NORTHWESTERN UNIVERSITY

Restructurating Science Learning with Emergent Systems Microworld (ESM)-based Learning Environments

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Learning Sciences

By

Sugat Dabholkar

EVANSTON, ILLINOIS

December 2021

Abstract

This dissertation is about designing learning environments that foreground students' epistemic agency as they *do science* in a science classroom. I build on the prior work that combines two important strands in learning sciences - agent-based modeling of complex systems and constructionism to design learning environments. I call such learning environments Emergent Systems Microworld (ESM)-based learning environments. ESMs are computational tools (a model or a collection of models) that are designed as constructionist microworlds. ESMs use agent-based representations to model emergent phenomena. Because an ESM has consistent underlying rules regarding agent properties and behaviors, it makes ESM an excellent experimental system for learners to investigate various aspects of the modeled phenomena. ESM-based curricula are designed for students to collectively construct knowledge of these phenomena by engaging in scientific inquiry practices. I use Wilensky and Papert's theory of restructurations to investigate the potential of agent-based representations in constructionist learning environments to restructurate science learning. I envision such restructuration of science similar to the restructuration of arithmetic that happened because of the invention and use of Hindu-Arabic numerals. In my dissertation, I conducted three studies to analyze how properties of restructurations change how students engage in science practices and teachers design technology-enhanced science curricula. The first study focuses on students' expansive learning from the perspective of their epistemic engagement in constructing knowledge of disciplinary ideas. The second study is about the reciprocity between a disciplinary context and science practices for students to use one as a generative space to learn the other. In the third study, I analyze a three-year-long partnership with a teacher that was about co-designing science

curricula integrated with Computational Thinking. I investigate how using ESMs for co-designing such curricula facilitated teacher involvement in the co-design process, changed her teaching practices, and increased the richness of the curricula. These three studies demonstrate the effectiveness of the ESM approach to design and co-design curricula to engage students in scientific inquiry practices in general and computational thinking practices in particular. They also contribute to understanding how properties of agent-based restructurations combined with constructionist design principles facilitate students' agentic participation in the process of collective knowledge construction.

Acknowledgements

A commonly used proverb that embodies the spirit of several African cultures goes “it takes a village to raise a child”. Taking inspiration from that proverb, I would like to say that it took a global village for my dissertation work to see the light of the day. I want to thank everyone who has contributed to making this dissertation what it is today.

I want to start by acknowledging the land and labor of people, which made it possible for me to have Northwestern University as my academic home during the dissertation. The Northwestern campus sits on the traditional homelands of the people of the Council of Three Fires, the Ojibwe, Potawatomi, and Odawa as well as the Menominee, Miami, and Ho-Chunk nations. I also respectfully acknowledge that the country, the United States, in which I conducted most of my dissertation work was built with exploited labor of enslaved people, primarily of African descent.

I would like to thank all the students and teachers, in the United States and in India, who I had the privilege to work with to conduct my dissertation research. My design partnerships with several school teachers including Teresa Granito, Lauren Bauknecht, Anagh Purandare, Karen Johnson, Susan Juhl, Tiffany Verbeke, Shirish Sant, Vidya Pillai, Santosh Bhandari, Sandip Lonkar, Ashok Swami, Tanali Bhattacharyay, Raadhika Oka, Manjusha Borikar, and Mrudula Adavadkar helped me grow academically as a Learning Sciences researcher.

I want to express my sincere gratitude to my advisor Uri Wilensky, who provided me design tools (NetLogo), theoretical tools, and guidance during my dissertation. I want to thank my dissertation committee members, Mike Horn and Joshua Danish for their thorough

comments, guidance, and encouragement through all the important milestones during the dissertation.

Members of the CCL lab and a nested group, CT-STEM, have contributed immensely to shaping my intellectual journey during my dissertation. Thank you, Gol, for teaching me ENA and mentoring me through the epistemic connections work. Thank you, Hillary, for the philosophical discussions about science, science education, and epistemologies. Thank you, Kevin, for our design partnerships and for thinking about the idea of a design partnership. Thank you, Sally, for the writing group which helped me keep me going through the pandemic-affected difficult periods of writing my dissertation. Thank you, Connor, for being a great CT-STEM, and office partner throughout the dissertation and helping with various parts of the dissertation, including coding data and early NetLogo models. Thank you, Gabby and Umit, for helping with GenEvo course design and data collection. Thank you, Mandy and Gabby, for helping with coding data for in the third study of the dissertation. Thank you, members of development team, Robert, Jason and Jeremy, for amazing NetLogo and NetTango updates. The long walks that I shared with many of you helped me develop, refine, and materialize many of my designs and analytical ideas.

I also want to thank members of several other Learning Sciences and other communities that I have been part of in the last few years – my Ph.D. cohort at Northwestern, Critical Contexts, GST, MSD, YOM, SALSRC, and LSGSC. My membership in these groups has helped me develop my identity as a learning sciences researcher. Several members of the School of Education and Social Policy at Northwestern, who I could take courses from and interact with helped me develop my academic scholarship. Especially the courses that were taught by Doug

Medin, Shirin Vossoughi, Kalonji Nzinga, and Alena Berry helped me expand my academic vision and develop a deeper relationship with my academic scholarship and writing.

English is not my first language and that still creates some challenges, especially with grammar! You might have spotted a few grammatical errors already! I want to thank the CCL lab OAs – Dana, Paul, RJ, Yaurie, and others for copy editing my research papers and dissertation chapters.

Several of my teachers from my academic life before the Learning Sciences program have helped me develop interests and insights into academic research areas such as evolutionary biology, and science education. I want to thank Milind Watve, Amitabh Joshi, Sutirth Dey, K. P. Mohanan, and many others for being great influences in my academic life.

This dissertation would have happened without the support from family - my parents, Chitra Dabholkar and Prasanna Dabholkar, and my wife, Saudamini Shevade. I thank all three of you for your emotional support and for being patient with me. My brother, Sujan, sister-in-law, Adrija, and nephews, Miro-Kian, thank you for being awesome; all our FaceTime calls reenergized me to keep going. Thank you, Mitawa and Muktaaatya, for our ongoing conversations about so many things. I also want to thank my grandfathers, who I fondly called Kuku-aajoba and Tapu-aajoba, for inspiring me to enter the educational research world.

This has become a long acknowledgment. I still don’t feel I have thanked everyone, and I have thanked enough. I truly want to thank all the amazing people in my life, past and present, who are helping me grow as an academic scholar.

I dedicate this dissertation to late Dr. Narendra Dabholkar, my granduncle, whose presence in my life inspired me to pursue science and science education.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Summary: My dissertation is about designing learning environments that reimagine *what* students learn in science classrooms and how they learn it. There has been an increased emphasis on engaging students in scientific inquiry practices to construct knowledge as opposed to presenting knowledge as established ideas in a disciplinary domain. In my dissertation, I present an approach to design for learning environments which are embedded with computational tools which I call Emergent Systems Microworlds (ESMs). ESM design approach is based on restructuration theory (Wilensky, 2020; Wilensky & Papert, 2010), which argues for the design and study of new representational forms for learning and expressing knowledge. I present three studies that focus on how the combination of agent-based modeling and constructionism in the ESM design framework supports student engagement in scientific practices and their learning of disciplinary ideas. The first study investigates how an ESM and an ESM-based curriculum mediate students' epistemically expansive learning in a science classroom. The second study focuses on how students' epistemic connections among practices and disciplinary ideas mature as they engage in an ESM-based curriculum. Lastly, in the third study, I present a qualitative analysis of a longitudinal case study which investigates how the ESM-based design approach supports co-designing curricula for student engagement in specific kinds of scientific practices called computational thinking practices.

In the digital age of learning, the importance of the role of technology-enhanced learning environments and curricula in engaging students in authentic scientific inquiry practice has been recognized early on in Learning Sciences research, commencing with Papert and colleagues

(Papert, 1972, 1980) and followed on by (e.g., Disessa, 1987; Edelson et al., 1999; Linn & Hsi, 2000; Wilensky, 1999). However, the nature of these authentic practices and our understanding of what counts as authentic practices is changing at a rapid pace with the incorporation of newer technological tools and research methods, and a deeper understanding of social and cultural aspects of learning. For example, the incorporation of computational modeling methods and an increasing focus on complex systems thinking has significantly changed the nature of research in biology, ranging from genetic to ecological networks (e.g., Kitano, 2002, 2017). In parallel, there has been considerable attention to complex systems frameworks in the learning sciences (Jacobson & Wilensky, 2006; Levy & Wilensky, 2011; Maroulis et al., 2010; Resnick & Wilensky, 1998; Sabelli, 2006). This calls for the design of newer learning environments that effectively and authentically incorporate these research practices of the community of inquirers, to foster students' legitimate peripheral participation in the disciplinary domains (Lave & Wenger, 1991). Additionally, scholars in the field of learning sciences are exploring ways to re-imagine disciplinary learning to support the onto-epistemic heterogeneity of students and their meaningful participation in science learning (Warren et al., 2020) and plurality in students' ways of engaging with a learning environment (Turkle & Papert, 1990). These researchers ask us, as educators and researchers of learning, to develop expansive forms of learning to support multiple ways of knowing and being by foregrounding interwoven sensibilities of multiplicity, horizontality, and dialogicality. This is an important and a huge ask. In my dissertation, I attempt to attend to such design considerations for creating expansive learning opportunities that aim to achieve a balance between students developing a deep understanding of established disciplinary ideas such as natural selection or gene regulation and their agency in devising and engaging in knowledge construction practices to investigate those ideas.

Following on early constructionist research, there has been increased emphasis on engaging students in the scientific practices in the recent educational reforms (NGSS Lead States, 2013; Schwarz et al., 2017). This requires reimagining the roles of students and teachers in the science classroom so that students become *doers of science* and not *receivers of facts* (De Jong et al., 2013; Miller et al., 2018; Papert, 1993; Schwartz et al., 2004). Learning scientists have pointed out the disparity in the way a science discipline such as biology is taught in school settings and practice of current research in biological sciences and have argued for the effectiveness of a computational modeling approach that allows students to engage contemporary disciplinary ideas by participating in contemporary disciplinary practices (e.g., Wilensky & Reisman, 2006). *Doing science* in the science classroom means engaging students in such contemporary science practices to construct disciplinary knowledge. Miller and colleagues (2018) ask science educators to revisit the question – how do we arrive at the scientific practices that we want students to engage in? The Next Generation Science Standards (NGSS) has recommended a set of science practices that are epistemically equivalent to those practices of scientists (NGSS Lead States, 2013). However, this framing of epistemic equivalence creates a contradiction for *doing science* using the NGSS framework (Miller et al., 2018). Having agency in shaping the community's shared knowledge construction work is an important aspect of science researcher practice. Engaging students in similar ways in the knowledge construction processes requires designing learning environments that support students' epistemic agency. The term epistemic agency was introduced into education literature in relation to the research on knowledge-building communities conducted by Scardamalia and Bereiter (1991). Epistemic agency refers to students' ability to shape and evaluate knowledge and knowledge-building practices in the classroom (Scardamalia & Bereiter, 1991; Stroupe, 2014).

Miller and colleagues (2018) argue that NGSS's focus on science practices is not sufficient for the move to envision students as *doers of science* and not *receivers of facts*. Having a set of practices chosen by others as important to learn and expecting students to mimic those practices does not position students with the epistemic agency - the power to shape the knowledge production and practices of a community. Positioning students as epistemic agents requires them to collectively evolve practices for knowledge construction. To truly support students' epistemic agency in a classroom, newer learning environments need to be designed that foster student engagement in shaping knowledge-building practices to construct and evaluate knowledge products.

Doing Science in the Science Classroom

Epistemically agentic learning by *doing science* would entail epistemologically meaningful engagement in science practice for sense-making, rather than merely knowing about scientific inquiry (Abd-El-Khalick et al., 2004; Berland et al., 2016; Lehrer & Schauble, 2012). In other words, science students should learn how to construct scientific knowledge about the world in the science classroom just like scientists do. In contemporary scientific practice, 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.

To achieve these goals of science education, the Next Generation Science Standards (NGSS) emphasize a three-dimensional way of learning science that supports students' development of proficiency across a set of science and engineering practices, disciplinary core ideas, and crosscutting concepts (NGSS Lead States, 2013). The first dimension is Science and Engineering Practices (SEP). These practices are the behaviors that scientists use to make sense of phenomena in the natural world and engineers engage in as they design and build models and systems. The second dimension, Crosscutting Concepts (CC), includes concepts that have application across all domains of science. The third dimension, Disciplinary Core Ideas (DCIs) are the key ideas that students should know in order to understand multiple science disciplines. NGSS promotes an integrated approach for these three dimensions, which means that each dimension should be strongly interconnected across different subjects throughout the school years. The Framework introductory chapter states "*The framework is designed to help realize a vision for education in the sciences and engineering in which students, over multiple years of school, actively engage in scientific and engineering practices and apply crosscutting concepts to deepen their understanding of the core ideas in these fields.*" (National Research Council, 2012).

In a classroom context, such a learning approach requires students – (1) to become familiar with the context of a phenomenon they are investigating, (2) to ask relevant research questions that they can investigate using a system, (3) to test and verify ideas by designing and performing investigations, and (4) to construct explanations regarding the phenomenon based on their investigations. Learning scientific inquiry practices in such an authentic way in a classroom context is often challenging because it takes a substantial amount of classroom time to engage students in cognitive processes and epistemic activities similar to those of scientists. Additionally, there are instructional challenges that teachers have to overcome while achieving the authentic engagement in processes and activities (Chinn & Malhotra, 2002). Chinn & Malhotra (2002) consider authenticity in terms of the similarities between classroom practices and the practices that scientists actually engage in. NGSS's emphasis on practices is aligned with this idea of authenticity by engaging students in activities similar to those of scientists (NGSS Lead States, 2013). Recent work in the field of science education that focuses on three-dimensional learning also underscores the challenges regarding engaging students in practices because of the tensions between teaching established disciplinary ideas and authentically participating in practices to construct knowledge (Russ & Berland, 2019; Schwarz et al., 2017). Another challenging aspect regarding engagement in authentic practices is supporting students' strategy-like ways of engaging in knowledge-building processes that Krist and colleagues refer to as epistemic heuristics (Krist et al., 2019). We argue that one approach to address these challenges is to use a learning environment that is designed for student engagement with the core aspects of disciplinary ideas and practices to investigate scientific phenomena from cognitive, affective, and social perspectives.

Designing and Using ESMs for *Doing Science*

My work addresses this issue of engaging students in *doing science* by combining two powerful design approaches in learning sciences: agent-based modeling of complex systems (Paulo Blikstein & Wilensky, 2010; Brady et al., 2015; Jacobson & Wilensky, 2006; Wagh et al., 2017; Wilensky, 2001) and constructionism (Papert, 1980; Papert & Harel, 1991). We call such computational learning environments Emergent Systems Microworlds (ESM) (Dabholkar et. al, 2018). Chapter 2 of this dissertation is dedicated to explaining the theoretical foundations of the ESM design framework. The ESM design framework is inspired by years of research on combining agent-based modeling and constructionism (For example, see Blikstein & Wilensky, 2010; Sengupta & Wilensky, 2009; Wilkerson-Jerde & Wilensky, 2010) and a related design framework called Emergent Systems Sandboxed (ESS) (Brady et al., 2015). In Chapter 2, I also discuss similarities and differences between these two frameworks and how my work builds on and contributes to the prior work using similar design approaches.

In my dissertation, I study how ESMs and ESM-based curricula engage students in scientific practices to construct knowledge of disciplinary ideas. I present three interconnected studies about characterizing students' learning with ESM-mediated curricula and supporting teachers' integration of authentic practices using computational tools in science classrooms. ESM design is based on Wilensky and Papert's theory of restructuration (Wilensky, 2020; Wilensky & Papert, 2010). ESMs use agent-based restructurations to engage students in investigating and learning emergent phenomena. In the following part, I first discuss the theory of restructurations and properties of restructurations. Then I explain agent-based restructurations and their use in ESMs.

Theory of Restructurations

The ESM design framework and design of ESM-based curricula draw their theoretical foundations from Wilensky and Papert's theory of restructurations (Wilensky, 2020; Wilensky & Papert, 2010). Restructurations are about using new representational forms to reformulate knowledge in various disciplines. Wilensky and Papert (2010) define structuration as the encoding of the knowledge in a domain as a function of the representational infrastructure used to express the knowledge. For example, Hindu-Arabic numerals (1, 2, 3…., 99, 100, etc.) is a representational infrastructure for arithmetic today. Before Hindu-Arabic numerals became common, roman numerals (I, II, III, …, XCIX, C, etc.) were the representational infrastructure for arithmetic in the western world. The move from Roman to Hindu-Arabic numerals in arithmetic is an example of restructuration. A shift from one structuration of a domain to another resulting from such a change in representational infrastructure is restructuration. Because of this restructuration, arithmetic operations like multiplication shifted from being a specialized skill that only a small number of trained people held to a part of elementary school curriculum.

PROPERTIES OF RESTRUCTURATIONS

In order to study and evaluate restructurations, Wilensky and Papert discuss five core properties of restructurations (Wilensky & Papert, 2010). These five properties fall into two general categories- power properties and learnability properties. Power properties are about the power of a new structuration to *do* what could be done by an old structuration and to go beyond what was possible before it. Power properties of a restructuration are more about advancing knowledge in a domain because of the restructuration. Learnability properties of a restructuration are more about the affordances of the restructuration for learning. Four core properties of restructuration fall under the category of learnability properties. These four properties are cognitive, social, affective, and diversity.

Power properties

A new structuration can restructurate the domain of knowledge if it has the power to do what an old structuration did and more. This power is about the use of a structuration for performing tasks and for developing analytical insights into the knowledge domain. For example, the Hindu-Arabic numerals can do basic operations in arithmetic such as addition, subtraction, division, and multiplication which could be done with Roman numerals as well. These operations with large numbers were much more cumbersome to do with Roman numerals. But think about how many ideas and theorems in Number Theory[1] can be pursued using Roman numerals! The performative and analytical power of Hindu-Arabic number structuration is much more than Roman number structuration for using numbers and investigating properties of numbers.

Learnability properties

The second high-level category of restructurations properties is learnability properties. These properties are about learning the knowledge domain. Papert and Wilensky (2010) discuss four core properties of restructuration as learnability properties. These properties are cognitive properties, social properties, affective properties, and diversity properties.

Cognitive properties

The cognitive properties are about the ease of learning the knowledge domain. For example, to compare cognitive properties of Hindu-Arabic and Roman number systems, one needs to compare the cognitive ease of learning operations such as multiplication with each of the structurations. Another example of a restructuration that Wilensky and Papert highlight from Disessa's book, *Changing Minds* (DiSessa, 2001) also resulted in dramatic gains in learnability. This example is about the historical restructuration of simple kinematics from a text-based to an algebraic representation. Galileo struggled to establish and share ideas regarding relations between entities such as time, distance, and velocity because he used text-based representations.

[1] https://en.wikipedia.org/wiki/Number_theory

However, the use of algebraic representations makes it very easy to express a relationship between time, distance, and velocity, which is d = vt.

Social properties

The social properties are about the ease of sharing the newly established or developing ideas as *knowledge products* in a domain. This requires a structuration being effective in supporting the expression of ideas in ways that make the ideas easily sharable and usable for others. The text-to-algebra restructuration referred to above is useful for thinking about social properties. The use of such algebraic representations has helped develop and share knowledge about kinematics and mechanics in general over several decades as algebraic representations restructurated the knowledge domain.

Affective properties

A restructuration can make the knowledge more or less engaging. Wilensky and Papert (2010) argue that the likability of domain knowledge can be influenced by structurations that are used to engage with the domain knowledge. Especially from the perspective of engaging students in investigating and learning ideas in a disciplinary domain, their affective engagement to engaging meaningfully with those ideas matters.

Diversity properties

Structurations of a discipline can differ in their match with diverse ways of learning and thinking. Wilensky and Papert (2010) argue that it is important to investigate how a structuration influences engagement with the domain knowledge of learners with different cultural, ethnic, gender, cognitive, and emotional identities, and backgrounds.

AGENT-BASED RESTRUCTURATIONS

In my work, I specifically focus on one type of restructurations called agent-based restructurations (Wilensky, 2001, 2020; Wilensky & Papert, 2010). One powerful methodology that has emerged from complex systems theory is agent-based modeling (Epstein & Axtell, 1996; Grimm & Railsback, 2013; Wilensky & Rand, 2015; Wilensky & Resnick, 1999). In contrast to more traditional mathematical modeling, which typically involves symbolic representations in the form of equations, agent-based modeling makes use of simple computational rules. The core elements in the model are computational objects or "agents". Each of these agents has state variables that describe its particular state, such as age, energy level, hunger, etc. The behavior of the agents is determined by the computational rules that tell each agent what to do. The rules are framed from the agent's point of view. For example, an agent could be a goose in a flock of geese. The computational rules followed by a computational agent goose are – collision avoidance (separation), speed and/or direction matching (alignment), and flock centering (cohesions) (Reynolds, 1987; Wilensky, 1998). As each agent follows these computational rules, complex patterns emerge, such as the V-shape of a flock.

Learnability Properties of Agent-Based Restructurations

Earlier in this chapter, I have discussed Wilensky and Papert's (2010) idea about properties of restructurations and the importance of investigating properties of restructurations. They argue for affordances of restructurations of traditional science content to increase the scientific power and learnability. The greater value of the agent-based approach lies in understanding any emergent phenomenon using computational tools. Once the basic computational model is set up, it is easy to explore a large set of configurations and to understand the possible trajectories of the system. And because the representation system is composed of simple, understandable micro-rules rather than aggregate level equations, it is easier to modify them to explore a host of other phenomena (Sengupta & Wilensky, 2009; Wilensky & Reisman, 2006). The agent-based restructurations enable students to reason about natural phenomena from the bottom up (P. Blikstein & Wilensky, 2005; Levy & Wilensky, 2008; Wilensky, 1999a; Wilensky & Novak, 2010). Whereas traditionally, students employ heuristics and formulae given to them by authority, they are now able to author their own heuristics and formulae derived from their modeling experience. Just as the restructuration of numerals enabled ordinary folks to do multiplication and division for themselves, the restructuration with computational agent-based representations allows students to explore, investigate and reason about complex systems phenomena using computational models. This has a tremendous potential to democratize powerful ideas in modern sciences (Wilensky & Papert, 2010).

The agent-based restructurations reduce cognitive and perceptual limitations and allow students to reason about emergent patterns at the system level by observing the behaviors of agents (Goldstone & Wilensky, 2008). Such restructurations have been demonstrated to be pedagogically effective to support the learning of several complex natural phenomena in science education (e.g., electric current and resistance, temperature, pressure, evolution) (Sengupta & Wilensky, 2009; Wagh et al., 2017; Wilensky, 2003). In my dissertation, I investigate the properties of agent-based restructurations in constructionist curricula which I call ESM-based curricula. I argue that agent-based restructurations and constructionist design principles in ESM-based curricula restructurate science learning in a science classroom.

AGENT-BASED RESTRUCTURATIONS OF BIOLOGICAL PHENOMENA

There is a significant disparity between how biologists study biological systems and how high school biology students learn about those systems (Wilensky & Reisman, 2006). Other than technical advances in molecular biology, one of the significant shifts in contemporary research in biological sciences is the use of systems theoretical perspectives to understand and investigate biological complexity using computational approaches (Kitano, 2002, 2017). From the molecular level to the cellular level to the organismic level to the population level, biological systems can be studied as complex systems comprised of interconnected constituent parts. Agent-based restructurations have been demonstrated to be effective in investigating such complex systems (Aslan et al., 2018; Dey et al., 2006; Wilensky, 2020) and teach emergent biological phenomena such as population dynamics and evolution (Wagh & Wilensky, 2018; Wilensky & Novak, 2010; Michelle Wilkerson-Jerde et al., 2015). From the educational standpoint, it is important to use agent-based restructurations to teach biological phenomena because – 1. restructuration properties of agent-based representations increase learnability of emergent phenomena in biology, which are otherwise difficult to understand (Wilensky & Reisman, 2006), 2. agent-based representations help students develop systems theoretical perspective which is increasingly becoming a part of contemporary biology research practice. In my dissertation, I designed and co-designed three restructurated biology curricula using the ESM design approach. I conducted three studies with those restructured curricula to investigate different aspects of learning and design.

Goal and Contributions

The three main goals of my dissertation are:

1. To develop a design framework for ESMs and ESM-based curricula and analyze how the introduction of a new tool like ESM mediates students' *epistemically expansive learning* in a science classroom

2. To analyze how an ESM-based curricular unit facilitates students learning by fostering their *epistemic connection* making among scientific inquiry practices and disciplinary ideas
3. To better understand how to co-design with science teachers can be supported in teaching ESM-based curricular units by increasing their involvement in designing ESMs and ESM-based curricula.

I have defined these epistemically expansive learning and epistemic connections in the following section. The three goals of my dissertation inform the following set of research questions that I pursue in this dissertation:

1. How do design features of an ESM support epistemically expansive learning in a science classroom?
2. How does the design of an ESM-based curricular unit support student connection-making among scientific inquiry practices and disciplinary ideas?
3. How does restructuration through ESM facilitate the co-design process for CT-integration into science units and its outcomes?

THREE INTERCONNECTED STUDIES

My dissertation focuses on investigating how the ESM design approach restructurates science learning. There are two aspects of the ESM design approach – designing ESMs and designing ESM-based curricula. Both these aspects are strongly rooted in constructionism (Papert, 1980) and agent-based restructurations (Wilensky, 2001, 2020; Wilensky & Papert, 2010). I study restructuration of student learning in terms of shifts in their roles in a science classroom in epistemic activities. These shifts are about their engagement in practices for building disciplinary knowledge. My work contributes to understanding how ESMs and ESM-based curricula mediate these shifts. I do this through three interconnected studies. The first study focuses on students' epistemic agency in the classroom and how ESMs support students' epistemically agentic participation in classroom activities. The second study is about how ESM helps students make connections among various disciplinary ideas and scientific practices. Since these connections are about knowledge construction, I call them epistemic connections. In the third study, I focus on how the ESM design approach supports teacher participation in co-designing ESM-based curricula for engaging students in specific kinds of scientific practices, called Computational Thinking practices.

I describe and analyze the three studies in a repeating pattern. Each study is divided into two chapters (Chapters 3 & 4; 5 & 6; and 7 & 8). The first chapter of each study describes the design of the ESM, and the second chapter presents an analysis of the data from the implementations of the restructurated unit.

Study 1: Epistemic Expansion

In the first study (Chapter 3 and 4), I investigate how design features of an ESM support epistemically expansive learning in a science classroom. The ESM, GenEvo, is a restructuration of mechanisms of evolution for describing molecular mechanisms of gene regulation and mechanisms of evolution – natural selection and genetic drift. Chapter 3 describes the design and design rationale for GenEvo using its restructurational properties. In Chapter 4, I use the activity theory framework (Engeström, 2001) to define epistemic expansion as a transformation of knowledge construction activities in a classroom activity system. I investigate how an ESM-based curriculum mediates epistemic expansion by involving a classroom community in collectively evolving epistemic practices of knowledge construction. I argue that agent-based restructurations in a constructionist learning environment can support students' epistemically expansive learning by facilitating their participation in shaping practices for evidence-based knowledge construction and evaluation. I specifically focus on analyzing how cognitive, affective, and social properties of restructurations in an ESM mediate students' epistemically agentic learning to understand emergent aspects of gene regulation and evolution in a biology classroom.

Activity theory focuses on understanding *mediation* by a tool when a subject uses that tool to achieve an *object* (Vygotsky, 1978). Engeström's conceptualization of the theory of expansive learning is within the framework of activity theory (Engeström, 2001). Expansive learning is about creating a transformation in the activity system to start producing new patterns of activity. I build on this idea to focus on expansive learning in a science classroom in terms of the transformation of epistemic activities. Such epistemic expansion of a classroom activity means the transformation of the classroom activity system to engage students in collectively constructing knowledge and evolving practices to produce that knowledge. I define *epistemic expansion* in a science classroom as the transformation of classroom activity from the one in which students are positioned as *receivers of facts* to the one in which they are positioned in epistemically agentic roles as *doers of science*.

Using micro-ethnographic methods for classroom interactions and discourse (Erickson, 1986, 2006), I analyze the dynamics of students' shifting participation in social intellectual activity and in shaping the knowledge construction practices of the classroom (Engeström, 2001; Rogoff, 2008). My framing of the GenEvo learning environment includes the ESM-based curriculum and ESM-enabled pedagogy. Because ESMs are designed as interactive microworlds with agent-based restructurations, ESMs enable certain pedagogical moves to support students' individual and collective epistemic engagement in disciplinary knowledge construction. I investigate how different design features of the GenEvo learning environment and accompanied pedagogical moves mediated shifts in students' epistemic agency by providing them with a system to generate, test, and share their ideas to develop a deep understanding of some of the fundamental ideas in modern biology.

STUDY 2: EPISTEMIC CONNECTIONS

In this study (Chapters 5 and 6), I delve deeper into investigating how an ESM-based curriculum supports student learning of science practices and disciplinary ideas in connection with each other. In chapter 5, I focus on the design of a restructuration of population biology. I present another ESM-based curriculum natural selection. In this curriculum, students investigate the evolution of a rock pocket mice population in a sandy (light) or a rocky (dark) background by designing and conducting experiments. This ESM-based curriculum is also designed to scaffold student engagement in NGSS- aligned science practices (NGSS Lead States, 2013)in the context of an ESM to learn how to construct knowledge of those disciplinary ideas. In chapter 6, I investigate how the ESM-based curriculum supported student making *epistemic connections* among Science and Engineering Practices (SEPs) and Disciplinary Core Ideas (DCIs). I define *epistemic connections* as connections among SEPs and DCIs when students use those SEPs to construct knowledge about the DCIs. As students engage in the epistemic activities using an ESM, they are expected to investigate disciplinary ideas using science practices. Since student learning in an ESM-based curriculum involves both practices and ideas, tracking student learning progression in terms of practices and ideas is of analytical importance. In the analysis, I focus on the co-occurrence of ideas and practices as I study students' learning progress through a curriculum. For example, in a curriculum unit about natural selection, students construct knowledge about disciplinary ideas, such as heredity, environment, and survival in the context of natural selection by engaging in scientific inquiry practices. I analyze the epistemic connections to characterize student learning with the ESM-based curriculum.

To analyze such connection making, I use 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, Chesler, & Shaffer, 2014; Bagley & Shaffer, 2015). These affordances furthermore allow researchers to assess student learning as they express their ideas (Arastoopour et al., 2016).

Previous work using ENA has demonstrated that students exhibited science and computational learning gains after engaging with a Computational Thinking integrated biology unit with computational models (Arastoopour et al., 2020). This line of research also investigated students' understanding of systems thinking practices as they participated in a chemistry unit using ENA and demonstrated how the design of the unit supported understanding of micro-macro relationships regarding emergent concepts such as pressure and temperature of gases (Arastoopour et al., 2019). In Study 2, I extend this work further by investigating how an ESM-based curriculum supports student learning of practices and disciplinary ideas by making connections among those. From the disciplinary learning perspective, this curriculum is designed for students to learn about how populations change over time in different environmental conditions (Dabholkar, Woods, et al., 2018).

Further, I use Collins and Ferguson's (1993) notions of epistemic forms and games to investigate how the NGSS-aligned epistemic form of an ESM-based Rock Pocket Mice Curriculum in a biology classroom supported students' learning of science practices and disciplinary ideas regarding natural selection. Using ENA (Shaffer et al., 2009) to characterize student learning progression in terms of practices and disciplinary ideas, I provide evidence that the ESM-based curriculum systematically created learning opportunities for both - (a) learning authentic science practices in a disciplinary context, and (b) constructing knowledge of disciplinary ideas using those practices.

STUDY 3: ESM-MEDIATED CO-DESIGN

The third study of my dissertation (Chapters 7 and 8) focuses on co-designing ESM-based curricula to integrate Computational Thinking (CT) in disciplinary contexts. In chapter 7, I present an ESM that I co-designed with a high school teacher, Ms. Tracy (a pseudonym) for engaging students in CT practices in different learning contexts in a biology classroom. This ESM restructurates a curriculum to learn about animal behavior and experimental design. The ESM is about the habitat preference behavior of isopods, commonly known as rollypollies. The ESM-based curricular activities are designed to engage students in - (1) computational investigation of rollypolly habitat preference behavior, (2) modification of a computational model to incorporate new experimental conditions, (3) computationally automated experiential setup and data collection, (4) construction of a computational model using coding blocks. In Chapter 8, I investigate how restructuration through ESM facilitates the co-design process for CT-integration into science units and its outcomes. I present a longitudinal case study of a researcher-practitioner design partnership for three years as it matured from using pre-designed CT-integrated curricula to co-designing new CT tools and curricula. We co-designed new ESMs and curricula activities using those ESM. Use of the ESM design approach for CT-integration supported the co-design partnership, as the cognitive and social properties of restructurations (Wilensky & Papert, 2010) in ESMs mediated Tracy's increased involvement in the co-design process. Reciprocally, Tracy's highly valuable contributions in identifying relevant biology learning contexts and devising pedagogically effective learning activities in the co-design process enriched the ESMs and ESM-based curricula.

In this study, I first investigate how the following aspects of the co-design process were mediated by the use of the ESM design approach - (a) changes in Tracy's involvement in the co-design process, (b) shifts in the curricular designs in terms of the richness of CT-integration, and (c) shifts in Tracy's classroom teaching practice. Then, I present an argument for the effectiveness of ESMs for designing CT-integrated curricula and the co-design approach that uses ESMs for co-designing such curricula.

The dissertation concludes with chapter 9, in which I summarize the conclusions and implications of each of the three studies and discuss overall contribution of the dissertation research for designing and using learning environments for students' epistemically agentic participation in scientific inquiry practices in general and computational thinking practices in particular.

Chapter 2: Design Framework: ESMs and ESM-based curricula

Summary: In this chapter, I discuss the theoretical foundations of the idea of Emergent Systems Microworlds (ESMs) and ESM-based curricula. ESM design uniquely combines the use of an agent-based approach of modeling complex phenomena and constructionist design principles. Agent-based models are designed to investigate and/or demonstrate how system-level properties emerge from the uncoordinated behavior of individual agents, as well as how, on the other hand, the system affects individuals. Agent-based representations are restructurations that have power and learnability properties. The learnability properties of agent-based representations in an ESM facilitate students' engagement in knowledge construction of modeled complex phenomenon. The ESM design attempts to incorporate the following three key ideas from the constructionist design framework: (a) personally meaningful engagement, (b) construction of public entities, (c) expression and validation of ideas through computational microworlds. These constructionist design features in an ESM and ESM-based curricula facilitate student engagement in practices to construct disciplinary knowledge.

INTRODUCTION

In the first chapter, I discussed the importance of engaging students in disciplinary practices to construct knowledge as opposed to presenting knowledge as established ideas and how I intend to use the theory of restructurations (Wilensky, 2020; Wilensky & Papert, 2010) to design learning environments for such epistemically agentic learning and evaluating effectiveness of these learning environments. By learning environments, I mean learning activities and pedagogical support to engage in those activities. To design learning environments to support students' epistemic agency while constructing knowledge about disciplinary ideas and practices, I draw on two design frameworks: (a) agent-based models of complex systems (Wilensky, 2001; Wilensky & Rand, 2015), and (b) constructionism (Papert, 1980). I call this design approach Emergent Systems Microworld (ESM) (Dabholkar, Anton, et al., 2018). In this chapter, I discuss the theoretical foundations of designing ESMs and ESM-based learning environments.

In this dissertation, there are four chapters dedicated to explaining ESMs and ESM-based curricula. The ESMs that I present and discuss in these chapters consist of computational tools, most commonly computational models. ESM-based curricula include learning activities that use ESMs and are designed using pedagogical principles of learning with agent-based models and constructionism. This chapter explains the theory behind the ESM design framework. Chapters 3, 5, and 7 explain three examples of ESMs and ESM-based curricula that I have designed and co-designed. In those chapters, I discuss in-depth the elements of these ESMs and the decisions that influenced the designs and organizations of those elements in particular ways. In this chapter, I focus on explaining the theoretical background and prior work with ESMs and how my dissertation work contributes to that work. Even though ESM is a term that I have coined during my dissertation work, there has been decades of work at The Center for Connected Learning and Computer-Based Modeling at Northwestern University that have used elements of agent-based modeling and constructionism in various ways (Paulo Blikstein & Wilensky, 2010; Guo & Wilensky, 2018; Hjorth & Wilensky, 2014; Levy & Wilensky, 2009; Sengupta & Wilensky, 2009; Wagh & Wilensky, 2018; M. Wilkerson-Jerde & Wilensky, 2010). In addition, Brady and colleagues in the ModelSim project, coined the term ESS to describe agent-based constructionist learning environments that are designed as constructionist sandboxes for model-based inquiry

(Brady et al., 2015). I build on this work and contribute to it by demonstrating the power properties of restructurations in the design of three ESMs (Chapters 3, 5, and 6) and investigating how learnability properties of restructurations (Wilensky, 2020; Wilensky & Papert, 2010) contribute to the effectiveness of ESMs and ESM-based curricula to engage students in knowledge building practices by supporting their epistemic agency. Additionally, the third study of my dissertation (Chapters 7 and 8) investigates how the use of ESMs support co-design of Computational Thinking (CT)-integrated curricula.

THEORETICAL FRAMEWORK

The following two ideas provide theoretical foundations for designing ESMs and ESM-based curricula:

a. agent-based restructurations (Wilensky, 2001, 2003, 2020; Wilensky & Papert, 2010)
b. constructionism (Papert, 1980).

ESMs are designed as agent-based models of emergent natural phenomena (Wilensky, 2001) in the form of easily manipulable computational microworlds (Papert, 1980).

AGENT-BASED RESTRUCTURATIONS IN ESMS

Agent-based modeling of emergent systems is one of the central design features of an ESM. Such dynamic computational agent-based representations are restructurations of emergent phenomena, which are typically taught with differential equations or static models (Wilensky & Papert, 2010). The power properties of agent-based restructurations are about the analytical power to express and investigate an emergent phenomenon. The learnability properties of agent-based restructurations include cognitive, social, affective, and diversity properties (See Chapter 1 for details about properties of restructuration). In my dissertation, I demonstrate the power properties of agent-based restructurations by designing three ESMs that restructurate three emergent phenomena – (1) stimulus-based regulation of protein production through gene regulation in *E. coli* bacterium (Chapter 3), (2) changes in a rock pocket mice population because of natural selection (Chapter 5), and (3) habitat preference behavior of rollypolly bugs (Chapter 7). In Chapters 4, 6, and 8, I investigate how the learnability properties of agent-based restructurations in these ESMs support student learning in science classrooms and teacher learning through a process of co-designing ESM-based curricula.

In this section, I first discuss the use of the agent-based modeling approach in science research. Then I review the literature regarding using the emergent systems perspective and agent-based modeling approach in science education. Finally, I discuss how agent-based restructurations are a core aspect of ESM design and how I investigate learnability properties of agent-based restructurations in ESMs in my dissertation.

Emergence

To understand the importance of agent-based models in science research and science education, it is important to understand the idea of emergence in the first place. Emergence is a transdisciplinary idea across philosophy, systems theory, science, and arts. Though the word *emergence* is used to denote different related conceptualizations regarding properties of a system (Goldstein, 1999), I use the idea of emergence in my dissertation as discussed by Wilensky and Rand in the context of agent-based modeling of complex systems (Wilensky & Rand, 2015). They define emergence *as the arising of novel and coherent structures, patterns, and properties through the interactions of multiple distributed elements*. The fundamental notion here about emergence is that emergent structures at the aggregate level cannot be deduced solely from the properties of constituent elements, rather they arise through the combination of properties of elements and their interactions with one another.

The idea of emergence is related to the study of complex systems and complex systems theory (Bar-Yam et al., 1998; Bar-Yam, 2004). Complex systems are systems that consist of many autonomously interacting parts. These autonomous interactions result in predictable emergent patterns at the system level. Many Learning Sciences researchers have discussed why it is important and cognitively difficult to learn about complex systems (Chi, 2005; Hmelo-Silver & Azevedo, 2006; Hmelo et al., 2000; Wilensky & Resnick, 1999). Wilensky and Rand (2015) discuss two main reasons to explain why it is difficult to learn about emergent properties of complex systems. The first reason for the difficulty in trying to figure out the aggregate pattern when one knows how individual elements behave. The authors refer to this as integrative understanding, as it parallels the cumulative integration of small differences in calculus. A second reason is related to understanding the properties and behaviors of elements when you know system-level aggregate patterns. The authors refer to this as differential understanding, or compositional understanding, as it parallels the search in calculus for the small elements that produce an aggregate graph when accumulated. So, the difficulty is because it is hard to imagine and reason about emergent properties or patterns at the systems level when you know about the properties of the elements and vice versa.

Agent-based models (ABMs) are effective in addressing these issues because of the power and learnability properties of agent-based restructurations (Wilensky, 2020; Wilensky & Papert, 2010). In the following section, I discuss the ABM approach to modeling complex systems to understand natural phenomena and the use of ABMs to teach about complex systems.

Since my work focuses on student learning about emergent properties of complex systems and how those arise, I refer to such systems as complex emergent systems or emergent systems in my dissertation.

ABMs in research

Agent-based models (Railsback & Grimm, 2019; Wilensky & Rand, 2015), also referred to as individual-based models (IBMs), or agent-based models and simulations (ABMS), are widely used in investigating emergent properties of complex systems in a variety of domains ranging from ecology (DeAngelis, 2018; DeAngelis & Mooij, 2005; Grimm & Railsback, 2005) to the social sciences (Axtell & Epstein, 1994; Epstein & Axtell, 1996; Maroulis et al., 2010), economics (Tesfatsion, 2002), demography (Billari & Prskawetz, 2012), and political sciences (Axelrod, 1997; Holman et al., 2018). ABMs allow researchers to study how system-level properties emerge from the uncoordinated behavior of individuals as well as how, on the other hand, the system affects individuals (Wilensky & Rand, 2015).

The rise of computation has tremendously supported the advancement of agent-based modeling and simulation (for example, see Bandini et al., 2009). Several computational tools have been developed to model complex systems using the ABM approach. NetLogo (Wilensky, 1999b) is perhaps the most widely used software for agent-based modeling (Wilensky & Rand, 2015). It is used both for scientific research and educational purposes. NetTango Web is a front-end interface to NetLogo Web, a domain-centric block-based coding environment used to make NetLogo models (Horn et al., 2020; Horn & Wilensky, 2012). The other most widely used general purpose software environments that are currently in use to model complex systems include: Swarm, developed at the Santa Fe Institute (Hiebeler, 1994), Repast, developed at

Argonne National Laboratory (North et al., 2006), and MASON, developed at George Mason University (Luke et al., 2005). In my dissertation work, I have used NetLogo and NetTango Web to design ESMs.

Emergent Systems and ABMs in science education

Learning about the emergent complex systems perspective involves understanding how uncoordinated interactions between autonomous elements can result in complex emergent patterns at the system level (Wilensky, 2001; Wilensky & Rand, 2015). In addition to becoming a focus of real-world scientific research in several fields as discussed before, the emergent systems perspective has been recognized to be important for science education (Jacobson & Wilensky, 2006; Yoon et al., 2017). Researchers of science education have argued for and demonstrated the effectiveness of an emergent systems perspective for understanding several natural phenomena ranging from prey-predator relationships to nectar collection by honeybees, to the kinetic molecular theory (Danish, 2014; Hmelo-Silver & Azevedo, 2006; Klopfer, 2003; Wilensky, 2003; Wilensky & Reisman, 2006). The Next Generation Science Standards in the United States has incorporated ‘systems and systems models’ as one of the seven key cross-cutting concepts (NGSS Lead States, 2013).

ABMs and ESMs

Are all agent-based models (ABMs) ESMs? A simple answer to this question is no. There are several ABMs that are designed to investigate a phenomenon and do not use constructionist pedagogical principles (for example, ABMs used disciplinary investigations in specific fields social sciences (Axtell & Epstein, 1994; Epstein & Axtell, 1996), economics (Tesfatsion, 2002), demography (Billari & Prskawetz, 2012), and political sciences (Axelrod, 1997; Holman et al., 2018)). Only those ABMs that are designed as constructionist microworlds are ESM. As such, most of the models in the NetLogo models library can be characterized as ESMs. I explain how constructionist design principles inform the theoretical foundations of ESM design in the next section. In this section, I discuss how agent-based representations are a central part of ESM design and how they improve the learnability of phenomena modeled in an ESM.

ESMs are designed by modeling behaviors of individual agents and representing emergent patterns through computational visualizations of systems, system-level data, and graphs. Figure 1 shows A Rock-Pocket-Mice ESM (Dabholkar & Wilensky, 2020), which I designed for students to learn about how populations change due to natural selection (see Chapters 5 and 6). All the coded behaviors in the ESM are of the mice agents. These behaviors are uncoordinated. The mice agents are coded to move randomly, reproduce by mating randomly with a neighboring mouse of the opposite sex, get predated if they do not camouflage well, and pass on traits based on Mendel's laws of inheritance. These uncoordinated actions of individual mice generate predictable emergent patterns at the population level – the population of mice in the dark rocky areas is predominantly of mice with dark fur coats and the population of mice in the light sandy areas is predominantly of mice with light fur coats (Figure 2-1). Students can observe how populations change over generations in these dynamic virtual microworlds and use computational visualizations to see system-level patterns as they emerge. The graph in this ESM shows how phenotype frequencies change over time in the population. Students also have access to data of phenotypes and genotypes of mice in the population (see the monitors AA males, AA females, etc.). These computational representations of agent-level and system-level changes are part of the agent-based restructurations in the ESM. Additionally, students can access the NetLogo computational code to see how agent behaviors are encoded. This provides them opportunities to engage more deeply to learn how such models are created, and critique and improve those.

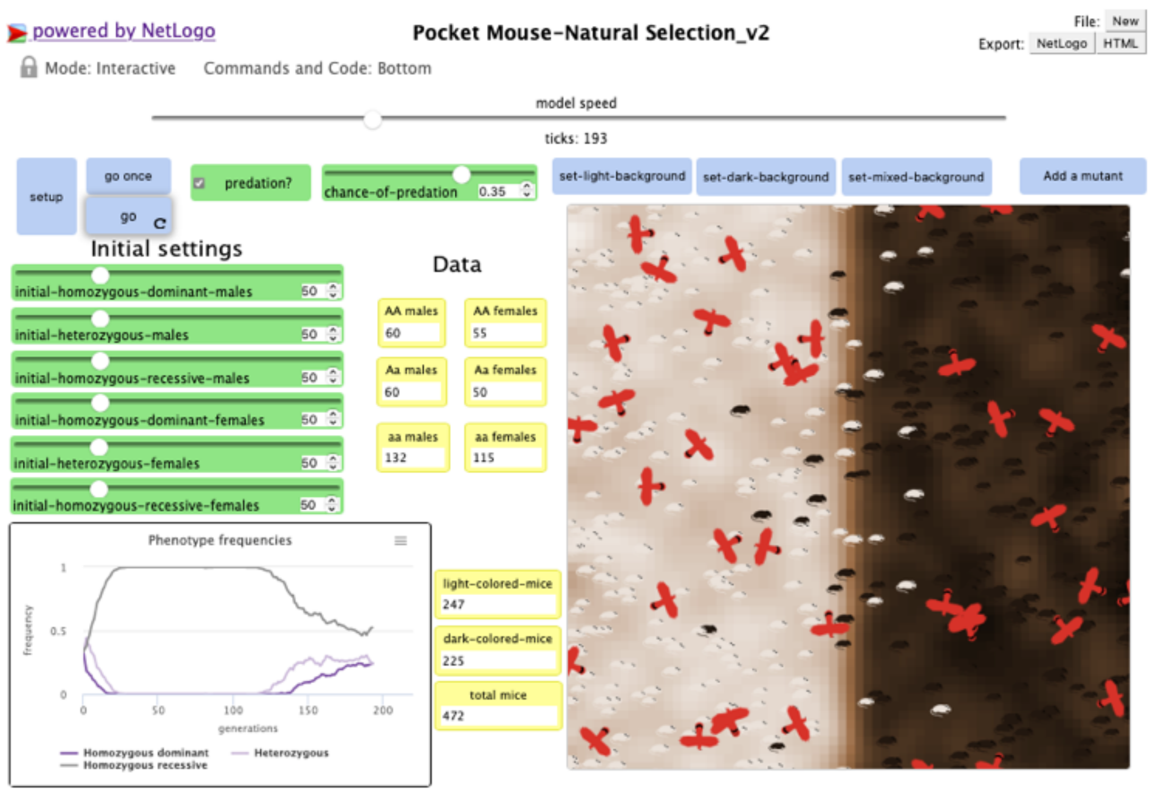


Figure 2-1 A Rock-Pocket-Mice ESM (Dabholkar & Wilensky, 2020) to learn about how populations change due to natural selection (see Chapters 5 and 6)

Whereas in traditional learning settings, students employ heuristics and formulae given to them by authority, agent-based representations in ESMs enable students to author their heuristics and formulae derived from their modeling experience (Wilensky, 2020). Agent-based restructurations have been demonstrated to be pedagogically effective in supporting the learning of several complex natural phenomena in science education (e.g., electric current, resistance, crystallization, temperature, pressure, evolution) (Levy & Wilensky, 2009; Sengupta & Wilensky, 2009; Wagh et al., 2017; Wilensky, 1999a; Wilensky & Novak, 2010). Agent-based representations provide visual access to agent behaviors and interaction as well as system-level emergent patterns. This reduces cognitive and perceptual limitations regarding thinking in levels (Wilensky & Resnick, 1999) and allows students to reason about emergent patterns at the system level by observing behaviors of agents (Goldstone & Wilensky, 2008). For example, Chapter 3 discusses an ESM which allows students to visualize DNA-protein interactions inside a bacterial cell and emergent patterns regarding gene-regulation and changes in the energy of a cell. The power properties of the agent-based restructurations (Wilensky & Papert, 2010) allow designers of learning environments to model an emergent phenomenon by encoding agent-level interactions. In the GenEvo ESM discussed in Chapter 3 and 4, I have modeled a phenomenon of gene regulation in bacterial cells (Lac Operon (Müller-Hill, 1996)), genetic drift and natural selection using agent-based restructurations. The other four core properties of restructurations – cognitive, affective, social and diversity affect the learnability of the modeled disciplinary ideas. In Chapters 4 and 6, I discuss how some of these learnability properties of restructurations in support student learning of scientific inquiry practices and disciplinary ideas in ESM-based curricula.

ESMs as Constructionist Microworlds

In this section, I first discuss Papert's powerful idea of constructionism (Papert, 1980) from which I derive inspiration for ESM design. In addition to using agent-based restructurations, ESMs are designed as microworlds for students to explore and investigate phenomena (Dabholkar, Anton, et al., 2018). A microworld, which is a constructionist design concept, is an encapsulated open-ended computational exploratory environment in which a set of ideas can be investigated through interactions that lead to knowledge construction (Edwards, 1995; Papert, 1980).

Constructionism

Constructionism is an epistemological paradigm, a learning theory, and a design framework that harnesses computational technologies for students' engagement in disciplinary meaning-making individually and collectively (Kynigos, 2015; Papert, 1980). As a learning theory, the constructionist paradigm builds on Piaget's theory of constructivism that the new knowledge is built on the foundations of prior knowledge (Ackermann, 2001; Papert, 1980; Piaget, 1970). Constructionism contributes to the theory of constructivism through its unique attention to the ways of facilitating personally meaningful engagement of students to construct knowledge. Constructionist learning environments are designed to support the creation of individual and collective bricolage with computationally supported artifacts, influenced by negotiated changes students make to these artifacts with an explicit emphasis on self-driven production and ownership (Ackermann, 2001; Kynigos, 2015). In Papert's own words – "*Constructionism -- ... --shares constructivism's connotation of learning as "building knowledge structures" irrespective of the circumstances of the learning. It then adds the idea that this happens especially felicitously in a context where the learner is consciously engaged in constructing a public entity, whether it's a sandcastle on the beach or a theory of the universe*." (Papert & Harel, 1991). The ESM design attempts to incorporate the following three key ideas from the constructionist design framework: (a) personally meaningful engagement, (b) construction of public entities, (c) expression and validation of ideas through computational microworlds.

Constructionism and ESMs

The Microworlds part of an ESM is inspired by Papert's idea of microworld. In my conceptualization of microworlds, I use the functional definition of microworlds (Edwards, 1995). From the functional perspective, microworlds are conceptualized as encapsulated open-ended exploratory learning environments in which a set of ideas can be explored through interactions that lead to knowledge construction (Papert, 1980; Edwards, 1995). In this conceptualization, a microworld could be (a) a computational model (such as Rock Pocket Mice ESM in Chapters 5 and 6), (b) a pedagogically linked set of computational models, (GenEvo models in Chapters 3 and 4), or (c) a pedagogically linked set of computational modeling platforms (Rollypolly Animal Behavior models that use block-based coding in NetTango Web Chapters 7 and 8). Even though many examples of microworlds are computational, microworlds, by definition, are not necessarily computational. However, all the Emergent Systems Microworlds that I discuss in this dissertation are designed in the form of computational learning environments.[2]

Constructionist learning environments in the form of microworlds have been demonstrated to be effective for learning in several contexts (Brandes & Wilensky, 1991; Edwards, 1995, 1997; Feurzeig, 1986; Noss & Hoyles, 2017; Roschelle, 1991). Early examples of microworlds include Logo-based microworlds for mathematics learning (Edwards, 1997; Feurzeig, 1986; Hoyles & Noss, 1987; Papert, 1980). 'House' is an example of a microworld that was developed to learn ratio and proportions (Hoyles & Noss, 1987, 1992). One of the first Logo based microworlds in a non-mathematics domain, the "dynaturtle", was developed to explore Newton's laws of motion (DiSessa, 1982). TEGO is another example of a microworld that included multiple representations to study transformative geometry. Wilensky (1993) expanded

[2] In the part that follows in this chapter and in the rest of the dissertation, when I mention a microworld, I mean a computational microworld.

the functional definition of microworlds to computational models developed for "playing with and exploring large ensemble behavior." GasLab (formerly called *LogoGas in (Wilensky, 1993)) is an early example of such microworlds (Wilensky, 1999a). Wilensky and Resnick further developed StarLogo and StartLogoT as massively parallel Logos to develop microworlds to explore aggregate behavior in several contexts such as termites, and traffic jams (Resnick, 1997; Wilensky, 1997b). Wilensky developed NetLogo (1999b) as an agent-based modeling platform that hosts hundreds of microworlds in different domain areas which have been developed to model emergent patterns for research as well as educational purposes (See a list of such microworlds in Table 2-1).

