcebook. SAGE Publications, Inc.

Nehm, R. H., & Schonfeld, I. S. (2008). Measuring knowledge of natural selection: A comparison of the CINS, an open-response instrument, and an oral interview. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 45*(10), 1131–1160.

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

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health,42*(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books.

Reiser, B. J. (2014). *Designing coherent storylines aligned with NGSS for the K-12 classroom*. National Science Education Leadership Association Meeting.

Reiser, B. J., Novak, M., & McGill, T. A. W. (2017). Coherence from the students' perspective: Why the vision of the framework for K-12 science requires more than simply "combining" three dimensions of science learning. *Board on Science Education Workshop "Instructional Materials for the Next Generation Science Standards*.

Ruis, A. R., Rosser, A. A., Quandt-Walle, C., Nathwani, J. N., Shaffer, D. W., & Pugh, C. M. (2018). The hands and head of a surgeon: Modeling operative competency with multimodal epistemic network analysis. *The American Journal of Surgery, 216*(5), 835–840.

Schwarz, C. V, Passmore, C., & Reiser, B. J. (2017). Moving beyond "knowing about" science to making sense of the world. *Helping Students Make Sense of the World Using next Generation Science and Engineering Practices*, 3–21.

Shaffer, D. W. (2017). Quantitative ethnography. Lulu.com.

Shaffer, D. W., Borden, F., Srinivasan, A., Saucerman, J., Arastoopour, G., Collier, W., Ruis, A. R., & Frank, K. A. (2015). The nCoder: A technique for improving the utility of inter-rater reliability statistics. *Games and Professional Simulations Technical Report, 1*.

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics,3*(3), 9–45. https://doi.org/10.18608/jla.2016.33.3

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Frank, K., Rupp, A. A., & Mislevy, R. (2009). Epistemic network analysis: A prototype for 21st-century assessment of learning. *International Journal of Learning and Media, 1*(2), 33–53. https://doi.org/10.1162/ijlm.2009.0013

Shaffer, D. W., spsampsps Ruis, A. R. (2021). How we code. *Advances in Quantitative Ethnography: Second International Conference, ICQE 2020, Malibu, CA, USA, February 1–3, 2021, Proceedings 2*, 62–77.

Siebert-Evenstone, A., & Shaffer, D. W. (2019). Cause and because: Using epistemic network analysis to model causality in the next generation science standards. *Advances in Quantitative Ethnography: First International Conference, ICQE 2019, Madison, WI, USA, October 20–22, 2019, Proceedings 1*, 223–233.

Swanson, H., Anton, G., Bain, C., Horn, M., & Wilensky, U. (2019). Introducing and assessing computational thinking in the secondary science classroom. *Computational Thinking Education*. https://doi.org/10.1007/978-981-13-6528-7_7

Waterman, K. P., Goldsmith, L., & Pasquale, M. (2020). Integrating computational thinking into elementary science curriculum: An examination of activities that support students' computational thinking in the service of disciplinary learning. *Journal of Science*

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Education and Technology, 29(1), 53–64. https://doi.org/10.1007/s10956-019-09801-y

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology,25*(1), 127. https://doi.org/10.1007/s10956-015-9581-5

Wilensky, U. (1999). NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo/

Wilensky, U. (2001). Modeling nature's emergent patterns with multi-agent languages. *Proceedings of EuroLogo*.

Wilensky, U. (2003). Statistical mechanics for secondary school: The GasLab multi-agent modeling toolkit. *International Journal of Computers for Mathematical Learning,8*(1), 1–41. https://doi.org/10.1023/A:1025651502936

Wilensky, U., Brady, C. E., & Horn, M. S. (2014). Fostering computational literacy in science classrooms. *Communications of the ACM*. https://doi.org/10.1145/2633031

Wilensky, U., & Papert, S. (2010). Restructurations: Reformulations of knowledge disciplines through new representational forms. *Constructionism*.

Wilensky, U., & Jacobson, M. J. (2014). Complex systems and the learning sciences. In *The Cambridge handbook of the learning sciences* (2nd ed., pp. 319–338). Cambridge University Press.

Wilensky, U., & Rand, W. (2015). *An introduction to agent-based modeling: Modeling natural, social, and engineered complex systems with NetLogo*. MIT Press.

Wilkerson, M. H., Shareff, R., Laina, V., & Gravel, B. (2018). Epistemic gameplay and discovery in computational model-based inquiry activities. *Instructional Science, 46*(1), 35–60.

Wilkerson-Jerde, M., & Wilensky, U. (2010). Restructuring change, interpreting changes: The deltatick modeling and analysis toolkit. *Proceedings of Constructionism*.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Content courtesy of Springer Nature, terms of use apply. Rights reserved.

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com