Table 2-1 Examples of ESM-based curricular units

ESM	**Disciplinary Domain**	**Citation**
MaterialSim	Crystallization, Casting, Grain Growth	(P. Blikstein & Wilensky, 2005)
Connected Chemistry	Kinetic Molecular Theory	(Levy & Wilensky, 2009; Stieff & Wilensky, 2003)
NIELS	Drude's free electron theory	(Sengupta & Wilensky, 2009)
PopBio	Population Dynamics	(Michelle Wilkerson-Jerde et al., 2015)
Redesigning Your City	Urban Planning	(Hjorth & Wilensky, 2014)
GenEvo	Genetic regulation, Genetic drift, Natural Selection	(Dabholkar & Wilensky, 2016a)
Mind The Gap	Economic Inequality	(Guo & Wilensky, 2018)
SimEvolution	Evolution	(Wilensky & Novak, 2010)
EvoBuild	Natural Selection	(Wagh & Wilensky, 2018)

In an ESM-based curriculum, a learner is expected to manipulate computational objects in the form of agents, and execute specific operations instantiated in a microworld in order to study specific aspects of a phenomenon. Such manipulations result in observable changes in the microworld. As learners observe those changes, they receive feedback about agent behaviors and changes in the system. Learners use this feedback to induce or discover the properties and functioning of the system as a whole. Through this process, they self-correct or 'debug' their understanding of the domain to develop new powerful ideas (Papert, 1980). In an ESM-based science unit, students are encouraged to actively construct knowledge in a computational microworld using science practices similar to those that scientists use to construct knowledge about the real world.

ESSs and ESMs

The ESM design approach is inspired by the ESS (Emergent Systems Sandbox) design approach (Brady et al., 2015). ESSs are agent-based computational models that are designed as sandboxes that provide learners with ''entity''-level construction primitives. Students can use these primitives to directly interact with entities in the sandbox space. They can then combine, arrange, and manipulate the primitives to construct complex systems and explore the emergent properties of those systems.

The ESS approach focuses on designing agent-based computational models of emergent systems by preserving the fundamental established scientific principles or paradigms. This makes the behaviors and patterns arising from interactions between computational agents in ESSs consistent with real-world systems. The disciplinary ideas that students learn using an ESS are therefore consistent with the real-world established understanding of those ideas. ESSs and ESS-based curricula are designed for students to use the models as open-ended sandboxes to discover some of those principles.

ESMs, like ESSs, are also agent-based computational models that preserve scientifically established principles regarding the modeled phenomenon. However, ESMs are designed for learners to engage in more structured investigations with scientific models compared to ESSs. The purpose of these structured investigations is to engage students in specific epistemic activities. The constructionist nature of an ESM allows students to conduct model-based investigations of their interests as they participate in the epistemic activities. In Chapters 4 and 6, I present analysis of student participation and their learning with ESMs.

Core Design Features of ESMs

Based on an analysis of previously designed ESMs, I have identified a list of seven core design features of ESMs. This list is not intended to be exhaustive or final. Rather, it is intended to serve as a set of guidelines. I have developed this list based on the design principles that I have used from the previously designed agent-based constructionist curricula (see table 2-1) for designing ESMs. The list includes ideas from constructionist design principles for designing microworlds (Edwards, 1995; Papert, 1980; Wilensky, 2003) which are about visualization and manipulation of computational objects-to-think-with and agent-based modeling principles (Wilensky, 2001; Wilensky & Rand, 2015) which are about designing agent-based representations to understand and investigate emergent phenomena. I have used this list of design features to guide the ESM designs that I present in this dissertation. Throughout my dissertation, you will find the implementation of these design features and discussion about their pedagogical affordances.

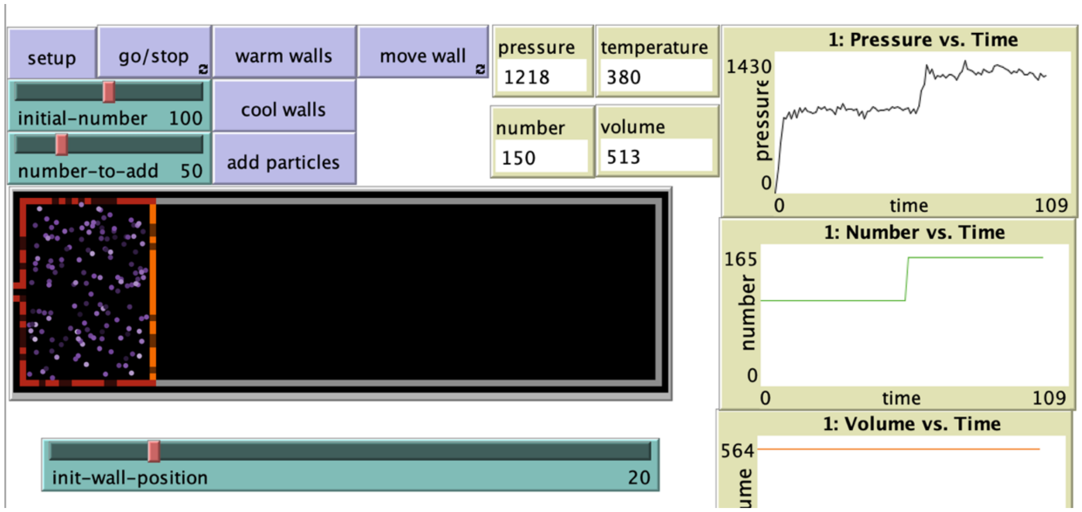


Figure 2-2 A NetLogo model (Wilensky, 2005) in an ESM-based curriculum – Connected Chemistry (Wilensky et al., 2004)

In the following part, I explain the core design features of ESMs and give examples of each feature from the Connected Chemistry curriculum (Wilensky et al., 2004), which is one of the early ESM-based curricula. Figure 2-2 shows a NetLogo model which is part of the Connected Chemistry curriculum. The ESM is designed for users to investigate and learn about the behavior of gases. The model in Figure 2-2 explores the relationship between the variables in the ideal gas law (number of particles, container volume, gas pressure, and gas temperature). Most of the models in this ESM-based curriculum use the same basic rules for simulating the behavior of gases. Each model highlights different features of how gas behavior is related behavior of gas particles. In all the models, the behavior of gas particles is coded such that they move and collide, both with each other and with objects such as walls based on basic laws of classical mechanics. In this model, users can change the volume of the gas container, the number of particles, and the temperature of walls, which in turn change the temperature of the gas by transferring heat to colliding particles. This model and the ESM-based curriculum are designed to help students study properties of gas – such as gas pressure by observing and manipulating the dynamics of the gas particles that lead to increases and decreases in pressure.

In the following table (Table 2-2), I explain each of the design features and provide examples from the ESM-based curriculum Connected Chemistry (Levy & Wilensky, 2009; Stieff & Wilensky, 2003; Wilensky et al., 2004).

Table 2-2 Core design features of ESMs

Feature	**Explanation**	**Example**
Visualization of agent properties and interactions	Users can see certain agent properties and interactions	Agents in Connected Chemistry ESM are gas particles. Users can choose to visualize the speed of gas particles as shades of different colors. Additionally, users can observe collisions of gas particles with the wall of the container, which is related to the pressure of the gas as an emergent property.
Visualization of system-level emergent structures, properties or patterns	Users can observe system level structures, properties or patterns	Pressure is an emergent property of a gas in the Connected Chemistry ESM. Users can visualize movement of the wall in the container and the change in the volume of the chamber containing the gas.
Ability to manipulate agent and system properties and interactions	Users can easily change agent and system properties that are related to emergent patterns using the widgets provided in the microworld [3]	Users can change the number of gas particles (system composition), the temperature of walls which affects the speed of each molecule, the volume of the container by setting the initial position of the wall or moving it.
Ability to track system-level changes	Users can track systems level changes by widgets such as monitors and graphs	Figure 2-2 shows four monitors – temperature, pressure, number, and volume. It also shows three graphs – pressure vs time, volume vs time, number vs time.
Agent-based coding	ESMs are coded by modeling agent behaviors	The underlying computational code in this model codes for the behavior of gas particles – by modeling the transfer

[3] ESM designers' choice of variables creates affordances and limitations for learners to manipulate the system and its constituents

	which result in emergent patterns at the system level[4]	of their energy, momentum as they collide. The relationship between Pressure, Volume, Temperature, and Number is not explicitly modeled in the code, rather it emerges because of interactions of gas particles with themselves and with the wall of the container.
Micro-level validity	An ESM designer should validate micro-level rules of agent properties and agent interactions by comparing model behavior with established scientific ideas.	An ESM designer needs to ensure that the computational implementation of classical mechanics regarding particle collisions is valid.
Macro-level validity	An ESM designer should validate micro-level emergent patterns	An ESM designer needs to ensure that the emergent patterns regarding the relationship between Pressure, Volume, Temperature, and Number are valid.

ESM-BASED CURRICULA

ESM-based curricula combine these two design ideas – agent-based restructurations and constructionism by, 1) focusing on emergent phenomena that are modeled using agent-based modeling and 2) using constructionist design principles with microworlds. I have integrated constructionist design principles in the design of ESM-based curricula discussed in this dissertation through the following three key ideas:

(a) Personally meaningful engagement,

(b) Construction of public entities, and

(c) Expression and validation of ideas through computational microworlds.

[4] For details about agent-based coding please refer to Wilensky and Rand's book on this topic (Wilensky & Rand, 2015)

For example, in the ESM-based curriculum about genetics and evolution discussed in Chapters 3 and 4, students were asked to come up with their research questions, design and conduct computational experiments to investigate those questions and share their research findings with their peers. The peers asked questions about students' research questions, experimental design, and the validity of the presented evidence to support the claims. The teacher then conducted discussions and the classroom, as a learning community, collectively established findings of emergent patterns and evolved epistemic practices of establishing such findings. This iterative sequence of questioning, investigating, sharing, and discussing allowed students to identify and investigate personally meaningful aspects of the system. Their research projects were the public entities that they shared with the classroom community. The GenEvo ESM (see Chapters 3 and 4) was designed such that students could conduct a variety of experiments to investigate different aspects regarding gene regulation in a bacterial cell, change in a bacterial population because of genetic drift and natural selection, and the importance of gene regulation for natural selection. This breadth of design of GenEvo ESM and the ESM-enabled constructionist pedagogical strategies of the teacher (see Chapter 4) allowed students to express and validate their ideas through the GenEvo microworld.

Restructurated science learning with ESM-based curricula

ESM-based curricula discussed in this dissertation are designed to restructurate learning of science in the classroom setting. These curricula are designed for students to explore and learn about scientific phenomena using ESMs. ESMs support students in exploring, investigating, and sharing virtual models of systems that exhibit emergent phenomena. These ESM-based curricula are designed for engaging students in actively constructing knowledge in a computational microworld using scientific inquiry practices similar to those that scientists use to construct knowledge about real-world phenomena (Figure 2-2). There are two ideas that are central to learning using ESMs and ESM-based curricula, which we call, 'big-M' Models and 'little-m' models (Brady et al., 2015).

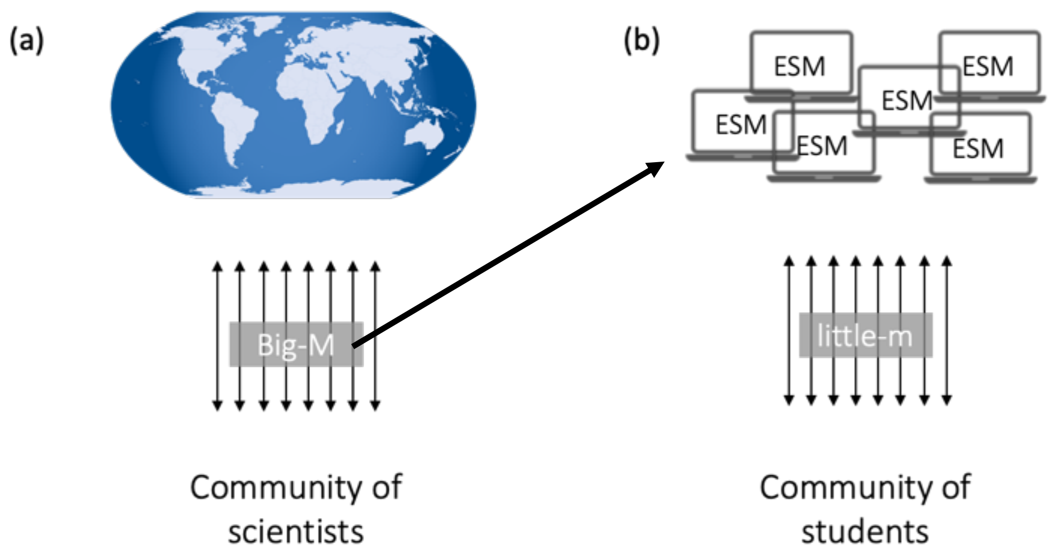


Figure 2-3 Knowledge co-construction using ESMs: (a) A community of scientists engaging in specific practices to construct knowledge about the world in the form of explanatory models (Big-M) (b) A community of students potentially engaging to construct knowledge (little-m) that is isomorphic to the process represented in (a). The arrow indicates that the Big-M models established by a community of scientists are used derive rules for designing ESMs.

Figure 1a represents the broad purpose of the endeavor of the community of scientists to, construct knowledge about general patterns and rules for how and why the natural world works in the ways that it does by generating explanatory models (Berland et al., 2016; Lehrer & Schauble, 2012; Russ et al., 2008). We call these models Big-M models. Big-M models are fundamental scientific paradigms (Kuhn, 2012) that form the basis for the design of ESMs. Every agent-level entity in the microworld follows the rules that are specified by the Big-M model (micro-level validity). Also, macro-level validity is established by verifying that the emergent patterns in an ESM are consistent with the Big-M model. An ESM is designed such that it captures core aspects of the Big-M principles for students to engage meaningfully and authentically in investigating the scientific phenomena that is under investigation. For example, since in the Rock Pocket Mice ESM is designed for students to learn about natural selection, the Big-M principles that are incorporated are Mendel's laws of inheritance, and camouflage affecting chances of predation (see Chapters 5 and 6), however the ESM does not incorporate other unrelated details such as sex determination mechanisms in mice.

ESM-based curricula are designed for engaging students in developing their little-m models by using scientific inquiry practices to investigate modeled emergent phenomena in ESMs, which gradually nudge their intuitions into alignment with the Big-M model. Little-m models can be thought of as students' contextual understandings in the form of personal hypotheses or theories about how systems function (Figure 2-2 (b)). Little-m models can be mental models or diagrammatic models, or computational models. As students construct their little-m models by exploring, observing, and experimenting with an ESM, the consequences of the Big-M rules become salient to them. For example, students investigate how populations change because of natural selection using the Rock Pocket Mice ESM. Students develop their contextual understanding in the form of little-m models regarding the phenomena of population change of mice due to natural selection. An example of Jane's investigation using the Rock Pocket Mice ESM (discussed in Chapter 6) illustrates how she corrected her understanding regarding the mechanism of inheritance of fur coat color of rock pocket mice in the ESM.

This argument is also evident in a series of prior studies started by Wilensky (Wilensky, 1999a, 2003), continued by Stieff and Wilensky (Stieff & Wilensky, 2003), and further developed by Levy, Wilensky, and colleagues (Gobert et al., 2011; Levy & Wilensky, 2008, 2009). Using agent-based computational models of chemical systems, the authors discuss the importance of developing conceptual knowledge that connects agent-level particle behaviors and interactions with emergent patterns formed at the aggregate level of the system. Though the authors do not call these learning environment ESMs, many of the learning environments designed and used in these studies fall into the category of ESMs because of the use of agent-based restructurations and constructionist design features incorporated in designing these learning environments as microworlds. The big-M model in these learning environments in the particulate nature of the matter and kinetic molecular theory regarding temperature and pressure of gases. These studies involve several pedagogical approaches using these ESMs for students to construct their context-specific little-m models and develop deeper levels of understanding of processes that fuel emergent behaviors and states as well as core ideas in the disciplinary domains.

Agent-based representations and constructionist design principles in ESMs restructurate science learning of various domains. My work builds on prior work of designing and teaching such restructurated curricula that have been referred to with different names, such as agent-based modeling inquiry or curriculum of emergent multi-agent-based computational models (Table 1). In my dissertation, I use the term ESM to capture these and other constructionist curricula that emphasize the key role of the emergent systems perspective using microworlds. Based on the constructionist pedagogical principles used in these curricula and other curricula that used constructionist microworlds, I have developed a set of pedagogical principles which form core design features of ESM-based curricula that I have designed for my dissertation work (Paulo

Blikstein & Wilensky, 2010; Edwards, 1995; Hoyles & Noss, 1987; Levy & Wilensky, 2008; Papert, 1980; Sengupta & Wilensky, 2009; Wagh et al., 2017; M. Wilkerson-Jerde & Wilensky, 2010).

Core design features of ESM-based curricula

Similar to the core design features for ESM design that I discussed before; I have used the following design features for designing ESM-based curricula:

1. **Expression and validation of ideas using an ESM**: In an ESM-based curriculum students first play around with the ESM and identify aspects related to the modeled phenomenon that they find puzzling and want to investigate systematically. They are asked to express their ideas concretely by stating a research question and developing an experimental design to investigate that question. For example, in Chapters 3 and 4, if a student is interested in investigating how a population changes because of natural selection, they need to formulate a research question about how a population of rock pocket mice living in specific environmental conditions would change over time. Then they are asked to design computational experiments to test their claims.
2. **Rigorous computational experimentation to construct explanations**: After designing experiments, students are asked to perform computational experiments and record data that would help them construct explanations regarding their research question.
3. **Reasoning across levels**: In an ESM-based curriculum, students investigate an emergent phenomenon, which requires constructing explanations about how system-level emergent patterns arise through agent interactions by reasoning across levels. Some ESM-based

curricula include questions to explicitly draw student attention to reason across levels to concretely construct agent-based explanations of emergent phenomena.

4. **Sharing ideas and collectively establishing new knowledge**: In an ESM-based curriculum, students are asked to present their research questions, experimental designs, analyses, and conclusions. Other students are encouraged to ask questions to seek explanations and challenge the sufficiency of evidence or issues with experimental designs to evaluate the evidence-based claims.

CONCLUSION

In this chapter, I discussed how agent-based representations and constructionist design principles are foundational to the design of ESMs and ESM-based curricula. Agent-based representations have restructuration properties (Wilensky & Papert, 2010) (discussed in detail in Chapter 1) which improves the learnability of a phenomenon from cognitive, social, affective, and diversity perspectives. Constructionist design principles influence the design of ESM-based learning activities and corresponding features of ESMs that enable those learning activities. ESMs and ESM-based curricula have a long history in the field of learning sciences (see Table 1). I discussed core design features of both ESMs and ESM-based curricula, which can serve as guidelines for designing new ESM-based curricula. I build on prior work of designing ESM-based curricula and contribute to it by investigating learning with ESM especially focusing on epistemic agency and epistemic expansiveness (Study 1 – Chapters 3 and 4), epistemic connections among practices and disciplinary ideas (Study 2 – Chapters 5 and 6), and co-designing CT-integrated curricula (Study 3 – Chapters 7 and 8).

In my dissertation, I attempt to characterize learning facilitated by the ESM design approach. My work contributes to understanding student learning, teacher learning, and researcher learning. ESMs and ESM-based curricula facilitate *student learning* through the creation of learning opportunities to support students' epistemic agency in ways that are meaningful to learn disciplinary ideas (Studies 1 and 2). ESM design approach also facilitates *teacher learning* of the following things – emergence, supporting students' epistemic agency, and supporting students' engagement in Computational Thinking (CT) practices (Studies 1 and 3). In Study 1, I present evidence of teaching practices that were enabled because of certain design features of the ESM. In Study 3, I present evidence of how a teacher developed a nuanced understanding of encoding agent behavior in a particular manner and the emergence of system-level patterns because of that. Additionally, Study 3 also discusses how a combination of a co-design approach and the use of ESMs for creating CT-integrated curricula resulted in shifts in the teaching practices of a teacher for supporting student engagement in CT practices. Lastly, ESMs and ESM-based curricula are designed for design-based research (Design-Based Research Collective, 2003). This design-based research also contributes to researcher learning of designing and co-designing curricula to support student and teacher learning of various kinds as discussed before.

Finally, I want to revisit the question - *Why does this design approach deserve a unique name?*

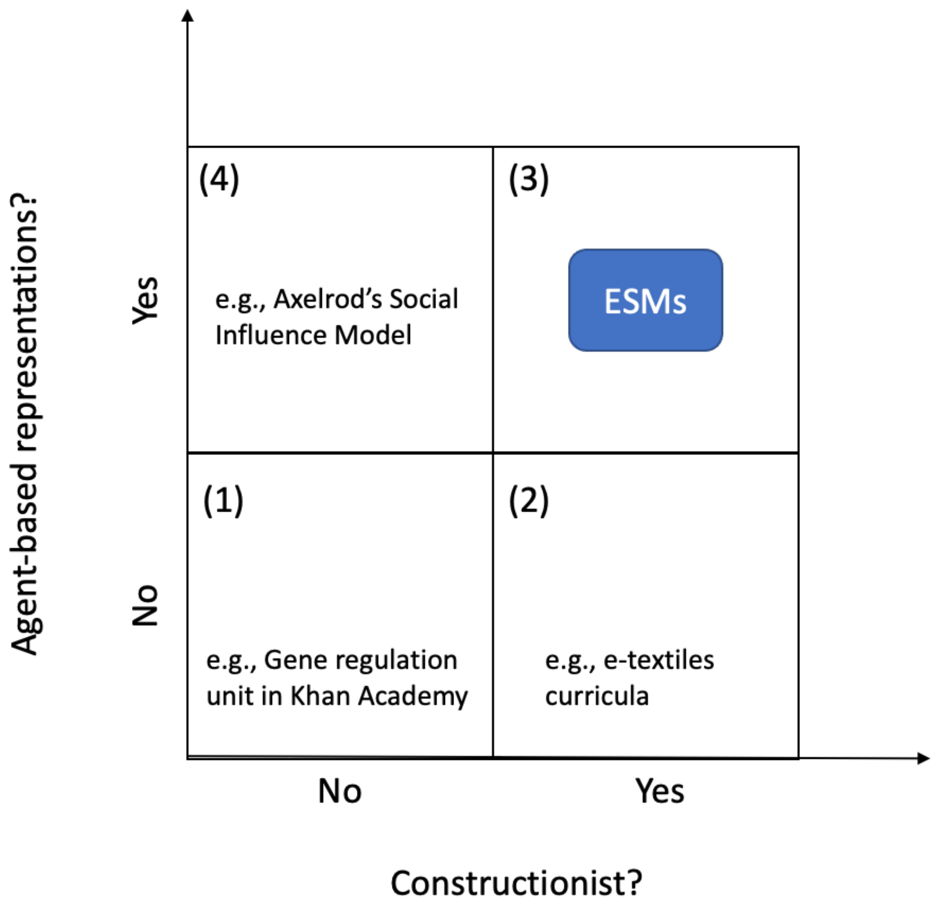


Figure 2-4 Mapping of ESMs in the design space of constructionist learning environments and agent-based models

Consider a two-by-two matrix (Figure 2-3) with one dimension (x-direction), indicating whether a model or a learning environment is constructionist or not, and another dimension (y-direction) indicating whether a model or a learning environment includes agent-based representations or not. I have constructed this space for models and learning environments. Some models are designed to be learning environments for classroom education, constructionist or otherwise, but there are many models that are specifically designed to understand and investigate a phenomenon. So, not all agent-based models are designed to be traditional classroom learning environments. For example, Axelrod's Social Influence Model (Axelrod, 1997) is designed for researchers to learn about a mechanism of convergent social influence that can result in global polarization.

Square (1) in Figure 3 consists of models and learning environments that are neither agent-based nor constructionist. An example of such learning environments is a unit about Operons and Gene regulation[5] in Khan Academy. This unit uses video that has diagrammatic representations and the voice of a person explaining the roles of various components of a gene regulatory system. I compare this unit and other similar units with a restructurated GenEvo curriculum in Chapter 3. Square (2) in this figure consists of constructionist learning environments that do not use agent-based representations. Examples of such learning environments include e-textiles, which are fabric-based items with electronics sewn into them with conductive thread (Buechley et al., 2007; Fields et al., 2021; Peppler, 2013). In such learning environments, students sew electric circuits consisting of different components such as microcontrollers like LilyPad Arduinos and LEDs and write computer programs to make their circuits behave in particular ways. Quadrant (3) in this figure uniquely includes constructionist microworlds that contain agent-based representations, which I call ESMs. Examples of prior ESMs are included in Table 1. In this dissertation, I discuss three newly designed ESMs (Chapters 3, 5, and 7). The square (4) in the figure consists of the models that use agent-based representations but are not designed as constructionist learning environments. Examples of such models include Axelrod's Social Influence Model that I mentioned before, as well as agent-based models in several disciplinary domains that I discussed before such as ecology (DeAngelis,

[5] https://www.khanacademy.org/science/biology/gene-regulation/gene-regulation-in-bacteria/v/operons-and-gene-regulation-in-bacteria

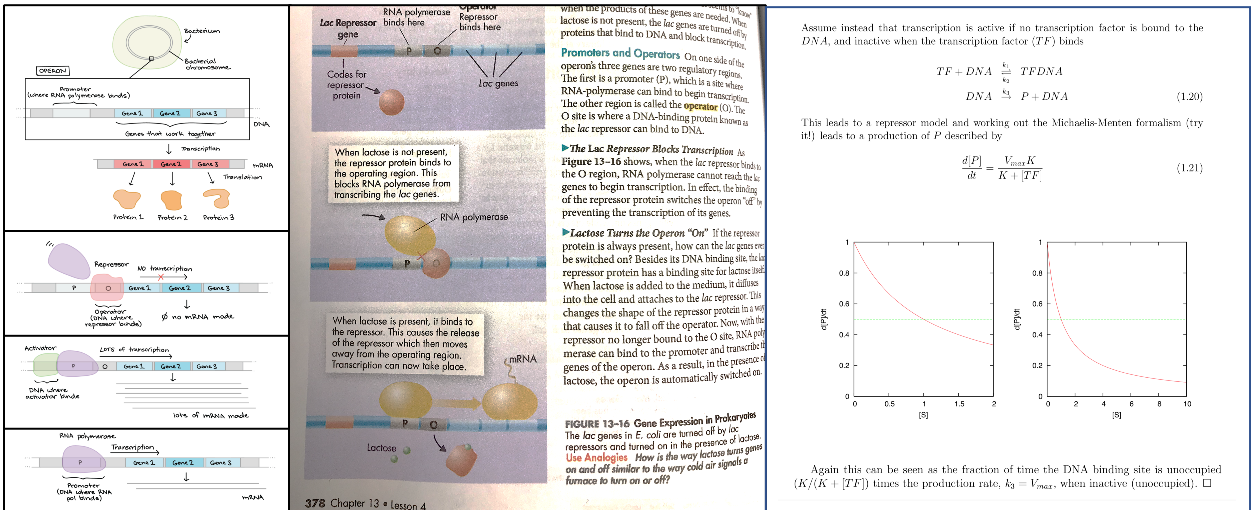
2018; DeAngelis & Mooij, 2005; Grimm & Railsback, 2005), the social sciences (Axtell & Epstein, 1994; Epstein & Axtell, 1996), and economics (Tesfatsion, 2002).

Chapter 3: GenEvo- An ESM to learn about genetics and evolution

Summary: This chapter describes an ESM that is designed to study molecular mechanisms of gene regulation and the evolution of populations. It is modeled based on an established understanding of the mechanisms of gene regulation of the *lac operon*. In this chapter, I first give an overview of the ESM that consists of four related models. These models are divided into three sets: (1) The first model, Genetic Switch, is about molecular interactions between proteins and regions of DNA inside a bacterial cell, (2) The second and third models are about changes in a population of bacterial cells in an environment with a limited supply of sugar, (3) The fourth model combines the cellular and the population models to allow simultaneous visualization across three levels of organization, molecular, organismic and population. In this chapter, I describe these models in detail, focusing on the agents, their behaviors and interactions, and the emergent patterns that users can investigate using this ESM. Finally, I present examples of curricular activities that use one of the ESMs for students to engage in investigating the phenomena related to gene regulation and evolution.

OVERVIEW

The nature of biology research has changed significantly with the incorporation of newer technological tools and computational research methods. However, most of the biology instruction in school and university courses involve the use of static models that explicate molecular interactions as deterministic processes (Figure 3-1 (a) and 3-1 (b)) or use differential equations-based representations that require mathematical sophistication to understand the core ideas (Figure 3-1 (c)). None of these existing structurations explain how simple biochemical interactions between these molecules allow a cell to make emergent complex decisions and perform complex functions, such as which molecular machinery (enzymes) to produce depending on change in environmental stimuli. Our restructuration of these cutting-edge ideas in modern biology with the ESM-based GenEvo curriculum involves simultaneous visualization of computational representations of biomolecules and their agent-level interactions and system-level decisions and behavior of an organism.



(a) Diagrams explaining molecular mechanisms of genetic regulation on Khan Academy (2018)	(b) Diagrams explaining molecular mechanisms of genetic regulation in Miller & Levin Biology Teacher's Edition (2018)	(c) Differential equations explaining molecular mechanisms of genetic regulation in a University level course (Jansson, 2018)

Figure 3-1 Molecular mechanisms of genetic regulation through existing structurations of molecular biology (A diagram from (Wilensky, 2020), pg. 296)

An explanatory video on Khan Academy ("Overview: Gene Regulation in Bacteria." Khan Academy, 2018) uses words like operon, trans-splicing, coordinately regulated genes, RNA polymerases, which are highly relevant but can be new jargon for students to understand. Other words in the video, like translation, gene expression, inducible, regulatory, promoter, are more familiar words, however, have highly specific contextual meaning. Most of the instruction is focused on the meanings of these words. Even though understanding such vocabulary is required, it is hardly sufficient to develop a deep understanding of these ideas which are highly central to modern biology. Also, such explanations make it sound like these processes are deterministic; whereas, in reality, these processes are highly stochastic, which results in emergent predictable patterns at the cellular and organismic level. In our work on this restructuration, we have created an ESM-based curriculum, called GenEvo (Dabholkar & Wilensky, 2016a).

THE GENEVO ESM

This curriculum incorporates a series of computational models designed using NetLogo (Wilensky, 1999a), an agent-based modeling platform that has been used for research work regarding emergent systems as well as for educational purposes such as designing educational curricular units or for student exploration and construction of models.

All of the computational models in the GenEvo curriculum are designed using the agent-based perspective of modeling emergent systems. In each model, the agents and their behaviors at the micro-level are computationally coded. As agents interact with each other and with their environment, it results in emergent patterns at the macro-level (Wilensky, 2001; Wilensky & Resnick, 1999). Students can observe both, the interactions at the agent-level and patterns at the system-level. In this curricular unit, the emergent properties of biological systems include genetic regulation, carrying capacity, genetic drift, and natural selection. All these models are created using agent-based modeling software, NetLogo (Wilensky, 1999b).

1. Genetic Switch

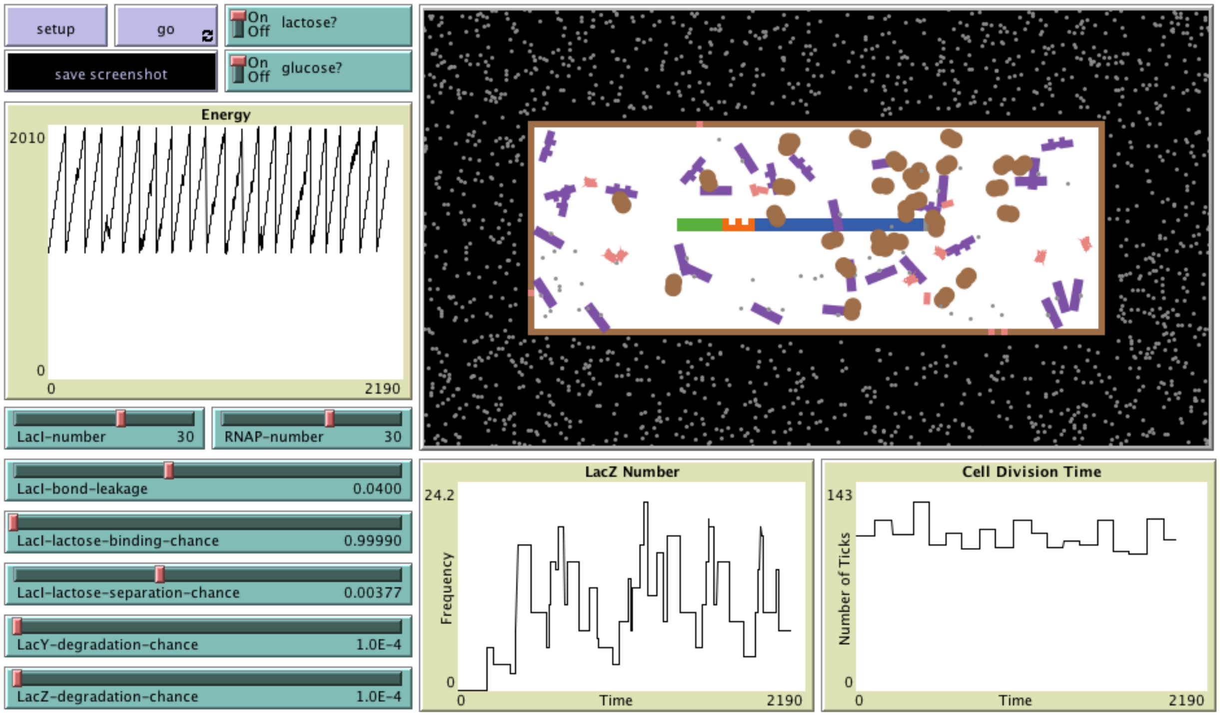


Figure 3-2 A screenshot of *GenEvo 1- Genetic Switch* model (Dabholkar et al., 2016)

The first model, called Genetic Switch, is a model of a bacterial cell that simulates a complex phenomenon in molecular biology, regarding regulating the production of certain proteins by "switching" (on and off) of genes depending on environmental conditions (Figure 3-2). Specifically, it is a model of the lac operon of a bacterium *E. coli*, in which the agent interactions at the molecular level are coded based on the established understanding of the functioning of lac operon (Müller-Hill, 1996).

2. Genetic Drift

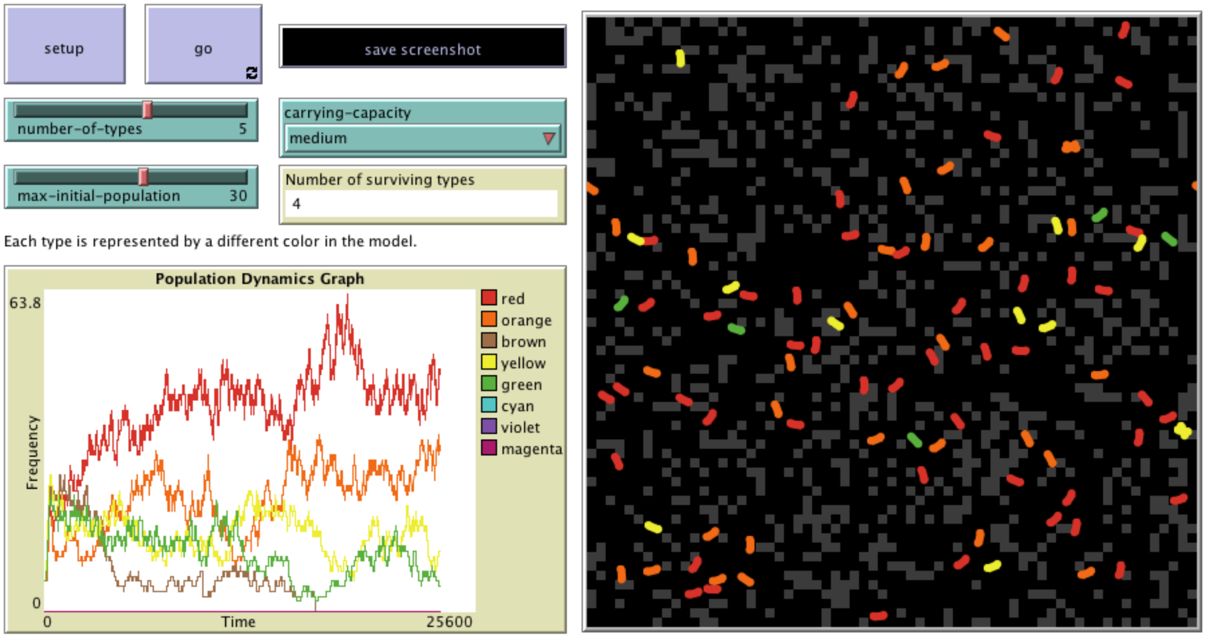


Figure 3-3 A screenshot of *GenEvo 2- Genetic Drift* model (Dabholkar & Wilensky, 2016b)

This model is an example of genetic drift in a population of asexually reproducing bacteria *E. coli*. It starts with a population of several types of *E. coli*, which are represented with unique colors (Figure 3-3). Each of these types or phenotypes, as they are referred to in the context of evolutionary biology, can be considered to have a unique allele of a gene. The model allows users to investigate the process of Genetic Drift[6] that competing phenotypes of *E. coli*, each reproducing with equal likelihood on each turn, will ultimately converge on one phenotype without any selection pressure forcing this convergence.

3. Genetic Drift and Natural Selection

[6] https://en.wikipedia.org/wiki/Genetic_drift

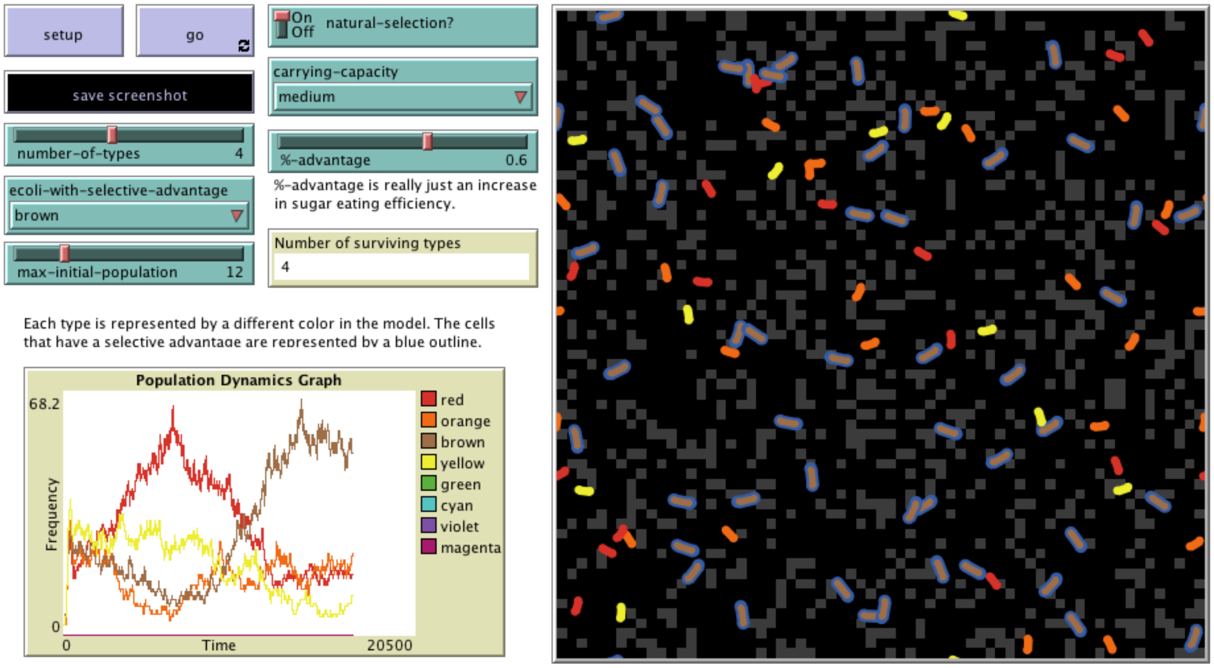


Figure 3-4 A screenshot of *GenEvo 3- Genetic Drift and Natural Selection* model (Dabholkar & Wilensky, 2016c)

This model is an extension of the previous model, which allows for the exploration and comparison of two different mechanisms of evolution: natural selection and genetic drift (Figure 3-4). Similar to the previous model, this model also starts with different phenotypes of *E. coli*, each with a different trait value, represented by different colors. However, a user can create a scenario in which natural selection can take place by (a) selecting a phenotype with a selective advantage, (b) determining the extent of the selective advantage, and (c) turning *natural-selection?* ON, which increases the efficiency of energy production of the selected phenotype based on the value of %-advantage. So, when *natural-selection?* is ON, the phenotype of *E. coli* with the selective advantage gains more energy from sugar in a given time unit. This results in faster reproduction by those cells, thus that phenotype becomes the most prominent phenotype in the population over time (See brown colored cells in Figure 3-4).

4. Competition

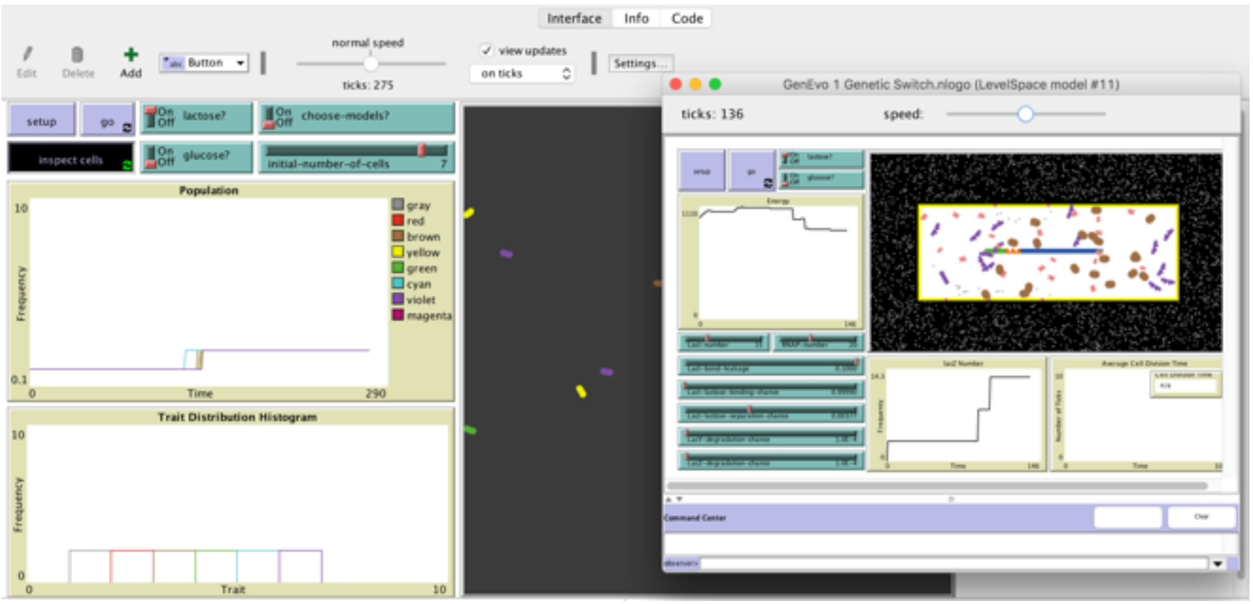


Figure 3-5 A screenshot of *GenEvo 4 – Competition* model. This model combines the Genetic Switch, and the population level models to simultaneous visualize and reason across three levels (Dabholkar & Wilensky, 2016d)

This is a population model of competition between cells in a population in which each cell is another NetLogo model (Figure 3-5). It uses the LevelSpace extension of NetLogo (Hjorth et al., 2020). Each cell in the main population model is controlled by the Genetic Switch model (Model 1 in the curriculum). In LevelSpace terminology, the population model is called a parent model and each cell is a child model. Depending on the selected options, each cell model begins with a different set of initial parameters. Then, each cell model simulates the DNA-Protein interactions in the lac operon of E. coli. In the population model, we see the competition for resources between these cells. When a cell's energy level doubles from its initial level, the cell produces two daughter cells that inherit the cell's genetic and epigenetic information. The cells become 'fitter' in terms of energy production using sugar and maintaining the energy cost of protein production by turning on and off the genes faster, resulting in a faster increase in the energy levels, and a faster growth rate. Because the molecular interactions in each cell model are stochastic, the population model displays the effects of both natural and statistical selection (genetic drift).

THE GENETIC SWITCH MODEL

In this section, I explain the first model in greater depth in terms of the agents in the model, their modeled behaviors and interactions, and the emergent patterns. This model simulates a complex phenomenon in molecular biology: the "switching" (on and off) of genes depending on environmental conditions. Through molecular interactions between specific regulatory proteins and specific DNA sequences, each regulated gene is turned on or off in response to environmental stimuli. The genetic switch mechanism of the lac operon is responsible for the uptake and digestion of lactose, when lactose is the preferred energy source in the environment (Alberts et al., 2002; Müller-Hill, 1996). It regulates the synthesis of the enzymes, lactose permease (LacY) and beta-galactosidase (LacZ).

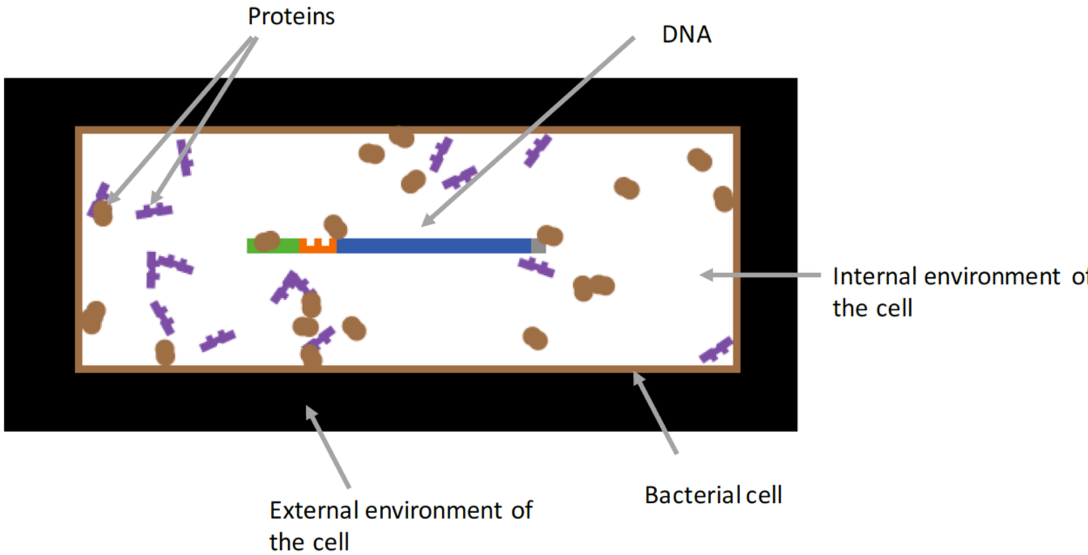


Figure 3-6 Computational representations of agents (DNA and proteins) and the environment in the *GenEvo 1 - Genetic Switch* model

AGENTS

There are four protein agents, which are dynamic moving agents in the model, typically referred to as turtles in NetLogo.

1. RNAP – These are RNA polymerases that synthesize mRNA from DNA. These are represented by brown blobs in the model. This model does not include mRNA.
2. LacI – The purple-colored shapes in the model represent a repressor (LacI proteins). They bind to the operator region (see below) of the DNA and do not let RNAP pass along the gene, thus stopping protein synthesis. When lactose binds to LacI, they form LacI-lactose complexes (shown by a purple shape with a grey dot attached to it). These complexes cannot bind to the operator region of the DNA.
3. LacY – These are shown in the model as pink rectangles. They are produced when an RNAP passes along the gene. When they hit the cell-wall, they get installed on the cell-

wall (shown by light patches). Lactose (grey pentagons) from the outside environment is transported inside the cell through these light red patches.

4. LacZ – These are shown as pink-colored proteins. They are present inside the cell. When they collide with a lactose molecule, the lactose molecule is digested, and energy is produced.

There are four regions on DNA which are static agents, defined as a set of patches that are spatial agents in NetLogo.

1. Promoter – This region is indicated by the color green. As an RNAP binds to the promoter region and if the operator is free, it moves along DNA to start transcription.
2. Operator – This region is indicated by the color orange. The repressor protein, LacI, binds to this region and prevents RNAP from moving along the DNA.
3. Operon – This is indicated by the color blue. This is where lacY and lacZ genes are. This model includes only these two genes of the operon.
4. Terminator – This is indicated by the color grey. RNAP separates from the DNA when it reaches this region.

AGENT PROPERTIES

Agent properties in this model are regarding the states of the agent in the system. For example, this model uses Booleans such as on-DNA? to track the state of RNAP. If an RNAP is on DNA, (after it attaches to a promoter), its state-property on-DNA? is set to TRUE. When on-DNA? is TRUE for RNAP, it is asked to move forward on DNA until it reaches the terminator region. After reaching the terminator region, the RNAP dissociates from DNA and its state-property, on-DNA? is changed to FALSE.

AGENT BEHAVIORS AND INTERACTIONS

As RNAP moves along the gene, LacY and LacZ proteins are produced (five molecules per transcription). Translation by ribosomes is not explicitly modeled. Producing and maintaining the protein machinery for a cell takes energy. So as a cell produces proteins and maintains those proteins (RNAPs, LacIs, and LacZs), its energy decreases.

The energy of the cell increases when lactose inside the cell is digested. When the energy of the cell doubles from its initial value, it splits into two daughter cells. Each of these cells has half of the energy of the original cell as well as half the number of each type of protein in the original cell.

The temporal progression in the model is represented in the form of ticks, similar to clock-ticks.

For example,

At each clock-tick,

Each RNAP,

1. Checks if it is near the 'Promoter' region of DNA
2. If it is not near Promoter, it moves a step in a random direction inside the cell
3. If it is near the 'Promoter', it latches on the DNA based on a probability, only if the 'Operator' region is not blocked
4. If it successfully latches on to Promoter, it moves along the DNA
5. If it is on DNA and it reaches the terminator region, it separates from DNA

Similarly, the behaviors of all the other agents in the model are coded. The rules of agent behavior are based on the established understanding of the molecular mechanism of gene regulation in lac operon (Müller-Hill, 1996). The implicit and explicit design choices mentioned earlier were made for pedagogical purposes. For example, only the proteins and DNA regions that are central to the gene regulatory behavior are included in the model. Since binding rates of biomolecules are critical in this process, parameters such as LacI-bond-leakage and LacI-lactose-binding-chance are included in the model.

EMERGENT PATTERNS

Turning on the genetic switch to produce the required proteins is an emergent process. It happens through the molecular interactions of proteins and regions of DNA when lactose is present, and glucose is absent in the environment. The most critical design aspect of this model is that the ‘genetic switch’ behavior of the cell is not directly coded in the model, rather it emerges through interactions of the protein and DNA agents. This genetic switch is responsible for regulating the synthesis of proteins (LacZ and LacY) required to conduct the uptake and digestion of sugar, lactose. In this model, there are two sugars: glucose and lactose. Glucose is “preferred” as an energy source over lactose. When there is glucose and/or no lactose in the surrounding environment, the genetic switch is at an off steady-state. This is because the repressor protein LacI prevents (mostly) the bacteria from producing the proteins, by binding to the operator site of the DNA. In this steady-state, relatively little permease (LacY) and beta-galactosidase (LacZ) are produced. When lactose is introduced to the outside environment, the lactose molecules enter into the bacterial cell through permease proteins (LacYs). Some lactose molecules that enter the cell bind to LacIs, preventing LacIs from binding to the operator site of the DNA. This, in turn, causes more LacYs to be produced. The LacYs get inserted into the cell-wall, which causes more lactose to enter the cell, thus creating a positive feedback loop. The LacZs, meanwhile digest lactose molecules inside the cell to produce energy. The regulatory effects of due to presence of glucose (through cAMP) is only implicitly modeled. The rate at which LacZ and LacY are produced reduces significantly when glucose is present.

Figures 3-7 and 3-8 demonstrate the power properties of agent-based restructurations to model the emergent behavior of the genetic switch. The state of the switch is monitored in terms of the number of LacZ molecules in the cell. The model run includes four environmental conditions that were sequentially created – only glucose, only lactose, only glucose, and glucose + lactose. The genetic switch gets turned as the environmental condition changes from ‘only glucose’ to ‘only lactose’ and it gets turned off when the condition returns to ‘only glucose’. The switch is unaffected when both sugars are present (condition 4).

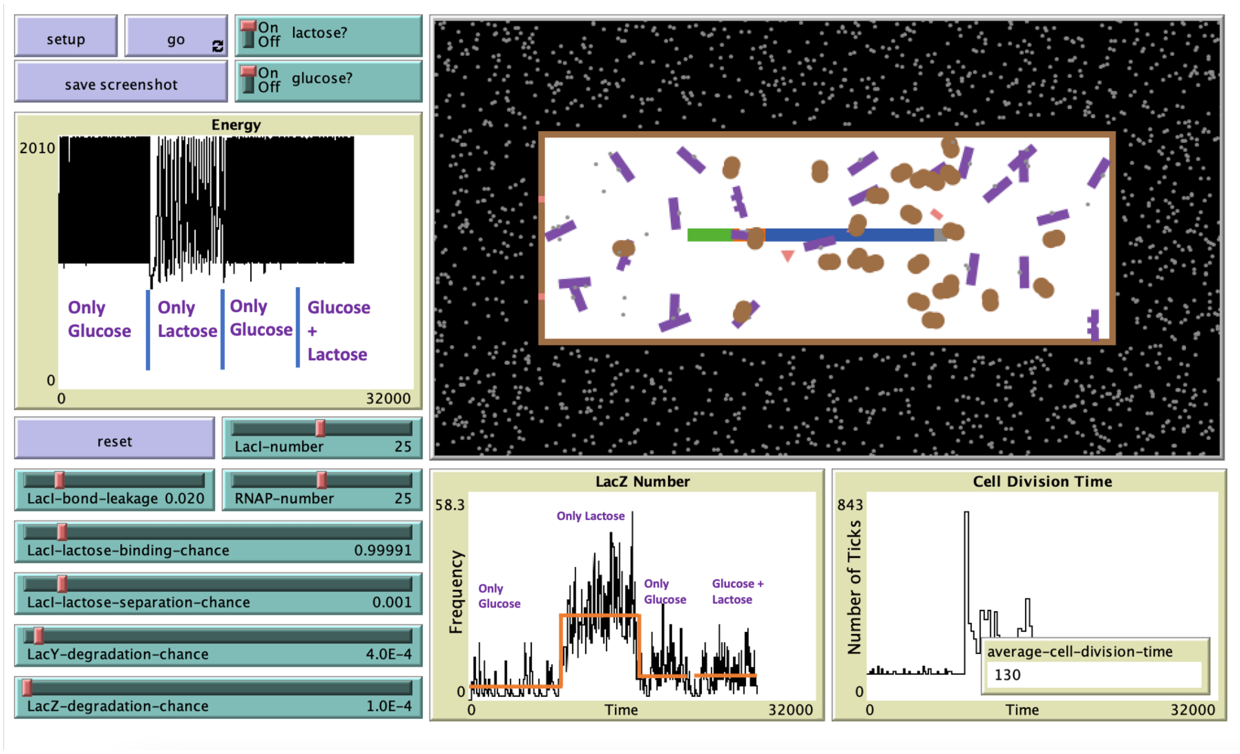


Figure 3-7 A screenshot of the *GenEvo 1 Genetic Switch* model. The model run includes four environmental conditions that were sequentially created – only glucose, only lactose, only glucose, and glucose + lactose. As observed in the graph of LacZ number, the number of LacZ proteins in the cell was high only when the environment contained only lactose.

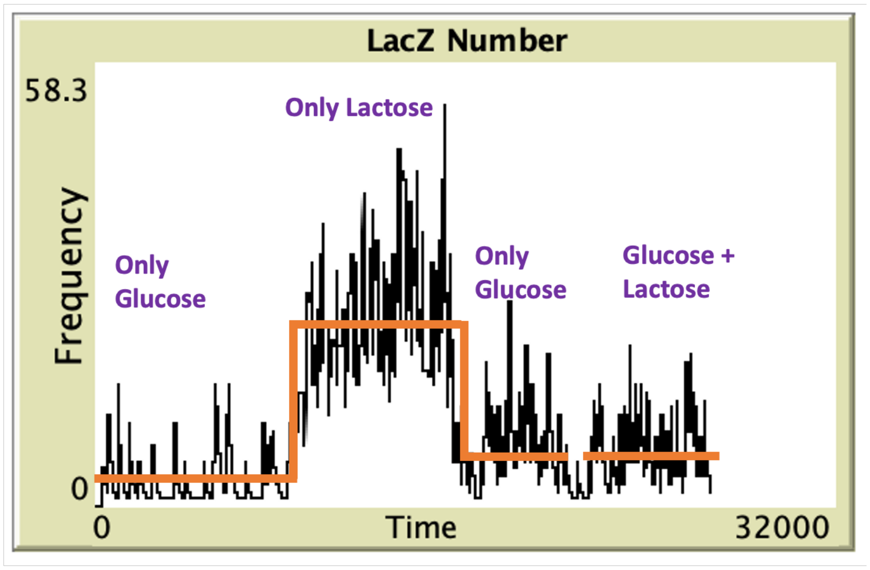
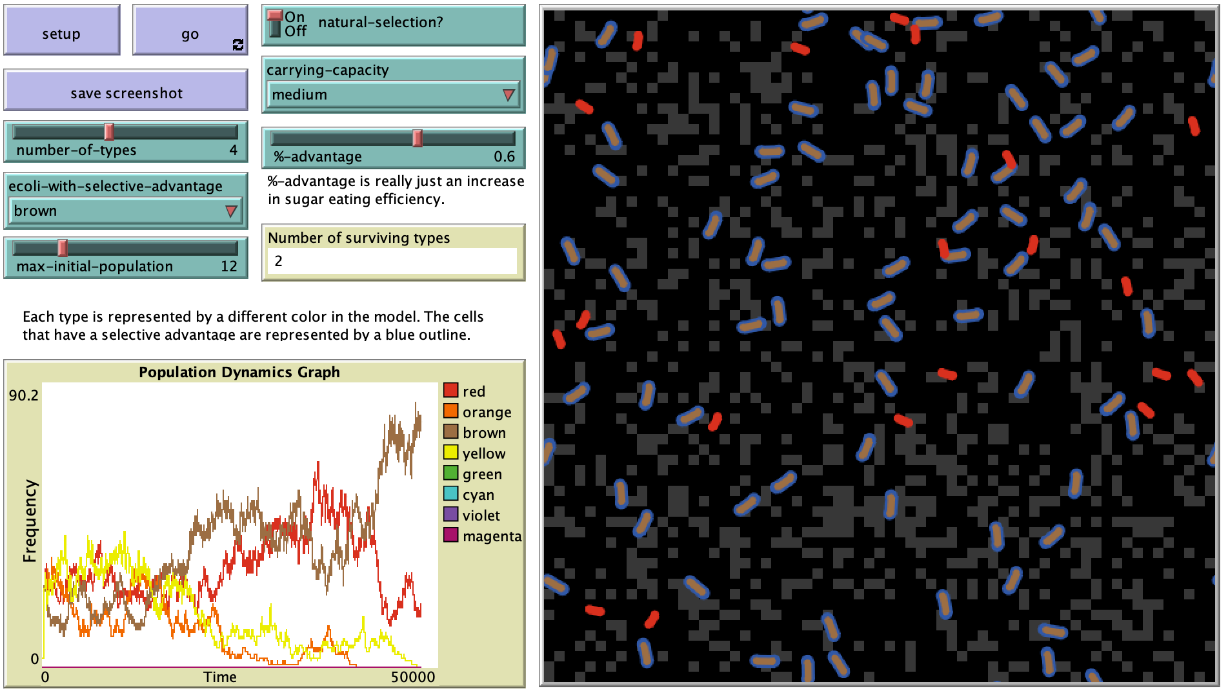


Figure 3-8 A screenshot of a graph of LacZ number in the *GenEvo 1 Genetic Switch* model. The emergent genetic switch behavior is observed in terms of the number of LacZ molecules in the cell. The genetic switch gets turned as the environmental condition changes from 'only glucose' to 'only lactose' and it gets turned off when the condition returns to 'only glucose'. The switch is unaffected when both sugars are present (condition 4).

The Genetic Drift and Natural Selection models

The next set of models combine two micro-evolutionary mechanisms – genetic drift and natural selection.



<u>Figure 3-9</u> A screenshot of *GenEvo 3- Genetic Drift and Natural Selection* model (Dabholkar & Wilensky, 2016c)

AGENTS

The agents in this model are bacterial cells and patches. In the screenshot of the model in Figure 3-9, two types of bacterial cell agents are visible, one type is of red color and the other type is of brown color. The brown color bacterial cells have a blue outline, indicating their selective advantage in this environment. The other types of agents, patches, are also of two colors in this screenshot (Figure 3-9). The patches containing sugar are grey in color and patches without sugar are black.

AGENT PROPERTIES

Bacterial cell agents in this model have two main properties:

1. color – The color of a cell indicates its phenotype. Since this is a population of asexually reproducing bacteria, the color is inherited from a parent. So, when a

bacterial cell divides into two cells, both the daughter cells are of the same color as the parent cell.

2. energy – Energy is another property of a bacterial cell that is tracked as a variable. When a cell happens to be on a patch containing sugar, its energy increases, and the sugar in the patch is depleted.

Patch agents in the model have a property that is tracked using a Boolean variable, *sugar?*.

1. sugar? – Initially, 10% of patches have sugar? set to TRUE, which means that those patches contain sugar. As the model progresses, depending on the carrying-capacity settings in the model, the sugar? state in a number of randomly chosen patches is set to be TRUE.

AGENT BEHAVIORS AND INTERACTIONS

The bacterial cells are the main agents in the model. When a user runs a model by pressing the button labeled GO, the model progresses temporally. The temporal progression of a model is shown by advancing the value of clock-ticks, which is an arbitrary time unit.

At each clock-tick,

each bacterial cell,

moves (in a random direction)

eats sugar (if it is on a patch that contains sugar)

reproduces (if its energy is double the value of its initial energy)

dies (if its energy is zero)

A cell moves in random directions. The movement costs energy. So, at every tick, a cell loses a certain amount of energy. It also gains, some amount of energy if it happens to be on a patch that contains sugar. If the *natural-selection?* is ON, a cell with selective-advantage gains more energy based on the value of *%-advantage* that a user can control. When the energy of a cell becomes twice its original energy, it divides into two cells. Each of the cells receives half of the dividing cell's energy. The daughter cells also inherit the color of their parent cell.

Environmental conditions

A user can set *carrying-capacity* to very high, high, medium, low, and very low. The carrying capacity of an environment is the maximum population that can be sustained in that environment. In this model, a *carrying-capacity* chooser allows users to choose the carrying capacity of the environment in the model. This chooser also changes sugar availability (rate of addition of sugar per tick) in the model.

Emergent Patterns

There are two main emergent patterns in this model, one is about population dynamics by controlling sugar availability and the other is about change in a population composition because of genetic drift or/and natural selection.

The carrying capacity of the system is modeled as an emergent property of the system. The model allows users to choose the carrying capacity of the system. Depending on the chosen carrying capacity, the rate at which the sugar gets added to the system is set. The sugar availability influences the growth rate and the death rate of the bacterial population, because sugar provides energy for a cell. As the energy of a cell doubles, it divides into two cells. So, the population growth rate is determined by the availability of sugar. Cells need the energy to move and when their energy becomes zero, they die. So, a lack of sugar availability results in the death of a cell, thus influencing the death rate of the population. As the bacterial population grows, sugar in the system gets depleted at a faster rate, because there are more cells that consume sugar. This in turn reduces sugar availability and causes the population to decline. The population decline again increases the sugar availability per cell and causes the population to grow. The cycle of positive and negative feedback loops results in the emergence of stable fluctuations in the population just below the carrying capacity of the system.

In the following part, I explain another emergent pattern regarding the changes in the composition of the bacterial population over several generations because of genetic drift or/and natural selection.

When *natural-selection?* is OFF:

The increase in energy by eating sugar is identical for each type (color) of an *E. coli* cell. By statistical advantage, a dominant color becomes more likely to 'win' the competition for survival and take over the population. However, because the process is random, there will usually be a series of dominant colors before one color finally wins, and even with identical model conditions, the winning color could be different in each model run. Figures 3-10 and 3-11 show two runs of the model under identical conditions. In a model run shown in Figure 3-10, the surviving phenotype is brown, whereas in another model run under identical conditions the surviving phenotype is of green color.

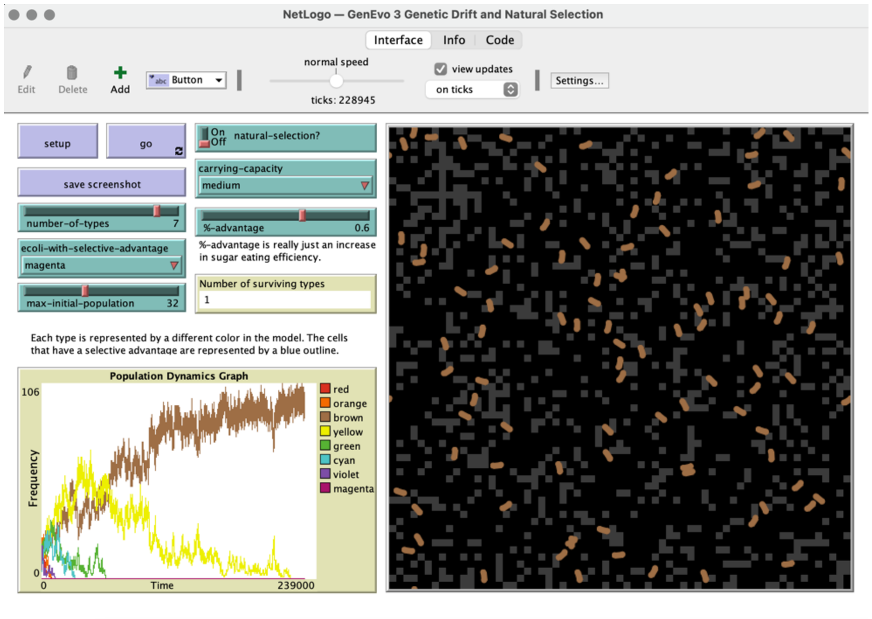


Figure 3-10 A screenshot of *GenEvo 3 Genetic Drift and Natural Selection* model when natural-selection? is OFF. The surviving phenotype after several generations is of brown color.

Figure 10A:

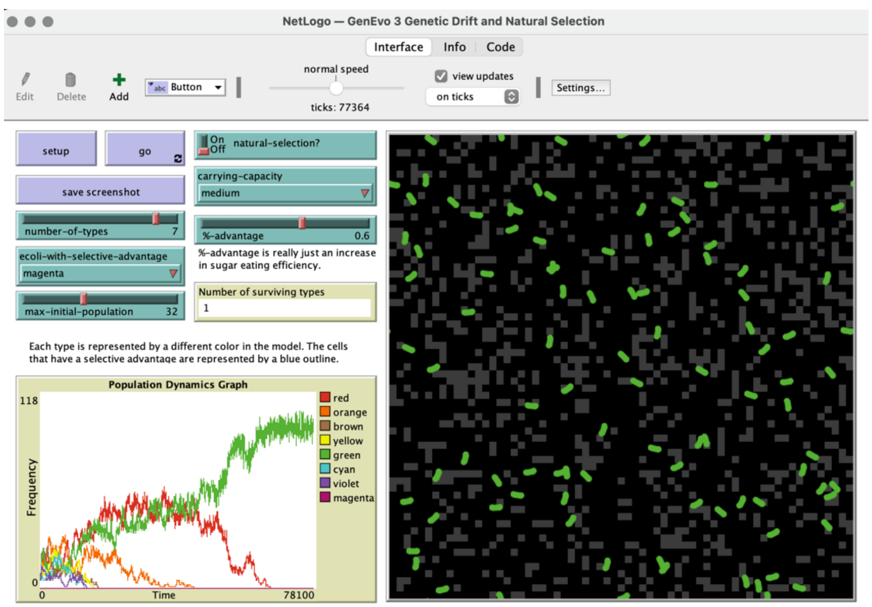


Figure 3-11 A screenshot of *GenEvo 3 Genetic Drift and Natural Selection* model when natural-selection? is OFF. The surviving phenotype after several generations is of green color.

When *natural-selection?* is ON:

A user can select which type (color) has a selective advantage in this world, causing it to gain more efficiently digest sugar and gain more energy from sugar at each time step. The cells with selective advantage are represented as cells with a blue outline in the model.

This in turn causes that particular type of E. coli to reproduce faster. The % advantage slider sets the percentage increase in energy gain by the cells with a selective advantage. Through this selective advantage, a dominant color becomes more likely to 'win'. However, if the selective advantage is low, the statistical advantage might still allow another color to 'win'.

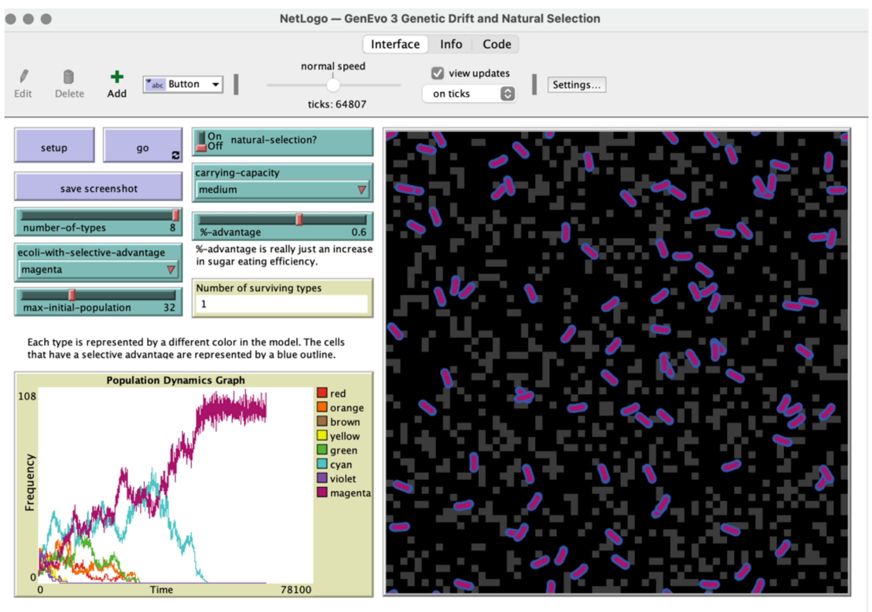


Figure 3-12 A screenshot of *GenEvo 3 Genetic Drift and Natural Selection* model when natural-selection? is ON. The surviving phenotype after several generations is of magenta color which has selective advantage.

THE RESTRUCTURATED GENEVO CURRICULUM

In the GenEvo curriculum, students are first presented with a computational model of a bacterial cell with a genetic circuit in which certain components such as proteins and parts of DNA interact in a specific manner (Model 1). Students use this model as a computational system to investigate how molecular interactions inside a cell result in complex emergent behavior at the cellular level. In the next two subunits, students use population-level models to learn about genetic drift and natural selection (Model 2 and Model 3). They observe competition between cells and reason about emergent patterns at the population level. Finally, students revisit the first model and engineer the genetic circuit to make their cells 'fitter' to reproduce (Model 4). The cells where genetic circuits are designed by the students 'compete for survival' in a limited resource environment. The fourth model allows students to learn about how mechanisms of genetic regulation affect the survival of cells in a resource-limited environment (For details of the curricular activities see Appendix 1).

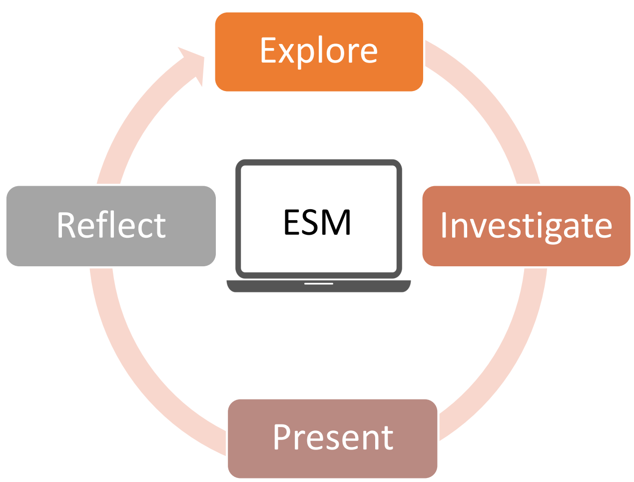


Figure 3-13 Pedagogical approach for an ESM-based curriculum

The GenEvo curriculum is designed for students to iteratively go through cycles of model exploration, investigation of specific aspects, sharing information by presenting their claims and evidence, and reflecting on the collective understanding of the microworld (Figure 3-13). Students first explore a model and talk about the observations that they find interesting. Students' investigations are scaffolded by guiding them to focus on specific aspects of agent behaviors such as sugar availability or DNA-proteins interactions. The primary observations often help students in identifying aspects of the system that they find interesting to investigate. They are asked to state their interests as a research question and state their preliminary answer as a testable hypothesis. Then they design and conduct computational experiments in the ESM learning environment to test their hypotheses and present their investigations. Their findings collectively build towards ideas about the emergent properties in the context of gene regulation and evolution.

CONCLUSIONS

In the GenEvo course, agent-based restructurations of biological systems, starting from molecular to cellular and then to organismic to the population level, are used to foster a deep understanding of emergent ideas about gene regulation and evolution. This course is designed for students to learn advanced ideas in molecular biology and their connections with evolutionary biology, which are usually reserved for graduate students. During our implementation of the GenEvo curriculum in middle school classrooms (details in Chapter 4), the students had no trouble engaging in scientific investigations of gene regulation and evolving populations and analyzing behaviors of these emergent systems. Similar to how numerical literacy or algebra literacy restructurated intellectual pursuits and learning of these disciplines, agent-based restructurations have the potential to transform research and learning of complex emergent systems (Wilensky, 2020). The traditional classroom structuration, that requires mathematical sophistication leaves out pre-college students from learning these powerful ideas in modern biology. In rare circumstances, when these are included in curricula, the students are told about these ideas authoritatively by their teachers using static models or animated videos to remember and explain those in the exams. Whereas in this GenEvo course, because of access to agent-based restructurations students learned these powerful ideas by collaboratively constructing knowledge through scientific investigations in the context of an ESM. I discuss in detail how different features of the ESM supported student learning in the next chapter.

Chapter 4: Designing ESMs (Emergent Systems Microworlds) and ESM-based curricula for Epistemically Expansive learning

Summary: To support students' agency in the process of constructing knowledge in a science classroom, it is important to design a learning environment that allows students to shape knowledge-producing practices to collectively develop knowledge products regarding a phenomenon under investigation. In this chapter, I present my work regarding an Emergent Systems Microworld (ESM)-based learning environment called GenEvo, which is designed to "restructurate" learning of fundamental ideas in modern biology, such as gene regulation. ESMs are learning environments that use agent-based representations and constructionist design principles. I build on earlier work of designing and using agent-based constructionist computational models with a specific focus on student engagement in knowledge construction practices and how properties of restructuration support such learning. I use a mixed-methods analysis to investigate student learning of complex disciplinary ideas and shifts in their perceptions of their roles in learning science in a classroom setting. Using Activity Theory as an analytical lens, I analyze student classroom participation and their pre- and post-interviews to demonstrate how properties of restructurations instantiated through different design features of the ESM mediated students' expansive learning. This work demonstrates the potential of ESM-based restructurated learning environments for supporting students' epistemic agency in a science classroom.

INTRODUCTION

Recent science education reforms emphasize engaging in and learning about practices that scientists use to make sense of the world rather than limiting science education to knowing scientifically established ideas (NGSS Lead States, 2013; Schwarz et al., 2017). This shift in science education requires reimagining the roles of students and teachers in the science classroom so that students become *doers of science* and not *receivers of facts* (Miller et al., 2018). *Doing science* in the science classroom means engaging students in science practices to construct disciplinary knowledge.

What are science practices that students should engage in? The Next Generation Science Standards (NGSS) has recommended a set of science practices that are epistemically equivalent to the practices of scientists (NGSS Lead States, 2013). However, this framing of epistemic equivalence creates a contradiction for *doing science* using the NGSS framework (Miller et al., 2018). Epistemic agency is an important construct to consider for supporting student engagement in *doing science.* The term epistemic agency was introduced into education literature in relation to the research on knowledge-building communities conducted by Scardamalia and Bereiter (1991). Epistemic agency refers to students' ability to shape and evaluate knowledge and knowledge-building practices in the classroom (Scardamalia & Bereiter, 1991; Stroupe, 2014). Miller and colleagues (2018) argue that NGSS's focus on science practices is not sufficient for the move to envision students as *doers of science* and not *receivers of facts*. Having a set of practices chosen by others as important to learn and expecting students to mimic those practices does not position students with the epistemic agency - the power to shape the knowledge production and practices of a community. Positioning students as epistemic agents requires them to collectively shape practices for knowledge construction. To truly support students' epistemic agency in a classroom, newer learning environments need to be designed that allow students to develop knowledge-building practices to construct and evaluate knowledge products.

Doing Science using Computational Models

My work in this chapter builds on the extensive earlier work of the last three decades by Wilensky and colleagues regarding using computational models for engaging students in learning about complex emergent phenomena (Paulo Blikstein & Wilensky, 2010; Levy & Wilensky, 2009; Sengupta & Wilensky, 2009; Stieff & Wilensky, 2003; Wagh & Wilensky, 2018; Wilensky, 2003; Wilensky & Novak, 2010; M. Wilkerson-Jerde & Wilensky, 2010). To *do science* in a science classroom, it is important to have a system for students to investigate a phenomenon and test their ideas to construct theories that provide mechanistic explanations of observed patterns related to the phenomenon. My work focuses on student learning of emergent phenomena. Emergent phenomena are the ones in which uncoordinated interactions between autonomous agents result in emergent patterns at the system-level (Wilensky & Rand, 2015). To learn about emergent phenomena by engaging in scientific practices, it is important to design learning environments to allow learners to manipulate and investigate agent behaviors and system-level aggregate patterns (Wilensky, 2001; Wilensky & Rand, 2015; Wilensky & Resnick, 1999). Researchers of science education have demonstrated the effectiveness of the emergent systems perspective for understanding several natural phenomena ranging from prey-predator relationships to nectar collection by honeybees, to the kinetic molecular theory (Danish, 2013; Hmelo-Silver & Azevedo, 2006; Klopfer et al., 2005; Wilensky, 2003; Wilensky & Jacobson, 2014; Wilensky & Reisman, 2006). Agent-based representations in these models serve as restructurations (Wilensky, 2020; Wilensky & Papert, 2010) which have cognitive, social, affective, and diversity properties that increase the learnability of the phenomenon represented using restructurated representations (See Chapter 1 and 2 for the theory of restructurations and restructuration properties). Agent-based restructurations have been demonstrated to be pedagogically effective in supporting the learning of several complex natural phenomena in science education (e.g., electric current, resistance, crystallization, temperature, pressure, evolution) (Levy & Wilensky, 2009; Sengupta & Wilensky, 2009; Wagh et al., 2017; Wilensky, 1999a; Wilensky & Novak, 2010).

My work is also strongly rooted in the theory of Constructionism (Papert, 1980; Papert & Harel, 1991). Constructionism is a learning theory, and a design framework that harnesses power of computational technologies for students' engagement in individual and collective meaning-making to make epistemically powerful ideas accessible to them (Kynigos, 2015; Papert, 1980). As a learning theory, constructionism builds on Piaget's theory of constructivism that the new knowledge is built on the foundations of prior knowledge (Ackermann, 2001; Papert, 1980; Piaget, 1970). Constructionism contributes to the theory of constructivism through its unique attention to the ways of facilitating engagement of students in constructing personally meaningful artifacts and learning about powerful ideas through such construction. Constructionist design framework provides guidelines for designing learning environments to support the creation of individual and collective bricolage with computationally supported artifacts, influenced by negotiated changes students make to these artifacts with an explicit emphasis on self-driven production and ownership (Ackermann, 2001; Kynigos, 2015). In Papert's own words – "*Constructionism -- ... --shares constructivism's connotation of learning as "building knowledge structures" irrespective of the circumstances of the learning. It then adds the idea that this happens especially felicitously in a context where the learner is consciously* *engaged in constructing a public entity, whether it's a sandcastle on the beach or a theory of the universe*." (Papert & Harel, 1991). The Emergent Systems Microworld (ESM) design framework that I use in this chapter incorporates the following three key ideas from the constructionist design framework: (a) personally meaningful engagement, (b) construction of public entities, (c) expression and validation of ideas through computational microworlds. My work in this chapter focuses on investigating how agent-based restructurations and constructionist design features in an (ESM)-based learning environment support students' epistemic agency.

Broader relevance of epistemic agency in a science classroom

I believe that epistemic agency is important not only in science education but also in a larger educational context. It is related to how students envision their role in the context of knowledge – as producers vs as receivers. In a larger context, I posit that epistemic agency could be a precursor for what Morales-Doyle (2017) operationalizes as Freire's Conscientização, or critical consciousness, in his justice-centered science pedagogy (Morales-Doyle, 2017). In the context of eliminating oppression, Conscientização is a process by which people come to view themselves as capable of transforming reality. The ability to view themselves as epistemic agents with the power to produce knowledge and shape knowledge-producing practices in a science classroom can potentially be a step for students towards developing Conscientização. This view of critical consciousness opens up a possibility of the notion of epistemic consciousness, which I consider to be people believing themselves capable of producing knowledge. In this study, I limit my focus to students' voiced perceptions of their role in a science classroom to investigate their perceptions regarding their epistemic agency. However, I contend that studying students' epistemic consciousness in a science classroom can be a generative construct to design for and analyze learning environments to support students' epistemic agency.

This view of epistemic agency and epistemic consciousness in the context of science education creates tension regarding the teaching of scientific practice and disciplinary ideas in a classroom. On one hand, (a) there are commonly used practices in scientific research, such as developing an argument based on evidence, as well as (b) settled disciplinary ideas, such as the particulate nature of matter or changes in population because of natural selection, which are important to learn. On the other hand, 'settled expectations' (Megan Bang et al., 2012) of how to teach these ideas and engage students in predetermined scientific practices do not make students *doers of science* in a science classroom. If these ideas and practices are presented to students as settled facts and ways of scientific knowledge construction, then that severely compromises students' development of epistemic agency and consciousness.

I argue that constructionist learning environments that are designed to foster student engagement in constructing knowledge of disciplinary ideas can support their epistemic agency by facilitating their participation in shaping practices for evidence-based knowledge construction and evaluation. In this chapter, I present one such learning environment that uses ESM and ESM-based pedagogical practices intended to support students' epistemic agency. I use Wilensky and Papert's theory of restructuration (Wilensky, 2020; Wilensky & Papert, 2010) to analyze how agent-based restructurations in an ESM support students' epistemically meaningful engagement to be agentic in shaping practices.

Theoretical Framework

Over the years, scientific communities across the globe have developed experimental model systems that have affordances to investigate specific aspects of natural phenomena. For example, fruitflies' (*Drosophila*) chromosomal organization and their short life span has made them a model system to study genetics. Similarly, the organization of a small number of neurons in roundworms (*C. elegans*) is beneficial to the study of neurobiology. I argue that using principles of Learning Sciences computational models can be designed to be pedagogically effective model systems that support students' self-driven investigations and therefore their epistemic agency within the constraints of a classroom. My work involves designing Emergent Systems Microworlds (ESMs) as computational model systems for students to investigate a modeled emergent phenomenon.

ESM design combines two design approaches in Learning Sciences, namely agent-based modeling of emergent systems and constructionism (Papert, 1980; Wilensky, 2001). Agent-based representations in ESMs create affordances for learners to engage deeply with a complex emergent phenomenon (Goldstone & Wilensky, 2008; Wilensky & Reisman, 2006). An ESM is designed as a microworld using constructionist design principles to mediate students' self-driven explorations to investigate various aspects of the represented disciplinary ideas (Edwards, 1995; Papert, 1980). An ESM-based curriculum uses an ESM to facilitate such student engagement in self-directed, interest-driven explorations and investigations. Students are encouraged to share their findings and participate in teacher-guided reflections to collaboratively construct knowledge about the modeled complex phenomenon in an ESM (See Chapters 1 and 2 for more details about the ESM design framework).

Theory of Restructurations

ESMs *restructurate* (Wilensky, 2020; Wilensky & Papert, 2010) disciplinary ideas through agent-based representations. Structuration is the encoding of knowledge in a domain, which is largely influenced by available representational infrastructure that can be used to express the knowledge. Hindu-Arabic (0, 1, 2, 3, ..) and Roman (I, II, III, IV,…) numerals are examples of representational infrastructures that support structurations in arithmetic. Arithmetic operations such as multiplication and division underwent a huge change because of restructuration from Roman to Hindu-Arabic numerals. Wilensky and Papert's theory of restructurations states that properties of restructurations influence the learnability of disciplinary ideas, especially from the point of view of democratizing access to powerful ideas to the wider and younger population. They describe five fundamental properties of restructurations, namely, power properties, cognitive properties, affective properties, social properties, and diversity properties (See chapter 1 and 2 for details). In this chapter, I analyze how cognitive, affective, and social properties of restructurations in an ESM mediate students' epistemically agentic learning to engage with emergent aspects of gene regulation and evolution in a biology classroom.

Activity Theory and Expansive Learning

To support students' epistemic agency in a classroom, the roles of students and teachers need to be reimagined. In this chapter, I analyze how a restructurated constructionist ESM-based curriculum supports the transformation of a classroom activity such that students take more epistemically agentic roles. I use Cultural Historical Activity Theory (CHAT) and its extension, called the theory of expansive learning by Engeström (Engeström, 2001), to investigate such transformation. Activity theory provides a lens to design and investigate learning environments that support student engagement in social knowledge-building activities in specific domain areas such as complex systems (eg., Danish, 2013, 2014; DeLiema, Enyedy, & Danish, 2019). The focus of such analysis is to investigate how different *tools* in a learning environment mediate individual and social transformation (Engeström, 1999), which is important to study for designing learning environments that mediate such transformations in a classroom.

Activity theory describes the context of any activity, such as classroom learning, in terms of the subject, object, tools, community, rules, and division of labor (Engeström, 2001). The original activity triangle or mediational triangle, proposed by Vygotsky, explains the relationships between subject, object, and a mediating tool (Vygotsky, 1978). Engeström (2001) developed his theory of expansive learning within the framework of activity theory. Expansive learning is about creating a transformation in the activity system to start producing new patterns of activity. Even though in Engeström's theory of expansive learning, community, rules, and division of labor are important aspects of an activity system, in the first part of my analysis in this chapter, I primarily limit my scope only to the three aspects that were first introduced by Vygotsky's (1978) mediational triangle (see Figure 1) (Cole, 1998). This is because the focus of this work is to identify design features that act as tools of transformation of epistemic activities. Such transformation of a classroom activity system changes it from the one in which students are positioned as *receivers of facts* to the one in which they are positioned in epistemically agentic roles as *doers of science*. Additionally, in the discussion section, I discuss how the other three aspects of an expansive activity system– community, rules, and division of labor, contributed to the transformation of students becoming *doers of science* in the ESM-based learning environment.

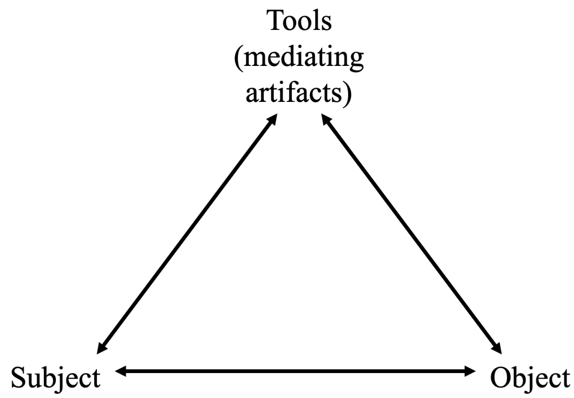


Figure 4-1 Vygotsky's mediational triangle (from (Cole, 1998), p. 119).

Activity theory's fundamental notion, as represented by the mediational triangle, is that humans use tools to mediate work toward the Object of Activity. These tools can be of different forms. Vygotsky (1978) mainly discussed physical tools, language, and signs as the mediators of human action. Cole (1998) expanded this concept to include representations that communicate the community's norms, beliefs, and understandings, as well as cognitive structures that guide action and thought. This extended concept of mediational tools becomes important in the context of designing a learning environment. Different features of a learning environment can be considered as tools to analyze how they support different intermediate activities that build towards change in an activity system. These features can act as tools to mediate student activity by supporting or constraining an action. Analysis of these features as mediational tools would allow deeper insights into learning activities in an environment designed to support certain forms of learning.

In his theory of expansive learning, Engeström (2001) discusses the possibility of expansive transformation of an activity system, in which the Object (goal/purpose) of Activity gets reimagined as subjects engage with a mediational tool. In the context of science education, such expansive transformations mean a shift in the role of students in a science class from listening and understanding to actively engaging in knowledge-building practices. The theory of constructionism has always foregrounded such students' epistemically meaningful participation in constructing knowledge (Papert, 1980; Turkle & Papert, 1992; Wagh et al., 2017; Wilensky, 2003). In this chapter, I use Engeström's theory of expansive learning for analyzing how agent-based restructurations and constructionist design features of a learning environment facilitate the epistemic expansion of an activity system in a science classroom.

The mediational triangle (Figure 4-1) helps in understanding the object-directed nature of human action. In a classroom setting, an object for a student is not necessarily the same as the object that is intended by a teacher for the student. The intended object is the purpose for which the activity is designed, and the enacted object is the object that a subject (student) uses a tool to achieve. This distinction becomes even more critical when one intends to use activity theory to analyze a learning environment. For example, when an ESM-based learning environment is designed for students to socially construct knowledge, the object 'to socially construct knowledge' is an intended object for the students by the designer of the ESM-based curriculum. For a learning environment designed for expansive learning, this intended object must become the enacted object of classroom activity. In other words, the designer's intention of the social construction of disciplinary knowledge needs to become students' object in the activity system. This framing is inspired by Engeström's theory of expansive learning. However, there is an important distinction in how I frame it and how Engeström discusses the transformation of the activity system during expansive learning. Engeström focuses on inherent contradictions resulting in the production of new cultural patterns of the activity system, which are often previously unimagined, as an activity system undergoes a transformation (Engeström, 1991, 1999); whereas an ESM-based curriculum is designed to guide a transformation of an activity system to make its enacted object to be 'social construction of disciplinary knowledge'. To underscore this distinction, I use the term *guided epistemic expansion* to refer to the ESM-mediated transformation of a classroom activity system. In the next section, I discuss this point in greater detail.

This kind of intended epistemic expansive learning in a science classroom may face a *problem of practice* (Russ & Berland, 2019). This problem arises due to a tension between learning correct ideas and constructing one's own ideas. Russ & Berland (2019) highlight a dichotomy between two intended objects of instruction in a science classroom (See Figure 4-2). This dichotomy is between "Figuring it out" vs "Learning about" activity systems, where the former is focused on using the scientific canon to figure out a natural phenomenon and the latter is focused on using a natural phenomenon as a tool to learn about the scientific cannon.

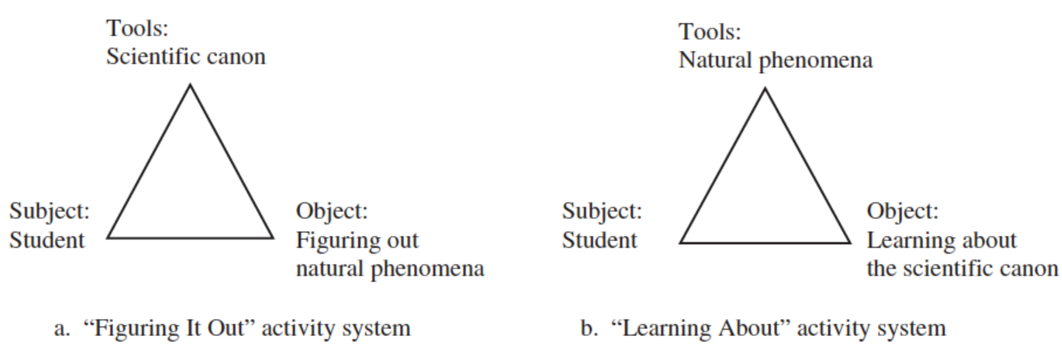


Figure 4-2 Mediational activity triangles in a science classroom (From (Russ & Berland, 2019), p. 286)

Russ and Berland (2019) propose a framework called Invented Science to address the problem of practice. The authors argue that children invent 'science' to satisfy their curiosity about why and how a phenomenon occurs. However, to achieve the object of satisfying curiosity, children need to have the appropriate cognitive resources to act as mediational tools (Elby & Hammer, 2010; Piaget, 1970; Smith III et al., 1994). In a classroom setting where teachers try to balance their time between "Figuring it out" vs "Learning about" activity systems, external tools that would provide students objects-to-think-with can serve to engage in epistemic pursuits to 'invent science' (Papert, 1980; Turkle & Papert, 1992). Analysis of mediational tools in a learning environment using activity theory provides insights into how students socially construct knowledge of a complex phenomenon (Danish, 2014). I argue that a computational model system in the form of an ESM provides students with such tools to collectively support the social construction of knowledge of the modeled phenomenon. I investigate how an ESM-based curriculum provides a computational model system to mediate guided epistemic expansion of a classroom activity system as shown in Figure 4-3.

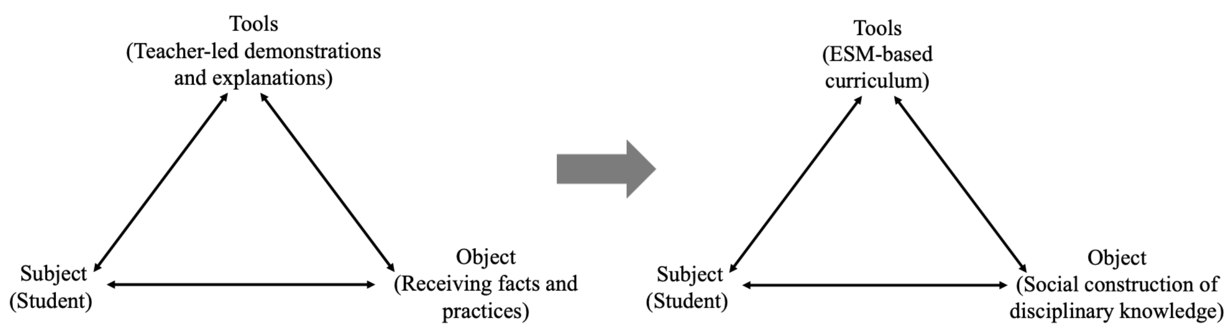


Figure 4-3 Intended epistemic expansion mediated by an ESM-based curriculum in a classroom

GUIDED EPISTEMIC EXPANSION

The ESM-based curriculum is designed to guide students' epistemically expansive learning as shown in Figure 4-3. This is a different take on Engeström's idea of expansive learning, in which a designer has an intended object for the transformed activity system. In Engeström's theory of expansive learning, the object of expansive learning activity is the entire activity system itself in which learners are engaged. For pedagogical reasons related to guiding student participation in specific kinds of practices, I take a somewhat restricted view while defining ESM-mediated epistemic expansion. The ESM-mediated epistemic expansion does not generate a completely unexpected form of an activity, instead, it is designed to guide the transformation of a classroom activity system into the one that has an intended object – social construction of disciplinary knowledge (Figure 4-3). To highlight this point, I have added the word guided before epistemic expansion. I investigate how the ESM-based curriculum supports guided epistemic expansion of a classroom activity system.

CONSTRUCTIONISM AND ACTIVITY THEORY

The learning environment discussed in this paper, Emergent Systems Microworld (ESM) – based curriculum and the pedagogical practices to support student engagement in epistemic activities is rooted in constructionism (Papert, 1980), complex systems theory (Bar-Yam, 2004; Jacobson & Wilensky, 2006), and in agent-based modeling (Railsback & Grimm, 2019; Wilensky & Rand, 2015; Wilensky & Resnick, 1999). From the perspective of designing learning environments for self-driven learning, constructionism is a big part of ESM design. A microworld, which is a design concept from Papert's theory of Constructionism, is an encapsulated open-ended computational exploratory environment in which a set of ideas can be investigated through interactions that lead to knowledge construction (Edwards, 1995; Papert, 1980). Learning activities in an ESM-based curriculum are designed to engage learners in manipulating computational objects and executing specific operations instantiated in a microworld. Such manipulations result in observable changes in the microworld that are related to the phenomenon that is modeled in a microworld. As learners investigate the effects of their manipulations and attempt to establish reliable patterns in the microworld, they induce or discover properties and functioning of the system as a whole. Through this process of manipulation, experimentation, prediction, and testing, they self-correct or 'debug' their understanding of the domain to develop new powerful ideas (Papert, 1980). In an ESM-based science unit that I present and discuss in this chapter, students are encouraged to actively develop practices to construct and evaluate knowledge regarding specific behavioral patterns in the computational microworld.

In the following part, I discuss two concepts of activity theory and constructionism that have distinct yet overlapping meanings and how I operationalize these concepts in my work. The two concepts are Object of Activity from activity theory, and objects-to-think-with from constructionism.

As I discussed earlier, an *Object of Activity* in activity theory is the goal or purpose of the activity (see Figures 4-1, 4-2, and 4-3). In the context of the science classroom, general examples of Objects of Activity include *figuring out a scientific phenomenon*, *learning about the scientific canon*, *receiving facts* and more specific examples include, *learning about the particulate nature of matter*, *establishing patterns based on evidence*. Sometimes these are intended objects of an activity system, that are not taken up by individuals or a group of participants as their objects. In my analysis, I discuss how some of the intended objects become participants' objects and how their engagement with those objects is mediated by the restructuration properties of an ESM.

Objects-to-think-with is a design concept in Constructionism (Papert, 1980; Turkle, 2011). In an ESM that I discuss in this paper, these objects-to-think-with are computational entities in a microworld. The word 'object' has two distinct meanings as it is operationalized in activity theory and in constructionism. Constructionism understands objects as physical objects like, gears that Papert discusses as the objects that helped him think about ideas since his childhood (Papert, 1980). By extension, in a constructionist microworld these are computational objects that can be manipulated and observed to think about various ideas related to behaviors of objects, their interactions and patterns that are generated through those behaviors and interactions. Examples of such computational objects-to-think-with in the ESM that I discuss in this paper are various proteins, and parts of DNA (see Figure 4 below). These *objects-to-think-with* can potentially mediate an object of an activity system. In the result section, I investigate how *objects-to-think-with* as one of the mediators of an activity system contribute to individuals' engagement in the object of an activity system.

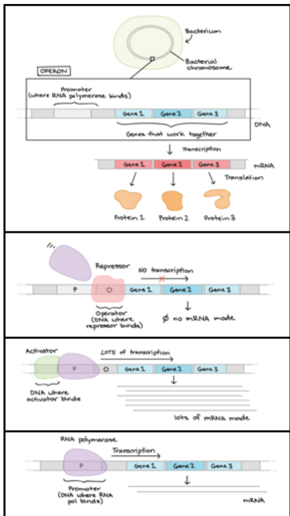
Another idea that is central to constructionist design and pedagogical principles is consciously engaging learners in constructing a public entity. Papert and Harel (1991) have famously given examples of such public entities as a sandcastle on the beach or a theory of the universe (Papert & Harel, 1991). Examples of such public entities in constructionist learning environments range from a block-based coding project in Scratch (Resnick et al., 2009) to an agent-based model in NetLogo (Wilensky, 1999a, 2001; Wilensky & Reisman, 2006) to electronic textiles (Buechley et al., 2007; Fields et al., 2021; Peppler, 2013). Researchers of embodied learning have argued that bodies can be interactional-syntonic resources that can individually or collectively generate unique forms of shared public artifacts (Danish & Enyedy, 2020). In the ESM-based learning environment that I discuss in this paper, students are asked to construct public entities for the classroom audience. These are their micro-theories and hypothesis with evidence that they collect by conducting computational experiments using the

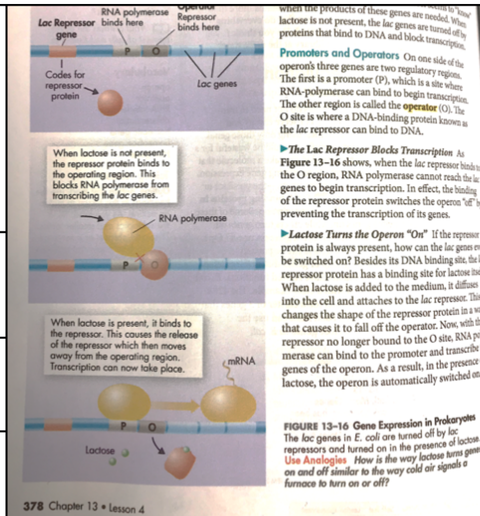
ESM. As they iteratively create and share their artifacts, they collectively build knowledge of disciplinary ideas related to gene regulation and the theory of evolution using the ESM. From an activity theoretical perspective, I do not treat these public entities as Objects of Activities, instead, I use the activity theory lens to identify different intermediate Objects of Activities that emerge as students engage in these epistemic activities.

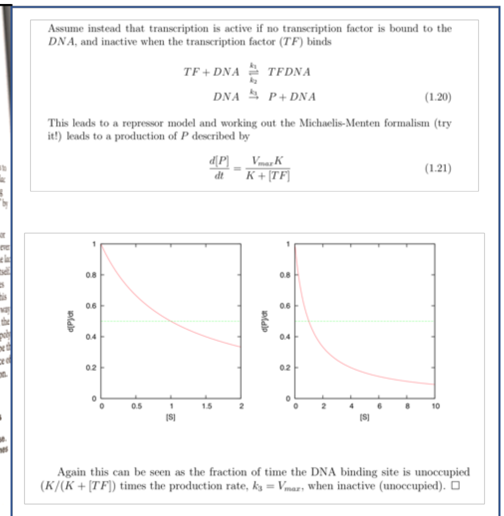
ESM-MEDIATED RESTRUCTURATION OF GENETICS AND EVOLUTION

To understand how restructuration properties of an ESM facilitate students' epistemically expansive participation to learn about emergent properties of a system, I investigate student learning with an ESM-based curriculum about genetics and evolution. In this section, I first discuss the need to use agent-based restructurations to learn fundamental ideas in modern biology. Then, I briefly introduce the ESM-based curriculum that I designed to learn these ideas. Details of this ESM and curricular activities are included in Chapter 3.

There has been a significant shift in biological research over the past few decades with the incorporation of newer technological tools, the use of computational modeling methods, and an increasing focus on understanding biological systems from the complex systems perspective (Kitano, 2002, 2017). For example, understanding the mechanism of control in gene regulatory networks that focuses on how stochastic chemical interactions between biomolecules such as proteins and DNA lead to complex cellular and organismic behaviors is fundamental to understanding advances in modern biology. From diagnosing a human disease, defining disease predilection, and developing individualized (personalized) treatment strategies (Loscalzo & Barabasi, 2011) to understanding how memory works in single cells and whole organisms (Kandel et al., 2014) the holistic systems biology approach has been effective for understanding biological systems and devising solutions to biological problems. However, in schools and universities, student exposure to central ideas about gene-regulatory biochemical mechanisms continue to rely on static models that depict molecular interactions as deterministic processes (Figure 4-4 (a) and 4-4 (b)) or mathematical models that use differential equations-based representations requiring mathematical sophistication to understand the fundamental aspects of the processes (Figure 4-4(c)). None of these existing structurations explain how simple biochemical interactions between these molecules allow a cell to make emergent complex decisions and perform complex functions. As a result, students don't gain access to the cutting-edge ideas of molecular biology. These cutting-edge ideas are not just the latest advances in biology, but more importantly from a learning point of view is that they allow learners to reason more deeply about biological processes. Our restructuration of these ideas in modern biology involves representing biomolecules as visualized agents with agent interactions forming system-level behavior of an organism.







(a) Diagrams explaining molecular mechanisms of genetic regulation on Khan Academy (2018)	(b) Diagrams explaining molecular mechanisms of genetic regulation in Miller & Levin Biology Teacher's Edition (2018)	(c) Differential equations explaining molecular mechanisms of genetic regulation in a University level course (2018)

Figure 4-4 Molecular mechanisms of genetic regulation through existing structurations of molecular biology (A diagram from (Wilensky, 2020), pg. 296, also in Chapter 3)

To engage students in constructing knowledge of disciplinary ideas regarding gene-regulatory mechanisms and evolution, I created an ESM-based curriculum, GenEvo (Dabholkar, Anton, et al., 2018; Dabholkar & Wilensky, 2016a) (See Chapter 3). In this chapter, I study students' epistemically expansive learning with the GenEvo curriculum (Dabholkar & Wilensky, 2016a). This curriculum incorporates a series of four interconnected computational models designed using Wilensky's agent-based programmable modeling environment NetLogo (Wilensky, 1999b). NetLogo was designed for both constructionist learning experiences and for use in scientific research. Since these four NetLogo models are strongly interconnected, they form an ESM because the underlying rules for agent behavior are consistent across the models. Using this curriculum, students can investigate the emergent properties of biological systems, including gene regulation, carrying capacity, genetic drift, and natural selection.

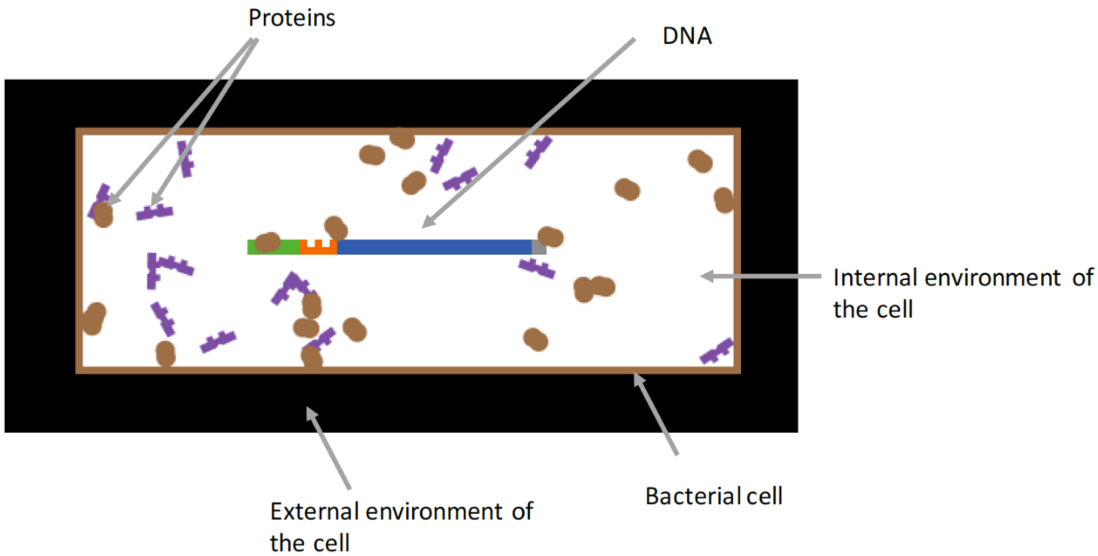


Figure 4-5 Computational representations of agents (DNA and proteins) and the environment in the GenEvo 1 - Genetic Switch model

The first model of the ESM, called Genetic Switch, is based on the lac operon[7], which is one of the most studied genetically regulated systems. This model allows students to investigate a molecular mechanism of genetic regulation, which ensures that the proteins required for intake and digitation of a type of sugar, lactose, are produced only when lactose is the preferred energy source in the environment (Figure 5). Understanding the regulation of protein production is fundamental to understanding modern biology. The second and third models are of a bacterial population, which students can use to investigate changes in the population because of genetic drift and natural selection in a limited-resource (sugar) environment. The fourth model combines these three models, which allows students to computationally engineer the genetic circuit and study its effect on the survival of cells.

Research Context and Methods

[7] https://en.wikipedia.org/wiki/Lac_operon

My research question is about *restructuration properties of an ESM instantiated through its design features supporting epistemically expansive learning in a science classroom.* To investigate this question, I use data from four iterations of an ESM-based curriculum about genetics and evolution. In this section, I first describe the research context of the ESM-based curriculum. Then I discuss a mixed-methods approach to investigate student learning of disciplinary ideas, students' perceptions of their epistemic agency in a science classroom, and how different features and an ESM-based learning environment supported student engagement in shaping inquiry practices to investigate an emergent phenomenon.

PARTICIPANTS AND SETTING

I was the lead designer of the ESM and the curricular unit and the lead teacher of these implementations. From the methodological perspective, this made me an "unusually observant participant" and a designer as well (Erickson, 2006). In his chapter about collaboration action ethnography, Erickson has coined this term *unusually observant participant,* to highlight the dichotomy between the insider/outsider role of a researcher. In the context of design-based research (Design-Based Research Collective, 2003) that I discuss in this paper, my role as a lead designer and a lead teacher puts me in the categories of the researched and the researchers.

I conducted these courses in two locations: twice during a weekend extra-school program for middle school students conducted by a talent-development center in a midwestern university in the United States, and twice in residential summer camps in a western city in India that students from all over the country attended. The students who participated in both these programs were ages 11 to 14. In the United States cohort, there were 6 female and 8 male students of mixed racial and ethnic backgrounds; the break-up of self-reported racial and ethnic backgrounds was 6 White non-Hispanics, 4 Asians, 1 White Hispanic, 1 American Indian or Alaskan Native, and 2 Others. In the summer residential program in India, 27 students participated, of which 14 were females and 13 were males. All the students were of Asian Indian origin. I collected data in various forms, namely videos of student discussions (around 150 hours), fieldnotes, workbooks in which students wrote their observations and explanations, the computational artifacts (models, screenshots, and presentations) that students created, pre- and post-tests, and pre- and post-interviews (See appendix 2 and 3).

Figure 4-6 Vidya[8] and Samir participating the ESM-based GenEvo curriculum

In the ESM-based GenEvo curriculum, the students used computational models (see Chapter 3 for details) designed as parts of an ESM to perform computational investigations regarding intracellular molecular interactions (Model 1) and survival of different cell types in a resource-constrained environment (Models 2 and 3). The investigations were scaffolded through questions hosted on a website (https://ct-stem.northwestern.edu). Students answered these questions by conducting computational investigations using the ESM (Figure 4-6). While some

[8] All the names used in the paper are pseudonyms.

questions asked students to perform open-ended investigations and record their observations, others asked students to focus on specific aspects of the ESM but did not direct the investigations explicitly (Appendix 1). These questions were designed for students to focus on aspects of the systems that they were investigating so that when they shared their findings, they could compare their results with others in the class. From the *practices* point of view, the students were asked to perform certain tasks, but the details of performing those tasks were intentionally left unspecified. For example, a set of questions about the energy of the cell was: (a) What are the effects of the presence or absence of sugar/s in the environment on the energy of the cell? (b) Upload the supporting material (experimental evidence for your answer) here. (c) Explain your scientific investigation process. (d) Part 1: What were the changes that you made in the model? (e) Part 2: What were your observations? (f) Part 3: How did you arrive at your answer using your observations?

These questions directly asked students to collect evidence to support their answers but did not specify what counted as evidence or how to collect it. The curriculum was designed in this way to guide student investigations and engage them in discussing and shaping epistemic practices to establish and evaluate knowledge claims. These discussions included topics like *what counts evidence for a particular claim*? *what are various ways to collect, analyze and present evidence*? *how does one establish a claim using evidence*? Since students were using an ESM to investigate specific aspects of a phenomenon, these discussions were strongly grounded in concrete aspects regarding the biological system under investigation. Throughout the curriculum, students iteratively explored specific aspects of the phenomenon, investigated specific questions, collected evidence, and presented how their evidence supported their claims regarding the research questions.

METHODS

To investigate student learning using the ESM-based curriculum, I conducted a mixed-methods analysis (Small, 2011), specifically focusing on the expansive aspects of student learning related to shifts in their epistemic agency. First, I used a quantitative approach to assess whether the students learned established disciplinary core ideas about genetics and evolution. The pre- and post-test questions were selected from the American Association for the Advancement of Science – Project 2061 website[9]. A question bank of 20 questions, focused on scientifically established ideas about molecular genetics and evolution, was created. Figure 4-7 shows two questions, the first about evolution and the second about molecular genetics, that were included in pre- and post-tests (See Appendix 2 for an example question set).

Which of the following is REQUIRED for the process of natural selection to occur?

A. Numerous species must have recently become extinct.
B. A food source must disappear.
C. There must be a sudden environmental change.
D. Traits must be inherited from one generation to the next.

What do DNA and proteins have to do with each other?

A. DNA is a type of protein.
B. Proteins are a type of DNA.
C. DNA provides information for making proteins.
D. DNA and proteins have nothing to do with each other.

Figure 4-7 Two example questions included in pre- and post-tests - the first one is about evolution and the second is about molecular genetics

[9] See more at http://assessment.aaas.org/pages/home

Each student was assigned a randomly selected set of 10 questions for the pre-test and received the remaining questions from the question bank for their post-test. I conducted a paired t-test to see if students' understanding of scientifically established ideas changed after they participated in the course.

Additionally, I conducted students' pre- and post-interviews to study the shift in their perceptions of science learning in the classroom and probe more about what they learned in the course and how they learned it. I focused on interview questions that were about practices those scientists follow to construct knowledge and students' perceptions regarding the process of learning science, especially from the perspective of understanding their agency in knowledge construction (See Appendix 3). The question prompts about work of scientists were: (1) Choose any topic that you learned in your science class/ in this course and explain how you learned it; (2) What do you think scientists do as their daily work?; (3) So, scientists construct knowledge about the world, right? How do they do that? How do they know that what they have figured out is right?. The question prompts to ask about student perception of their science learning were: (1) Can you mention some of the topics that you learned in this course? (2) Pick one topic and explain how you learned it? (3) Recall and describe how your learned it in the class.

The bottom-up, open coding was done using the constant comparative method to arrive at categories (Glaser & Strauss, 2017). Table 1 shows the coding categories regarding student perceptions about learning science and scientific processes. All the student responses were then coded by two researchers. Any disagreements between the researchers were discussed and resolved until Cohen's Kappa value was greater than 0.7 for each category. The codes for each response to each question category, their description, and examples are in the table below.

Table 4-1: Coding scheme

		Code description	Exemplar Coded Response
Learning of science		**Teacher directed:** Students talk about their learning of science being directed by a teacher	"We were taught from diagrams and all. Our teacher is not dependent on the book, she gives us extra information, she showed us some videos, extra notes she gave us."
Learning of science		**Active role:** Students talk about taking an active role while learning science	"I learned like a scientist mostly because you didn't tell us anything. You gave us no answers, so we all had to think for ourselves, experiment ourselves…."
Learning of science	Process of science	**Questioning:** Asking questions about unknown natural phenomena	"When we were asking questions about 'what is what' you were not answering us. And then we came to answer our own question by observing the model so well that…."
Learning of science	Process of science	**Experimenting:** Performing experiments to observe effects of a change/ manipulation	"We made experiments based on one variable, to understand the changes and to understand how the variable works."
Learning of science	Process of science	**Testing hypotheses:** Specific mention of testing preliminary ideas about observed patterns	"we figured out that, LacZ was the triangle, when we saw that LacZ graph go up, every time the triangles were made up by RNAP rolling over the DNA and we did that by reducing LacY degradation chance and increasing it and increasing and decreasing LacZ degradation chance and through that we came to the conclusion that LacZ is the pink triangles."
Learning of science	Process of science	**Sharing ideas/ communication:** Sharing newly learned ideas or observations with others in the class	"we came to know what function was happening to form a protein... we showed powerpoints to prove our points. We even showed evidence and all."
Learning of science	Process of science	**Community aspect of knowledge construction:** Talking about classroom community being engaged in knowledge construction	"Also, if I were not able to make some observations, others were making observations and telling, so I could use them in my presentations"

	Positive affect: Students describe learning of science being an enjoyable process	"And you were teaching, but not telling us anything so we had to figure out. I liked it."

Using micro-ethnographic methods, I analyzed the classroom discourse (Erickson, 1986, 2010), focusing on the dynamics of students' shifting participation in the activity and their shaping of the sociocultural practices of knowledge construction and evaluation in the classroom. Using a top-down approach, I identified all the instances in the field notes that indicated student engagement in knowledge construction. I performed a micro-ethnographic analysis of videos of these instances to investigate how restructuration properties of the ESM instantiated through ESM design features (mediators) and complementing pedagogical moves by the teacher (mediation) supported student engagement in knowledge construction.

For example, the following section in the field notes is coded positive for student engagement in knowledge construction:

"Sagar tries to answer the questions by talking about a fluctuating relationship, to which Shaurin disagrees. [The teacher] places both these opinions before the class and asks them if they have supporting evidence for their points. Shaurin comes to the front of the class to show his evidence so that everyone can have a look. [The teacher] ties that back by saying that we are trying to understand how to use evidence for argument. Mohan meanwhile has a different evidence for this. Meera points out that the conditions of the experiment are different in Shaurin's case and [The teacher] asks whether with the exact same parameters the same results will be observed in the model or not."

[Fieldnotes, May 15, 2018]

The micro-ethnographic analysis revealed that, in this episode, students were trying to establish a pattern regarding the changes in the energy of the cell. A student, Sagar, shared his answer to the question of what was causing the fluctuations. Another student Shaurin expressed his disagreement and shared his argument. The fieldnote mentions a pedagogical move made by the teacher to restate points made by the two students and ask the class about supporting evidence for the claims made. The teacher connected Shaurin's computer to the projector so that Shaurin could explain his recorded evidence to the class. Shaurin, explained the evidence that he had captured as a screenshot (Figure 4-8) to the class. Later, other students also presented their evidence. The teacher made another pedagogical move by asking students to decide if the evidence was comparable or not by comparing the experimental conditions that were used to gather the evidence. These pedagogical moves - a) asking a student to present recorded evidence from an investigation and, b) conducting a discussion about sufficiency of the evidence are considered ESM-enabled pedagogical moves because all the students were using the ESM as an experimental model system to investigate the same phenomenon, which allowed the teacher to make these moves to engage students in sharing and evaluation of evidence.

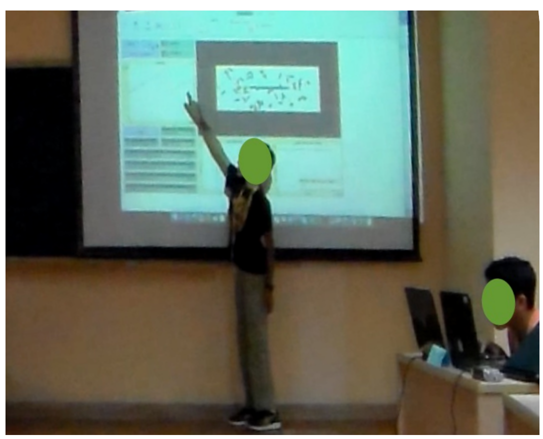


Figure 4-8 Shaurin explaining his evidence to the class that he recorded as a screenshot of the model to support his claim

All episodes related to knowledge construction and evaluation (n = 92) were further coded using a bottom-up approach to identify the objects (the intended goals of the activity), the design features as mediating tool (mediators), and accompanying pedagogical moves (mediation). Even though the ESM is a primary mediating tool for all the activities, I coded for specific design features of the ESM to further investigate how those design features mediated different aspects of knowledge construction. I identified how these design features are connected with restructuration properties of the ESM. For example, in the episode above, the object is 'to support a claim with evidence' and the mediating tool is 'experiments conducted using the ESM'. In this example, cognitive and social properties of restructuration are at play to support students' collective engagement in investigating a phenomenon.

Using this analytical approach, I identified three different objects of activity systems that were related to the process of knowledge construction, as well as the features of ESMs that mediated those objects, to construct a vignette. Since the research question that I investigate in this chapter is about design features of the ESM and restructuration properties instantiated through those design features, my analysis focuses on mediators and not the process of mediation. However, to explain how the mediators mediated subject engagement in achieving objects of the activities, I briefly discuss the process of mediation in the classroom community in the discussion section. I analyzed videos, student artifacts, and responses from students' post-interviews to perform triangulated micro-ethnographic analysis of the guided expansive learning in the classroom (Erickson, 1986; Small, 2011).

RESULTS

In the results section, I first present a quantitative analysis of the change in students' understanding of established ideas about genetics and evolution. Then, I present an analysis of students' pre- and post-interviews to discuss shifts in their perceptions about learning of science, specifically regarding their role and epistemic agency in the classroom. Finally, I will present a vignette that combines learning episodes and post-interview responses to highlight how different features of the ESM and ESM-based curriculum supported students' guided expansive learning.

Learning Genetics and Evolution

In the ESM-based curriculum students investigated molecular mechanisms of gene regulation in a cell and changes in a population because of evolutionary mechanisms – natural selection and genetic drift. There is a significant difference ($p < 0.005$) between students' pre-post scores in both the US and India ($n = 41$) (See Figure 4-9) (Dabholkar, Anton, et al., 2018). These differences indicate that students learned established scientific ideas regarding molecular genetics and evolution using the ESM-based GenEvo curriculum. This indicates that their investigations of phenomena related to genetics and evolution using the GenEvo ESM helped them develop a better understanding of scientifically established ideas, resulting in the post-test increase.

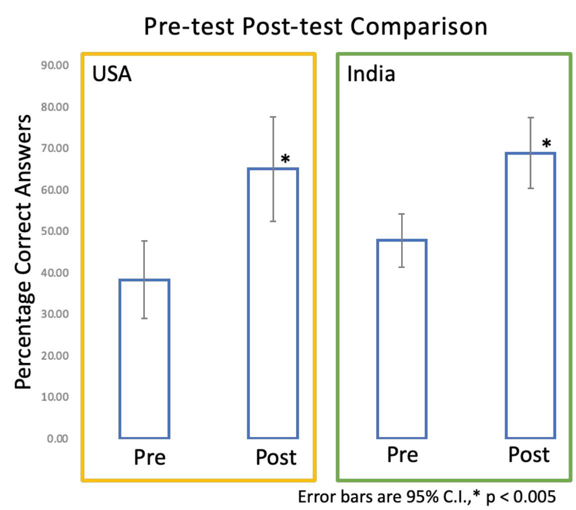


Figure 4-9 Students' performance in the pre- and post-tests about established ideas regarding molecular genetics and evolution

Shifts in Perception of Their Epistemic Agency

Using the codes mentioned in table 1, I identified a shift in students' perceptions about their agency in the process of constructing knowledge in a science classroom (Figure 4-10) (Dabholkar & Wilensky, 2019).

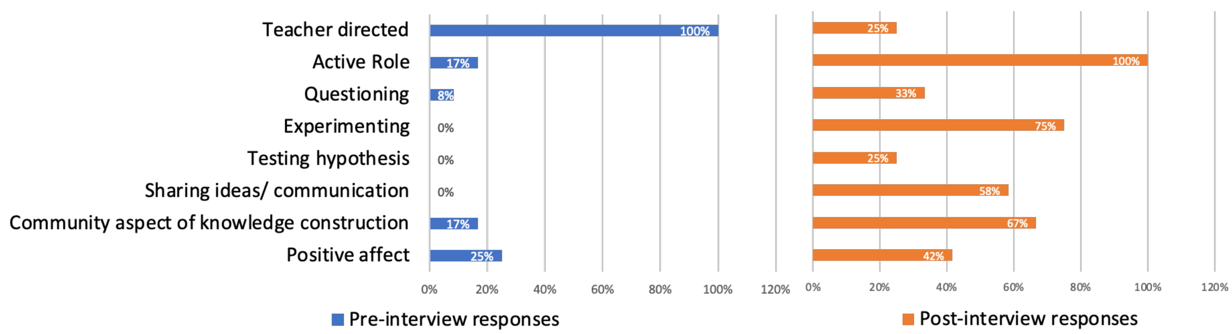


Figure 4-10 Students' perceptions about learning of science in a classroom

When asked about their past learning experiences in science classrooms, students viewed learning as a teacher-directed process, and they did not see themselves as having an active role in learning and knowledge construction (Figure 4-10). However, after participating in the ESM-mediated curriculum, 100% of the students talked about having an active role in the process of learning. The students specifically talked about the process of science that they engaged in when they discussed their learning in the course. It is important to note that none of the students even mentioned the process of science when they talked about their classroom learning before the course in response to the same question prompt.

More students indicated their positive affective engagement in the GenEvo learning experience. The following student responses are indicative of how affective and social properties of the restructuration collectively were at play to support student engagement in the ESM mediated learning activities.

When asked about how he learned what he learned in the GevEvo course, Pradeep responded as follows:

"*We got to play around with the bacterial cell. You on the first day didn't help us at all. You gave us the model and said figure out and think what you can. Then slowly we started to get answers. Then you helped us connect our thoughts. That's how we discovered what is what in that model.*"

[Pradeep, Post-interview, May-2018]

Pradeep described his experience of learning as a *playful* experience. His response also indicates how the ESM-enabled pedagogy of encouraging students to use an ESM as an experimental system resulted in collective knowledge construction. In his response, Pradeep talked about how initially he felt not being helped (through direct answers by a teacher) and compelled to figure out answers using the ESM. Pradeep also talked about the effectiveness of ESM-enabled pedagogical practice of asking students to share and collectively synthesize knowledge by *connecting thoughts* to *discover* modeled biological processes in the model. Samir mentioned something similar in his response about ESM-enabled pedagogy of a teacher *not telling answers* and having liked the experience of *figuring out answers together* [Samir, Post-interview, May-2018].

When Meera was asked about her learning experience in the GenEvo course, she said the following:

"*I learned it by [in a] very interesting way. You don't get to learn like that anywhere. We were ourselves trying to do experiments, and we were ourselves were trying to see what would happen if the cell lived in [a] certain kind of environment. I tried to play with the cell. I tried different environments, that's how I found out about things.*"

[Meera, Post-interview, May-2018]

Meera talked about this being a unique experience of learning in a *very interesting way*. She expressed her epistemically agentic experience of varying experimental conditions in the ESM and learning about the system by conducting computational experiments. Meera, similar to Pradeep, describes her learning experience as a playful experience. The playful perception of manipulating agent behaviors and environments to observe the effects and learn about the phenomenon are indicative of affective properties of agent-based restructurations in an ESM.

ESM-mediated Guided expansive learning

In this section, I present a micro-ethnographic analysis based on the episodes of student learning that were identified to be about the process of knowledge construction and on student responses about their learning in the post-interviews. I use the Activity Theory Lens to analyze mediation by different features of the ESM. The analysis focuses on how the restructuration properties of the ESM reflected through these design features mediated *objects* that were related to the process of knowledge construction. These episodes and interview responses are discussed through a vignette of a group of students - Vidya and Samir.

Careful evaluation of evidence to establish patterns

On the first day of the course, students explored a computational model of a cell and shared their observations regarding proteins, DNA regions, environmental conditions (availability of sugar), and the energy of the cell. On the second day, the teacher asked them to systematically investigate specific aspects of the model and to collect evidence to support their claims. This session turned extremely engaging for the students, in which they argued about the validity of their claims and observations using evidence that they collected. Vidya and Samir were members of a group that they named The Mad Scientists, which was perhaps the least vocal group in the class. I present a micro-ethnographic analysis of this session by focusing on the involvement of Vidya and Samir and their post-interview responses to illustrate how different cognitive and social properties of the ESM mediated students' knowledge construction regarding fundamental aspects in modern biology.

At the beginning of the session, the teacher asked students to answer a question regarding factors affecting changes in the energy of the cell. Using the ESM, students could change the environmental conditions, such as the availability of sugar, and observe changes in the energy of the cell as time progressed. Each student group presented their claims regarding the energy of the cell, based on their observations and collected evidence. Samir and Vidya performed a simple experiment. The following conversation took place between Samir and the teacher about it.

Teacher: Ok, let's start talking about the first question.
Samir: Page 2 or page 3?
Teacher: Page 2.
[The question was 'What are the effects of the presence or absence of sugar/s in the environment on the energy of the cell?'.]
Samir: When there is no sugar, at first, the cell will live comfortably, but then the energy will drop down drastically, and it will die.
The teacher nods. He walks towards the board and writes.
Teacher: So, no sugar… the cell will die. Does everybody agree?
All students: Yes
Teacher (to Samir): What is the second [condition that you tried?]…. Is that all? Or do you have anything else?
Samir nods negatively.
Teacher: That's all? Does any other group want to add anything?
[Transcription from video data, May 2018]

In this conversation, the teacher asked Samir to share his answer to the first question of the activity which was about an observation regarding the energy of the cell and sugar availability. Samir explained a simple experiment that Vidya and he performed regarding the survival of the cell when there is no sugar. The teacher asked other students if they agreed with Samir's statement that in absence of any sugar the cell would die. Since the students agreed with Samir's claim and the claim was valid, the teacher did not press for the evidence. This episode shows two pedagogical moves of the teacher – recording a claim made by a group (the teacher wrote the claim on the board) and encouraging sharing and evaluation of a claim (teacher asked other students about their agreement and if they wanted to add anything).

As the class discussion progressed, student groups shared their different environmental conditions and changes in the energy of the cell. The next group shared their observations regarding *cell division time* when both types of sugar – glucose and lactose, are present. Mitali, a member of that group, said that the cell division time was 108 ticks. Her partner Manav corrected her and said it was between 105 and 115 ticks. This generated a *heated discussion* in the classroom [Fieldnotes, May 2018]. Each group started saying that their observed number of ticks was different when a cell divided in the presence of both sugars. The teacher asked if the groups kept all the other parameters (genetic parameters) the same during the experiments.

To resolve the issue, the teacher asked all the groups to perform the same experiment and recorded their results on the blackboard. As the session progressed, the teacher asked the groups to repeat the experiment by changing the sugar availability. Since there are two sugars in the model, glucose and lactose, there were 4 possible conditions regarding sugar availability: no sugar, only glucose, only lactose, and both glucose and lactose. Though there was variation among the readings for a particular condition, the differences across the conditions were quite large. When asked for their conclusions the next day, Shaurin said, "*Even with the same parameters, you won't get the same results every single time…But there is some pattern*". When asked further about the pattern, Shaurin and other students added their conclusions regarding a shorter cell division time when glucose is present. In this exercise, students collectively developed an important insight regarding a practice of science, which is that there can be variability in observed data within the same experimental conditions. The ESM was designed specifically to have such variability in data because of randomness in agent behaviors to reflect how these molecules behave in nature. In spite of such variability in agent behaviors, and emergent outcomes of those behaviors, there are consistent patterns at the system level. For example, even though there is variability in cell-division time with lactose every time one conducts an experiment, cell-division time on an-average is more for lactose condition than that for glucose condition. With a well-designed experimental setup, one can establish such patterns through careful collection and evaluation of evidence.

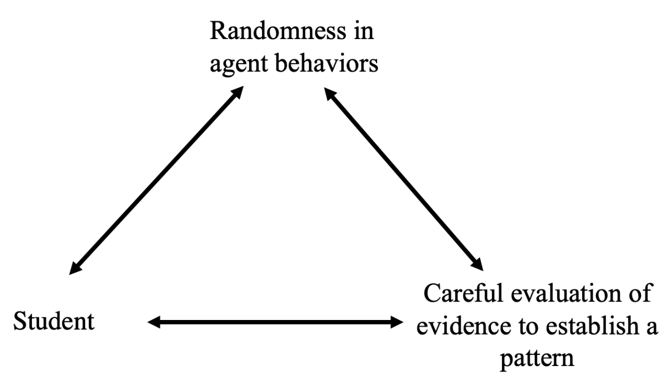


Figure 4-11 An activity theory triangle showing randomness in agent behaviors, a feature of an ESM, which required students to carefully evaluate evidence by repeated experimental trials to establish an observed pattern

In the Genetic Switch model of the GenEvo ESM, the behavior of agents (proteins) is random, which causes variability in results under the same experimental conditions. This feature is designed for students to learn about data variability in experimental systems and the robustness of certain observed patterns despite underlying variability. Because of this, the classroom community, including the teacher and the students, realized that they need to conduct multiple experimental trials to establish an observed pattern. The pedagogical move by the teacher to ask students to conduct the same experiment and recording the results enabled students to compare variations and identify a robust pattern.

Samir and Vidya did not contribute to these discussions about variability in experimental results and establishing robust patterns. However, on the next day, during a discussion about collecting evidence, Vidya mentioned creating a table as a different form of collecting evidence in comparison to taking a screenshot. This indicates that Vidya's indirect participation in the co-shaping investigatory practices, in this case, data collection practices, helped her appreciate *making a table* as an evidence-gathering practice.

Students continued to engage in discussions and arguments to establish what caused fluctuations in cell energy and how those fluctuations were connected to environmental conditions and protein production. Over the span of the next seven days, they collectively conducted more than 145 computational experiments, which were documented in their presentations. The activity and discussions analyzed in this section so far demonstrate how one design feature of the ESM – randomness in agent behaviors – mediated the object of careful evaluation of evidence to establish a pattern (Figure 11). This led to this community developing two practices for carefully evaluating evidence: conducting multiple trials and carefully observing the mechanistic details by slowing down a model. More rigorous forms of these practices are prevalent in modern biological research: statistical analysis and time-lapse microscopy.

Development of shared vocabulary to construct knowledge

The next part of the course involved understanding different types of proteins and regions of DNA and their functions. Computational objects in the ESM, proteins and regions of DNA serve as objects-to-think-with (Papert, 1980; Turkle, 2011), which students can manipulate and investigate. Since, these are computational agents, I refer to them as agents-to-think-with. In order to share their findings in ways that made sense to others, students needed to name these computational agents in the model. Students needed to arrive at a shared language and to establish the properties and function of the agents. The representational features (the shapes and colors of these objects) were chosen for easy identification and description; the behaviors and functions of these agents were based on the established ideas of molecular genetics (for example, promoter, operator regions of DNA, changes in binding affinity of LacI protein after lactose binds to it). For example, Samir established that the potato-shaped things inside the cells were special proteins, called RNA polymerases (RNAP), which moved on DNA to make other proteins.

> "*The first think that I discovered was that RNAPs were the potato-shaped things. So when I put RNAP to zero, potato-shaped things disappeared. So, I concluded from that the potato-shaped things are RNAP. Next, I figured out about LacI. Next, I figured how the pink triangles and reactangles are formed. When the potato-shaped things roll over the DNA and the bond is open, they produce LacZ and LacY. I was observing [potato-shaped things]. So first I observed that it was just random movement. Then I saw that it was going on a straight line (on DNA), so I saw that it was rolling along the DNA. And then suddenly, when it went off pink triangles and rectangles were produced. I did this experiment 2 or 3 times and then I figured out that the RNAP produced LacZ and LacY and when one RNAP rolls it produced 5 LacZs, from the graph I figured out*."
>
> [An excerpt from post-interview, May 2018]

In his response to a question in the post-interview, Samir described how he learned about the function of a protein called RNA polymerase. Using the ESM, Samir investigated the movement of RNA polymerase (represented as a “potato-shaped thing”). During the class discussions earlier, Samir had come up with the name “potato-shaped things”, when he talked about RNA polymerases. He observed a pattern that that was related to the production of other proteins, which were represented as pink triangles and rectangles. Samir hypothesized that the movement of RNA polymerase on DNA is related to protein production. Samir carefully established this pattern of agent behavior by repeating the experiment a few times under the same conditions.

In the class, as students developed practices for careful evaluation of evidence, they also developed shared vocabulary, such as calling computational agents “potato-shaped things” and “pink triangles” to talk about biomolecules and their interactions inside a cell. This vocabulary also evolved from “pink and purple stuff” to pink triangles, pink rectangles, purple keys, and so on. The pint tringles and rectangles in the models are genes whose scientific names as LacY and LacZ. In his answer, Samir used these words interchangeably. When Samir was talking about his observation of a phenomenon of protein production, he mentioned pink triangles and rectangles. When talking about the role of RNAP in protein production, Samir mentioned the scientific names, LacY and LacZ (see the excerpt from his interview).

The development of shared vocabulary is an important aspect of science practice. For example, naming conventions of organisms such as SARS-CoV-2, allows scientists across the world to easily share their findings. Sometimes, scientists also use unusual or funny names for naming a species, a star, or a gene. For example, there are Drosophila genes named Swiss

Cheese, or Cheap Date, or Boss gene (bride of sevenless – because of its connection with another gene called sevenless), or INDY (I'm Not Dead Yet)[10].

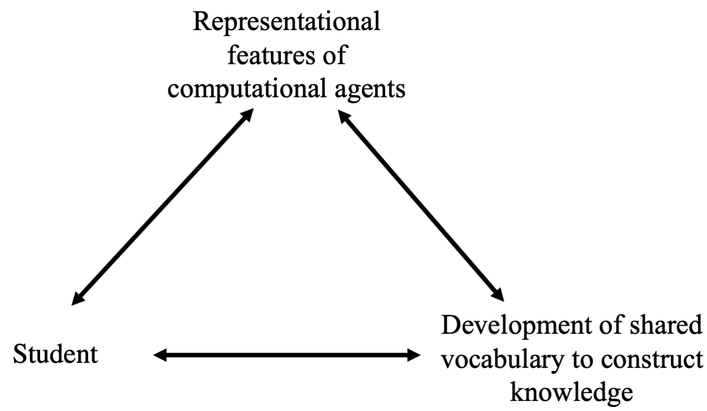


Figure 4-12 Representational features of computational agents in the ESM supported student developing a shared vocabulary to construct knowledge

Representational features of computational agents in the ESM helped them become agents-to-think-with because students could easily reference those to discuss their properties, interactions, and mechanistic involvements in a phenomenon under investigation. The teacher supported students' use of colloquial words as long as they established a shared understanding of those words in the context of the ESM. In a high school classroom in general, in a science classroom in particular, and in a biology classroom most specifically, vocabulary is considered extremely important. In this ESM-based class, the *representational features of computational agents* helped students to develop shared vocabulary as they investigated the function of these proteins and uncovered how genetic regulation worked in this system (Figure 4-12). The ESM

[10] https://www.lubio.ch/fruitfly-gene-names
These names are based on the physical or behavioral characteristics of mutants of these genes. Brains of fruitflies with a mutation in the Swiss Cheese gene look like swiss cheese. Fruitflies with a mutation in the gene Cheap Date are very susceptible to alcohol.

also had scientific names for these molecules. As seen in Samir's response, students established mapping between the *given names* and scientific names of protein agents in the model.

Reasoning about complex emergent patterns

Samir's carefully verified pattern became a piece in the puzzle that Vidya used to understand and explain the emergent nature of the molecular mechanisms of genetic regulation. Figuring out the effect of environmental conditions on cell growth and protein functioning was just the beginning. The core of the GenEvo course was about putting these ideas together to understand how and why cells regulate the production of certain proteins depending on environmental conditions. Production of proteins requires energy, so from an evolutionary perspective, it makes sense to produce proteins that are required in a particular environmental condition. The system of genetic regulation modeled in the ESM (lac operon) is evolved to regulate the production of proteins required for taking lactose (a type of sugar) into a cell and digesting it to produce energy. These proteins are produced only when lactose is the most preferred energy source in the environment. Using the Genetic Switch model in the ESM-based curriculum, Vidya could simultaneously observe changes at the intracellular interactions (micro-level) between protein, molecular, and DNA regions as well as the cellular level (macro), such as the energy of the cell and cell division time. In order to reason about complex emergent patterns, thinking across levels is important (Wilensky & Resnick, 1999). The visualization of processes across levels in the ESM allowed Vidya to reason about emergent patterns regarding genetic regulation.

When asked about her learning, Vidya responded as follows:

"(I learned) the cell's way of regulating the production of specific proteins that are needed because they eat up some energy. Because every protein has its cost, so a cell has to know when it is necessary to make it and not just make it when it's not needed….. Because, it also degrades, so it's of no use….. So, the cell's way of doing that is to produce LacI, which is.... when there is no lactose, it can join with the DNA and it can prevent the formation of LacY and LacZ by RNAP, but when there is lactose, it is unable to do so, because it is blocked by the presence of lactose. (I learned it) by piecing something together. It just came to me, I guess! Before that we were discussing, the LacI and lactose binding thing…. I was wondering why this happened. And then Samir [her partner] found out that when LacI is bound, the RNAP doesn't roll. Then I just thought of it."

[An excerpt from post-interview, May 2018]

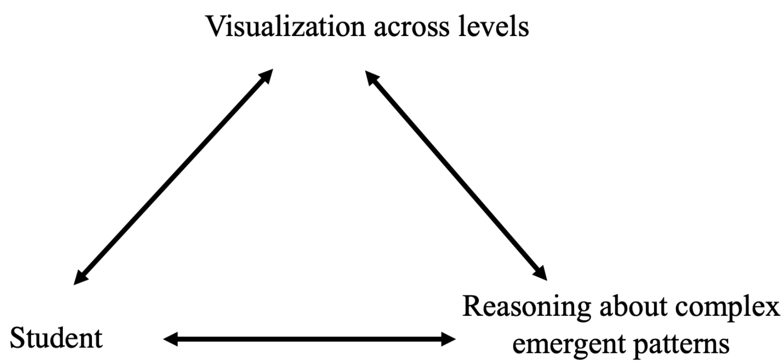


Figure 4-13 Visualization across micro (agent) and macro (system) levels in an ESM mediated students' reasoning about complex emergent patterns regarding genetic regulation

Vidya explained how she learned a highly advanced emergent phenomenon of molecular genetic regulation by observing interactions between DNA and proteins as computational agents in the model. These agent-level observations of objects-to-think-with and system-level changes in energy allowed students to reason about genetic regulation (Figure 4-13). Vidya mentioned in her response that the production of specific proteins needs to be regulated because they eat up some energy, which is costly for a cell. Vidya and Samir could observe protein behaviors and corresponding changes in cellular energy, which allowed them to reason about the molecular mechanism of LacI regulating the production of LacY and LacZ proteins. Reasoning about complex emergent patterns is an intended as well as enacted object of the activity system. It was an intended object because the designer of the learning environment wanted this to be an object for the students. The evidence demonstrates how it became an enacted object, because the ESM-based curriculum created opportunities for students to identify emergent patterns as puzzling patterns that demanded explanations and engaged them in designing experiments to construct those explanations. To develop complex systems perspectives about the processes of gene regulation and mechanisms of evolution was one of the intended learning outcomes of the ESM-based curriculum. To explain her mechanistic reasoning about the phenomenon under investigation – stimulus-based regulation of protein production, Vidya *pieced together* her reason about the complex pattern of cellular behavior that emerged through agent-level interactions.

The cognitive properties of agent-based restructurations allow learners to engage simultaneously with micro-level interactions, such as DNA-protein interactions, and macro-level patterns, such as stimulus-based regulation of protein production. Vidya and Samir's participation discussed through the vignette provides an example of how social properties of restructuration support collective knowledge construction through easy sharing, evaluation, and incorporation of ideas.

DISCUSSION

In this chapter, I demonstrated how cognitive, social, and affective properties of restructuration in an ESM-based curriculum mediated students' guided epistemically expansive learning in a science classroom. To shift student roles from *receivers of facts* to *doers of sciences,* it is important to design learning environments that consider and attempt the problem of practice (Miller et al., 2018; Russ & Berland, 2019). To foreground such epistemically agentic participation, based on the analysis presented in this chapter, I conceptualize *epistemic expansiveness* of a learning environment is in terms of providing students opportunities and ways to shape epistemic practices to investigate a phenomenon. I designed an Emergent Systems Microworld (ESM) as an experimental model system for students to that supported epistemically expansive learning of phenomena related to gene regulation and evolution. I analyzed how restructuration properties of an ESM and ESM-enabled pedagogy supported students' engagement in shaping science practices to investigate a phenomenon and establish knowledge of various aspects of the phenomenon as a classroom learning community.

Using Cultural Historical Activity Theory (CHAT), I demonstrated how restructuration properties instantiated through design features of an ESM-based curriculum mediated the transformation of classroom activity. As the activity system transformed the enacted object to align it with the intended object, the classroom activity shifted to the social construction of disciplinary knowledge. The cognitive, social, and affective properties of restructuration mediated this shift. In many traditional biology classrooms, when students are positioned as *receivers of facts*, the intended and enacted object for students is to listen to disciplinary ideas explained by a teacher using static representations. The ESM-based curriculum mediated a transformation of this activity system to position students as *doers of science* by engaging them in the social construction of disciplinary knowledge and shaping epistemic practices using an interactive experimental system that included agent-based restructurations (Figure 4-14).

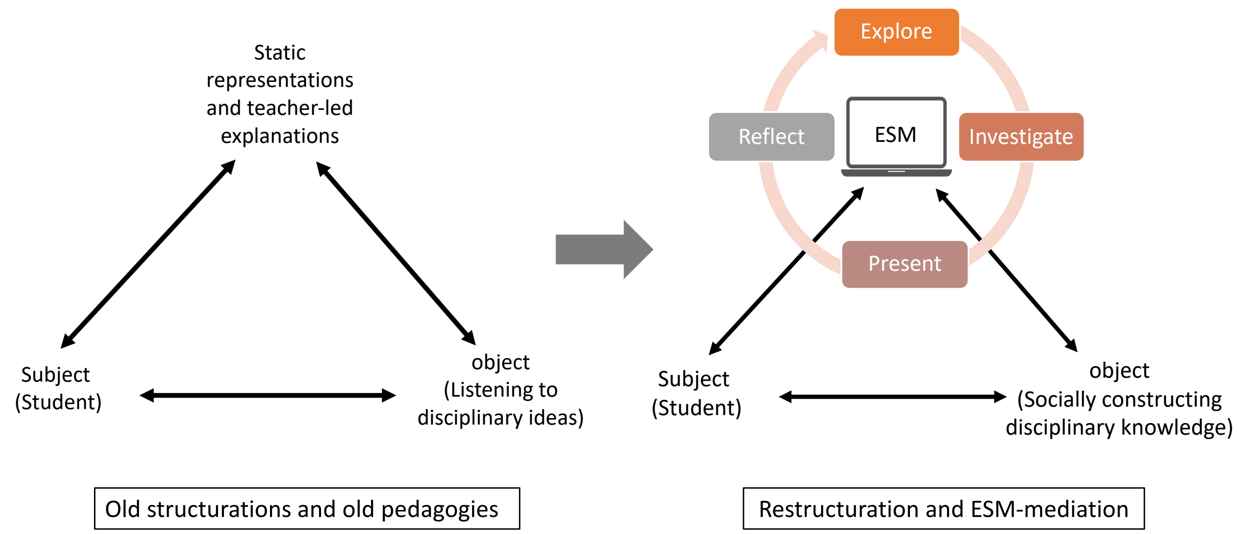


Figure 4-14 The change of an object of an activity system. The figure on the left represents a subject-object-tool activity system triangle for a traditional pedagogy. The figure on the right shows the change in the object of a classroom activity system through exploration, investigation, presentation, and reflection, mediated by an ESM

Socially constructing disciplinary knowledge in a science classroom is an epistemic process that is centered in educational reforms, including the Next Generation Science Standards (Abd-El-Khalick et al., 2004; Duschl, 2008; NGSS Lead States, 2013). However, it is challenging to support students' epistemic agency so that they construct disciplinary knowledge about not only the core ideas but also about the practices that professionals engage in when constructing those ideas (Miller et al., 2018; Russ & Berland, 2019). Using a mixed-methods analysis, I provided evidence for students' improved understanding of established disciplinary ideas, perceived a shift in their epistemic agency, affectively positive engagement, and social construction of disciplinary knowledge. Students collectively constructed disciplinary knowledge by shaping science practices because of the restructuration properties of an ESM-based curriculum instantiated through its design features.

An ESM has computational objects-to-think-with in an interactive microworld which is designed using agent-based representations (Papert, 1980; Wilensky, 2001; Wilensky & Papert, 2010). The GenEvo ESM discussed in the paper allowed students to investigate the natural phenomena of genetic regulation and change in a population because of evolution. The ESM presented to students mediated a shift in classroom activity as students *explored* the model system, *investigated* specific aspects of it, *presented* their findings to others, and *reflected* on the process of learning and figuring out (Figure 4-14). The cognitive properties of agent-based restructuration in an ESM reduced perceptual limitations by providing visual access to agent behaviors and emergent patterns (Goldstone & Wilensky, 2008; Levy & Wilensky, 2009; Wilensky, 2003). The learners manipulated agent-level behaviors of proteins and parts of DNA to investigate the effects of those manipulations at the system-level regarding gene regulation, which allows learners to overcome the confusion of 'levels' (Wilensky & Resnick, 1999). Using agent-based models students can easily ask and investigate what-if questions by changing agent behaviors in the code (Wilensky & Reisman, 2006) or through interface elements as the students of the GenEvo course did. Vidya, Samir, and others in the classroom community learned about the fundamental aspects of gene regulation by starting with simple experiments and observations using the ESM and then developing more sophisticated epistemic practices. As witnessed in their interview responses and in the analysis of their classroom participation, they displayed affectively positive engagement in the social construction of knowledge using the GenEvo ESM. The playful nature of agent-based restructurations and ease of sharing, testing, and incorporating each other's ideas were important aspects of social and affective properties of ESM that supported student learning.

TRANSFORMATION OF THE CLASSROOM ACTIVITY SYSTEM

The finding shows how three features of the ESM, (1) randomness in agent behaviors, (2) representational features of computational objects, and (3) visualization across levels, mediated students' engagement in intermediate objects as they progressed with constructing knowledge of disciplinary ideas and shaping scientific inquiry practices. The three objects of intermediate activity systems in the ESM-based learning environment are - (1) careful evaluation of evidence to establish a pattern, (2) development of shared vocabulary to construct knowledge, and (3) reasoning about complex emergent patterns. Student enactment towards achieving these objects, in turn, mediated the transformation of the classroom activity system through epistemic expansion. The object of the classroom activity system became the social construction of knowledge about complex disciplinary ideas.

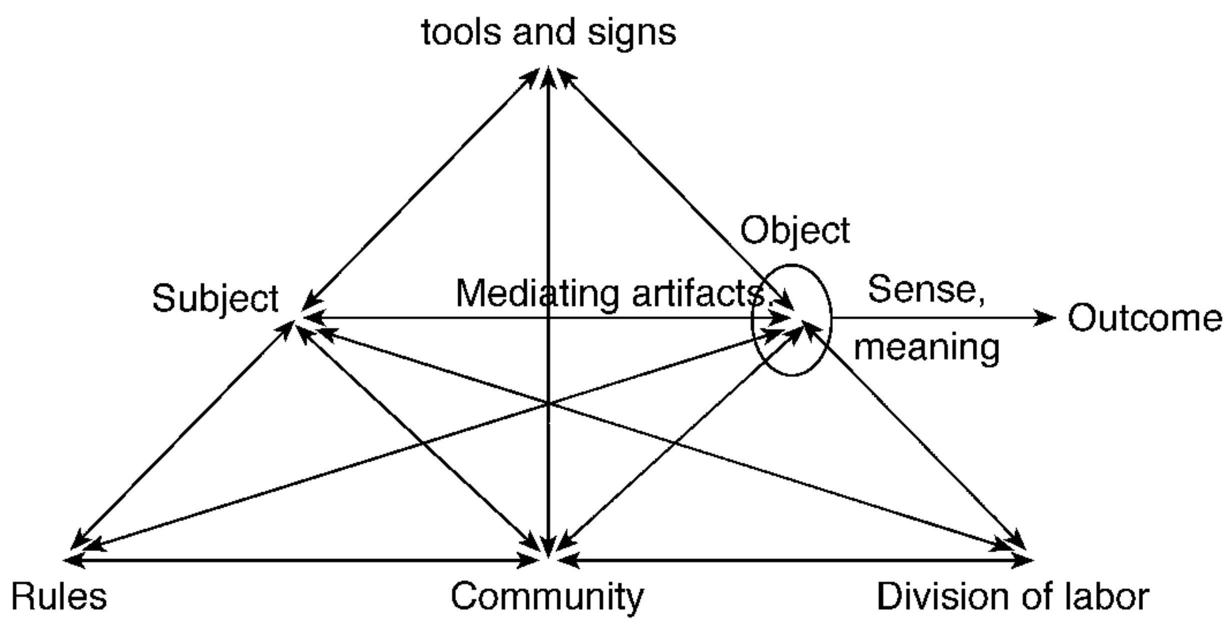


Figure 4-15 Engeström's structure of a human activity system (From (Engeström, 2001), pg,135) detailing how rules, community, and division of labor influence and are influenced by the subject, object, and mediational tools

There are four additional components in Engeström's structure of a human activity system that are important to consider in the context of epistemic expansion. The first component is an *Outcome* of learning. As the students engage in the social construction of knowledge about a complex phenomenon in the GenEvo learning environment, their learning outcomes are three-fold – they develop knowledge about disciplinary ideas about genetics and evolution, they develop knowledge about *doing science* by shaping scientific inquiry practices to investigate a phenomenon and they shift their perceptions about their agency in the process of knowledge construction. The third learning outcome is about their *epistemic consciousness,* the realization that they can participate in constructing knowledge and devising ways to construct knowledge.

In the ESM-mediated epistemic expansion discussed in this paper, the shifts in the other components of the activity system – rules, community, and division of labor, are also salient. Students and the teacher are part of the classroom community. Different features of the ESM-based learning environment and ESM-enabled pedagogical moves shaped interactions of the community as different rules regarding student engagement and divisions of labor were formulated in the classroom learning environment. For example, in the vignette discussed in the paper, inherent randomness in agent behaviors required students to conduct multiple trials, collect and analyze data systematically to establish a pattern. In the classroom community, student groups performed investigations separately – dividing the labor. The ESM-enabled pedagogical moves of the teacher, such as encouraging sharing and evaluation of evidence, encouraging the use of collectively emerged colloquial words for identifying and addressing computational objects, supported students' active role in the evolution of rules of interactions in the classroom. For example, the analysis demonstrates how certain rules evolved in the community regarding what counts as evidence (one trial vs multiple trials), how to present and evaluate the evidence. This shift in terms of establishing rules regarding evidence-gathering practices took place without the teacher directly telling students how they should collect evidence.

Conclusion

To support students' epistemic agency in a science classroom, it is important to design an environment that can provide them with opportunities to engage in epistemically expansive learning. The GenEvo learning environment, which includes the ESM-based curriculum and ESM-enabled pedagogy, facilitated students' engagement in epistemically expansive learning experiences by socially constructing knowledge of disciplinary ideas and scientific practices. The evidence presented in this paper demonstrates that the students learned about genetics and evolution and perceived shifts in their epistemic agency in the classroom. The cognitive, social, and affective properties of restructurations instantiated through design features of the ESM enabled such student engagement in learning and learning outcomes. The analysis presented in the paper highlights design features of the ESM-based GenEvo curriculum that acted as mediational tools for supporting students' epistemic expansive learning. The identified design features and complementing pedagogical practices serve as guiding ideas to design for and support students' epistemically expansive learning.

Chapter 5: ESMs to learn about population dynamics and evolution

Summary: This chapter describes three related Emergent Systems Microworlds (ESMs) and an ESM-based curricular activities that are designed to engage students in investigating and learning about population dynamics, population genetics, and the evolution of populations. The ESMs are modelled based on published research and curricula about a population of rock pocket mice in the desert of New Mexico. In this chapter, I first give an overview of the system that is modeled in this set of three related ESMs of rock pocket mice. Then I describe the agents, their behaviors and interactions, and the emergent patterns that users can investigate using this ESM. Finally, I present examples of ESM-based curricular activities that use one of the ESMs for students to engage in investigating the phenomenon of natural selection.

OVERVIEW

The designs of the ESMs discussed in this chapter are inspired by a short film designed by Howard Hughes Medical Institute (HHMI)[11] and an Advanced Placement (AP) Biology Lab ("Lab 8," 2001). The HHMI film is about natural selection and adaptation in populations of rock pocket mice living in the American Southwest. Mice living on light-colored sand tend to have light-colored coats, while mice living on patches of dark-colored rock have mostly dark-colored coats. Michael Nachman, a population geneticist, studied the evolutionary processes that led to these differences in fur-coat-colors in rock pocket mouse populations (Nachman et al., 2003). The HHMI film gives a geological history of American Southwest and discusses Nachman's

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

research about changes in populations of rock pocket mice. Rock pocket mice belong to a species of mice *Chaetodipus intermedius*, which are mainly found in rocky outcrops in the deserts of the southwestern United States and Mexico. Mice that camouflage well with the surroundings have a higher chance of surviving predation by avian predators such as hawks and owls. The mice that are found in sandy areas typically have light fur-coat whereas the ones found in rocky areas have dark colored fur-coat. Evolutionary biologists and molecular geneticists who have studied these populations in New Mexico have attributed this adaptive variation in the color of fur coat primarily to mutations in a gene called the melanocortin-1-receptor gene, *Mc1r* (Nachman et al., 2003).

In this chapter, I first explain the design of the ESM and they discuss curricular activities in the ESM-based unit. These ESMs are designed as computational experimental systems for students to learn about phenomena related to population dynamics, population genetics and natural selection. Students are expected to use the ESMs to design and conduct computational investigations. Each of the three ESMs is in the form of a NetLogo model which consist of a population of rock pocket mice. In this population, the color of the fur coat is determined by a single gene with two alleles, *A* and *a*. Presence of a dominant allele (*A*) results in a mouse with a dark fur color. A homozygous recessive mouse (*aa*) has light-colored fur. Inheritance of these alleles is modeled based on the Mendelian mechanism of inheritance. Mating between individuals is spatially restricted by allowing an individual to mate with another individual of the opposite sex that is within a specified distance. Other than this, there are no mating preferences, which makes the mating random. With no natural selection or predation, this population should attain Hardy-Weinberg Equilibrium[12], with the resultant ratio of the light and dark fur populations.

I have designed three related ESMs of rock pocket mice populations in order to allow a user to investigate specific aspects of population genetics and evolution.

1. Hardy-Weinberg Equilibrium

This ESM is designed to study changes in allele frequencies in a population when a population obeys Hardy-Weinberg assumptions: (1) random mating, (2) the absence of natural selection, (3) a very large population size (i.e., genetic drift is negligible), (4) no gene flow or migration, (5) no mutation, and (6) the locus is autosomal. These assumptions are incorporated into the model design. The ESM does not have predators. Because of the absence of any other evolutionary influences than those because of the simple Mendelian inheritance, when a population size is reasonably large, it reaches Hardy-Weinberg Equilibrium after a few generations. The design of this ESM is inspired by an activity in an Advanced Placement Biology Lab ("Lab 8," 2001). This activity is designed to simulate a breeding population and track changes in allele frequencies. In this activity, a class represents a breeding population, Students use cards with allele type 'A', or 'a' written on those. Each student receives four such cards. Students are asked to form random pairs. Each member of a pair shuffles the cards and picks a card at random. The two such randomly picked cards make the genotype of an offspring. For example, I pick a card 'a' and my partner picks a card 'A',

[12] https://en.wikipedia.org/wiki/Hardy–Weinberg_principle

the genotype of ‘our offspring’ is ‘aA’. My partnering teacher, Ms. Tracy[13], who I co-designed this curriculum with (See Chapter 8 for more details of the co-design work), had conducted this AP lab activity for several years in her class. This is an agent-based activity. After Tracy shared this activity with me, I designed an ESM based on the rules of the activity.

Using this ESM (Figure 5-1), users can investigate how the allele frequency values (p and q) and the additional Hardy-Weinberg equation vales (p2, q2, and 2pq) vary as populations change over time. They can also observe phenotype frequencies of light and dark-colored mice. Each mouse has a pair of alleles – *a* and *A*. aa mice are homozygous recessive, which have light fur-coat, whereas *aA* and *AA* mice have dark fur-coats. Since the assumptions of Hardy-Weinberg Equilibrium are obeyed, the equilibrium emerges in the population after a few generations: The genotype and phenotype frequencies match with Hardy-Weinberg equation predicted values.

[13] A pseudonym

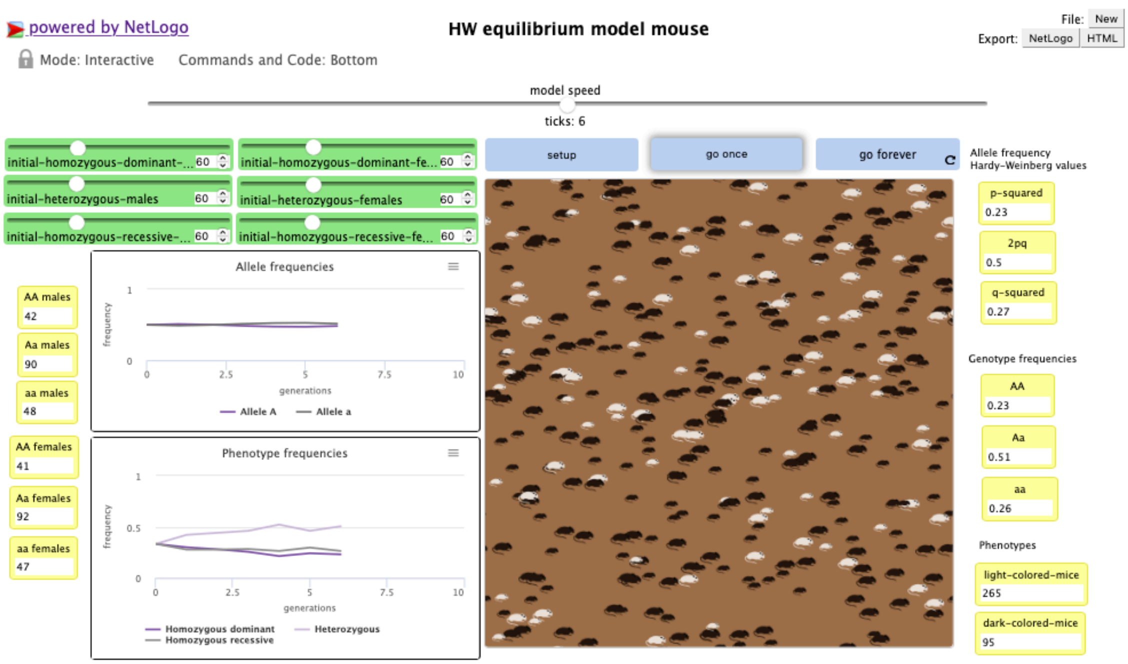


Figure 5-1 A rock pocket mice ESM to study Hardy-Weinberg Equilibrium (Dabholkar & Wilensky, 2020b)

2. Natural Selection – camouflage (simple)

The next ESM about Rock-Pocket-Mice adds predators to the model (Figure 5-2). This ESM is designed to use for investigating the phenomenon of population change of rock pocket mice in the dessert the desert of New Mexico, specifically in the area of Valley of Fire, based on the research of population geneticists and evolutionary biologist, the current understanding about the population change in the mice population in as follows. The rock pocket mice in this region were in the light-colored and they blended in with the sands of the dessert. About a thousand years ago a volcano erupted and deposited hardened lava in the Valley of Fire, which was of a dark color. With this new background, the light-colored RPM no longer blended in and were easy targets for predatory birds.

Even though the underlying model of genetic inheritance is the same as the previous ESM, Hardy-Weinberg values (p^2, q^2, and 2pq) are neither calculated nor displayed. This is because is ESM is designed to be an introductory model system that focuses on population change because of natural selection. *Selection pressure* in this model system is because of predation. The probability of a mouse depends on how well it camouflages and how many predators are present. There is a chance that any mouse gets predated at each time-step. This chance is modeled by two factors: (1) how well the mouse camouflages with its surrounding patches, and (2) the chance-of-predation slider, which can be adjusted by users. The chance-of-predation parameter models the number and hunting ability of predators. Users can set initial population composition, background color, and predation values to study emergent patterns regarding the evolution of a population in different conditions. Users can also use a button to add a mutant (a mutant is a mouse with a mutation in an allele that determines fur coat color), which are dark-colored mice, and investigate their survival in different environmental conditions.

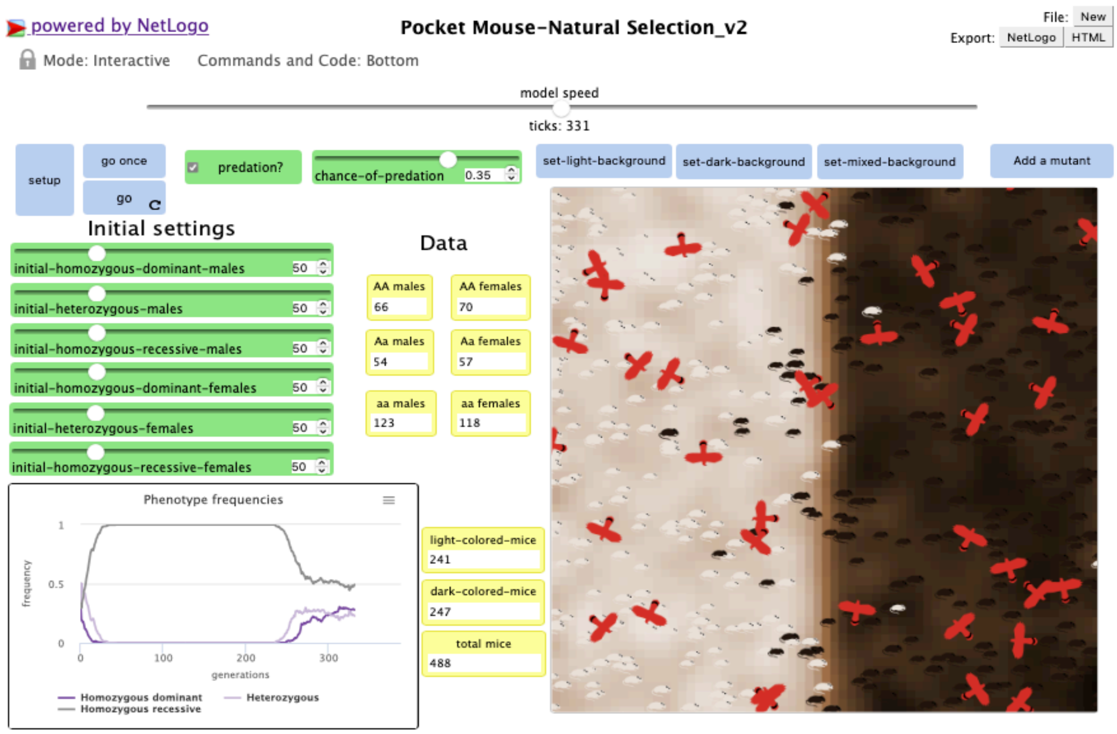


Figure 5-2 A second rock pocket mice ESM to study changes in a population because of natural selection (Dabholkar & Wilensky, 2020a)

3. Natural Selection – camouflage (advanced)

This advanced version of a natural selection ESM combines features of the first and second ESMs (Figure 5-3). Users can investigate how Hardy-Weinberg equation values change when there are other evolutionary influences, such as natural selection because of predation.

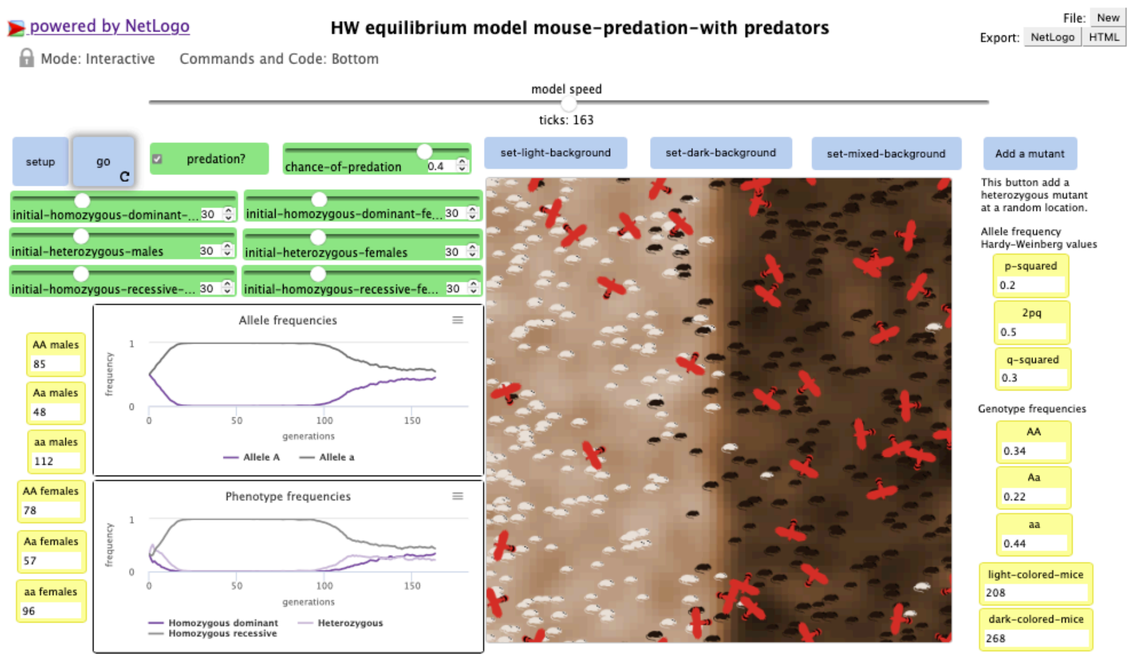


Figure 5-3 A third rock pocket mice ESM to study population genetics using the Hardy-Weinberg Equation, as a population changes because of natural selection (Dabholkar & Wilensky, in preparation)

I have co-designed two different ESM-based units that use these ESMs (Dabholkar, Granito, et al., 2019; Dabholkar, Woods, et al., 2018). In the next chapter, I analyze student learning of scientific inquiry practices and disciplinary ideas related to natural selection with an ESM-based unit that uses the second ESM. In the following sections, I discuss the design of the second ESM in detail and explain design of some of the ESM-based curricular activities. In this section, I explain the components of the ESM – agents, their properties, their behaviors and interactions, and emergent patterns.

Agents

There are only two groups of agents in this ESM: mice and birds (predators). Mice are the main agents in this ESM in terms of their modelled behaviors and interactions. Since the

ESM is designed to study changes in population of mice, in the initial versions of the ESM mice were the only agents in the ESM. Their predation was modeled abstractly by computing the probability of death-by-predation for every mouse based on how well it camouflaged. Based on feedback from a teacher doing the co-design process, birds were added as agents in the model for pedagogical purposes- for students to visualize presence or absence of predators in the ecosystem. This, the bird agents are used only for visualization purposes so that a user can see if predation is taking place or not. The number of birds in the model is based on the change-of-predation value set by the user. Even in the current version of the model (Dabholkar & Wilensky, 2020), the predatory behavior of birds is modelled indirectly by computing the probability of predation of a mouse based on camouflage and presence-of-predators (values set for predation? and chance-of-predation in the model).

Agent Properties

In this subsection, I discuss the properties of mice agents. Fur color is the most important property of a mouse in this ESM. The fur color of a mouse is determined by its genotype: homozygous dominant and heterozygous mice have dark fur color, whereas homozygous recessive mice have light fur color. The fur color is determined by genes at a particular locus for which there are two alleles: *A* and *a*. *A* is a dominant allele, whereas *a* is a recessive allele. So, *AA* and *Aa* mice are dark-colored, whereas *aa* mice are light-colored.

Sex is another property of a mouse. A mouse can be male or female. Users can set the number of male and female mice in the initial population. For simplicity, after mating, a mouse pair will produce two male and two female offspring. There are no overlapping generations in the model, which means that parents die after giving birth to offspring. Since the population declines due to predation, to avoid unnatural fluctuation in the population and roughly maintaining sex-ratio two male and two female offspring are produced as a result of each mating event. Users can potentially changes this to make the generations overlap or determine the number and sex of offspring probabilistically to investigate effects of these changes on population dynamics in presence and absence of natural selection.

Age is the third property of a mouse. If a mouse is not predated, it ages and eventually dies.

Agent Behaviors and Interactions

When a user runs a model by pressing the button labeled GO, the model progresses temporally. The temporal progression of a model is shown by advancing the value of clock-ticks, which is an arbitrary time unit.

At each clock-tick,

Each mouse,

Moves (in a random direction)

Ages (*age* is increased by 1)

Dies if predated or if age is high (5% chance of death at every tick after age > 10)

If not dead, looks for a partner (within a specified area) and

Reproduces (if it finds a partner)

In the following section, I explain agent behaviors in more depth, as well as the reasons for modeling choices regarding agent behaviors.

Movement: At each clock-tick, a mouse moves in a random direction. The movement of mice is important for two reasons. First, mixing of the population is important to obey the Hardy-Weinberg assumption regarding random mating. Secondly, since users can set the background color to observe the survival of mice in dark or light background colors, the movement of mice creates a chance for a light-colored mouse to move to light-colored areas and vice versa.

Age: At each clock-tick, the age of a mouse is increased by one time-unit. After a mouse reaches the age of 10 units, there is a high probability (0.95) that it dies in any subsequent clock-tick.

Death by predation: If predators are present, the value of the *chance-of-predation* slider and the ability of a mouse to camouflage with the background determines the probability of its death by predation. A camouflage factor is calculated for each mouse. Since NetLogo uses a number to represent a color[14], the numerical values of color of a mouse and its surrounding patches are used to compute the camouflage factor. It a mouse camouflages well with the surroundings the value of this factor is high for the mouse. While determining if a mouse dies because of predation, its camouflage factor is subtracted from the chance-of-predation value. The color values for mice color and patch color (color of NetLogo patches in the surrounding environment of mice) are chosen such that the camouflage factor is in a comparable range with the *chance-of-predation* slider value. When a mouse camouflages perfectly with the background, the chance of death because of predation is zero.

[14] http://ccl.northwestern.edu/netlogo/docs/programming.html#colors

Search for a partner: At every clock-tick, each mouse searches for a partner of the opposite sex within a specified distance. Once a mouse is partnered, it cannot be partnered with other mice in the same clock-tick. In the model implementation, all mice first look for partners. After all the partner-pairs are formed, they reproduce. It is implemented in the model this way to avoid the same mouse being paired again after it has reproduced.

Reproduction and inheritance: A pair of mice, a male and a female, produce four offspring, two males and two females. The inheritance of genes determining fur coat color is modeled based on Mendel's laws. Each offspring receives an allele *A* (dominant) or *a* (recessive) randomly chosen from each parent based on their genotype. Each clock-tick is thus a generation of mice. To avoid overlapping generations, the parents die after the offspring are produced. Overlapping generations are avoided because of the simplicity, based on the design of the original AP biology lab ("Lab 8," 2001), which did not have overlapping generations.

EMERGENT PATTERNS

In this section, I discuss emergent patterns regarding changes in a population because of natural selection in the second ESM (Natural Selection – camouflage (simple)). I present an example of a model run, divided into 6 steps, to demonstrate how a user can observe and investigate these patterns.

1. A user can set up their initial population by selecting values for sliders *initial-homozygous-dominant-males*, *initial-homozygous-dominant-females, initial-homozygous-recessive-males, initial-homozygous-recessive-females, initial-heterozygous-males,* and *initial-heterozygous-females*. They can set the background color. They can choose

whether to have predators or not by checking the box *predation?* and can set the *chance-of-predation*. In Figure 5-4, an example of an initial setup is shown with a mixed population in a light background with predators (*chance-of-predation* = 0.35).

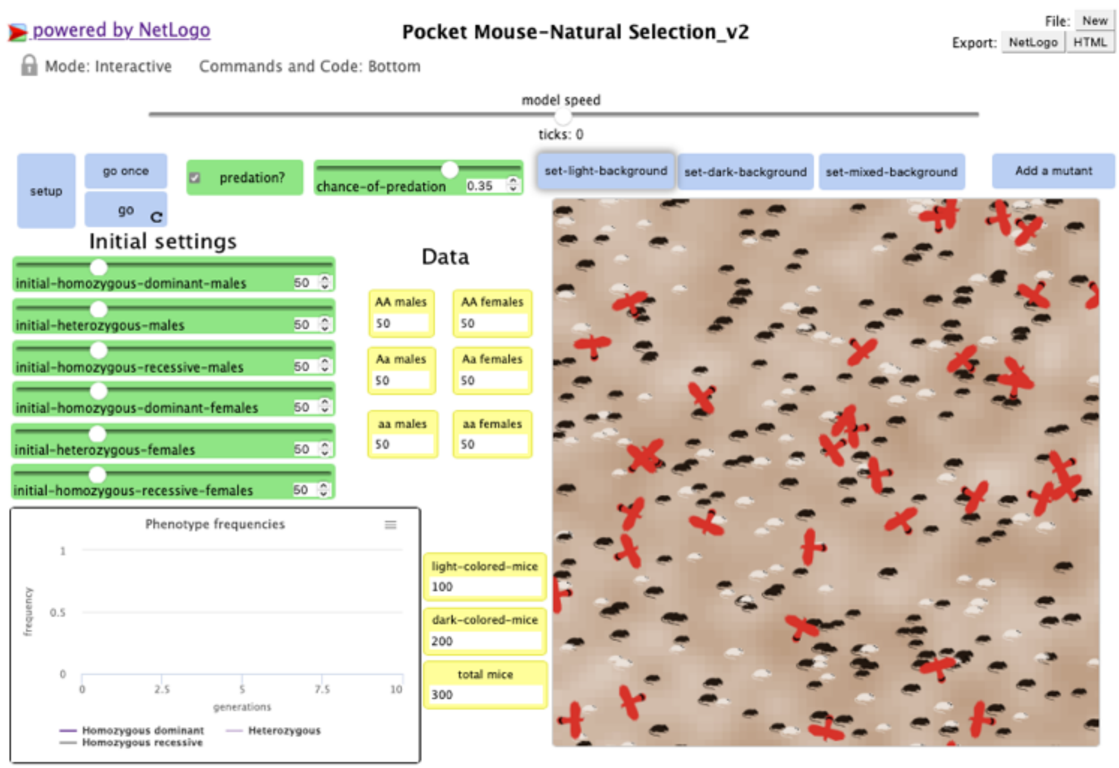


Figure 5-4 An example initial setup to investigate emergent changes in a population because of natural selection (Light background, mixed population)

2. Since dark-colored mice do not camouflage with a light background, after a few generations, the population consists of only light-colored mice (Figure 5-5). This is an example of an emergent pattern that arises without having assigned a direct survival advantage to any particular trait. At each clock-tick, dark mice are likely to be predated. This results in fewer dark-colored mice reproducing to make the next generation. Eventually, the entire population consists of light-colored mice.

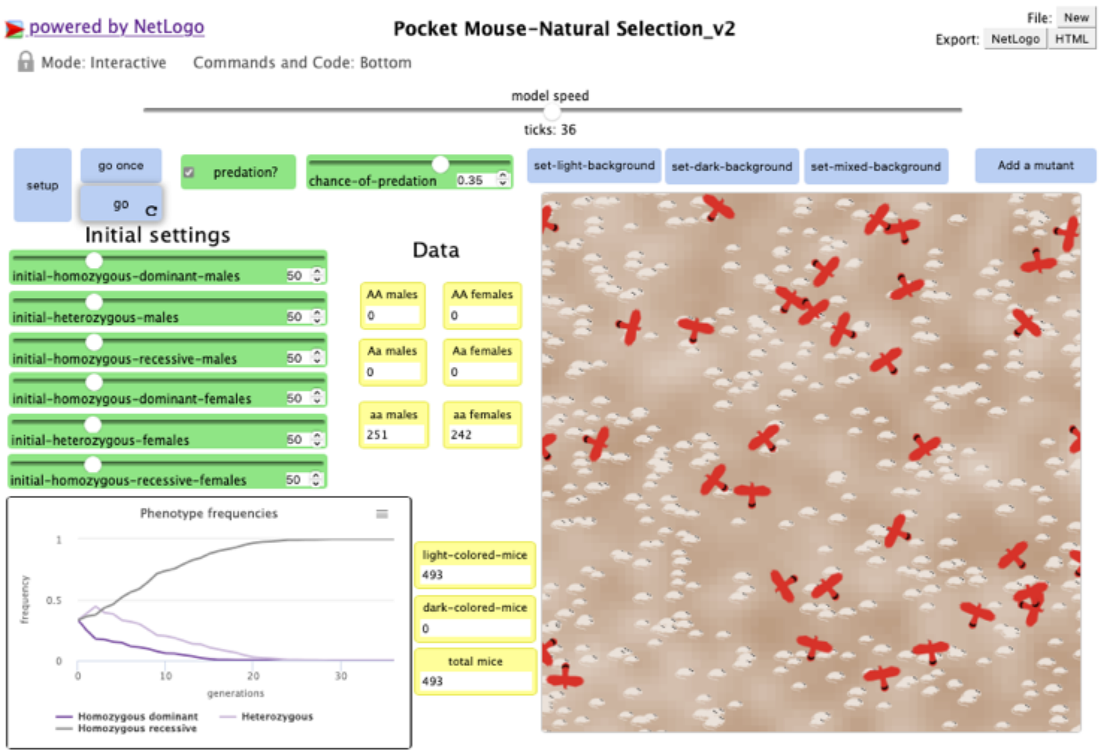


Figure 5-5 Changes in the populations after a few generations (generation number = 36). Mice with dark fur are predated and the population consists of light-colored mice

3. In the next step, the background is changed to a mixed-colored environment background (Figure 5-6). This background change is similar to what happened in the desert of New Mexico a thousand years ago. Because of a volcanic eruption, a part of the desert became rocky after the lava cooled down.

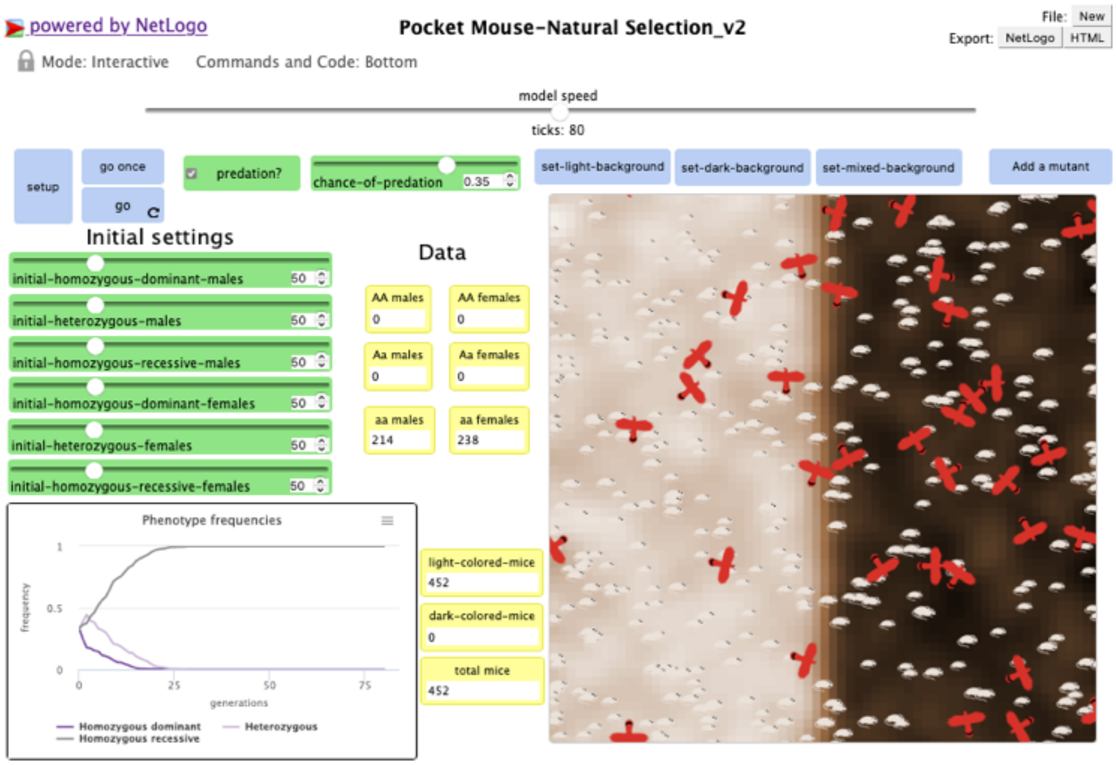


Figure 5-6 The background is changed after 80 generations to light (on the left) and dark (on the right).

4. Since light color is a recessive trait, the population consists of all homozygous recessive individuals. A mutation in one of the alleles (*a*) that makes it dominant (*A*), would result in a mutant mouse being a heterozygous dark mouse. A button in the ESM, '*Add a mutant*', introduces one heterozygous dark mouse in a random location in the environment. In the figure below (Figure 5-7) a mouse happens to be introduced on the dark side.

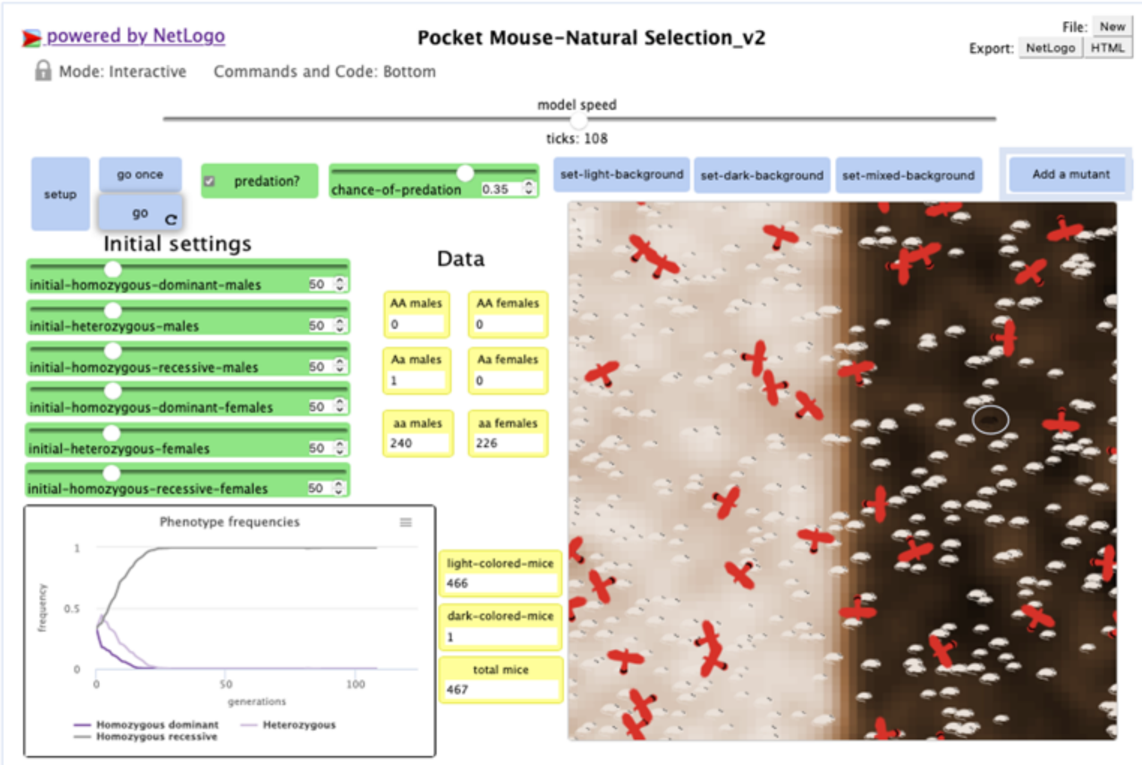


Figure 5-7 A dark colored (heterozygous) mutant is introduced to the population randomly. In this case, it happens to be on the dark side. The mutant mouse is circled for ease of viewing.

5. A user can now observe another emergent pattern regarding the spread of the mutation in the population. After a few (21) generations, the number of dark mice increases on the dark side (Figure 5-8).

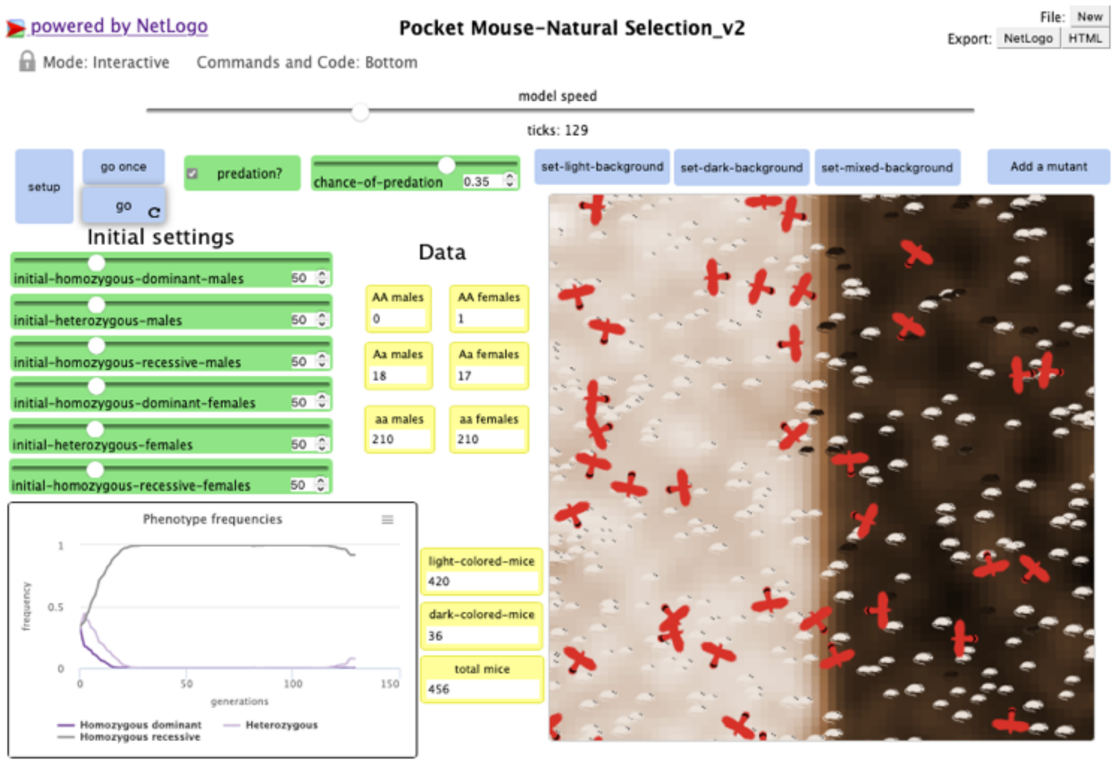


Figure 5-8 The subpopulation of dark colored mice spreads in the rocky area (dark side)

6. Finally, almost all the mice on the light side are light-colored mice and the ones on the dark side are dark-colored mice (Figure 5-9). This division of the population into two interbreeding subpopulations is also an emergent pattern. At the boundary of light and dark regions, one can observe dark mice in the light region and light mice in the dark region.

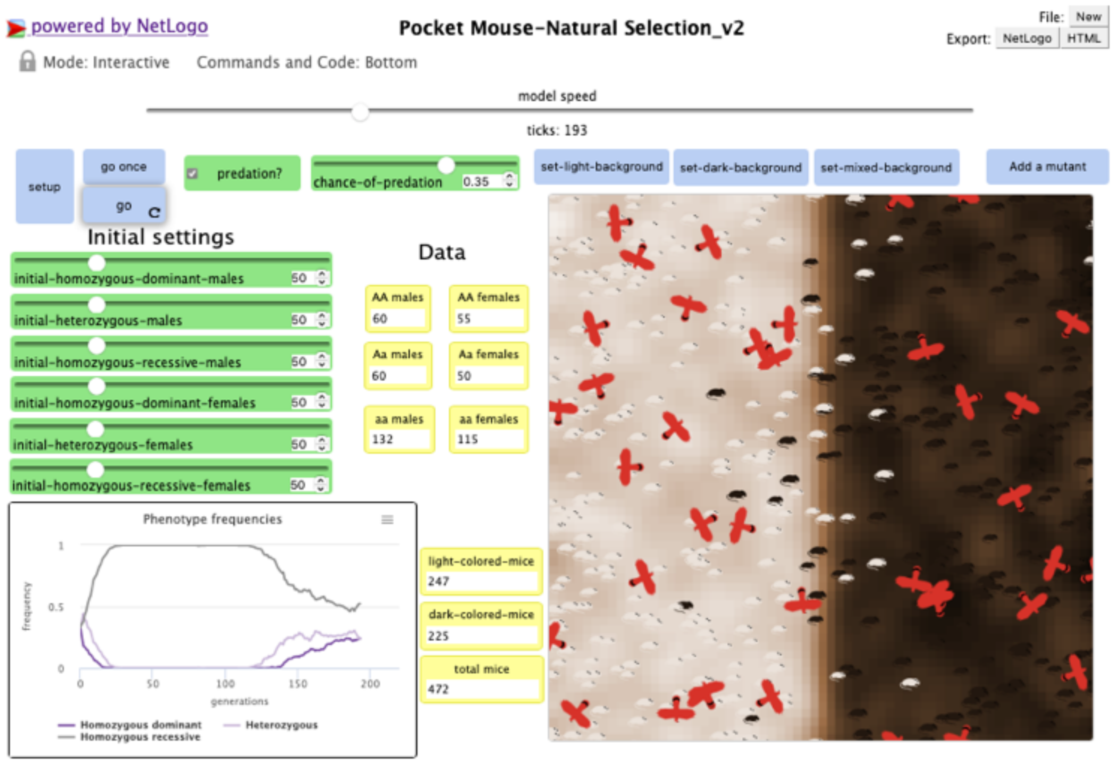


Figure 5-9 The population of rock pocket mice geographically divides into two subpopulations, with light-colored mice in the light background and dark-colored in the dark background

A RESTRUCTURATED CURRICULUM

Purpose of teaching population genetics in undergraduate introductory biology courses is to connect student understanding of Mendelian inheritance to more abstract principles of evolutionary processes, such as natural selection and genetic drift (Brewer & Gardner, 2013; Smith & Baldwin, 2015). One common issue that biology educators encounter with current methods of teaching HW equilibrium is that the emphasis on equation-based representations and calculations undermines student understanding of the biological ideas related to mechanisms of population change, in other words, biology educators are concerned about students getting caught up in the math and miss the biology (Brewer & Gardner, 2013; Williams et al., 2021).

Additionally, biology educators have identified and discussed difficulties that students face while engaging with phenomena related to micro-evolutionary changes because of natural selection or genetic drift (Ferrari & Chi, 1998; Rosengren et al., 2012; Sinatra et al., 2008). My work presented in this chapter builds on earlier work to use agent-based modeling to teach disciplinary ideas related to evolution (Wagh & Wilensky, 2018; Wilensky & Novak, 2010).

The Rock-Pocket-Mice ESM curriculum uses agent-based restructurations (Wilensky & Papert, 2010; Wilensky, 2020), which provide students computational objects-to-think-with (Papert, 1980) to learn about population genetics and evolution by engaging in scientific inquiry practices. The agent-based restructurations allow students to visualize agent behaviors such as movement, reproduction, predation etc. and population level patterns such as distribution of mice in dark and light regions, changes in allele frequencies. The curricular activities in the ESM-based curricula that use these models are designed for students to design and conduct experiments using the ESM to investigate various aspects of the phenomena of change in the rock pocket mice population and learn about Hardy-Weinberg equilibrium.

The versions of this ESM that I discussed before in this chapter are embedded in two curricular units, one for regular high school biology classes (Dabholkar, Woods, et al., 2018) and the other for college-level biology classes and advanced placement biology classes in the United States (Dabholkar, Granito, et al., 2019). The regular unit has the second ESM without Hardy-Weinberg equations, whereas the advanced unit has the first and the third ESMs, which allow students to more deeply learn about population genetics in relation to natural selection. In this section, I discuss a few examples of curricular activities in the regular biology unit since in the next chapter I present analysis of student learning using this unit.

Figure 5-10 A page from a CT-STEM lesson with the rock pocket mice ESM

Figure 5-10 shows one page of a lesson in a CT-STEM curriculum (Dabholkar, Woods, et al., 2018) in which students are expected to explore a model (using the drop-down menu and sliders to change parameters) and to engage in scaffolded and self-driven investigations regarding natural selection.

Scaffolded Engagement in DCIs and SEPs

This ESM-based curriculum about natural selection is designed for students to learn disciplinary ideas regarding natural selection in NGSS (Figure 5-11) and Science and

Engineering Practices (SEPs) (NGSS Lead States, 2013). This unit incorporates the following disciplinary ideas from the NGSS:

MS-LS4-4.	**Construct an explanation based on evidence that describes how genetic variations of traits in a population increase some individuals' probability of surviving and reproducing in a specific environment.** [Clarification Statement: Emphasis is on using simple probability statements and proportional reasoning to construct explanations.]
MS-LS4-6.	**Use mathematical representations to support explanations of how natural selection may lead to increases and decreases of specific traits in populations over time.** [Clarification Statement: Emphasis is on using mathematical models, probability statements, and proportional reasoning to support explanations of trends in changes to populations over time.] [*Assessment Boundary: Assessment does not include Hardy Weinberg calculations.*]

Figure 5-11 Disciplinary core ideas about Natural Selection in NGSS that are incorporated in the ESM-based unit

In this section, I explain how the unit scaffolds student engagement in three learning activities that include NGSS-recommended SEPs: asking questions, learning about and with a computational model, and performing scientific investigations.

Asking questions.

Asking questions is the first science practice in the NGSS list of Science and Engineering Practices. The *Evolution of Populations* curriculum had three questions that directly scaffolded students' learning of this practice.

First, students are shown a video developed by Harvard Hughes Medical Institute (HHMI) BioInteractive to describe the ongoing research at the desert of New Mexico to investigate mice population evolution[15]. Then, they are asked to write down some things that they found interesting in the video:

[15] https://youtu.be/sjeSEngKGrg

1. Write down at least two questions that you would like to investigate about the pocket mice in the desert of New Mexico (in Lesson 1).

After that, students are asked to explore the model and perform simple investigations to get to know how it works. Then, they are asked to refine the question that they would like to investigate using the model:

2. Based on your investigation using this model, modify your questions that you wrote before. Specifically, you want to try to change your questions so that you could actually investigate them using a computational model like the one you explored in this lesson (in Lesson 1).

Finally, in Lesson 3, they are again asked to come up with a research question to investigate using the model in question (C).

3. Come up with a question about natural selection in the case of the pocket mice which can be answered using this model (in Lesson 3).

> One example of a question that could be answered using the model: If we introduce a mutant (with a dark-fur-coat) into a population of mice with light-fur-coats that are living in a mixed background environment, how will the population change after 500 generations? (in Lesson 3).

The progress of students from question (A) to question (C) in this ESM-based curriculum helps students ask questions in the context of natural selection. There are two important aspects of this practice. The first aspect is curiosity, that is identifying parts of a phenomenon or system that arouse your curiosity. The second aspect is feasibility, which involves understanding the feasibility of answering a question with the tools at your disposal. Both of these aspects are crucial for a science practitioner. We considered these aspects while designing questions A, B, and C. Question (A) in Lesson 1 asks students to write questions based on their *curiosity* after watching a video about the anchoring phenomenon. Question (B) in the same lesson asks students to refine their questions after exploring the model based on the *feasibility* of answering them. After having done a more thorough investigation and exploring the model in Lesson 2, question (C) asks students to come up with a question that they are curious about and that is feasible to investigate with the model.

Learning about and with a computational model.

In an ESM-based curriculum, students first learn about a computational model and then they use that model to learn about a natural phenomenon. Learning about a model involves learning about the agents, environment, and parameters. For example, in this particular ESM-based curriculum, students first learn how genotypes and phenotypes are related before they start preforming investigations regarding natural selection.

An ESM is a computational model, so learning with ESMs naturally involve the science practice of using models. However, ESMs can be used to engage in other science practices such as planning and carrying out investigations, analyzing and interpreting data, and more. Throughout the ESM-based curriculum, activities are designed to scaffold student engagement in these practices in the context of gradually progressing disciplinary ideas. We expect students to engage in these practices to establish knowledge by inventing their own methods of scientific validation. ESMs make it easy for students to make predictions regarding the disciplinary phenomenon under investigation, propose mini experiments that can be quickly performed, and perform those experiments to establish reliable patterns. We expect that involvement in such

ESM-based learning activities would help them learn disciplinary ideas as well as the ways (science practices) they devise to establish them.

The following set of questions at the beginning of Lesson 2 are designed to engage students in inquiry-based learning by exploring the model and performing simple experiments. Specifically, we want them to understand how genetic composition (genotype) regarding the fur color is related to the actual fur color (phenotype) of a mouse and how inheritance of these traits works in the model. For example, homozygous recessive mice (aa) have light-colored fur. Instead of telling students this genotype-phenotype relationship in the model, they were asked to change the settings to identify it themselves. After they set up the initial population of a particular phenotype composition, they were asked to run the model for several generations to notice how fur color traits are inherited in subsequent generations.

MQ1	Change the sliders under the "Initial Settings" in the model. Make sure every time you change the sliders that you press SETUP afterwards so that you can actually see the effects of your new settings. Try to change the settings such that all the mice have light-colored fur.
MQ2	Once you get all mice with light fur, describe the initial settings you used. What will happen after lots of generations if the initial population of mice all have light-colored fur?

MQ3	Run an experiment to prove or disprove your answer to the previous question and explain your observations.
MQ4	What will happen after lots of generations if the initial population of mice all have dark-colored fur?
MQ5	Run an experiment to prove or disprove your answer to the previous question and explain your observations.

Performing scientific investigations.

Once students learn about the workings of the model, they can use it to investigate the modeled phenomenon. For example, in this unit, students can investigate how the frequencies of mouse fur colors change and evolve in populations under different environmental conditions.

After a student decides on a question they want to investigate, they are asked to guess an answer to that question based on their exploration of the model and state their guess as a testable hypothesis (Question set 1). Then they are asked to design and conduct an experiment to test that hypothesis (Question set 2). Students are encouraged to collect data and use their observations to argue whether the data support their hypotheses or not (Question set 3). We have divided these questions into three sets for the analytical purpose of tracking their epistemic connections as they progress through independent investigations:

Question set 1:

Come up with a question about natural selection in case of the pocket mice which can be answered using this model. One example of a question that could be answered using the model: If we introduce a mutant (with a dark-fur-coat) in a population of mice with light-fur-coats that are living in a mixed background environment, how will the population change after 500 generations?
Based on your earlier exploration of the model, try to guess the answer to your question and state it in the form of a testable statement (hypothesis) - something that you can test using the model.
Question set 2:
Design an experiment to test your hypothesis. Explain your design.
Question set 3:
Collect data from the experiment in an excel or word file. Describe your observations and explain whether those support your hypothesis or not.
Explain the conclusion of your experiment.

In sum, the ESM-based curriculum first introduces students to the phenomenon of natural selection in a population of rock pocket mice in New Mexico. Students are then asked to identify aspects of the phenomenon that they find interesting and want to investigate further. They are introduced to the ESM as an experimental model system to investigate their questions. Students' self-driven investigations are scaffolded so that they engage systematically in science inquiry practices. Since students use an ESM to design and conduct these investigations, agent-based restructurations help them in identifying how individual properties and behaviors result in population level changes over generations.

In the next chapter, I discuss how agent-based restrictions cognitive affordances for students to observe, manipulate, and interpret properties, behaviors of agents, and system level aggregate patterns. These cognitive affordances supported student engagement in science practices and deeper aspects of disciplinary ideas related to the emergent properties regarding evolution of populations due to natural selection.

Chapter 6: Students' epistemic connections between practices and ideas in an Emergent Systems Microworld (ESM) based curriculum

Summary: Calls for science education reforms emphasize that students should learn practices that scientists use to make sense of natural phenomena in disciplinary contexts. Effective integration of science practices in disciplinary contexts, especially the more contemporary ones, such as computational thinking, requires redesigning learning environments and appropriate pedagogical scaffolds. In this chapter, I present my work regarding the analysis of students' epistemic connections among practices and disciplinary ideas as they learn with an Emergent Systems Microworlds (ESM) based curriculum. An ESM-based curriculum is designed to create a discourse between students and the curriculum, with a series of guiding questions and embedded computational models to conduct computational investigations to answer those questions. In this analysis, I used Epistemic Network Analysis and qualitative analysis of the discourse between the students and the curriculum to investigate students' epistemic connections. The results suggest that the designed properties of an ESM supported students to make epistemic connections. Students learned about practices and disciplinary ideas in an integrated manner by iteratively refining their practices and ideas. This work demonstrates the potential of restructured learning environments and aligned curricular activities for supporting students' epistemic connections between science inquiry practices and disciplinary ideas.

INTRODUCTION

In my dissertation, I argue for the effectiveness of the ESM approach for designing learning environments in curricula. In previous chapters, I have discussed the theoretical foundations of this approach and how it supports student learning of molecular mechanisms of genetic regulation and engagement in epistemically expansive learning. In this chapter, I focus on how an ESM-based curriculum supports student learning of science practices and disciplinary ideas in connection with each other. Design features of an ESM, such as agent-based representations, have cognitive affordances for learners to engage deeply with a complex phenomenon (Wilensky & Papert, 2010; Wilensky & Reisman, 2006). These cognitive affordances enable students to engage in epistemic activities of investigating and constructing explanations regarding a natural phenomenon in a similar manner as scientists do.

To achieve these goals of science education, the Next Generation Science Standards (NGSS) emphasize a three-dimensional way of learning science that supports students' development of proficiency across a set of science and engineering practices, disciplinary core ideas and crosscutting concepts (NGSS Lead States, 2013). The first dimension, Science and Engineering Practices (SEP) describes 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. The second dimension, Crosscutting Concepts (CC), includes concepts that have application across all domains of science. The third dimension, Disciplinary Core Ideas (DCIs) are the key ideas that students should know in order to understand multiple science disciplines.

NGSS promotes an integrated approach for these three dimensions, which means that each dimension should be strongly interconnected across different subjects throughout the school years. The Framework introductory chapter states "*The framework is designed to help realize a vision for education in the sciences and engineering in which students, over multiple years of school, actively engage in scientific and engineering practices and apply crosscutting concepts to deepen their understanding of the core ideas in these fields.*" (National Research Council, 2012).

In a classroom context, such a learning approach requires students to become familiar with the context of a phenomenon they are investigating, to ask relevant questions that they can investigate using a system, to test and verify ideas by designing and performing investigations, and to construct explanations regarding the phenomenon based on their investigations. Learning scientific inquiry practices in such an authentic way in a classroom context is often challenging because (1) it takes a substantial amount of classroom time to engage students in cognitive processes and epistemic activities similar to those of scientists, and (2) there are instructional challenges that teachers have to overcome while achieving the authentic engagement in processes and activities (Chinn & Malhotra, 2002). Chinn & Malhotra (2002) consider authenticity in terms of the similarities between classroom practices and the practices that scientists actually engage in. NGSS's emphasis on practices is aligned with this idea authenticity by engaging students in similar activities to scientists (NGSS Lead States, 2013). Recent work in the field of science education that focuses on three-dimensional learning also underscores the challenges regarding engaging students in practices because of the tensions between teaching established disciplinary ideas and authentically participating in practices to construct knowledge (Russ & Berland, 2019; Schwarz et al., 2017). Another challenging aspect regarding engagement in authentic practices is supporting students' strategy-like ways of engaging in knowledge-building processes that Krist and colleagues refer to as epistemic heuristics (Krist et al., 2019). I argue that one approach to address these challenges is to use a learning environment that is designed for cognitive ease for engaging quickly with the core aspects of disciplinary ideas and practices to investigate scientific phenomena.

I use Collins and Ferguson's (1993) notions of epistemic forms and games to investigate how the NGSS-aligned epistemic form of an ESM-based curriculum in a biology classroom supported students' learning of science practices and disciplinary ideas regarding natural selection. I use Epistemic Network Analysis (Shaffer et al., 2009) to characterize student learning progression in terms of practices and disciplinary ideas. I provide evidence that the ESM-based curriculum systematically created learning opportunities for both - (a) learning authentic science practices in a disciplinary context, and (b) constructing knowledge of disciplinary ideas using those practices.

THEORETICAL FRAMEWORK

EPISTEMIC FORMS, EPISTEMIC GAMES AND EPISTEMIC CONNECTIONS

Computational modeling, simulations and games are increasingly being used in science classrooms for learning (Clark et al., 2009; Danish, 2014; De Jong et al., 2013; Gobert et al., 2011; Levy & Wilensky, 2009; Schwarz & White, 2005). Such computational model-based learning activities take various forms. Some activities focus on students using a pre-designed computational model, some focus on students refining a pre-designed model, whereas some activities engage students in building new computational models. Learners can struggle when they attempt to translate their scientific ideas into modeling contexts (Basu et al., 2016;

VanLehn, 2013). Such struggles can potentially reduce the power of model-based learning activities in engaging students authentically in epistemic pursuits (Hmelo-Silver & Azevedo, 2006; Xiang & Passmore, 2015). Wilkerson and colleagues argue that a major cause for such difficulties is a misalignment between the epistemic games that learners play, and the epistemic forms a given modeling activity is designed to support (Wilkerson et al., 2018). Effectively engaging students in using authentic science and engineering practices (SEPs) for constructing knowledge about disciplinary core ideas (DCIs) requires alignment between what the learning environment is designed for, and the activities students actually engage in when they interact with the learning environment. This is even more important in the context of an exploratory computational modeling environment, in which teachers have even less access to in-the-moment student thinking and participation.

Wilkerson and colleagues' study (2018) draws on Collins and Ferguson's notion of epistemic forms and games (Collins & Ferguson, 1993). Epistemic forms are the target structures that guide a scientific inquiry, and epistemic games are the rules and strategies that researchers use to generate epistemic forms. A metaphor that Collins & Ferguson (1993) use is of the game tic-tac-toe. The nine squares form a target structure to be filled and thus determine the epistemic form. The rules that players follow and strategies they use to place crosses and noughts are part of the epistemic game. Epistemic forms are generative frameworks with slots and constraints regarding filling those slots. A desired result of an epistemic game is to complete a target structure that corresponds to a particular epistemic form. For example, making a list of invasive species in a particular habitat is an epistemic game and the target epistemic form is *the list* itself. Collins and Ferguson's (1993) notions of epistemic forms and games inform how practitioners of various disciplines engage in knowledge construction. They suggest that forms and games should be taught to students of science in addition to facts, concepts, methods and theories. I use these notions of form and games to inform the design of learning environments and curricula.

In the context of designing a learning environment, a learning environment can be designed for students to engage in epistemic games to generate a desired epistemic form. NGSS-recommended learning of practices in disciplinary contexts can take various epistemic forms. I have developed one such form for pedagogical purposes, which I have used to design an ESM-based curriculum (Figure 6-1). This target structure in an ESM-based curriculum often takes the form of what Collins and Fergusson refer to as process analysis (Collins & Ferguson, 1993). Process analyses attempt to characterize behaviors of dynamical systems. ESMs are models of a dynamic complex systems which provide easy access to visualization and investigation of emergent patterns that arise because of agent interactions. Users have visual access to agent-level interactions and system-level emergent patterns. As evidenced in the data that I present later in this chapter, when students use an ESM to engage in this epistemic form (Figure 6-1) their explanations about emergent properties of a system (such as change in a population because of natural selection) are often based on the mechanistic details of agent integrations in the system (such as inheritance, predation, reproduction).

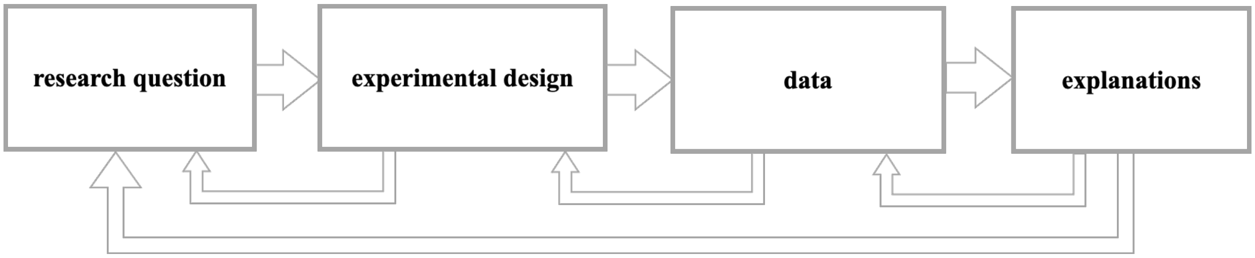


Figure 6-1 A target epistemic form of an ESM-based curriculum that is aligned with NGSS recommended practices.

These four slots are the epistemic form that guide student inquiry in an ESM-based curriculum. My conceptualization of an epistemic game to create this epistemic form is not a linear process, starting with a research question and ending with an explanation. Rather, this is a complex and iterative process. However, there is a general sense of temporality in terms of progression in this epistemic form, which is representative of authentic scientific practice. I highlight this temporality using the large arrows that flow left to right (Figure 6-1). For example, an experiment is designed to answer a particular research question. However, a research question is often modified and becomes more specific as one advances in the epistemic game. The temporal aspect serves a pedagogical purpose as well. Learning activities can be designed for students to develop this epistemic form, as well as allowing for iteratively refining each step, the research questions, and strategies for investigation, data collection and analysis.

In an ESM-based curriculum, these four slots provide a grammar for students to engage in a causal inquiry epistemic form (Collins & Ferguson, 1993) using an ESM. Constraints with respect to each of the slots exist because of a learning environment: the kinds of questions that can be asked and answered in a particular learning environment, the methods that can be used to answer those questions, the ways in which data are collected, and so on. Careful consideration of these constraints while designing a learning environment is vital to scaffold student engagement in the desired epistemic games.

Students participate in epistemic games through learning activities that are designed to help them construct new knowledge. In this chapter, I focus on games that are supported by a computational model designed in the form of an ESM. In an ESM-based curriculum, students can participate in the following *epistemic game* aligned with the *epistemic form* shown in Figure 6-1:

- Understand the context of an anchoring phenomenon,
- Explore the ESM to arrive at a research question of interest related to the anchoring phenomenon that can be experimentally investigated,
- Design an experimental setup - identification of the parameters to be varied (independent variables) and to be kept constant (control variables), experimental duration, parameters/ conditions to be observed (dependent variables), number of trials, etc.,
- Conduct a computational experiment using an ESM to collect data,
- Notice patterns in the data that are related to the research question,
- And construct an explanation about an emergent based on the experimental evidence (analyzed and interpreted data).

In an ESM, agent-based restructurations and constructionist design features provide unique affordances to support student engagement in an epistemic game to investigate and learn about an emergent phenomenon. The agent-based restructurations in an ESM reduce cognitive and perceptual limitations by providing visual access to emergent patterns at the system level as well as interactions of agents (Goldstone & Wilensky, 2008). Because an ESM is designed an interactive microworld (Papert, 1980), students can easily manipulate agent interactions and system compositions and devise various investigation strategies as parts of their epistemic games. For example, in an ESM designed to learn about the role of natural selection in the process of evolution of populations, students can change the system composition by introducing a mutant organism in a population and investigating survival of mutants under specific environmental conditions. This is a critical-event analysis game (Collins & Ferguson, 1993). An advanced version of this epistemic game can take a form of controlling variable game, which would involve investigating spread of a mutation in a population by systematically varying a variable that affects survival, such as predation rate.

Epistemic Connections with ESMs

As students engage in the epistemic activities using an ESM, they are expected to investigate disciplinary ideas using science practices. Since student learning in an ESM-based curriculum involves both practices and ideas, tracking student learning progression in terms of practices and ideas is of analytical importance. In the analysis, I focus on the co-occurrence of ideas and practices as I study students' learning progress through a curriculum. For example, in a curriculum unit about natural selection, students construct knowledge about disciplinary ideas, such as heredity, environment, and survival in the context of natural selection by engaging in science inquiry practices. To characterize student learning with the ESM-based curriculum, I analyze the connections that students make between practices and ideas. I call these connections *epistemic connections*. Epistemic connections are connections between: Practices ↔ Practices; Ideas ↔ Ideas; and Practices ↔ Ideas.

To analyze such connection making, I use 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, Chesler, & Shaffer, 2014; Bagley & Shaffer, 2015). These affordances furthermore allow researchers to assess student learning as they express their ideas (Arastoopour et al., 2016).

The previous work using ENA has demonstrated that students exhibited science and computational learning gains after engaging with a Computational Thinking integrated biology unit with computational models and that the use of embedded assessments and discourse analytics tools revealed how students learned using computational tools throughout the unit (Arastoopour et al., 2020). This line of research also investigated students' understanding of systems thinking practices as they participated in a chemistry unit using ENA and demonstrated how the design of the unit supported understanding of micro-macro relationships regarding emergent concepts such as pressure and temperature of gases (Arastoopour et al., 2019).

In this study, I extend this work further by investigating how an ESM-based curriculum supports student learning of practices and disciplinary ideas by making connections among the two. This curriculum is about natural selection. In this curriculum, students investigate the evolution of a rock pocket mice population in a sandy (light) or a rocky (dark) background by designing and conducting experiments. From the disciplinary learning perspective, this curriculum is designed for students learn about how populations change over time in different environmental conditions (Dabholkar, Woods, et al., 2018). The ESM-based curriculum is also designed to scaffold student engagement in science practices in the context of an ESM to learn construct knowledge of those disciplinary ideas. My investigations into student learning with an ESM-based curriculum are guided by the following research question:

How does an *ESM*-based curriculum support student making *epistemic connections* among Science and Engineering Practices (SEPs) and Disciplinary Core Ideas (DCIs)?

RESEARCH CONTEXT AND METHODS

PARTICIPANTS AND SETTING

Evolution of Populations is a ten-day biology unit designed by the lead author in consultation with high school biology teachers. The unit focuses on predator-prey dynamics, competition among individuals, and natural selection (See Appendix 4). The unit was taught by a biology teacher, Ms. Lydia (pseudonym), in a large Midwestern city's public school. 100% 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 balanced gender distribution (Figure 6-2).

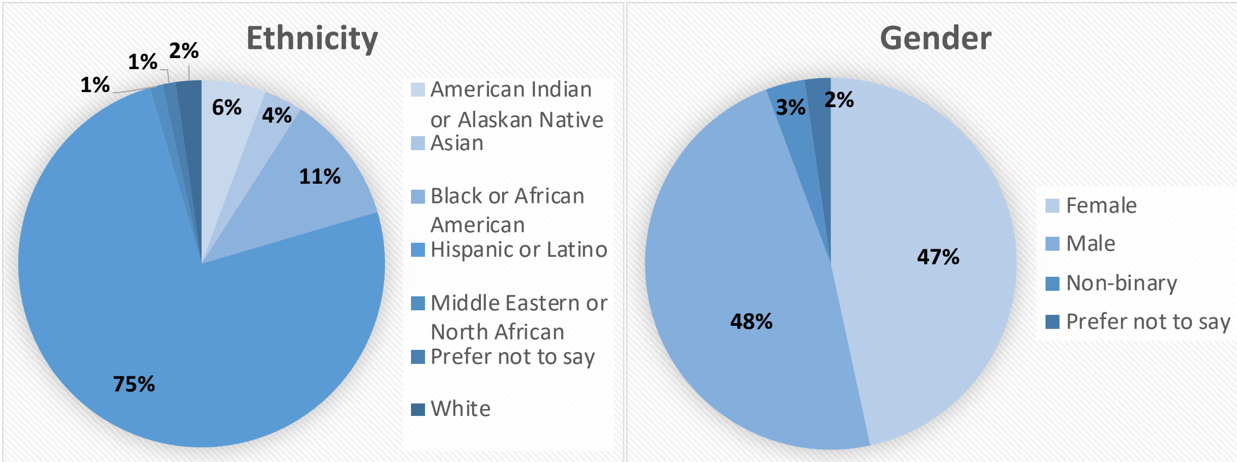


Figure 6-2 Demographic distribution of participating students (total students = 88)

ESM-BASED CURRICULUM

The ESM-based curriculum was designed for high school biology students to learn about natural selection. Activities were delivered through an online curriculum portal (https://ct-stem.northwestern.edu/curriculum/preview/681/) and were split into lessons. Each lesson consisted of 3-4 pages. Typically, on each page, students read a prompt with a description of a computational model and suggestions for exploration. Then, students answered 2-5 embedded assessment questions on the same page. The ESMs in the unit are built using NetLogo (Wilensky, 1999b), an agent-based modeling platform which is intentionally designed to foreground emergent systems modeling for educational and research purposes. In this chapter, I analyze the last three lessons in this curriculum which I call Lesson 1, Lesson 2, and Lesson 3. These lessons use the rock pocket mice ESM discussed in detail in the previous chapter (See Chapter 5).

IDENTIFICATION OF DCIS, SEPS AND EPISTEMIC CONNECTIONS

To study student engagement in Science and Engineering Practices (SEPs) and Disciplinary Core Ideas (DCIs) that the Next Generation Science Standards (NGSS) recommends (NGSS Lead States, 2013), I examined students' responses to embedded questions in three different lessons of the unit. I coded for students' explicit engagement in SEP such as using a model or analyzing data, as well as explicit mentions of DCIs such as adaptation or inheritance. I used both a top-down and a bottom-up approach to develop codes (Hashimov, 2015).

I have used the term Science and Engineering Practices (SEPs) as it is coined and operationalized in the context of science learning by NGSS. Though I primarily focused on science practices, I use the term SEP and not SP because SEP is a more common term amongst the researchers and practitioners who use the NGSS framework. Computational and Mathematical Thinking, which is listed as one of the eight SEPs by NGSS, is especially emphasized in my work. Advancement in computational tools and the availability of computational power has significantly changed the way scientists approach problems (Weintrop et al., 2016). Across a wide variety of domains, the application of statistical and mathematical approaches that rely on computation, such as Machine Learning and Artificial Neural Networks, have proved essential for opening new avenues of exploration and yielded advances in numerous fields as diverse as the origins of the universe in computational astronomy (Vogelsberger et al., 2014) to conservation biology (Kwok, 2019). Because of the increased use of computational tools and methods, Computational Thinking (CT) has impacted other science practices as well. 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). 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. In the work, I define CT in the context of a specific disciplinary domain such as science or mathematics (Weintrop et al., 2016).

To identify student responses that show student engagement in science practices in the context of a computational model, I used top-down codes derived from NGSS science practices (NGSS Lead States, 2013) and from Weintrop et al.'s (2016) taxonomy of CT practices. I coded for five practices that overlap with these two sets of practices and are relevant to learning with an ESM-based curriculum: Asking Questions, Developing and Using Models, Planning and Carrying Out Investigations, Analyzing and Interpreting Data, and Constructing Explanations.

I used a bottom-up coding approach to devise codes for characterizing students' knowledge of DCIs. Based on iterative analysis of students' responses I developed the following codes: Populations/Individuals/agents, Phenotypic Properties or Characteristics (Phenotype), Genotypic Properties or Characteristics (Genotype), Environments, Heritability, Survival, Adaptation mechanism, and Change/Mutation/Variation (See Appendix 6A and 6B). Because the data contained a large number (n = 2,026) of responses, I developed an automated coding algorithm using keywords and regular expressions (see Arastoopour, et al., 2019 and 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).

Table 6-1: 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*

To analyze student connection-making among science practices and disciplinary ideas, I used Epistemic Network Analysis (ENA) (Shaffer, 2017; Shaffer et al., 2016). In prior work, ENA has been used to assess and visualize learners' connections among CT-STEM practices and specific disciplinary knowledge (Arastoopour et al., 2019, 2020). In this study, I extend this approach further by analyzing epistemic connections in conjunction with design features of a learning environment. In my analysis, I have operationalized epistemic connections in terms of co-occurrences among the codes in each student response. 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 knowledge. In a weighted network, thicker lines indicate that a student made that connection often and thinner lines indicate that a student made that connection less often. In addition to a network representation, I 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 a broad, macro-level analysis, I used the centroid representations which displayed all student networks as points in one space together. In a detailed, micro-level analysis, I examined individual student's network representations for three students, Alejandro, a Hispanic male, Emma, a Hispanic female, and Jane, an Asian female (all pseudonyms) to further investigate student learning in terms of the particular connections that they made among science practices and disciplinary core ideas.

RESULTS

In this section, I first present a broad, macro-level epistemic network analysis of student progression using the centroid representation. In this representation, the centroid of each student's network is projected into a 2-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. I then present a micro-level discourse analysis of student responses in which I focus on three students. This analysis illustrates students' learning progression regarding practices and disciplinary ideas. The first case illustrates how the disciplinary context provided rich ways to refine a science practice.

The second case illustrates the reverse--how a combination of a few science practices helped a student refine her disciplinary ideas. Finally, the third case illustrates the reciprocal relationship among practices and ideas and how the disciplinary context and scaffolded engagement in science practices contributed to a student's scientific learning experience.

Aggregate Analysis of Student progression through ESM lessons

The ESM-based unit was designed for the students to progress through science practices sequentially from (1) introduction to the phenomenon and *exploring the model*, to (2) *using the model* for a scientific investigation, to (3) *proposing a question* and *investigating* a hypothesis using the model. The same is true about progression of disciplinary ideas– from genotypes and phenotypes of the mice to understanding survival, heredity, and change in the population across generations in different environments. I expected that the progression of practices and disciplinary ideas that the unit was designed for would be reflected in students' epistemic connections as they progressed from Lesson 1 to Lesson 3.

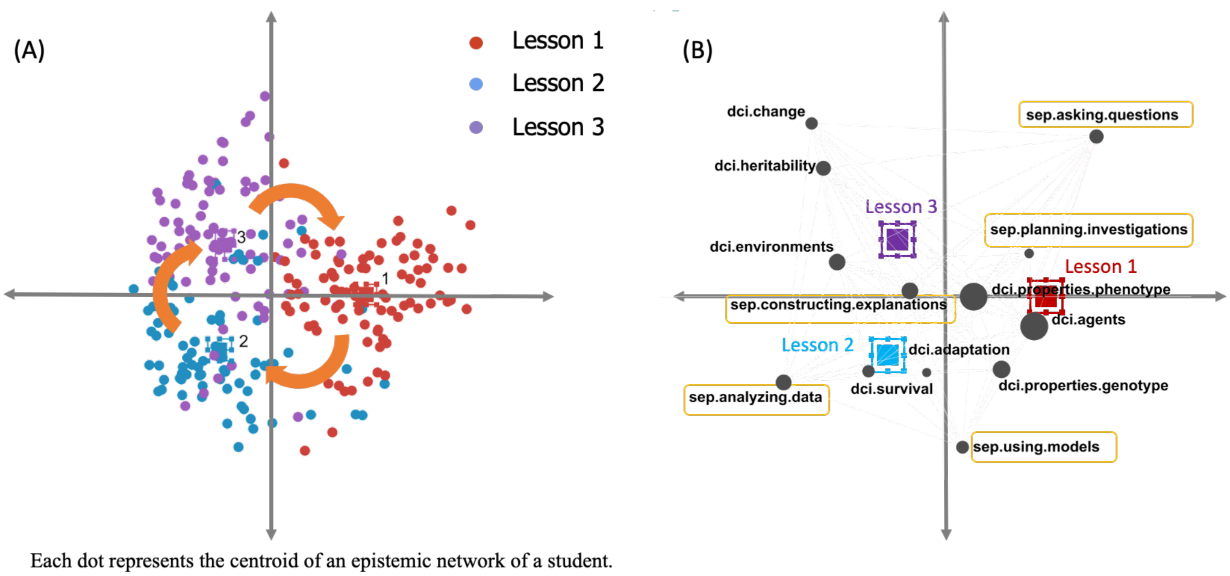


Figure 6-3 (A) Centroids of networks for all students for lesson 1 (red), 2 (blue), and 3 (purple). The average is represented as a square with confidence intervals. Axes represent the first two dimensions of the multi-dimensional scaling in ENA to maximize variance in the data. (B) Mapping of DCI and 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.

For each lesson, students had statistically distinct epistemic connections between disciplinary core ideas (DCIs) and science practices (SEPs), as represented by the means (Figure 6-3 A, Table 6-2). This means that the students' expressed engagement in SEPs and DCIs in their responses to the embedded questions was significantly different for each lesson (Figure 6-3 B). For example, the centroids of student networks for Lesson 1 (red dots) are located in the same cartesian space where specific DCIs and SEPs are situated. These DCIs are related to the properties of the agent mice (dci.agents, dci.properties.phenotype, dci.properties.genotype) and SEPs are about asking questions (sep.asking.questions) and planning investigations (sep.planning.investigations). This means that these nodes had higher weights in the student epistemic networks in Lesson 1, which is aligned with the learning goals of coming up with research questions that can be answered using the ESM. To investigate the modeled emergent phenomenon about natural selection, these questions are about changes in agent population in terms of their phenotypes and genotypes.

Table 6-2: 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)

In Lesson 2, students moved on from posing questions and progressed towards investigating more fundamental ideas required to understand natural selection. Students used the ESM to construct explanations regarding the change in mice populations across several generations under different environmental conditions. They used the model (sep.using.models) to design investigations (sep.planning.investigations), and they collected and analyzed data (sep.analyzing.data). Students connected related disciplinary ideas - dci.survival, dci.environments and dci.heritability - and they engaged in additional science practices - sep.using.models and sep.analyzing data. In this lesson, students used the model to investigate natural selection by testing their hypotheses through data analysis.

In Lesson 3, students posed their own questions and investigated various aspects of natural selection that they were interested in. The Lesson 3 centroids (purple dots) are in the cartesian space where the DCI nodes for dci.change, dci.heritability, and dci.environment are situated. This indicates that these nodes had higher weights in the student epistemic networks in Lesson 3. The investigations included changing the initial mice populations, varying predation, and changing environment in terms of the background color (dci.heritability, dci.change, and dci.environments) and explaining the phenomena (sep.constructing.explanations). In this lesson, students made connections across the most practices and disciplinary ideas.

This aggregate-level analysis shows that as students progressed through the lessons, they engaged in different science practices and learned aspects of disciplinary ideas in a sequential manner. However, the practice of science is not strictly linear or sequential. A more in-depth analysis of average student networks reveals that certain DCI and SEP nodes became more prominent in student networks over lessons (Figure 6-4). DCIs dci.change, dci.heritability, and dci.environments became progressively prominent in Lesson 2 and Lesson 3; similarly sep.analyzing.data and sep.constructing.explanations have stronger connections in later lessons. Other DCIs and SEPs are also present, if less prominent.

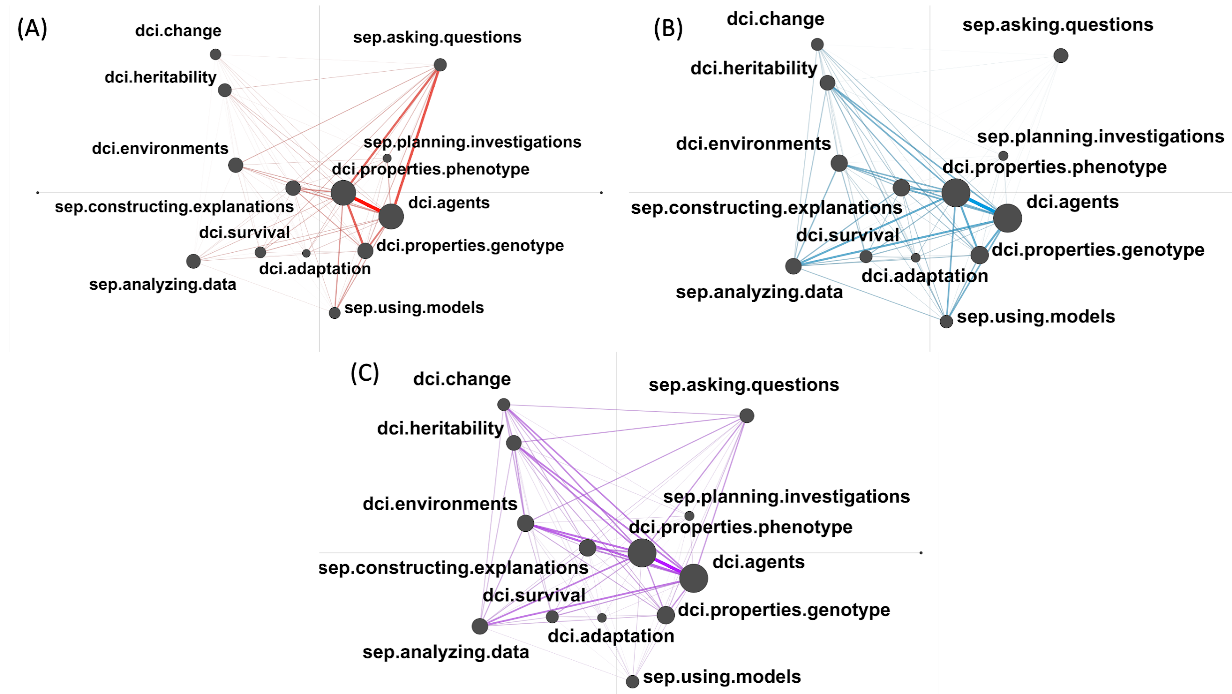


Figure 6-4 Average networks for all students for (A) Lesson 1 (red), (B) Lesson 2 (blue), and (C) Lesson 3 (purple).

This demonstrates that student learning was sequential in terms of the practices and ideas. Additionally, as the new ideas and practices became more prominent in student networks, students also made connections with the ideas and practices that were prominent earlier. The qualitative analysis of individual students in the following section further illustrates how ESM supported students' learning of SEPs and DCIs by making epistemic connections.

QUALITATIVE ANALYSIS OF STUDENT PROGRESSIONS

As students progressed through the natural selection lessons, they engaged in different science practices. In this section, I present the learning progressions of Alejandro, Jane, and Emma. I first discuss Alejandro's learning of the SEP of *asking questions*, focusing on how he refined his research question in the context of an experimental system. Then, I discuss Jane's learning progression, highlighting the way in which she learned several important aspects of DCIs by engaging in the SEP of *using a model*. Last, I discuss how the DCIs modeled in the ESM provided a rich context for Emma to engage in SEPs, and also how she constructed knowledge about DCIs through her engagement in SEPs. This last case illustrates the reciprocal nature between SEPs and DCIs for student learning.

Refining an SEP in a Disciplinary Context

As described in the methods section, the curriculum asked students three times to propose and refine their research questions. 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, creative questions reveal Alejandro's curiosity about the contents of the video. His question 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 blackmice 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 coat color. Between question (A) and question (B), Alejandro performed a series of scaffolded learning activities about investigating the ESM model. He played around with the model parameters to deduce how mice genotype and phenotype are modeled in the ESM. Alejandro's new question was about several specific aspects of DCIs such as environment (color of the background) and mouse phenotypes (color of mice fur). Compared to the first question, Alejandro's refined question indicates learning the practice of *asking questions* because this was more specific, more feasible, and more aligned to the context of ESM.

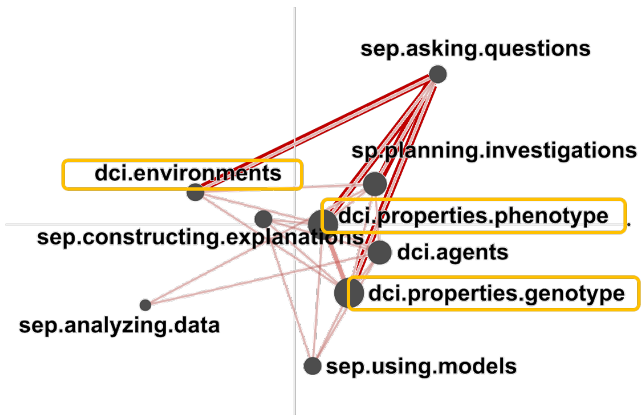


Figure 6-5 Alejandro's Network (Lesson 1).

Alejandro's network in Lesson 1 (Figure 6-5) shows the prominent epistemic connections in his responses to all the questions in this lesson. The connections between sep.asking.questions and other nodes are highlighted. The four connections that Alejandro makes with sep.asking.questions were dci.environments, dci.properties.phenotype, sep.planning.investigation, and dci.properties.genotype. 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"), 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.

In Lesson 2, students were asked to conduct computational mini experiments to investigate the ways that the mechanism of inheritance and the rate of predation affected changes in the modeled population. After having performed those activities 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 was considering 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 important aspects of DCIs: *change* in the population, *inheritance* of fur coat color, *genotype* (homozygous dominant) and *phenotype* (dark fur coat color) of mice.

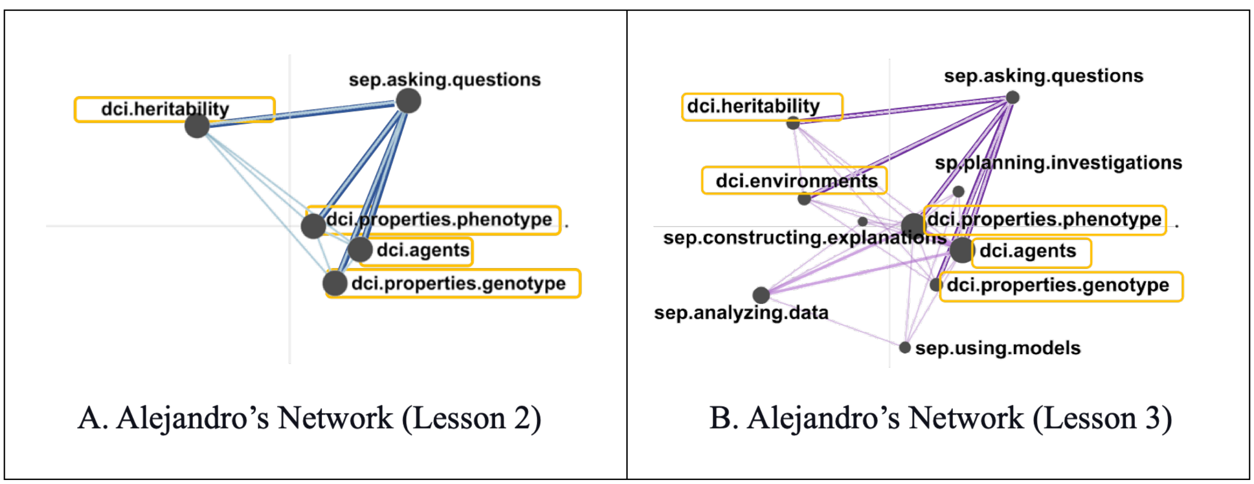
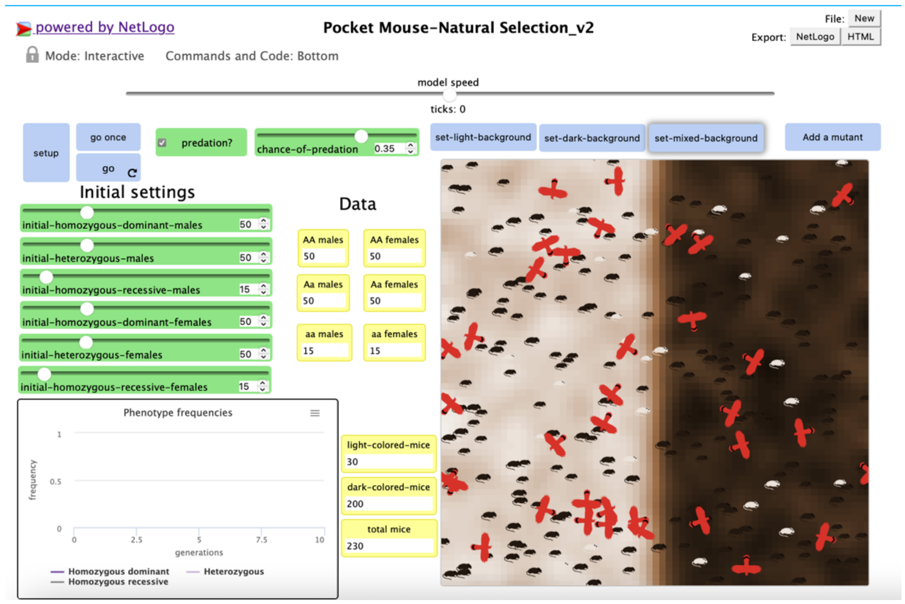


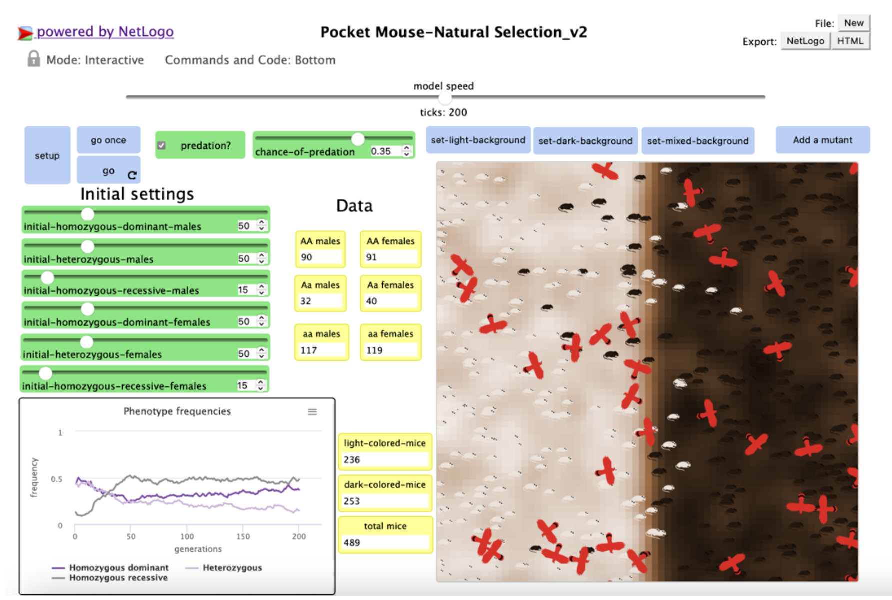
Figure 6-6 Alejandro's epistemic networks in Lesson 2 and Lesson 3.

In Lesson 2, Alejandro made epistemic connections between sep.asking.questions and dci.heritability, dci.agents, dci.properties.phenotype, and dci.properties.genotype (highlighted connections in Figure 6-6 A). This indicates that in Lesson 2, Alejandro connected an important aspect of DCI, namely, *heritability of a trait* in the context of the practice of asking questions. 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 change in the mouse population with respect to 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 understanding of both dci.heritability and dci.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 (Figure 6-6 B), 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).



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



B. ESM state after running the model for 200 generations

Figure 6-7 Initial and final states of the ESM corresponding to Alejandro's research question

In sum, Alejandro's initial framing of questions was driven by his curiosity and then refined as he progressed through the unit. There is no doubt that curiosity is an important driver of asking questions. However, the scientific practice of asking questions also entails considering how to feasibly investigate a phenomenon with an available experimental system. The experimental system of the rock-pocket-mice ESM contained agent-based restructurations. Agent-based restructurations create cognitive affordances for a user to observe, manipulate, and interpret properties, behaviors of agents, and system level aggregate patterns (Wilensky & Papert, 2010). Alejandro's iterative refinement of the questions was mediated by his cognitive engagement with the agents (the mice population), their properties and behaviors (dci.properties.genotype, dci.properties.phenotypes, dci.heritability), and their surroundings (dci.environment). His questions became more specific and included more aspects of DCIs related to natural selection over time. His final question considered two very critical aspects of natural selection, heritability and environment, that he could investigate using the ESM. The states of ESMs corresponding to Alejandro's research question (Figure 6-7) 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. In the next sections, we see how Jane and Emma used the ESM to engage in SEPs for learning about DCIs related to natural selection.

Refining DCIs by engaging in SEPs.

This section illustrates how engagement in science practices in the context of an ESM helped students refine their DCIs. I discuss Jane's learning progression as she learned how to use the computational model to perform micro-experiments. She conducted these micro-experiments to learn about relationships between genotype and phenotype in the context of the model. Jane's responses to five modeling questions (MQs) (see the methods section) illustrate how she corrected her ideas regarding inheritance of mice fur coat color.

MQ2 was about understanding inheritance in a population of homozygous recessive (light fur-colored) mice. Jane's prediction for MQ2 was that she would have ***mostly*** light color mice. However, after performing the learning activity using the ESM, she refined her idea when 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.

Jane predicted in her response to MQ4 ("What will happen after lots of generations if the initial population of mice all have dark-colored fur?") that after a lot of generations, the population will be of all dark-furred mice. This was a somewhat complex question because the initial population of dark-furred mice can be set up in multiple ways: All homozygous, all heterozygous, or a mix of the two. Answering this question correctly required nuanced understanding of genetic inheritance and how it was modeled in the ESM. The exploration and investigation with an ESM helped Jane make the required epistemic connections between using a model, heredity, and agent properties of phenotype and genotype. Jane seems to have understood that the inheritance of light fur-color and dark fur-color work in different ways, which is an important aspect of learning how genotype and phenotype are related in the model. Because light fur-color is a recessive trait, light fur-colored parents will only produce light fur-colored offspring. However, if parents with dark fur-color phenotype have heterozygous genotype, they are likely to produce some (on average 25%) light fur-colored offspring. Jane's answer to MQ5 ("Run an experiment to prove or disprove your answer to the previous question and explain your observations.") indicates that she indeed learned the relationships between genotype and phenotype by observing generations 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.

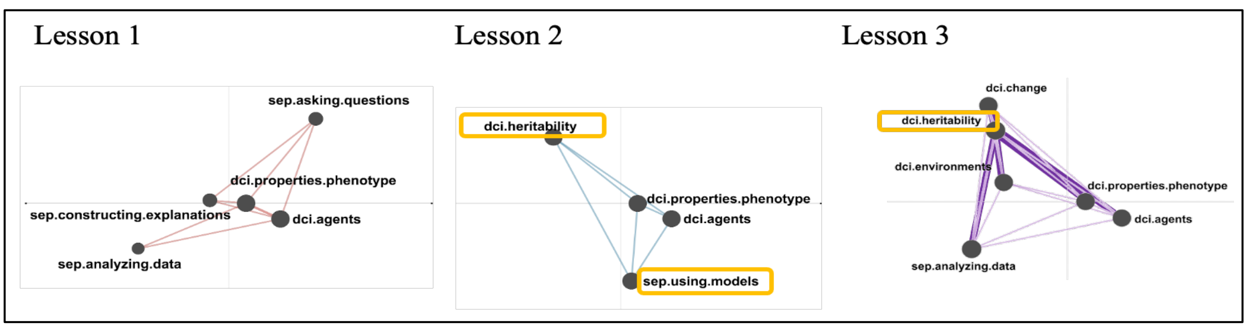


Figure 6-8 Jane's Epistemic networks in three lessons.

Network representation of Jane's epistemic connections (Figure 6-8) provides insights into how Jane's learning of inheritance mechanisms (dci.heritability) in Lesson 2 helped her overall learning of Natural Selection and her engagement in the science practices. In Lesson 2, Jane made a connection between sep.using.models and di.heritability. Both of 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 inheritance of the fur color over many generations.

Analysis of Jane's learning progression illustrates how the agent-based restructurations in the ESM provided an experimental system for Jane to test her predictions by performing mini-experiments. This also allowed her to correct her understanding about how inheritance worked in the context of the ESM. Jane's learning progression demonstrates how her engagement in ESM-based learning activities taught her nuances of the mechanism of inheritance, which she further used in her final experimental investigation in Lesson 3.

Reciprocity between DCIs and SEPs.

In this section, I discuss how the ESM-based curriculum supported Emma's engagement in science practices and learning about disciplinary ideas regarding Natural Selection. In Lesson 3, students were asked to propose a research question and state its possible answer in the form of a testable hypothesis (see methods section for more details). Then they designed and performed a scientific investigation. Finally, they wrote the conclusions of their investigations. Each student group was asked to come up with a different research question. Each student wrote their research question, hypothesis, experimental design, data analysis, and conclusion. Emma's responses in Lesson 3 were reflective of the scientific investigation that her group performed. To discuss how the design of this curriculum scaffolded Emma's learning progression, I divide the questions in Lesson 3 into three question sets (QS).

Question Set 1 (QS1) which included two questions, asked students to state (a) their research question and (b) a hypothesis as an answer to the question that they can test using the model. Emma stated her group's research question and hypothesis as follows:

Emma's Research Question -

If we introduce 10mutants in a population of homozygous recessive mice that are living in a light background environment, how will the population change after 400 generations?

Emma's Hypothesis -

If we add the 10 mutants with dark colored fur to the population of homozygous recessive mice that are living in a light background environment then the population of the dark colored mice will be less and less becuase they will be the ones standing outin the light enviroment not being able to camoflauge.

Emma's group wanted to investigate the change in the population of homozygous recessive (light colored) mice after 400 generations when 10 mutants that were heterozygous (dark colored) were introduced. Emma specified the environmental conditions (light background) in her hypothesis and stated that the population of dark colored mice will be "less and less" because they will not be able to camouflage. Emma's group chose to focus on the change in population when mutants have a disadvantage for surviving in the given environmental conditions.

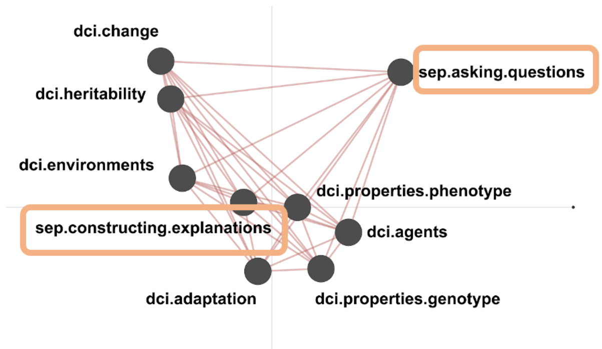


Figure 6-9 Emma's Epistemic Network for Question Set 1 in Lesson 3

In QS1, she asked a question (sep.asking.questions) to construct an explanation (sep.constructing.explanations) about change in a mouse population after several generations when mutants are introduced in the population (Figure 6-9). Emma's carefully worded question and hypothesis include several fundamental aspects related to natural selection DCIs. In her network, those two SEP nodes are connected with all the following DCI nodes:

dci.change - change in the mice population,

dci.heritability - inheritance of the trait (fur coat color) in the population over 400 generations,

dci.environment - light background environment,

dci.phenotype - dark colored fur of the mutants,

dci.genotype - homozygous recessive genotype of the native mice population,

dci.adaption - being able to camouflage as an adaptive advantage to not stand out,

dci.agent - change in the population of the agents (mice).

Emma's question was not a typical natural selection question that an evolutionary biologist would ask or that an evolution curriculum would include. In a typical natural selection experiment, a new mutant phenotype would have a selective advantage and after several generations the mutant phenotype would become more prominent in the population. Even though Emma's question was not a typical question, she could use the ESM to investigate her question, and engage in investing it which helped her make connections with several important aspects of DCIs (see Figures 6-9 and 6-10).

QS2 in the lesson asked students to design an experiment and test their hypothesis. To do so, Emma's group designed the following experiment:

we kept the predation on at 0.35 *we set the background environment color to light* *we set all the other sliders to zero except for homozygous recessive males and homozygous recessive females.* *we set thehomozygous recessive males and homozygous recessive females sliders to 100* *we then would add 10 mutants* *we would keep the number of mutants the same in each experiment.*

Even though Emma did not state some details in her previous responses, such as the presence or absence of predators, she explicitly mentioned all the experimental conditions in her experimental design. She chose to set the chance-of-predation value to 0.35, which meant that predator density was such that if a mouse does not camouflage, there is a 35% chance that it will be eaten. Emma decided to set the background color to light and the initial population to 200 total/100 each gender homozygous recessive male and female mice. Though Emma did not specify how many trials she would conduct, she mentioned that in *each experiment* she would keep the number of mutants the same. This indicates that the investigation involved multiple trials.

The experimental design by Emma and her group was strongly aligned with the research question that they wanted to investigate survival of mutants with a disadvantageous phenotype. One advantage of ESM design is that it allows students to perform several different kinds of investigations. Since the underlying rules in the microworld regarding inheritance, camouflage, and predation are consistent with the established scientific understanding, all investigations are likely to lead to conclusions that are consistent with established disciplinary ideas regarding natural selection.

Emma also mentioned that they would *keep the number of mutants the same in each experiment*. Emma's answer indicates their intention of performing multiple experiments. Interactions between agents in an ESM follow constrained randomness. For example, the chance-of-predation parameter determines the probability of a mouse getting predated if it does not camouflage well in the environment. This constrained randomness in an ESM makes it a realistic system which requires multiple carefully designed experimental trials to establish a pattern. Performing multiple trials is a very important aspect of the practice of planning and performing investigations.

QS3 asked students to collect experimental data, describe their observations, and conclude if their hypothesis is supported by the data they collected.

Our observation does support our hypothesis. This is because we observed that the black furred mice died off really quickly after just a few generations. This meant that the black furred mice did not reproduce mixed babies with the light furred mice. The light furred mice increased in population in a light environment and the black furred mice were the ones that died and did not come back.

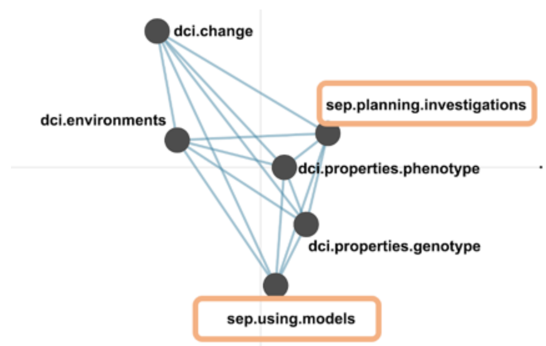
Emma explained that the observations by her group supported their hypothesis and explained her reasoning. Her answer demonstrates her consideration of multiple aspects of DCIs that are important to understanding natural selection, such as survival, reproduction, and inheritance. Emma's mention of "*black furred mice died off really quickly after just a few generations*" indicates that she is tracking changes in the population after several generations to see the effects of natural selection. Emma had set predation rate of 35%, which is high in this model for dark mice to survive and reproduce in the light background since they do not camouflage. Emma was able to learn more about reproduction and heritability of traits based on her experimental observations. The last part of her answer indicates that she and her group investigated and constructed an explanation regarding how environmental conditions affect survival and extinction. Emma's constructed explanation was based on the data in the following table that she and her group recorded and analyzed.

This is how Emma tabulated her data:

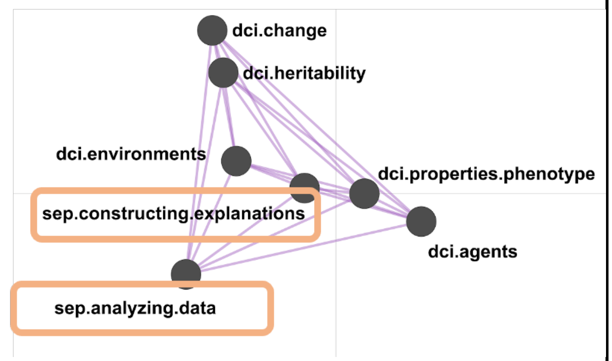
Trials	**Number of white mice at start**	**Number of mutants at start**	**Color of Background**	**Predation**	**Number of heterozygous mice at the end**	**Number of white mice at the end**	

1	200	10	Light	0.35	0	490	At 30 ticks the mutants died off the number of white mice kepy [kept] on increasing
2	200	10	Light	0.35	0	519	After 7 ticks the mutants died off. The number of white mice stayed in the 500 range.
3	200	10	Light	0.35	0	492	At 17 ticks the mutants died off. The number of white mice decreased by 10 into the 490 range
4	200	10	Light	0.35	0	521	At 30 ticks the mutants died off [,] but the white mice still increased
5	200	10	Light	0.35	0	524	At 23 ticks the mutants died off. The number of white mice stays consistant [consistent] in the range of 500
6	200	10	Light	0.35	0	507	The mice died off at 14 ticks. The white mice were steady
7	200	10	Light	0.35	0	487	The mice died off at 12 ticks. The number of black mice decreased vey [very] fast
8	200	10	Light	0.35	0	515	the black mice died off at 23 ticks. The number of white mice decreased to around 480
9	200	10	Light	0.35	0	508	The mice died off at 17 ticks the number of white mice kept increasing untill [until] it stayed in the 500 range
10	200	10	Light	0.35	0	507	After 12 ticks the mutants died off. The white mice were between 480 and 500

Emma and her group members designed the format of this table, designed the experimental setup, and decided what to measure and what additional observations to record. Their measurements included the numbers of heterozygous and white mice at the end of the experiment. In their notes for each trial, they also recorded how many ticks (generations) it took for the mutant mice to die out.



A. Emma's Network (QS2)



A. Emma's Network (QS3)

Figure 6-10 Emma's epistemic connections in QS2 and 3 of Lesson 3.

In QS2, Emma planned an investigation (sep.planning.investigations) using the computational model (sep.using.models) (Figure 6-10 A). Finally, in the last question set, QS3, she analyzed data (sep.analyzing.data) to construct explanations (sep.constructing.explanations) (Figure 6-10 B). The progression of Emma's epistemic networks in Lesson 3 demonstrates how the ESM-based curriculum supported Emma's progressive engagement in the practices as she constructed knowledge about the DCIs.

Based on the patterns they observed in the data, students were asked to draw a conclusion.

Emma stated her conclusion as follows:

In conclusion or hypothesis was correct. We asked about how adding 10 mutants to the population of mice would chnage it . We thought that the black furred mice would decrease in population because they would be the ones tanding out in the light environment. We were right *becuase in all of our trials the black furred mice ended up dieing off after a couple of generations. The mutants did not affect the population of light furred mice becuase probably predators would eat them before they got a chnace to reproduce and produce mixed offspring.*

In her conclusion, Emma elaborated on her thoughts regarding the mechanism of natural selection. She mentioned why the mutant phenotype did not spread in the population: The mutant mice did not get a chance to reproduce and produce mixed offspring because they were eaten by predators too soon. Emma and her group first asked a question regarding the change (dci.change) in the mouse population that they could investigate using the ESM. They wanted to study the survival of a mutant phenotype when mutant mice are introduced in specific environmental conditions in which they have selective disadvantage. Then, they designed a systematic experimental investigation that considered agent properties - genotypes, phenotypes, and the environment - to track specific aspects of population change. They collected data about the genotypes and phenotypes and arrived at a conclusion based on the patterns they observed in the data. Their conclusion was about sustenance of a heritable trait, rather the lack of it, as the population changed over several generations. Their constructed explanation for the extinction of the mutant trait was very specific in terms of predation and reproduction in a given environment.

Analysis of Emma's responses revealed her progression through the unit, from asking a question to constructing an explanation to planning investigations using a model and finally to analyzing data and constructing explanations. Emma and her group were able to engage with several DCI aspects in a scaffolded, yet self-driven manner.

DISCUSSION

In this paper, I investigated how an ESM-based curriculum supported students' epistemic connection-making among science practices and disciplinary ideas as they constructed knowledge about an emergent phenomenon. The results regarding student learning at an aggregate level suggest that the ESM-based curriculum supported students' engagement in making these connections in a sequential and integrated manner. The macro-level analysis illustrates a temporal progression of practices as students moved from asking questions to constructing explanations about emergent patterns, such as changes in a population because of introduction of a mutant phenotype. In addition, our analysis suggests that there was a progression of disciplinary ideas from genotypic and phenotypic properties of individuals to heredity and change in a population. As new ideas and practices became prominent nodes in student networks, they also contained and were connected to earlier nodes, which supports the claim regarding the integrated manner of student learning. This sequential and integrated approach to designing learning environments is aligned with the NGSS three-dimensional framework.

NGSS also recommends designing learning environments that provide meaningful and authentic learning opportunities for students. However, such engagement in authentic scientific inquiry in classroom contexts is instructionally challenging (Chinn & Malhotra, 2002), mainly because designing for such authentic learning experiences requires creating an experimental system that students can use to construct and validate explanations regarding a disciplinary phenomenon by engaging in science practices. This also requires making experimental systems and anchoring phenomena cognitively accessible to students. To design for cognitive accessibility of the experimental system of rock-pocket-mice, I incorporated agent-based restructurations in the ESM design. Agent-based restructurations create cognitive affordances for a user to observe, manipulate, and interpret properties and behaviors of agents and system-level aggregate patterns (Wilensky & Papert, 2010). These cognitive affordances supported student engagement in science practices and deeper aspects of disciplinary ideas related to the emergent properties of the system under investigation.

The micro-level analysis revealed that Alejandro's iterative refinement of the questions was mediated by his cognitive engagement with the agents (mice population), their properties and behaviors (dci.properties.genotype, dci.properties.phenotypes, dci.heritability), and their surroundings (dci.environment). His questions became more specific and included more aspects of DCIs related to natural selection over the lessons. His final question considered two very critical aspects of natural selection, dci.heritability and dci.environment, that he could investigate using the ESM. Similarly, the interactive constructionist microworld features of the ESM provided an experimental system for Jane to test her predictions regarding effects of agent properties (phenotypes) and interactions (reproduction, mendelian inheritance) on emergent patterns by performing mini experiments. Agent-based restructurations in the ESM allowed Jane to visualize and interpret her experimental results. This led to correction of her understanding about mendelian principles of inheritance. In the final lesson, Emma and her group enagaged systematically from asking a question (sep.asking.questions) to planning investigations (sep.planning.investigations) to using a model (sep.using.models) and finally to analyzing data and constructing explanations (sep.construcing.explantions). All the students engaged with SEPs and several DCI aspects in a scaffolded, yet self-driven manner.

These student case studies illustrate how agent-based restructurations provided cognitive ease for students in making *epistemic connections* between practices and ideas. In order to make epistemic connections, there should be alignment between an epistemic form that a learning environment is designed for and the epistemic games that students can play using the learning environment (Wilkerson et al., 2018). Alejandro and Jane's cases illustrate how the alignment between epistemic forms and games created learning opportunities for the students to refine and connect their practices and disciplinary ideas. For example, Alejandro refined his question twice, first after he explored the ESM as an experimental system and then after he practiced data collection and analysis with the ESM. As he refined his questions, he operationalized his initial curiosity to ask more specific questions addressing the disciplinary ideas. This created an alignment of his epistemic game with the epistemic form of the curriculum, because he could meaningfully investigate those aspects by engaging in science inquiry practices. Jane refined her disciplinary ideas regarding heredity after attempting to construct explanations based on the data that she collected using the ESM. Emma and her group's epistemic game was about manipulating a system by adding agents with specific properties and investigating the effects of those manipulations. Their experimental investigation focused on change in a mice population in a particular environmental setting. Their integration of ideas and practices (Figures 6-9 and 6-10) in Lesson 3 followed a similar progression to the aggregate student progression across all lessons. This indicates that students' iterative engagement in practices and ideas shaped their epistemic game in Lesson 3.

Our results demonstrate how the ESM-based curriculum aligned student epistemic games with the epistemic forms and supported student learning of practices and ideas in a sequential, yet integrated manner through iterative refinement (Figure 6-11).

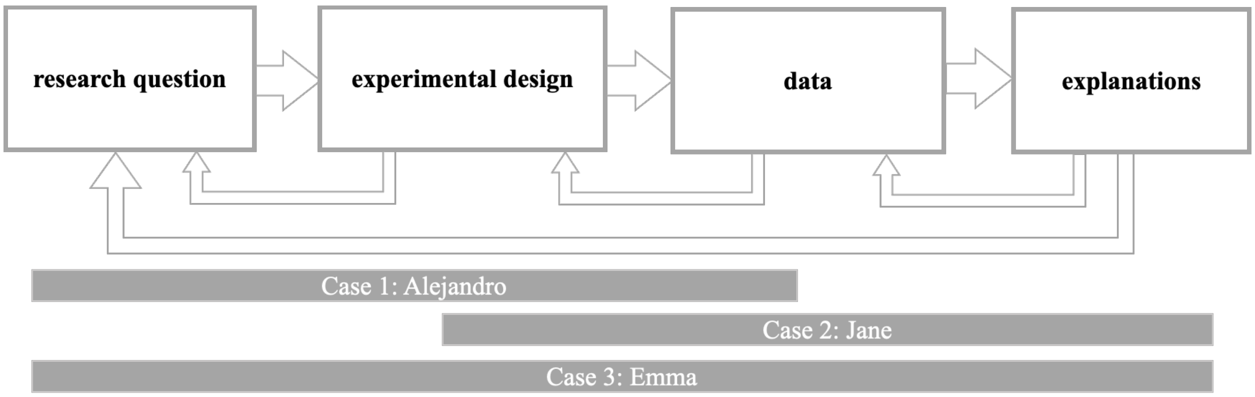


Figure 6-11 Mapping between the case studies and the epistemic form of the ESM-based learning environment

The ability to engage in more refined and sophisticated aspects of a practice is an important part of the epistemic game that the students were scaffolded to participate in. Alejandro's refinement of a practice, Jane's refinement of disciplinary aspects, and Emma's engagement in practices and ideas in the final lesson all demonstrate how the curriculum leverages disciplinary context and practices in a reciprocal manner to support learning of each other.

The ESM-based curriculum discussed in the paper provided multiple pathways for students to investigate various aspects of modeled disciplinary ideas regarding changes in a population because of natural selection. However, some of the students' investigations lacked important details about natural selection. For example, the investigation of Emma's group did not account for the effects of the predation rate. Their results about elimination of a mutant phenotype with survival disadvantage are valid only when predation rates are high. The predation rate in this model is what evolutionary biologists refer to as the selection pressure. If Emma's group had conducted additional experiments with different predation rates, they would have seen different results and learned about how selection pressure affects the process of natural selection. To leverage such learning opportunities, it is important to ask students to share their investigations and conclusions. Questions and feedback from their peers and the teacher can help students refine their experimental designs and engage in more thorough investigations. Additionally, such discussions will allow students to learn from others about different but related aspects of the phenomenon compared to the ones that they investigated. Our future work will involve designing specifically for such sharing between peers and teachers to take advantage of such learning opportunities in ESM-based curricula.

IMPLICATIONS

The operationalization of the notions of epistemic forms and games (Collins & Ferguson, 1993) to design for and assess learning of practices and disciplinary ideas provides a structure for developing NGSS-aligned curricula for engaging students in authentic science practices. In this paper, I demonstrated how an ESM-based curriculum effectively created learning opportunities for students to engage them in creating an epistemic form (Figure 6-1) aligned with NGSS recommended practices. This form of the canonical science practice is much more complex than the form of tic-tac-toe. It requires sequential and iterative engagement in the practices in the context of the epistemic game that a student chooses to play. In an ESM-based curriculum, students can engage in different variations of an epistemic game to generate the target epistemic form. This way they engage in knowledge construction of different aspects of disciplinary ideas modeled in the ESM. As witnessed in Emma's case, the variations of the epistemic game are possible because of the structure of an ESM.

The ESM-based curriculum design approach provides ways to iteratively engage students in learning and refining practices and disciplinary ideas. Epistemic form of an ESM-based curriculum and ways to scaffold possible epistemic games provide a framework for increasing alignment between the games that students participate in and the form that the curriculum intends them to generate. Computational models designed as ESMs have agent-based representations that create cognitive ease for students to engage in an epistemic game that uses practices to construct disciplinary ideas. The work also provides guidelines for designing ESM-based curricula that align students' epistemic games with the epistemic form that the curriculum is designed for.

In this chapter, I have discussed an ESM-based curriculum that is based on an epistemic form aligned with NGSS-recommended science practices. However, it is important to acknowledge that engagement in these practices is only one way to construct knowledge about the world. Pluralistic considerations regarding students' and practitioners' epistemological orientations are important to embrace diversity regarding the ways of knowing and making sense of natural phenomena in the contexts of science education (M Bang & Medin, 2010; Warren et al., 2020) and learning with computational learning environments (Turkle & Papert, 1990). Although in this chapter I investigate learning of NGSS recommended practices that are aligned with Western modern scientific ways, I do not explicitly or implicitly value Western modern scientific ways of knowing over the other ways. Nevertheless, I assert the importance of learning the practice of constructing and validating knowledge as specified by NGSS by defining a set of

SEP as simply *a* practice of knowledge construction, not *the* practice of knowledge construction. In future studies, ESM-based curricula can be designed that are aligned with different epistemic forms, which can support multiple epistemologies in disciplinary contexts (M Bang & Medin, 2010). For example, an ESM-based curriculum can be designed so that students can use indigenous ways of exploring and knowing. Students' epistemic connections can be studied in different contexts to examine how practice and ideas are connected in multiple ways of doing science.

To expand the goal of science education from *knowing about science* to *practicing science* requires the creation of learning opportunities that effectively support students in doing both. Designing ESMs and ESM-based curricula for different disciplinary contexts is a way of achieving this goal. Analyzing students' epistemic connections can be an effective way to assess student learning and the effectiveness of restructured curricula that support student learning of authentic scientific practices in disciplinary contexts.

Chapter 7: An ESM to learn about habitat preference behavior and experimentation

Summary: This chapter describes an Emergent systems microworld (ESM) that is designed for high school students to learn about both animal behavior and experimental design. The ESM is about the habitat preference behavior of isopods, commonly known as rollypollies. The inspiration for designing this ESM is an Advanced Placement (AP) Biology Unit about experimental design that uses rollypolly habitat preference behavior as a context to teach experimental design. Ms. Tracy[16], a high school biology teacher and I co-designed an ESM that models rollypolly habitat preference behavior and a curricular unit that uses the ESM for students to learn experimental design and specifically engage in Computational Thinking (CT) practices. In this chapter, I first give an overview of the ESM-based curricular unit. Then I discuss the ESM design in detail, focusing on the agents, their behaviors and interactions, and the emergent patterns regarding habitat preference that users can investigate using this ESM. Finally, I present four curricular activities that use the ESM – (1) computational investigation of rollypolly habitat preference behavior, (2) modification of a computational model to incorporate new experimental conditions, (3) computationally automated experiential setup and data collection, (4) construction of a computational model using coding blocks.

OVERVIEW

There are three design chapters in this thesis. They describe the designs of ESMs that I designed or co-designed as a part of the dissertation work. This chapter is the third of these

[16] pseudonym

design chapters. Each of the ESMs discussed in these chapters is designed for students to investigate and learn about natural phenomena based on an established scientific understanding of that phenomenon. Chapter 3, the first design chapter, is about an ESM of a genetic regulatory mechanism Lac Operon in bacteria called *E. coli*. This ESM is designed to study molecular mechanisms of gene regulation and the connection between genetic regulation and evolution. The underlying coded behaviors of agents in the ESM are based on the scientifically established understanding of the Lac Operon (Müller-Hill, 1996). Chapter 5, the second design chapter, is about an ESM of rock pocket mice. The rock pocket mice ESM is designed to study changes in a population that happen because of natural selection and genetic drift. The underlying coded behaviors of agents in the ESM are based on the research on the molecular mechanism of inheritance studied in the pocket mice (Nachman et al., 2003).

The ESM that I discuss in this chapter is about rollypollies. Rollypollies are isopods also known as woodlice or pillbugs[17]. Their common name is spelled differently – roly polies, roly-polies, rolly pollies, etc. I spell it as 'rollypolly/ rollypollies' because that is what Ms. Tracy and I decided during the co-design process and I want to be consistent with it. The rollypolly ESM is different from the previous two ESMs for the following three reasons: (1) it expands the definition of an ESM from a computational platform consisting of a NetLogo model or models to a computational platform consisting of NetLogo and NetTango Web (Horn et al., 2020) models with additional features such as block-based coding; (2) this ESM and the ESM-based curriculum are specifically designed for students to engage in Computational Thinking (CT) practices in a disciplinary context (Weintrop et al., 2016); (3)) the underlying behavioral

[17] https://en.wikipedia.org/wiki/Armadillidiidae

mechanisms of agents in this ESM go beyond the ones that are discussed in pre-existing curricular activities. The second point is discussed in detail in the next chapter, which explains a co-design approach for developing CT-integrated curricula using ESMs. In this chapter, I focus on the first and the third point. I attempt to explain why and how this expanded definition of ESM is useful for designing technology-enhanced restructurated curricula. I also demonstrate how agent-based restructurations allow the easy incorporation and testing of a behavioral mechanism of agents that generates an observed emergent pattern. This is evidenced in one of the curricular activities which asks students to code the behavioral mechanism using block-based coding and test if they can code the observed emergent behavior. I argue that incorporating block-based coding is a powerful way to engage students in (a) expressing their understanding of an underlying agent-level behavioral mechanism using a computational medium, and (b) testing if it results in the observed system-level emergent patterns and continue to refine the model until it does.

THE ROLLYPOLLY ESM UNIT

The rollypolly ESM unit was designed as a part of the CT-STEM project, which focused on re-designing existing science curricula by incorporating CT activities (Swanson et al., 2019). The CT-STEM project involved teacher professional development during summers and classroom implementations during the school years (Peel et al., 2020). Ms. Tracy was a participant teacher in the CT-STEM project for three years.

The rollypolly ESM unit design was inspired by an Advanced Placement (AP) Biology Lab (Lab 11: Animal Behavior) that Ms. Tracy had taught for several years. The AP biology lab involves the investigation of habitat preference behavior of rollypolly bugs. Students are asked to collect these bugs and bring those to school. A two-chamber setup is created by cutting and connecting plastic Petri Dishes (Figure 7-1). Each chamber has a specific condition, such as moist/dry, acidic/neutral, or light/dark. The bugs are released in the chambers and their distribution is observed after regular time intervals. Using a Chi-square test, students investigate if there is a significant difference in the distribution to determine if the bugs prefer one chamber habitat over the other. The curricular activities are intended for students to learn how to design and conduct an experiment to study animal behavior. They are also expected to learn hypothesis testing and data analysis.

Figure 7-1 Rollypolly physical lab setup in Ms. Tracy's class

The rollypolly lab is typically conducted to determine if the behavior of the bugs is taxis or kinesis (for example, an Honors biology lab, named Pillbug Behavior Lab: Kinesis vs Taxis.)[18] Taxis response is defined as the movement of an animal either towards or away from a

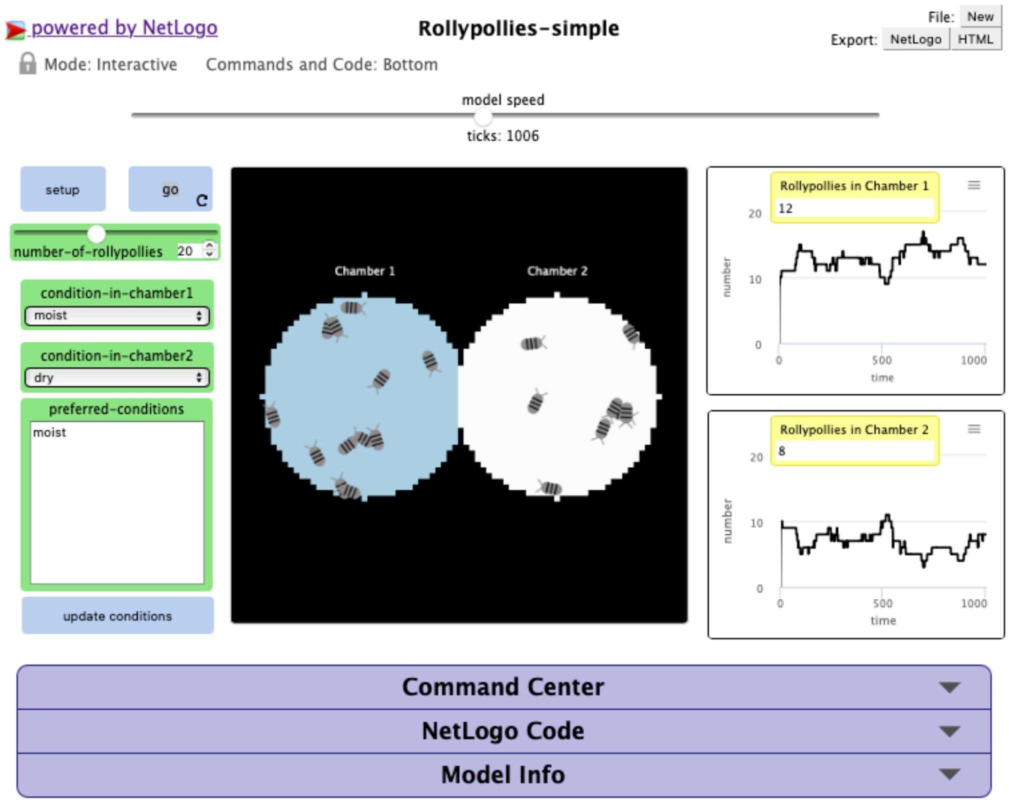
[18] http://meganfuellingbiology.blogspot.com/2014/04/pillbug-behavior-lab-kinesis-vs-taxis.html

stimulus, whereas kinesis is defined as a random response, in which animals move randomly without any preference to which direction to go to. The AP Biology Pillbug Behavior Lab is designed to test if the movement of pillbugs/rollypollies is a kinesis response or a taxis response. The kinesis vs taxis dichotomy does not consider the third type of response which is to spend more time in a preferred environment and thus creating a pattern that resembles a taxis response. This *delayed-movement* response would generate similar population distribution pattern as expected with a taxis response. For example, in a moist/dry chamber-setup rollypollies move randomly, however if a rollypolly is in a moist chamber the probability of its movement is lower than the probability of its movement when it is in a dry chamber. We designed our ESM based on this third type of response. This response is not included in the pre-existing AP Biology Lab. However, it is something that can easily be tested experimentally with the rollypolly bugs, by tracking individual bugs and noting how much time each bug spends in a specific chamber condition before it moves into the other chamber condition.

Ms. Tracy and I co-designed an ESM about rollypolly behavior to integrate it into Ms. Tracy's regular curricular unit to complement the physical lab. The focus on the integration was to engage students in Computational Thinking (CT) practices - *modeling and simulation*, *data analysis*, *computational problem solving*, and *systems thinking* practices which are part of the taxonomy of Computational Thinking (CT) practices (Weintrop et al., 2016). The details of design choices for student engagement in these practices are discussed in the next chapter. The rollypolly ESM was designed in three distinct forms to focus student engagement in specific CT practices. In the following part, I discuss these three forms.

Habitat Preference Behavior Model

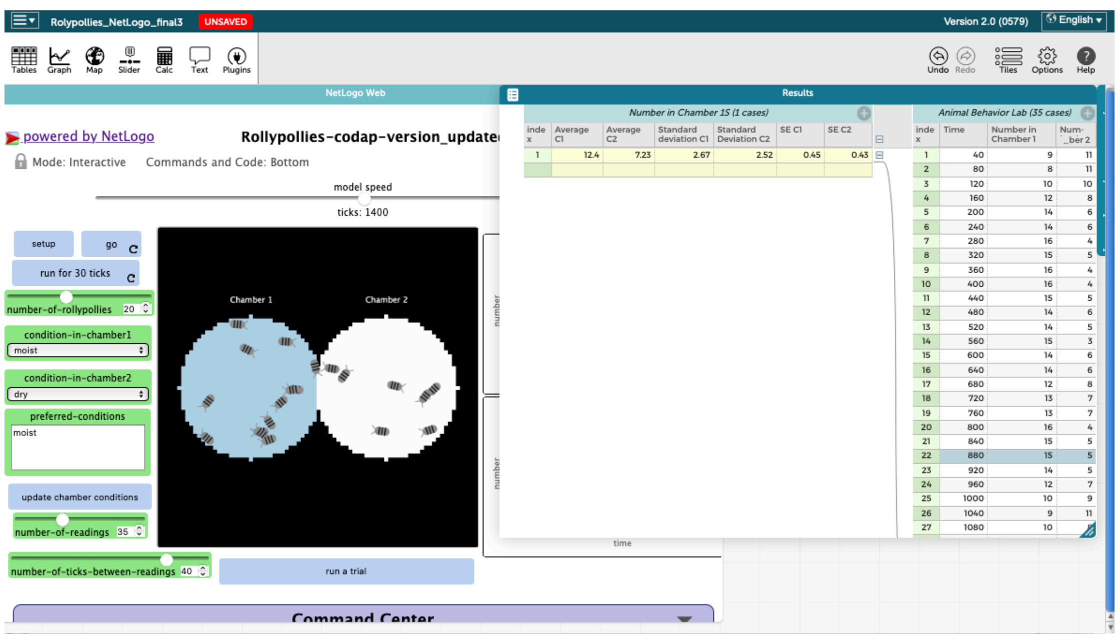
The habitat preference behavior model (Dabholkar, Granito, et al., 2020) is a model that is a fundamental component of this rollypolly ESM. This model is similar in design to the ESMs discussed in Chapter 3 and Chapter 5. It contains a two-chamber setup with rollypollies. Students can decide the number of rollypollies and conditions in the chambers, such as dry or moist. They can also set a preferred condition or multiple conditions. Even though students can test two conditions at a time, the model is designed such that students can add more conditions. These conditions are represented with different colors. Figure 7-2 shows two conditions; Chamber 1 is moist, and Chamber 2 is dry. As the model runs, the rollypolly bugs move around randomly. The randomness in their movement is constrained randomness depending on the environmental condition they are in. This is explained in detail in the next section. Users can observe the temporal progression of the numbers of rollypollies in each chamber in the graphs.



<u>Figure 7-2</u> Rollypolly computational lab setup in the co-designed ESM-based curricular unit (Granito et al., 2020). The figure shows computational representations of agents (rollypollies) and the environment (two chambers with specified conditions) in the Habitat Preference Behavior Model (Dabholkar, Granito, et al., 2020)

COMPUTATIONAL AUTOMATION OF DATA CREATION, COLLECTION AND VISUALIZATION

The next form of ESM is created by integrating the NetLogo (Wilensky, 1999b) Habitat Preference Behavior Model (Dabholkar, Granito, et al., 2020) with a data representation and analysis tool, CODAP (Finzer, 2016; Konold et al., 2017) to allow computationally enhanced data collection and analysis (Figure 7-3). The purpose of this integration of a NetLogo model with a CODAP platform is to use data collection and data analysis facilities in the CODAP platform. This integration is to engage students in computational *data practices* and *modeling* *and simulation practices* together. In the form of the ESM, users can set experimental conditions such as environmental conditions in the chambers and number of rollypollies. Additionally, in this model, they can set data collection parameters such as the number of data points to be collected (number-of-readings), and the time duration between two readings (number-of-ticks-between-readings). 'RUN A TRIAL' button runs a complete trial and collects data in the CODAP file based on the settings. After a trial is complete, the CODAP file also displays summary statistics.



<u>Figure 7-3</u> The NetLogo Habitat Preference Behavior Model (S. Dabholkar, Granito, et al., 2020) integrated in a CODAP platform to enable computationally enhanced data collection and analysis

CODING BLOCKS TO CONSTRUCT A MODEL

The next form of ESM is designed using block-based programming software, a blocks-based interface to NetLogo Web, NetTango Web (Horn et al., 2020). NetTango Web is a computational platform for domain-centric block-based coding. It is developed by building on earlier software platform called DeltaTick which used domain-centric blocks for agent-based modeling (M. Wilkerson-Jerde & Wilensky, 2010; Michelle Wilkerson-Jerde et al., 2015). The NetTango Web Rollypolly ESM allows students access to programming blocks specifically designed to construct the rollypolly habitat preference model (Figures 7-4 and 7-5).

Figure 7-4 Coding blacks for students to create a computational model of the rollypolly experimental setup

For this computational learning activity, we designed two variants of this NetTango Web ESM. The first variant provides programming blocks to create the experimental setup and to model rollypolly movement in the chambers (Figure 7-4). The ESM form has highly specific coding blocks designed for specific kinds of programming that are sometimes referred to as phenomenological programming (Aslan et al., 2020). Users can drag, drop, and connect these blocks in the empty white space to code the SETUP and GO procedures in the model. The unit has an embedded video for students to see how to put the blocks together and construct a model.[19]

[19] https://www.youtube.com/watch?v=CDutYTJ8c4Q&t=16s

The second variant asks students to model the behavior of rollypollies related to habitat preference to get the emergent patterns that they experimentally observed regarding rollypolly distribution in the two chambers. Users are asked to use programming blocks to ask the computational agents to 'behave like rollypollies' by encoding the rules of rollypolly movement based on their habitat preference (Figure 7-5). This is determined probabilistically. In a programming block, users can set the 'chance' that a rollypolly would move at every clock-tick (Figure 7-5). This chance can be set differently depending on where they are at. The chance that a rollypolly would move when it is at a preferred patch can be set lower than the chance of its movement when it is not on a preferred patch. This is how students can code preference-based probability of movement. When a bug is in a preferred environment the probability of movement can be set to be lower than the one when it is in a non-preferred environment. For example, in Figure 7-5, the probability of movement is 0.3 or 30% when in a preferred environment and 0.6 or 60% when in a non-preferred environment.

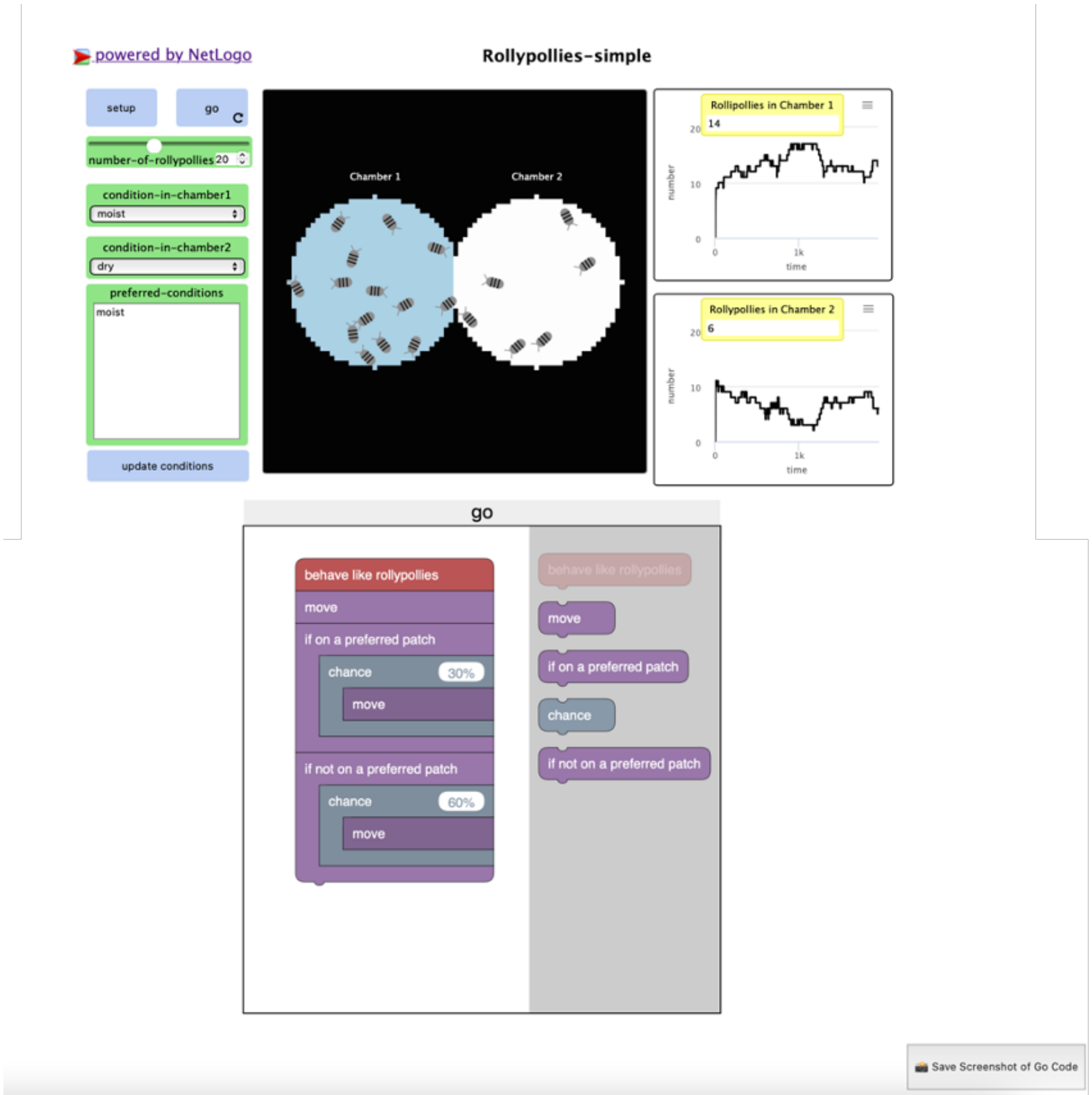


Figure 7-5 A NetTango Web ESM with coding blocks for students to code habitat preference behavior of rollypollies

The Habitat Preference Behavior Model

In this section, I explain the main computational model of rollypolly ESM in greater depth describing the agents in the model, their modeled behaviors and interactions, and the emergent patterns. This model is designed to simulate the habitat preference behavior of rollypollies.

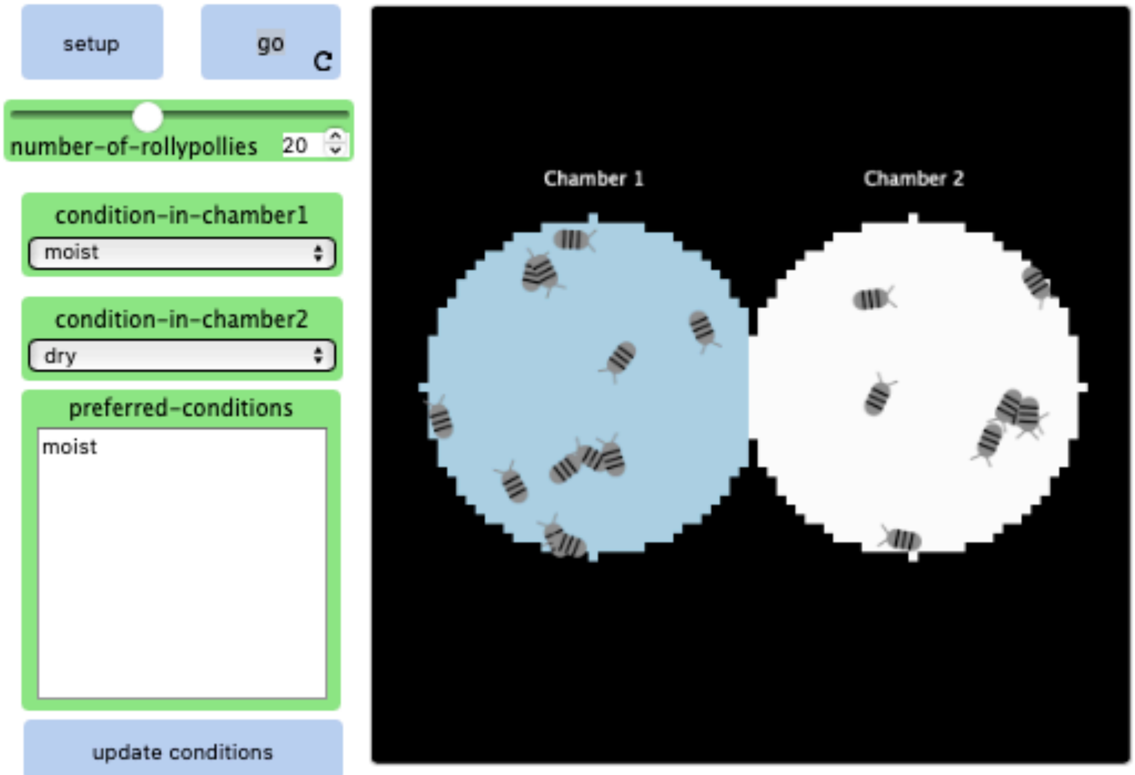


Figure 7-6 Habitat Preference Behavior Model (S. Dabholkar, Granito, et al., 2020) to investigate behavior preference of rollypolly bugs using a two-chamber experimental setup

AGENTS

Agents in this model are rollypolly bugs.

AGENT PROPERTIES

There is only one agent property in the model that is related to the emergent behavior of rollypollies. This property is '*preferred-condition*'. Users can set this property using the model interface. It is set as a global variable, which means that all the rollypollies in the model (NetLogo turtle agents) have the same preferred-condition.

When a rollypolly is in an environment with a *preferred-condition,* the probability of its moving is less than the probability when it is not in that condition. Users can input a list of multiple conditions that are preferred by rollypollies. The model assumes the same weights for all the preferred conditions incorporated in the model. For example, if moist and dark are listed as preferred conditions, then the probabilities of a bug moving when in a chamber with a moist or dark condition are the same. However, the model code can be easily modified to have a weighted preference for different conditions.

AGENT BEHAVIORS AND INTERACTIONS

The temporal progression in the model is represented in the form of ticks, similar to clock-ticks.

At each clock-tick,

Each rollypolly bug,

Checks the environment condition of the patch that it is on

If it is in a patch with a preferred-condition ->

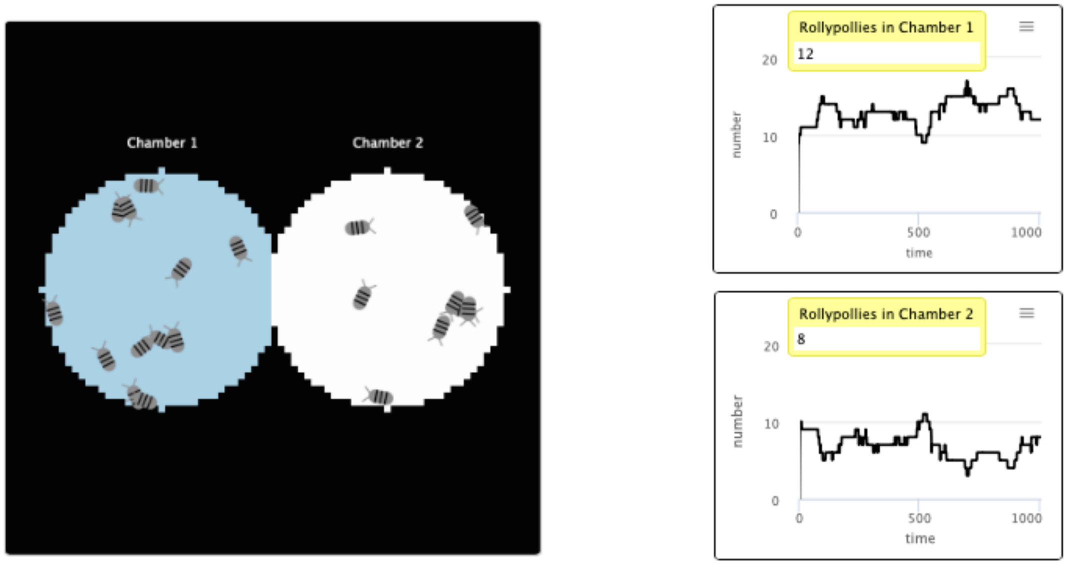
moves in a random direction within the boundaries of the chambers (based on a probability conditional to current environmental conditions)

otherwise ->

moves in a random direction within the boundaries of the chamber (based on a different (higher) probability value)

EMERGENT PATTERNS

The emergent pattern is the distribution of rollypollies in the two chambers. The chamber which is set to have the *preferred-condition* has more rollypollies on average over time. This can be observed visually in the computational microworld (Figure 7-7 A) as well as in the graphical representation (Figure 7-7 B).



(A) (B)

Figure 7-7 (A) Visual representation of rollypolly distribution in two chambers (b) Graphical representation of rollypollies in two chambers over time.

THE RESTRUCTURATED EXPERIMENTAL-DESIGN UNIT

The newly co-designed curricular unit includes physical experiments with rollypolly bugs, followed by computational experiments (Figure 7-8). In the following part, I discuss curricular activities in the computational lab part of the unit. It starts with students exploring the habitat preference behavior of rollypollies using a computational model. Then they are asked to investigate the model by designing and conducting experiments and modifying the model to incorporate conditions that they have physically experimented with. Finally, the curricular activities are designed to engage students in constructing the model using programming blocks.

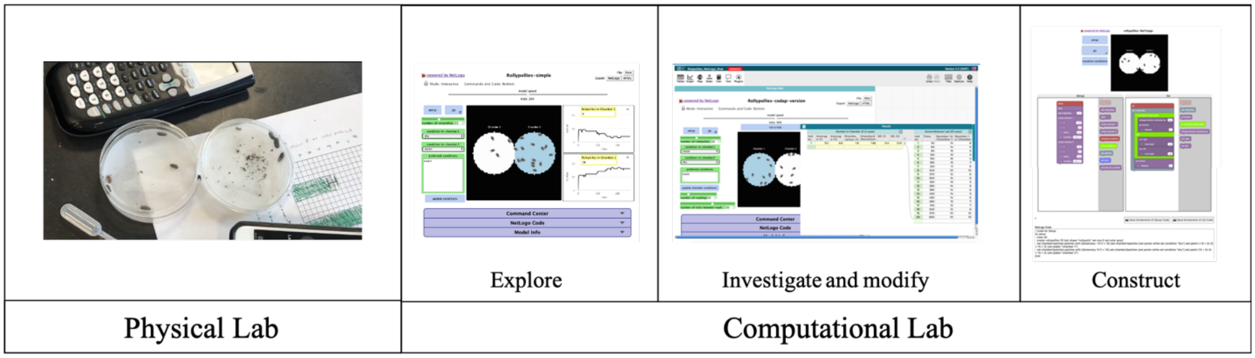


Figure 7-8 Unit components of Animal Behavior Lab - It starts with a physical lab and continues with a computational lab, in which students are asked to explore, investigate, modify and construct computational models.

COMPUTATIONAL EXPLORATION AND INVESTIGATION OF ROLLYPOLLY HABITAT PREFERENCE BEHAVIOR

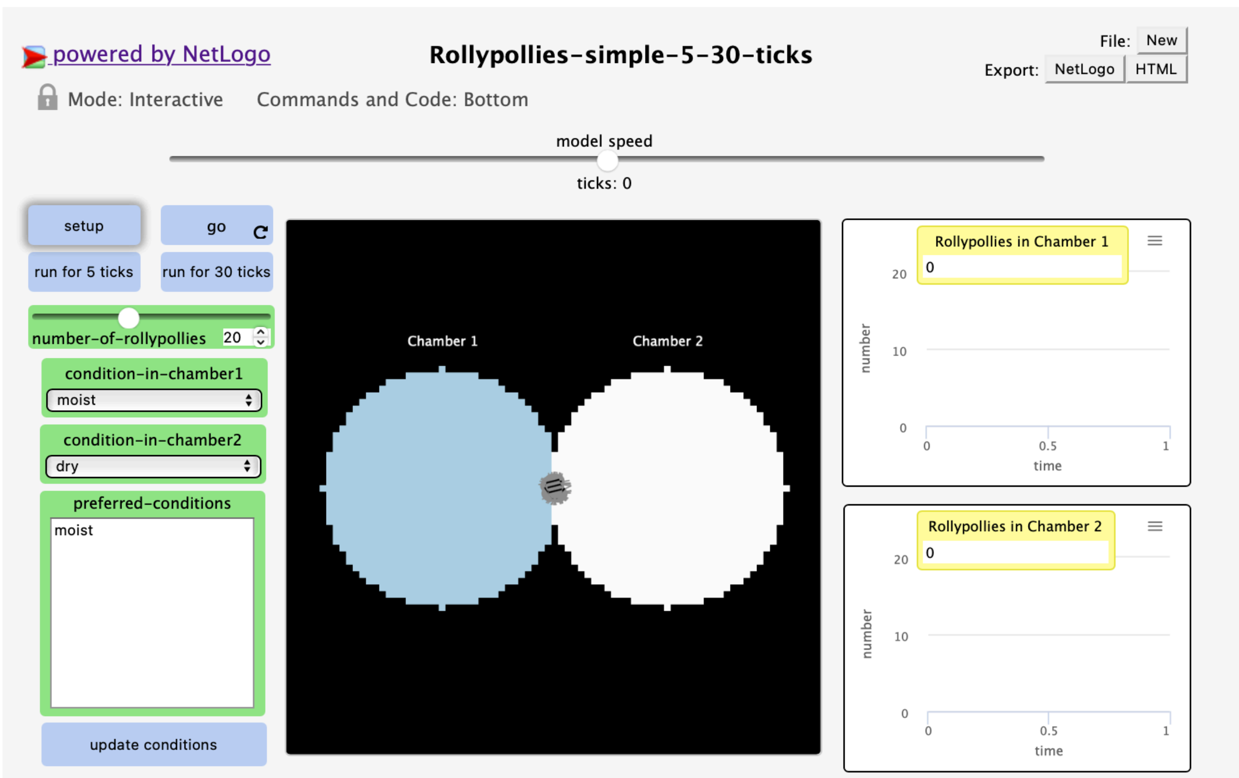


Figure 7-9 An initial model for students to learn about the setup of the computational experiment

The initial model exploration activity asks students to play with the model and note observations that they find interesting or surprising (Figure 7-9). This activity is designed for students to get familiar with the setup of the computational experiment, learn to notice, and start thinking about how they can use this model to perform scientific investigations. Further in the activity, students are asked to make predictions with specified experimental settings, and they run a computational experiment to test their predictions. These questions are designed for students to understand the use of an experimental system to make and test predictions. Specific model features such as the buttons 'run for 5 ticks' and 'run for 30 ticks' allow students to have better control over running computational experiments and comparing results from those. The questions in these activities ask students to repeat their experiments under the same conditions and compare results. These questions are designed for students to investigate the inherent randomness in agent behavior in the model and the emergence of predictable global patterns.

MODIFICATION OF A COMPUTATIONAL MODEL TO INCORPORATE NEW EXPERIMENTAL CONDITIONS

The next activity asks students to modify the NetLogo model to incorporate new conditions. Students are asked to incorporate the new conditions that they have previously experimented with. A physical lab activity that occurs before the computational lab asks students to conduct habitat preference experiments with different environmental conditions such as light/dark, acidic/neutral/basic pH, leaf litter/no leaf litter, and salt/no salt. The curriculum provides students a series of screenshots to guide the model modification activity. Figure 7-10 shows two initial screenshots for students to learn about how to modify the computational model by getting into an authoring mode.

Figure 7-10 Screenshots in a guiding document for student to modify a NetLogo model to incorporate new conditions

This activity is designed for students to learn about the epistemic nature of models as tools to think about a phenomenon. The lesson also includes questions such as - *How do the results from the computational models compare with the rollypolly experiment you conducted? What are some advantages of using computational models to predict animal behavior vs the physical lab experiments we conducted at the beginning of the unit? What are some limitations of the computation models used in this unit? What are some limitations to the physical lab? How can a computational model aid a scientist with their work?*

These questions are included for students to think and write about their understanding of the usefulness and limitations of computational models. Our prior work has demonstrated that ESMs and ESM-based curricular activities support students' meaningful engagement in thinking about epistemic considerations of modeling related to their usefulness and limitations (Dabholkar, Swanson, et al., 2019). After having performed physical and computational experiments, these questions ask students to compare these approaches in terms of their usefulness and limitations. Also, after having modified a model, these questions invite students to think about the usefulness of computational modeling for professionals such as scientists.

COMPUTATIONALLY AUTOMATED EXPERIENTIAL SETUP AND DATA COLLECTION

The next activity, computationally automated data collection, is designed for students to learn more about designing experiments to investigate a phenomenon. It also focuses on student engagement in *data practices*. Students are asked to design their experiments to investigate their research questions about rollypolly habitat preference behavior. Students are asked to specify data collection conditions in their protocol. The protocol design should include the initial conditions, the number of trials, and the time interval between two readings. The model allows students to choose the number of data points (*number-of-readings*) and the time interval between two data points (*number-of-ticks-between-reading*) (Figure 7-11).

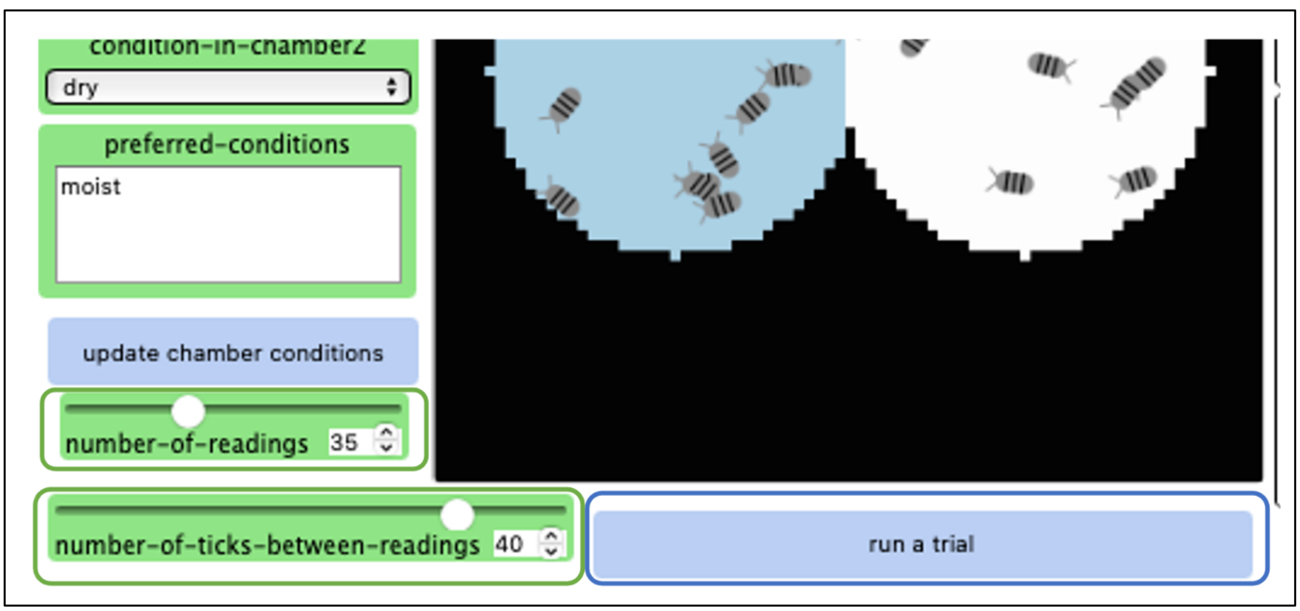
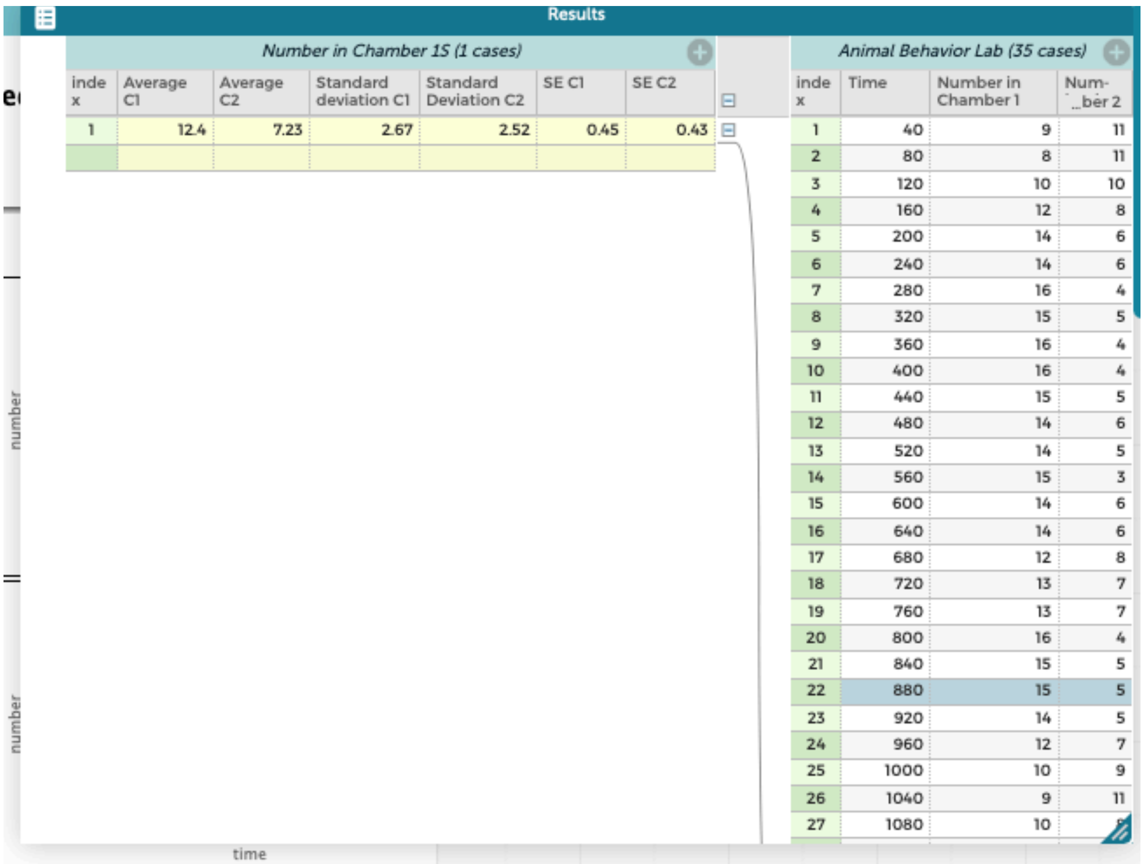


Figure 7-11 Additional sliders and buttons in the model for students to specify data collection conditions and run a complete automated experimental trial

Students can set an automated data collection protocol using the sliders *number-of-readings* and *number-of-ticks-between-readings*. Then they can click on 'run a trial' button and a complete data set is generated and exported in a CODAP data table (Figure 7-12). An automatically generated statistical summary is also available for students.


Results

Number in Chamber 1S (1 cases)

index	Average C1	Average C2	Standard deviation C1	Standard Deviation C2	SE C1	SE C2
1	12.4	7.23	2.67	2.52	0.45	0.43

Animal Behavior Lab (35 cases)

index	Time	Number in Chamber 1	Num-_ber 2
1	40	9	11
2	80	8	11
3	120	10	10
4	160	12	8
5	200	14	6
6	240	14	6
7	280	16	4
8	320	15	5
9	360	16	4
10	400	16	4
11	440	15	5
12	480	14	6
13	520	14	5
14	560	15	3
15	600	14	6
16	640	14	6
17	680	12	8
18	720	13	7
19	760	13	7
20	800	16	4
21	840	15	5
22	880	15	5
23	920	14	5
24	960	12	7
25	1000	10	9
26	1040	9	11
27	1080	10	

Figure 7-12 A CODAP data table showing a computationally generated dataset of pre-specified data collection conditions

This activity is designed for students to learn about ways to automate data collection using computational tools. Additionally, they are also prompted to think about the data collection protocol influencing the quality of data which potentially results in limitations for drawing conclusions based on a dataset. Questions towards the end of the lesson include: *Describe the* *benefits of computationally enhanced data collection and analysis tools. For what purpose would a scientist use computationally enhanced automated data collection and analysis tools?* These questions are designed to encourage students to think about computational automation related to *data practices* and their affordances.

CONSTRUCTION OF A COMPUTATIONAL MODEL USING CODING BLOCKS

The curricular activity of constructing a computational model is designed in parts. In the first part, students are asked to use programming blacks to create the experimental setup. The activity includes a YouTube video[20] that demonstrates how to add blocks and connect those. The NetTango Web interface provides students with pre-specified programming blocks (See Figures 7-4 and 7-6).

There are two sections for the first activity using the NetTango Web model - one is to set up the model (Setup) and the other is to run the model (Go). The blocks in the 'Setup' section include *setup*, *add-rollypollies*, *create a chamber*, and *clear*. The blocks in the 'Go' section include *go*, *ask rollypollies*, *go forward*, *if I am facing a dark patch*, *change direction clockwise by*, *turn right*, and *turn left*. Students are required to decide the location (*x* and *y* coordinates) and size (radius) of the blocks to connect them in a manner that looks similar to the physical experimental setup (See Figure 7-1). The questions in this activity are posed as a set of

[20] https://www.youtube.com/watch?v=CDutYTJ8c4Q

challenges for students to try out. The list of challenges in the lesson that Ms. Tracy taught are given below:

Try the **challenges** below:

1. Drag blocks over to make a chamber.
2. Increase the size of the chamber.
3. Move the chamber to the left, and to the right.
4. Move the chamber up, down.
5. "add-rollypollies"
6. Make 2 Chambers that are connected.
7. Add the 'clear' block and see how it works.
8. Setup again and run the model by pressing the 'go' button.

After students *add-rollypollies* and run the model using the setup and go procedures (see Figure 7-12), they see all the computational bugs move out of the chamber (see Figure 7-13).

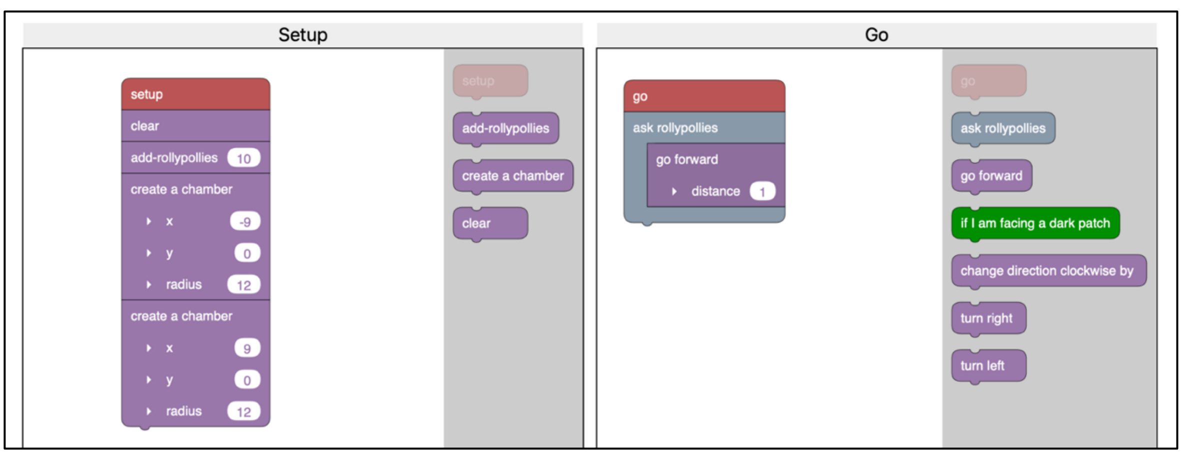


Figure 7-13 The assembled programming blocks to design an experimental setup and define agent behavior when the model is run using a button 'go'

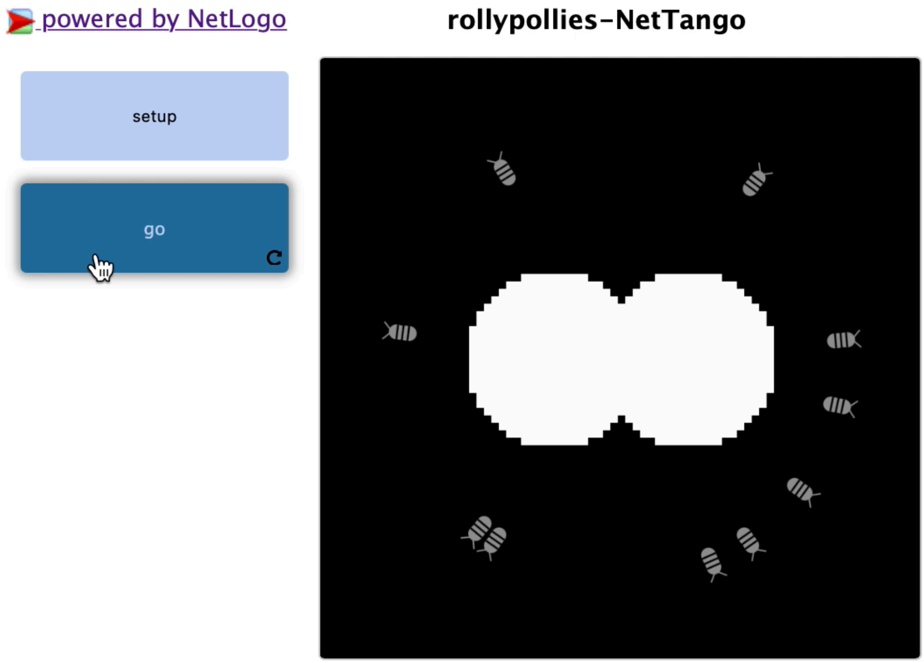


Figure 7-14 The computational rollypollies move out of the chamber

The next activity poses another challenge to students.

Challenge: Add rollypollies and have them move in a realistic way.

1. Drag blocks over to make two chambers that are connected. Add rollypollies. Tell the rollypollies how to behave.
2. Get the rollypollies to move around the field.
3. Get the rollypollies to move around the field more realistically.
4. Get the rollypollies to stay within the chambers.

There are two aspects to make rollypollies move realistically in this model. First, they should not always move in a straight line; they should wiggle. Second, they should not jump out of the chambers. Students need to set a computational mechanism for the bugs to understand the chamber boundary. One set of blocks for making rollypolly bugs move realistically is shown in Figure 7-15, and the resultant realistic bug movement as shown in Figure 7-16.

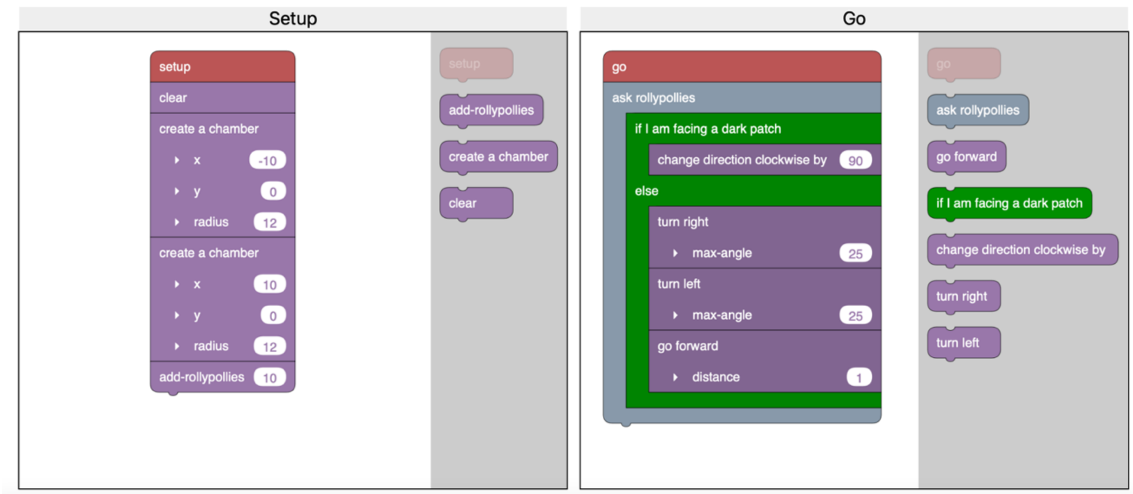


Figure 7-15 Assembled programming blocks to create an experimental setup and define more realistic rollypolly behavior when the model is run using a button 'go'

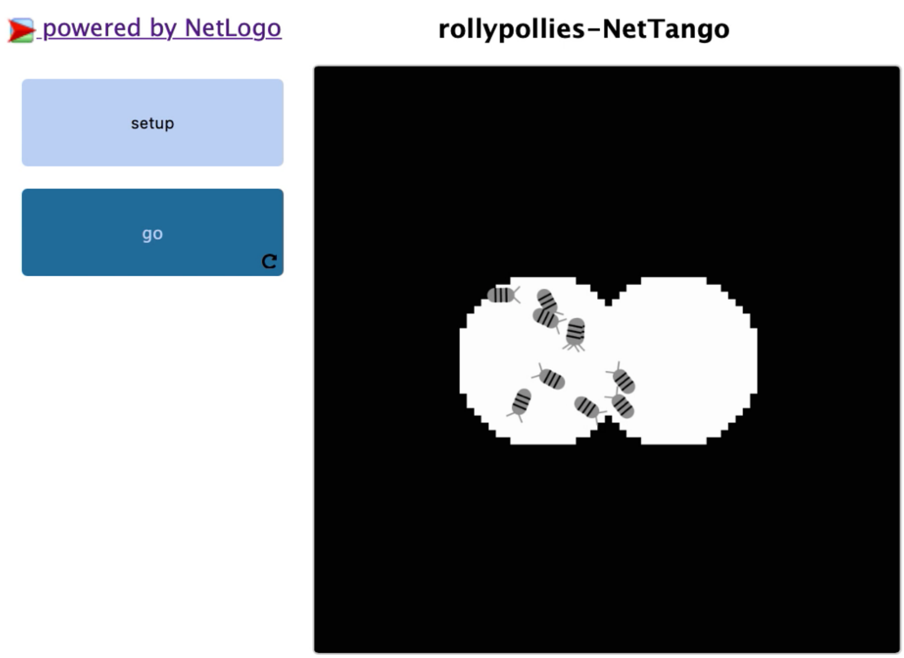


Figure 7-16 Computational rollypollies stay inside the chamber when a conditional block (Green block in Figure 7-15) is used to enable sensing of chamber boundaries

The next set of activities involve students setting up and representing environmental conditions in each chamber. Finally, students are provided with a NetLogo model with NetTango Web blocks (Figure 7-17). The students are asked to use the blocks to make the computational bugs behave like rollypollies by using the blocks and specifying parameter values in the blocks. The activity asks students to "*keep changing the parameter in the blocks until the rollypolly behavior in the model more closely matches the rollypolly behavior observed in the experiment we conducted*". Students are then asked to upload the screenshots of the NetTango Web block assembly and the graphs of temporal variation in rollypolly distribution in the two chambers.

This activity is inspired from Blikstein and Wilensky's bifocal modeling approach to curricular design (P. Blikstein & Wilensky, 2007). Till this point in the curriculum, students have already conducted experiments with rollypollies using a physical apparatus. This activity is designed for students to use their observations in a real-world experimental setup and encode those computationally. Additionally, this activity serves the following to purposes to engage students in specific CT practices. First, it asks students to construct a model to produce a reference pattern that they have experimentally observed. This requires students to engage in programming practices that are part of *computational problem solving practices*. Secondly, this activity requires students to think about computationally encoding agent behavior such that would result in an emergent system-level pattern. This aspect requires them to engage in *systems thinking practices*.

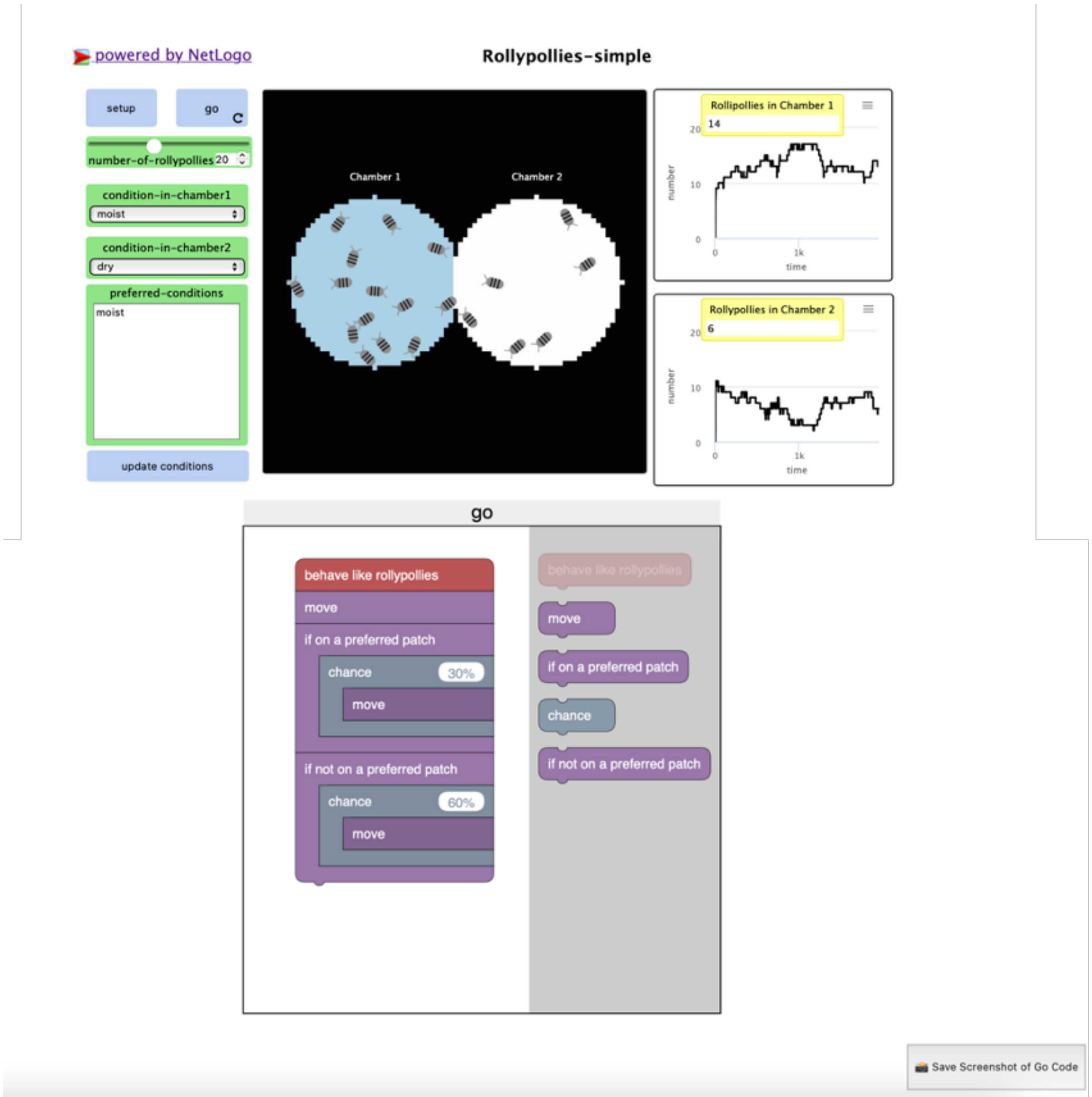


Figure 7-17 A NetTango Web interface with an integrated NetLogo model for students to specifically code the agent behavior

The questions towards the end of the lesson ask students: 1) *A few lessons ago, you changed a model by typing lines that setup "new-condition" in a model. Now you changed a model by moving blocks to set up conditions. What are the benefits of using block-based coding* *like this NetTango model, compared to typing lines of code?* 2) *What other blocks would you want to add to this model?*

These questions are designed for students to engage in thinking about various modalities of programming and modifying a computational model.

CONCLUSION

This ESM-based restructurated curriculum unit is designed to achieve the following three goals – (1) use computational tools to engage students in nuances of experimental design, such as the number of data points, sample size, the number of replications, etc., (2) engage students in a physical lab and a complementary computational lab to appreciate the usefulness and limitations of computational approaches, (3) engage students in Computational Thinking (CT) practices. In this chapter, I have discussed in detail how the curricular unit is designed to address these goals.

Power properties of a restructuration (Wilensky & Papert, 2010) are the properties that make a restructuration do what could be done before and more (See Chapter 1 for details). The original rollypolly unit is intended to teach students about experimental design. Curricular activities in the ESM-based rollypolly unit are designed for students to engage in CT practices help the learn nuances of experimental design such as sample size or number of trials. Reciprocally, the disciplinary context of experimental investigation of animal behavior provides opportunities for students to meaningfully engage in CT practices. For example, creating a model using block-based coding is not necessarily required to learn experimental design as it is taught in school, however, the disciplinary context of experimental investigation of animal behavior provides a meaningful scenario for student to engage in *computational problem solving* and *systems thinking practices*.

Throughout this unit, ESM-based activities are designed to engage students in all four categories of Computational Thinking practices (Weintrop et al., 2016). Initially, students engage in *modeling and simulation* practices to understand and investigate a phenomenon. Then, they engage in *data practices* to design data collection protocol in an experiment, collect data and analyze data. *Computational problem solving* and *systems thinking* practices are integrated throughout the unit. However, these practices are foregrounded in the last activity when students are asked to create a computational model and code the behavior of individual computational agents to produce system-level emergent behavioral patterns.

The restructurated curriculum discussed in this chapter is inspired from the bifocal modeling approach (P. Blikstein & Wilensky, 2007). The rollypolly curriculum uses both a physical system and a computational system for students to learn about experiment design and CT practices. However, the rollypolly unit does not have a crucial component of Biikstein & Wilensky's bifocal modeling approach. Bifocal modeling focuses on a sensory apparatus that can feed data from a physical apparatus into a computational apparatus. Such sensory inputs from real-world physical systems into modeled virtual computational systems can be powerful for students to learn about the physical system as well as its computational model. Future work with this curriculum will involve measurement of rollypolly behavior in the physical apparatus using various sensors and feeding of that data into the ESM.

In addition to being effective in supporting student engagement in CT practices, the ESM-based design approach also supports the process of co-designing such CT-integrated science curricula. In the next chapter, I investigate the effectiveness of a co-design approach for creating CT-integrated curricula using the ESM design. I demonstrate that the use of ESM-based curricular design approach supports shifts in the teacher involvement in the co-design process in meaningful ways. It also changes teachers' classroom practices to support student learning. Finally, such co-design approach increases richness of curricula in terms of CT-integration.

Chapter 8: An ESM-mediated co-design approach for integrating Computational Thinking in a science classroom

Summary: Integrating computational thinking (CT) activities in the K-12 science curriculum can be an effective way to engage students in learning contemporary science practices. However, this requires designing effective CT-integrated curricula and supporting teachers in adopting new teaching practices. I argue for involving teachers as partners for co-designing such a curriculum to increase their ownership and engagement in crafting effective pedagogical practices. I present a qualitative analysis of the involvement of a science teacher in the co-design process who participated in a Design-Based Implementation Research project aimed at creating CT-integrated curricular materials. I discuss how the underlying design approach of using ESM and ESM-based activities mediated an increase in teacher involvement, the outcomes of teacher involvement in the form of co-designed units, and changes in the classroom practice. Findings yield implications for how best to support teachers in curricular CT-integration.

INTRODUCTION

While many researchers and educators agree that there is a need to integrate computational thinking (CT) in high school curricula, they differ in their conceptualization of CT, ways to achieve integration, and the rationale for doing so (Grover, 2019; Weintrop et al., 2016; Wing, 2006). In line with the ongoing efforts for curricular reforms to engage students in authentic disciplinary practices (e.g., Goldman, Ko, Greenleaf, & Brown, 2018), the focus on integrated CT in disciplinary contexts has also been increased (Lee et al., 2020). Wilensky, Horn, and colleagues (2014) argue for CT integration in science and mathematics classrooms for the following reasons: (a) for students to learn authentic contemporary disciplinary practices, (b) pedagogical effectiveness of thoughtfully integrated computational tools, and (c) to reach the widest possible audience, especially students from backgrounds that have been historically marginalized in computational fields (Weintrop et al., 2016; Wilensky et al., 2014).

Research in the learning sciences over the past couple of decades have demonstrated pedagogical effectiveness of computational tools for learning disciplinary ideas in K-12 education (Clark et al., 2009; Guzdial, 1994; Klopfer et al., 2009; Levy & Wilensky, 2009; Quintana et al., 2004; Sengupta & Wilensky, 2009; Wilensky, 2003; Michelle Wilkerson-Jerde et al., 2015). Other work about integration of CT into disciplinary contexts has also demonstrated the effectiveness of CT integrated curricula for engaging students in scientific inquiry practices, and more specifically CT practices (Arastoopour et al., 2020; Dabholkar, Arastoopour, et al., 2020; Hutchins et al., 2020). This approach of creating CT integrated curricula involves increased focus on student engagement in *practices,* specifically Computational Thinking (Weintrop et al., 2016) which is included as one of the Science and Engineering Practices by the Next Generation Science Standards (NGSS Lead States, 2013). To increase the focus on student engagement in practices to construct knowledge about disciplinary ideas requires designing curricular activities for making the practices meaningful in curricular contexts and balancing the tension between teaching practices vs disciplinary ideas (Berland et al., 2016; Russ & Berland, 2019; Michelle Wilkerson-Jerde et al., 2015).

Especially, for student engagement in CT practices in science classrooms, it is important to design CT-integrated curricula that are balanced in terms of the learning goals regarding CT practices and disciplinary ideas. Additionally, from the perspective of the third goal of CT-integration, to broaden participation in CT learning opportunities, it is important to design appropriate curricula and pedagogical strategies that effectively support student learning of CT practices. In this chapter, I discuss a co-design approach for creating CT-integrated curricular materials and investigate the effectiveness of a technology-based design framework of co-designing CT-integrated curricula.

CO-DESIGNING CT-INTEGRATED CURRICULA

In the field of design research in general, and the learning sciences in particular, the value of co-designing solutions is increasingly being appreciated for creating innovative and sustainable designs (Blomkamp, 2018; Mitchell et al., 2016; Trischler et al., 2019). Design researchers have effectively employed co-design methodologies for designing sustainable travel solutions to reduce single-occupancy car journey in a UK university (Mitchell et al., 2016), and for designing policies regarding driver-licensing in South Auckland, New Zealand for local families predominantly from indigenous (Māori) and Pacific Island cultures (Blomkamp, 2018). Learning scientists also have used participatory and co-design methods to create innovations in science education focusing on culturally relevant practices (M Bang et al., 2010), to social design experiments (Gutiérrez et al., 2020), to school district reforms (Kwon et al., 2014). Similar co-design methodologies has also been demonstrated to be effective for teacher professional development (Voogt et al., 2015) and designing for curricular materials (Peters & Slotta, 2009).

Designing for effective technology-integrated curricula requires the development of appropriate technological tools and novel methodological approaches such as specific ways to involve teachers in co-designing curricula (e.g., Groff & Mouza, 2008; Jeong & Hmelo-Silver, 2016). To support student learning with newly designed curricular materials requires teachers' adaptiveness to student needs and thinking (Windschitl et al., 2012). Teachers need to adapt their pedagogical practices such that they support student learning aligned with the new educational goals, such as learning of CT practices. Inviting teachers to be design partners and facilitating their agency in the co-design of curricular materials can be an effective approach to increase their ownership and engagement in the appropriate pedagogical practices (Kyza & Georgiou, 2014; Severance et al., 2016).

In this chapter, I investigate how the ESM (Emergent Systems Microworlds)-based design approach mediated such co-design efforts. ESMs combine constructionist (Papert, 1980) and agent-based (Wilensky, 2001, 2003) design principles to offer an interactive learning environment for students to explore and investigate phenomena (Dabholkar, Anton, et al., 2018; Dabholkar, Arastoopour, et al., 2020). In chapters 4 and 6, I have discussed how cognitive and social properties of agent-based restructurations (Wilensky, 2020; Wilensky & Papert, 2010) in ESMs support students' collective knowledge construction by engaging in scientific inquiry practices. In this chapter, I discuss how the ESM-based curricular design approach support co-design efforts to create CT-integrated curricula. I present a case study of a researcher-practitioner design partnership for three years as it matured from using pre-designed CT-integrated curricula to co-designing new CT tools and curricula. I co-designed ESM-based curricular units with a high school biology teacher, Ms. Tracy (a pseudonym) for engaging students in CT practices in different biology contexts. We co-designed new ESMs and curricula activities using those ESM. Use of the ESM design approach for CT-integration supported the co-design partnership, as the cognitive and social properties of restructurations (Wilensky & Papert, 2010) in ESMs mediated Tracy's increased involvement in the co-design process. Reciprocally, Tracy's highly valuable contributions in identifying relevant biological contexts and devising pedagogically effective learning activities in the co-design process enriched the ESMs and ESM-based curricula.

In this chapter, I first investigate how the following aspects of the co-design process were mediated by the ESM design approach - (a) changes in Tracy's involvement in the co-design process, (b) shifts in the curricular designs in terms of the richness of CT-integration, and (c) shifts in Tracy's classroom teaching practice. Then, I present an argument for the effectiveness of ESMs for designing CT-integrated curricula and the co-design approach that uses ESMs for co-designing such curricula.

THEORETICAL FRAMEWORKS

The approach to CT-integration in a science curriculum discussed in this chapter is based on a theoretical framework for integrating CT in science and mathematics (Weintrop et al., 2016). This framework is in the form of a taxonomy of practices that are recommended to be integrated in science and mathematics curricula. Whereas the design approach for creating computational tools that were embedded in CT-integrated curricula, such as models or data analysis tools, and learning activities that use those tools is based on the ESM design framework.

CT-STEM PRACTICES

This work is framed with the operational definition of CT in science and mathematics as a taxonomy of practices (Weintrop et al., 2016). This taxonomy categorizes CT-STEM practices in terms of four major categories: *data practices, modeling and simulation practices, computational problem-solving practices, and systems thinking practices* (Table 8-1). This project focused on integrating these CT-STEM practices into high school science curricula. We developed new computational models and activities that are rooted in the disciplinary contexts for students to learn disciplinary ideas as well as CT practices. We used two technology platforms to create computational models for students - (1) NetLogo (Wilensky, 1999b, 2001), an agent-based modeling software, and (2) NetTango Web (Horn et al., 2020; Horn & Wilensky, 2012), a block-based programming interface to NetLogo which uses semantically meaningful blocks tuned to the content domain built on a previous software platform and methodology, DeltaTick (M. Wilkerson-Jerde & Wilensky, 2010; Michelle Wilkerson-Jerde et al., 2015).

Table 8-1 The CT-STEM project's Taxonomy of CT-STEM Practices

Data Practices	**Modeling and Simulation Practices**	**Computational Problem Solving Practices**	**Systems Thinking practices**
Collecting Data	Using Computational Models to Understand a Concept	Preparing problems for computational solutions	Investigating a complex system as a whole
Creating Data	Using Computational Models to Find and Test Solutions	Programming	Understanding the relationships within a system
Manipulating Data	Assessing Computational Models	Choosing effective computational tools	Thinking in levels
Analyzing Data	Designing Computational Models	Assessing different approaches	Communicating information about a system
Visualizing Data	Constructing Computational Models	Developing modular solutions	Defining a system and measuring complexity
		Creating computational abstractions	
		Troubleshooting and debugging	

RESTRUCTURATIONS WITH EMERGENT SYSTEMS MICROWORLDS (ESM)

I have discussed in the earlier chapters that agent-based models in general, and ESMs in particular have been demonstrated to be effective for learning fundamental ideas in sciences and mathematics such as particulate nature of matter, electric current, the evolution of populations (Dabholkar, Anton, et al., 2018; Levy & Wilensky, 2009; Sengupta & Wilensky, 2009; Wagh et al., 2017; Yoon & Hmelo-Silver, 2017). The restructuration properties (Wilensky, 2020; Wilensky & Papert, 2010) of agent-based restructurations in ESMs provide cognitive and social affordances among others to support student learning (see chapter 1 for the detailed explanation of restructuration properties). However, the ESM-based curricular design approach has not been investigated for its effectiveness in co-designing CT-integrated curricular units. In this chapter, I argue that the co-design approach that uses ESM framework for curricular design allows a teacher to design pedagogically effective representations and devise appropriate pedagogical strategies to support student learning of CT practices. I focus on the outcomes of the co-design process in the form of the created CT-integrated units and enacted pedagogical practices to support student learning in the classroom. To understand the effectiveness of the ESM-mediated co-design approach I investigated the following research question:

How does restructuration through ESM facilitate the co-design process for CT-integration into science units and its outcomes?

METHODS & DATA SOURCES

I investigated ESM-mediated co-design in a longitudinal case study of a design partnership of three years between me and a high school biology teacher, Ms. Tracy, who was a participant teacher in the CT-STEM project, which involved teacher professional development during summers and classroom implementations during the school years (Peel et al., 2020). The

CT-STEM project evolved over the three years and so did the co-design partnership and Tracy's involvement in the project. Over the three years, Tracy taught three CT-integrated units that are published and publicly available open-source units (Dabholkar, Granito, et al., 2019; Dabholkar, Hall, et al., 2018; Granito et al., 2020). These units have embedded computational models and tools, and questions to scaffold students' learning of disciplinary ideas as well as CT practices. The units and models were designed based on the ESM design framework. Tracy taught these curricular units in her biology classrooms at Greenville High School (pseudonym). See Table 8-2 for school demographics.

Table 8-2 Demographics of the School

School	**Race Demographics**	**Free/Reduced Price Lunch**	**English Language Learners**	**Individualized Education Plans**
Greenville	44% White, 30% Black, 18% Hispanic, 5% Asian, 0.4% Native American	39% free or reduced-price lunch	4.2% ELL	12% IEP's

In year one, Tracy was given a CT integrated biology unit that was previously designed by the research team. In year two, Tracy provided me directions as we designed a new CT integrated biology unit. She chose the biology content, provided her lesson plans, and gave feedback me as I designed the lessons. In year three, Tracy worked side-by-side with me to co-design a new CT integrated unit during a Computational Thinking Summer Institute (CTSI). Each year Tracy taught the curricular units in eight to ten class periods of 45-50 minutes.

Through this multi-year process, I collected data to characterize Tracy's experiences. The data sources include several interviews with the teacher conducted every year (See Appendix 7 for an example of an interview protocol), weekly reflections during the co-design periods, session recordings of co-design sessions. Additionally, I analyzed the unit designs and classroom video of implementations to support the analysis. I used a case-study method (Yin, 2012) to analyze how the ESM-based design approach supported and increased Tracy's involvement and changed her pedagogical approach while teaching in the classroom. I used the top-down coding approach to code the interview data when Tracy talked about her involvement in co-design, classroom teaching, and her views regarding student learning. Two other coders and I coded a sample of the data. Cohen's Kappa was calculated for each of the categories of practices for each author-pair. Disagreements between authors were resolved through discussion until a reasonable agreement was reached for each category. Kappa > 0.7 was considered as a cut-off for reasonable agreement between the coders. One-third of the rest of the data was coded by each coder. Additionally, I coded Tracy's responses that referred to her involvement or student learning with an ESM to analyze how Tracy viewed the ESM design approach in the context of co-design, student learning, and her pedagogy. To triangulate the qualitative analysis with the video data (Small, 2011), I used the activity-logging approach to identify episodes that illustrated Tracy's pedagogical practices while using computational tools. The claims and codes were also triangulated where possible with other data sources, such as the intermediate and final co-design artifacts.

FINDINGS

In this section, I present a qualitative analysis of how the agent-based restructurated representations increased Tracy's involvement in the co-design process across three years. I also discuss how ESM-mediated pedagogical discussions during the co-design process resulted in changes in Tracy's pedagogical practices to support students' learning of CT practices.

SHIFTS IN DESIGN CONTRIBUTIONS

Even though the research project evolved towards being more focused on the co-design process, Tracy's reflection about her role in the co-design process helps in understanding how she perceived the shifts in her role and design contributions.

Year 1. Since this unit used pre-designed ESMs, Tracy did not have any role in designing the ESMs nor did she design the pedagogical activities using the ESMs to support student learning. Tracy's pedagogical practices to support students' engagement in the embedded CT activities changed over the next two years. In her post-implementation interview, when asked about her experience regarding her partnership with the research team, she mentioned, *"... just fantastic and like, Sugat went through like all the lessons with me so I would know what to expect and no pressure like we had made a schedule.*" Tracy was appreciative of the fact that a researcher explained all the lessons to her thoroughly and she had the flexibility to progress the unit at a pace that she was comfortable with.

Year 2. In the second year, Tracy was more involved in identifying ways to integrate CT into her biology curriculum. During a short co-design time of 4-5 hours in total, Tracy participated in deciding which activities from her biology curriculum can be converted into computational activities.

In her interview, Tracy mentioned the following when she was asked about her co-design experience:

"... we met at least a couple of hours at least once a week...We started from the activities from [last] year working on the ones that were already made… he already asked me, so what do you do? What are the activities that you already do? … and then [he] talked about, well how can we turn these into the computational?"

[Post interview, March 2019]

Tracy mentioned her involvement in designing computational activities based on the activities that she used in the class. Tracy played a pivotal role in co-designing the activities with her partner in the Year 2 Curriculum.

Year 3. In the third year, Tracy played more active role in both designing the ESMs and the learning activities. Over the course of two weeks, Tracy and I discussed the designs of ESM computational models, as I coded the models. During that time, the ESM design platforms, NetLogo and NetTango Web, allowed Tracy to view the models as they were being built to give feedback for making it more pedagogically effective and imagine pedagogical activities using the model. I could quickly incorporate her design suggestions. We co-designed pedagogical activities and often discussed how to support students' learning using computational tools, which potentially impacted Tracy's teaching practices. A vignette in the next section demonstrates how Tracy used a pedagogical strategy based on her own experiences in the co-design process.

When asked about her learning through the co-design experience, Tracy mentioned:

"I learned because behind the code… I didn't understand what made the agents work the way they work. I wouldn't even know what an agent was… You have to tell program to… they're not moving naturally. And you have to tell them to do that. You have to tell them what the preferences and how did you figure that out? Well, you actually did the research, right? So we knew this is a model of like real behavior."

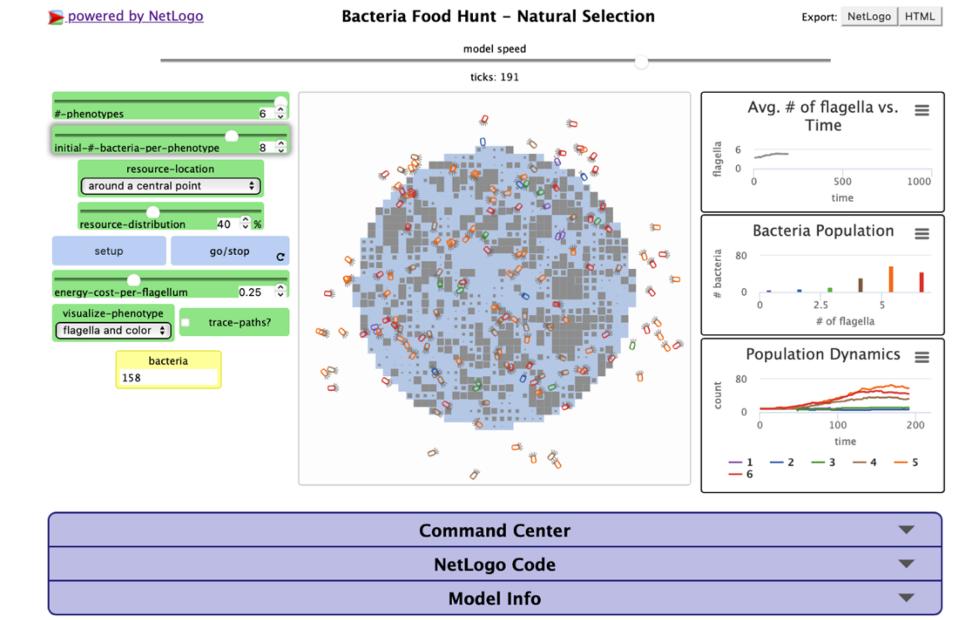
In this response, Tracy explains how learning about the agent-based code in an ESM helped her understand the relationship between computationally coded agent behavior and system-level behaviors observed in the real-world contexts. NetLogo makes the code easily accessible for the users. The NetLogo language is designed for users to easily understand how agent behaviors are coded (Wilensky, 2001). Despite this, in the first two years of her involvement, Tracy did not look at the code to understand agent behaviors. She mentions here in her response that she did not understand what made the agent work the way they work. This changed in the Year 3, when we co-designed the agent code and learning activities about coding agent behavior. For our co-designed unit in Year 3, Tracy and I developed the agent-based models in NetLogo and NetTango Web and computational activities for students that involved among other things coding agent behavior to match observed system-level behavior. Tracy's response indicates that she understood how behaviors of agents were coded and how the computational logic to code the behavior in a particular way was based on the real-world observations.

SHIFTS IN CURRICULAR DESIGNS

Over the three years, the curricular units became richer in terms of integration of CT-practices. Both the variety of integrated CT practices and the depth of engagement in CT practices improved over the years.

Year 1: A pre-designed ESM-based curriculum

The year 1 curriculum was embedded with three ESMs, which were about a wolf-moose prey-predator ecosystem, changes in a population of bacteria due to natural selection and genetic drift (Figure 8-1), and speciation in a plant population.



<u>Figure 8-1</u> A NetLogo model about natural selection in a bacterial population (modified from Novak & Wilensky, 2015)

The curricular activities in this unit were highly scaffolded. Students were asked to explore and learn about an ESM, asked to set up certain conditions, make predictions and engage in interpretations of computational visualizations of data to make sense of a phenomenon. Agent-based restructurations in these ESMs allowed learners to have visual and cognitive access to agent-level representations as well as system-level patterns. For example, learners could visualize behaviors of individual moose and wolves (adapted from Wilensky, 1997), as well as the population changes in the NetLogo model of prey-predator interactions and how random events related to the survival of individuals created predicted patterns at the system-level because of natural selection. Tracy appreciated how learning using these ESM contributed to a deeper understanding of the disciplinary content, she mentioned in her interview:

“I think this is the first time that kids have a real understanding that evolution can happen because of random forces, and because of natural selection.”

[Post interview, May 2018]

It is likely that Tracy's appreciation of the power of computational tools and learning activities helped her decide a topic for CT-integration for the next year and design computational learning activities. However, Tracy's view of learning of CT practices with these units were limited to how models can be used for "*predicting what would happen in an actual ecosystem*" and "*using computer models to also teach the concepts*". Her ideas of learning CT in biology context were limited to using computational models and simulations to understand a phenomenon, which is a part of one of the four CT practices, *modeling and simulation* (Weintrop et al., 2016).

Year 2: A moderately co-designed unit

The topic that Tracy chose for CT-integration in year 2 was Hardy-Weinberg equilibrium in population genetics and natural selection. Tracy had been using a research study about the population of rock pocket mice in the desert of New Mexico as an example of natural selection. Tracy's co-design partner, Author 1, designed an ESM consisting of a series of NetLogo computational models that allowed students to study both Hardy-Weinberg equilibrium as well as natural selection in a population of rock pocket mice (Figure 8-2) (Another version of this unit is discussed in detail in Chapter 5). Tracy and I designed curricular activities to engage students in *modeling and simulation* and *data* practices to learn about population dynamics and natural selection.

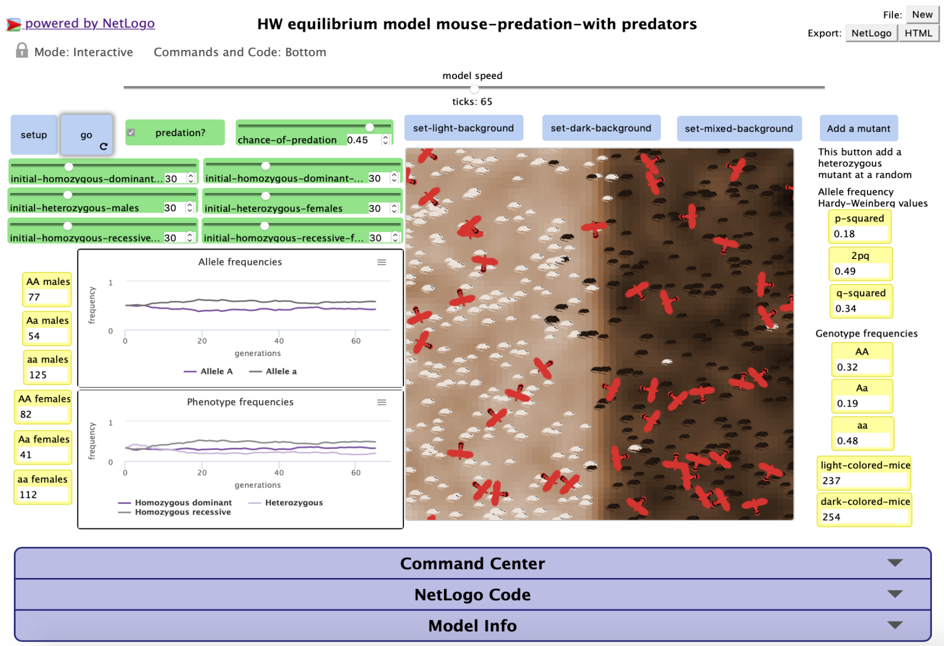


Figure 8-2 A NetLogo model for learning about Hardy-Weinberg equilibrium and natural selection (modified from Dabholkar & Wilensky, 2020)[21]

Tracy not only engaged deeply with co-designing the computational activities, but Tracy also appreciated their pedagogical effectiveness after she taught the unit.

"I mean it definitely helped them get a deeper understanding of the content. They got to see for themselves what would happen when the variables changed. They got to see the environments and the predators and all that in action. I just, I just feel like it made the equations more real to the students. Um, they meant something to the students rather than just this thing that they had to memorize for class."

[Post interview, March 2019]

Tracy mentioned how student engagement in *modeling and simulation* and *data* practices, helped them learn disciplinary ideas. A cognitive property, which is one of the five properties of

[21] NetLogo models that I designed and co-designed for CT-STEM curricula have many embedded versions. Based on the focus of the learning activities on a CT-STEM page/lesson, the most pedagogically appropriate version is embedded on the page/lesson. A canonical version of this Rock Pocket Mice Model is published in NetLogo Models Library.

the agent-based restructuration (Wilensky & Papert, 2010) (see chapter 1 for the details) increases the learnability of the content. Tracy mentioned that the agent-based restructurations, "*the predators and all that in action*", made the Hardy-Weinberg equations real for the students. Programming for behaviors of computational agents and investigating system-level emergent patterns is an important aspect of an ESM and of the *systems thinking practices* in the CT taxonomy. Tracy's appreciation of the effectiveness of agent-based restructurations in an ESM for conducting computational experiments and understanding system-level patterns is likely to have influenced her choices in designing a curriculum with more thorough computational data practices and agent-based coding activities in the next year.

Year 3: A fully co-designed unit

Tracy chose a topic for CT-integration in the third year and was fully involved in designing ESMs and ESM-based CT activities. The curricular unit was designed to teach experiment design in an Advanced Placement biology class that Tracy taught. It started with conducting a real-world experiment regarding the habitat preference of isopods and testing a hypothesis using inferential statistics (Figure 8-3). The co-designed ESM models rollypolly habitat preference behavior. The computational activities in the curricular unit are designed to learn more nuances of experimental design, such as sample size, multiple trials, etc., and specifically engage in Computational Thinking (CT) practices. In the Computational Lab lessons, CT activities were designed using a NetLogo (Wilensky, 1999b) model to engage in *modeling and simulation*, and *data practices*. To further learn about *data practices*, students conducted automated computational trials to generate large data sets to study the effects of sample size and the number of trials. To engage more deeply in *modeling and simulation* and *computational problem-solving* practices the later activities in the unit involved modifying the

NetLogo model to add new factors and incorporate student observations from the real-world experiments. Finally, students created a new model using block-based programming in the NetTango Web interface (Horn et al., 2020), in which they coded the behavior of agents to observe system-level changes, thus engaging in *systems thinking practices*.

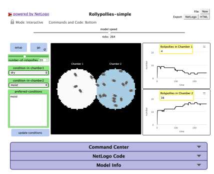
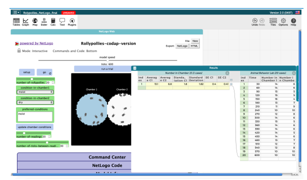
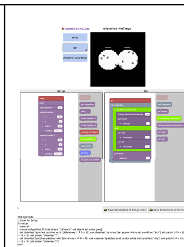
	 Explore	 Investigate and modify	 Construct
Physical Lab	Computational Lab		

Figure 8-3 Curricular flow of the co-designed unit in 2019 about Experimental Design to investigate habitat preference behavior

For students to learn nuances of experimental design, Tracy specifically designed activities for students to learn the importance of large sample size and multiple trials from the perspective of experimental design. Because students were using the ESM, they could easily conduct multiple experiments, each with different sample size and a different number of experimental trials. When asked about student learning with the computational tools, Tracy mentioned in her interview:

“Just, but I also had them use the model on purpose to show the point that the value of having large sample size. Remember we made them do that over and over again. Now add 10 rollypollies, now add 20 rollypollies. So, you know, they were learning about experimental design.”

[Post interview, September 2019]

In the unit, students conducted real-world as well as computational experiments with isopods, commonly known as rollypollies. Questions in the unit asked students to change the number of rollypollies in the model and the number of experimental trials while investigating their habitat preference behavior using two chambers with different environmental conditions.

Tracy also appreciated student learning of *systems thinking* by coding the behavior of animal agents and observing emergent patterns regarding their habitat preference. While talking about student learning of CT practices, Tracy said:

"So is that like systems thinking is like with the agents yeah. And the, the, what the behavior of the rollypollies and telling rollypollies how to behave in the chambers and yeah.."

[Post interview, September 2019]

In this response, Tracy explained how students engaged in *systems thinking practices* during the unit. In one of the activities with the NetTango blocks (Figure 8-4), students were asked to construct a computational model of rollypolly behavior. They were asked to code for the behavior of individual rollypollies and observe the emergent behavior of a population of rollypolllies regarding their preference in terms of a moist or a dry chamber. Difference in the probabilities of movement, when on a preferred-patch vs *not* on a preferred patch, result in an emergent pattern of more rollypollies being in a chamber with a preferred habitat over time.

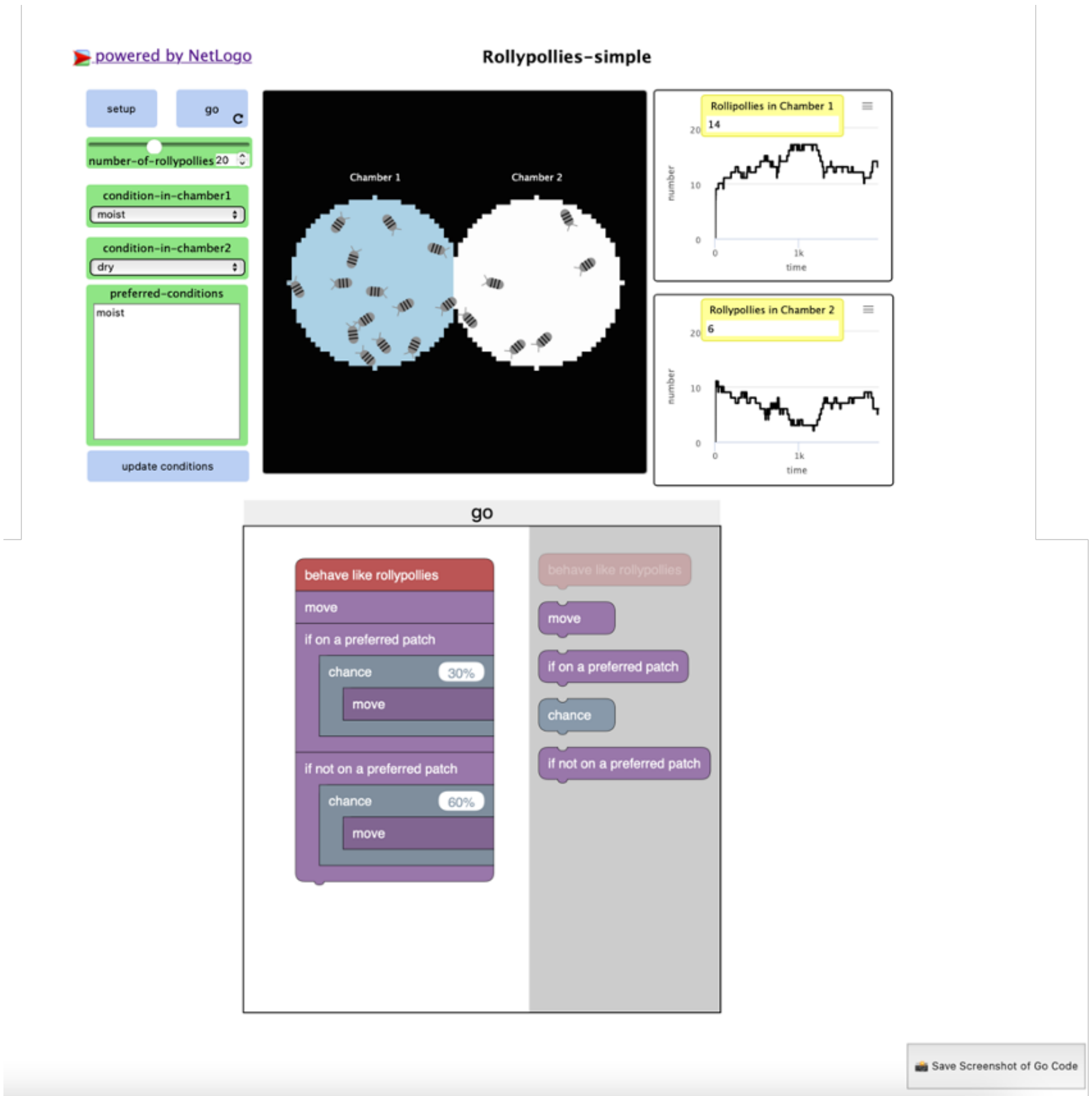


Figure 8-4 A block-based coding activity in NetTango Web which involved students coding habitat preference behavior of rollypollies and studying system-level emergent patterns

SHIFTS IN PEDAGOGICAL STRATEGIES

Over the three years, Tracy's pedagogical practices shifted from checking if students were on task and learning about disciplinary ideas, to scaffolding their computational experimentation so that they can engage meaningfully in *modeling and simulation* and *data* practices, to guiding them to figure out how to debug a computational program and meaningfully engaging in *computational problem solving* and *systems thinking* practices.

Year 1: Checking on student progress

While teaching this unit, Tracy mainly encouraged students to explore and learn from the agent-based models, but she did not provide any specific suggestions to guide their explorations. Her focus was to use computational models to teach the concepts. Her questions regarding student progress were non-specific and they rarely were about CT practices other than using a computational model to understand a phenomenon.

Tracy: *"*What are some things that you are noticing? Can you tell me what you added?*"*
Student: *[inaudible]*
Tracy: *"*And you kept everything else the same?*"*
Student points at her screen.
Tracy: Ok, ok. And what happens when you decrease the amount of grass?
Student says something about the carrying capacity.
Tracy: *"Yeah" [inaudible]*
Both the student and Tracy laugh and Tracy moves to another student.
[classroom video, May 2018]

In this conversation, Tracy asked a student about the changes that the student had made in the model and what she had noticed after making those changes. Tracy appeared to be satisfied after noticing that the student had used the model to learn about a disciplinary idea (carrying capacity) and moved to the next student.

Year 2: Making specific connections between computational learning activities and disciplinary ideas

In year 2, Tracy took a much more active role in guiding students' CT activities regarding *modeling and simulation*, and *data practices*. In her post-implementation interview, Tracy said that because the students were actively engaged in these practices, they had a "*much deeper* *understanding*" of the content. The questions that Tracy asked student groups as she moved around in the classroom were more specific than the previous year. She asked questions like, '*how many generations are you running your experiment for?*', '*what is the predation rate are you setting?*' [classroom video, February 2019]. Without giving direct answers, Tracy asked students to specify the experimental conditions so that they will develop insights into the change in the population due to natural selection. Tracy also conducted lectures during the unit, in which she discussed her regular slides and made direct connections between the disciplinary ideas mentioned on the slide and the computational learning activities student engaged in using the ESM. Discussing a slide about how different genotypes create variation in phenotypes in a population for natural selection to act on, Tracy asked students to explain how genetic differences and related phenotypic differences were modeled in the ESM. The students answered correctly, mentioning that the ESM had an 'either or' [dark fur color or light fur color] phenotype that was controlled by a single gene. [classroom video, February 2019]

Year 3: Facilitating student engagement in CT Practices

The agent-based coding blocks in NetTango Web allow an easy visualization of changes in agent behaviors based on the changes in their properties. In this vignette, Tracy and students are talking about two computational agents, *chambers,* that are used for an experimental setup (Figure 8-4). The dialogue between Tracy and a student transcribed from a video illustrates how Tracy used her prior experience of debugging during the co-design process to encourage a student to debug her model without giving a direct answer to her computational problem.

Student: *"In my chamber…. I am changing the sizes, but it won't get bigger"*
Tracy: *"This is how Sugat told me. He never told me answers either. He made me figure it out. When you are taking a math class and when you are making a graph, what does x mean?"*

Student: *"I mean, it's like….." (makes a horizontal movement with hands)* (see Figure 8-5)
Tracy: *"So what does Y mean?"*
Student: *Students "It's that". (Makes a vertical gesture)*
Tracy: *"So when you are changing the x and the y, are you changing the size of the chamber?"*

[classroom video, September 2019]

After Tracy asked this question, the student gestured positively indicating that she figured out the solution to her coding problem (Figure 8-5). The *bug* in the program was because the student was changing the value of *x* in the program and expecting that it would change the size of a chamber. The *x* and *y* are the coordinates that determine the position of a chamber. These variables do not affect the size of a chamber. Tracy identified the *bug*, when the student told her that she is changing the sizes and the chamber was not getting bigger. Even though the student thought that she was changing the size, she was actually changing the position of a chamber. Tracy helped the student debug her code and her understanding of what *x* and *y* meant in the context of the model by asking her questions about *x* in a math class and making graphs. The student indicated that she understood what Tracy was helping her get at, by making a horizonal-hand-movement gesture (Figure 8-5) for *x* and a vertical-hand-movement gesture for *y*. Then Tracy asked a next question to the student to help her to realize that changing values of *x* and *y* variables in the model would not change the size of a chamber.

Figure 8-5 Tracy discussing a debugging strategy with a student. The white arrow shows the horizontal motion gesture made by the student while answering Tracy's question.

What this vignette illustrates is how Tracy facilitated students' debugging of a computational model in the classroom. During the co-design process, Tracy had faced a similar issue when she had attempted constructing this model. She not only remembered the issue and how to debug it, but also, she used the same pedagogical approach that she had experienced during co-design. Tracy did not give an answer directly, instead she provided a hint to help the student debug the issue herself. Computational models designed as ESMs allow students to engage in computational thinking practices by quickly trying out, reasoning, and debugging code as they modify and built new models. Tracy's approach of encouraging students to engage in *computational problem solving practices* by asking relevant questions is an important pedagogical strategy for helping students bridge the science and the CT.

DISCUSSION

In this chapter, I demonstrate how agent-based representations and constructionist design features of computational models designed as ESMs make them effective to design CT-integrated science curricula. The analysis presented in this chapter illustrates how the ESM-based design approach mediated the co-design process and its outcomes, and the co-design approach reciprocally helped development of new ESMs, and increasingly enriched CT-integrated curricular units. Over three years, Tracy's involvement in curricular design changed from providing minor suggestions for modifications, to choosing a topic for CT-integration and co-designing curricular activities, to choosing a topic, co-designing models and curricular activities. Agent-based representations helped Tracy not only in supporting student learning, but also in thinking about curricular topics for CT-integration with ESMs and co-designing those ESMs and ESM-based curricula. Since ESMs are designed as microworlds, based on the constructionist design approach (Papert, 1980), they are easy to manipulate to engage in *computational data* and *problem solving* practices. Tracy and her co-design team used these constructionist design features to design activities for students to engage in CT practices in biological contexts. Over the three years, The CT-integrated curricula became richer in terms of the depth of CT-integration and adequate coverage of all four categories of CT practices included in the taxonomy (Weintrop et al., 2016). Starting from the focus on *modeling and simulation* and *data* practices in the first year, the curricula became more thorough in terms of integration of all the four CT practices, with a specific focus on *systems thinking* and *computational problem solving practices* in the third year.

These shifts in Tracy's participation in the co-design process and unit designs were mediated by the ESM design framework. In the first year, Tracy realized the power of agent-based restructurations (Wilensky & Papert, 2010) in ESMs as she saw that her students learn deeply about disciplinary ideas using the ESMs through constructionist investigations of the microworlds. In the second and third year, Tracy chose topics for CT-integration and co-designed even more advanced self-directed learning activities using an ESM. The learning activities in Year 2 and Year 3 Curricula prompted students to engage in *data* practices more deeply and investigate emergent patterns in the system by engaging in *systems thinking* practices.

ESMs are restructurated agent-based representations, which include computational visualization of agent-behaviors and the computational code that is used to encode those behaviors. Computational visualizations that allow students to see agent-level interactions and system-level emergent patterns are highly effective to understand an emergent phenomenon (eg, Levy & Wilensky, 2008; Wilensky & Resnick, 1999). Whereas changing, debugging, and creating the code for agent behavior is useful to learn computational problem solving and systems thinking. Tracy's involvement in designing both visualizations and code in the co-design process influenced her design of CT activities. Tracy's co-designed activities included students learning about robust experimental procedures to account for randomness. The agent behavior in an ESM can be specifically coded to have random variations, which still results in robust system-level patterns. Because of randomness in encoded agent behaviors, experiments need to be designed methodically to establish system-level patterns. This requires careful considerations regarding *data practices* such as sample size, multiple trials, etc. Tracy used this feature of ESM to design learning activities for students to learn deeply about *computational data practices* in the context of experimental design.

ESM enables easy manipulation of underlying code as well as changes to interface elements. Tracy and her co-design team used this feature to design an activity in Year 3 Curriculum that involved modify a computational model to add new factors based on the experiments that students conducted in the real-world. This activity involved student engaging in *computational practices.* Another activity in the same unit included student coding behaviors of rollypolly agents and visualizing emergent patterns regarding their habitat preference, thus engaging in *systems thinking practices*.

Tracy's pedagogical practices in the classroom also shifted as her participation in the ESM-mediated co-design process increased. Tracy's interactions with students became more specific in terms of facilitating their engagement in and learning from different CT practices. She made more explicit connections between the CT activities and disciplinary ideas. She facilitated student engagement in CT practices by asking questions to help them think through and arrive at solutions. The interaction analysis of Tracy and her student provides an illustration of how Tracy facilitated engagement in debugging practice of a student while constructing a computational model. Tracy herself had engaged in debugging a model during the co-design process. This example illustrates how Tracy's co-design experiences in the ESM-based co-design approach mediated shifts in her pedagogical practices.

Scholarly Significance

Even though the learning sciences research community is increasingly engaging in designing for integrating CT in disciplinary contexts (Lee et al., 2020), the research community still has a lot to learn about how to teach CT in science contexts, and even less is known about how to support teachers in designing and implementing this integration (Sands, Yadav, & Good, 2018). In this chapter, I described a novel co-design approach for creating CT-integrated science units. I discussed how an emergent systems microworlds (ESM)- based design approach supported CT-integration into biology curricula, and how the ESM design framework mediated a co-design partnership for such CT-integration. Visual and cognitive access to agent-level representations in an ESM facilitated deeper and meaningful participation of the teacher in the co-design process. Shifts in the teacher's participation resulted in the curricular units becoming richer in terms of the depth of the content and CT integration. Increased teacher participation in the co-design process also resulted in shifts in her pedagogical practices in terms of supporting student learning of CT in a biology context. This work demonstrates the effectiveness of our ESM-mediated co-design approach to increase teacher ownership and engagement in designing CT-integrated curricula and crafting effective pedagogical practices. As illustrated in this work, computational visualization plays an important role in engaging learners deeply with investigating and computationally representing a phenomenon. The current version of CT taxonomy (Weintrop et al., 2016) includes Computational Visualization in *Modeling and Simulation*, and *data practices.* The work discussed in this chapter supports the newly proposed version of CT Taxonomy of Practices which has Computational Visualization Practices as a separate category (Peel et al., 2021).

Chapter 9: Conclusions

Summary: In this chapter, I discuss the conclusions and implications of the three studies of my dissertation. The three studies are about investigating how restructuration properties (Wilensky, 2020; Wilensky & Papert, 2010) in Emergent Systems Microworld (ESM) design support student engagement in *doing science* by shaping and using scientific inquiry practices to learn about disciplinary ideas. In my dissertation, I define and study the ESM design approach for creating learning environments, which uniquely combines principles of agent-based modeling and constructionism. The contributions of my dissertation work are as follows:

1. The first main contribution of my dissertation work is a design contribution in the form of the three ESMs and ESM-based curricula (Chapters 3, 5, and 7).
2. The theoretical contributions of my dissertation build on restructuration theory and explain how restructuration properties agent-based restructurations in a constructionist curriculum support student learning and facilitate the co-design partnership with a teacher.
    a. In the first study, I identified and analyzed design features of a restructurated curriculum about genetics and evolution that created *epistemically expansive learning* opportunities for students. These design features enabled students to collectively shape their inquiry practices to construct knowledge about emergent aspects related to gene regulation and evolution.
    b. The second study demonstrated how an ESM-based curriculum supported students' learning of scientific inquiry practices and disciplinary ideas in a high school biology classroom. The alignment between students' *epistemic games* using an ESM and the

desired epistemic form (aligned with NGSS recommended practices) of the curriculum facilitated the progression of students' *epistemic connections* among practices and ideas through iterative refinement.

c. In the third study, I analyzed how cognitive and social properties of restructuration facilitated a co-design partnership with a teacher that used the ESM approach to create science curricula integrated with Computational Thinking (CT) activities. The use of the ESM approach for the co-design work increased teacher involvement in the design process and richness of CT-integration in the co-designed curricula and shifted in her pedagogical practices to foster student learning of CT practices.

My dissertation work focused on designing technology-enhanced science curricula to support students' epistemically agentic participation in a science classroom and analyzing design features of those curricula. In this dissertation, I have presented three design-based research studies that I conducted to investigate the effectiveness of a specific approach to designing computational learning environments called Emergent Systems Microworlds (ESMs). These studies investigated how restructuration properties (Wilensky & Papert, 2010) instantiated through different design features of an ESM mediated student learning and teacher participation in co-designing curricula. I conducted these studies to investigate the following research questions:

1. How do design features of an ESM support epistemically expansive learning in a science classroom?

2. How does the design of an ESM-based curricular unit support student connection-making among scientific inquiry practices and disciplinary ideas?
3. How does restructuration through ESM facilitate the co-design process for CT-integration into science units and its outcomes?

In this chapter, I first discuss the overall conclusions of my dissertation work. Then, I explain the conclusions and implications of each of the studies. I discuss a thread across these three studies and how my dissertation contributes to designing learning environments and developing theoretical insights into how specific design features of those learning environments facilitate student learning.

The ESM design work in my dissertation builds on the extensive earlier work by Wilensky and colleagues regarding agent-based computational models for engaging students in learning about complex emergent phenomena (Paulo Blikstein & Wilensky, 2010; Levy & Wilensky, 2009; Sengupta & Wilensky, 2009; Stieff & Wilensky, 2003; Wagh & Wilensky, 2018; Wilensky, 2003; Wilensky & Novak, 2010; M. Wilkerson-Jerde & Wilensky, 2010). I investigate how computational tools designed as ESMs facilitate students' participation in *doing science* to learn about complex emergent phenomena in science classrooms. In my dissertation, I demonstrate how ESMs serve as such experimental model systems which allow students to express, investigate, validate, and share their ideas to collectively construct disciplinary knowledge. Agent-based representations in ESMs serve as restructurations (Wilensky, 2020; Wilensky & Papert, 2010) which have cognitive, social, affective, and diversity properties that increase the learnability of the phenomenon represented using restructurated representations (See Chapter 1 and 2 for the theory of restructurations and restructuration properties). My dissertation work demonstrates that agent-based restructurations make ESMs pedagogically effective in supporting the learning of several complex natural phenomena related to fundamental ideas in modern biology such as gene regulation, natural selection, and habitat preference behavior (See Chapters 4, 6, and 8).

My work is also strongly rooted in the theory of Constructionism (Papert, 1980; Papert & Harel, 1991). As a design theory, constructionism provides design principles to harnesses the power of computational technologies for engaging students in individual and collective meaning-making (Kynigos, 2015; Papert, 1980). As a learning theory, constructionism contributes to the theory of constructivism through its unique attention to explaining how student engagement in the creation of computationally supported artifacts with an explicit emphasis on self-driven production and ownership support their learning (Ackermann, 2001; Kynigos, 2015; Papert, 1980; Piaget, 1970). The ESM design framework incorporates the following three key ideas from the constructionist design framework: (a) personally meaningful engagement, (b) construction of public entities, (c) expression and validation of ideas through computational microworlds. I conducted three studies to investigate how agent-based restructurations and constructionist design principles in ESM-based curricula facilitate student learning.

The first and second studies of my dissertation focus on student learning. ESM-based curricula in these studies enabled students to take agentic roles in designing and performing computational experiments to express and validate their claims regarding aspects of a modeled phenomenon in an ESM that they found meaningful to investigate. Such constructionist engagement in arriving at and investigating their own research questions made these learning experiences personally meaningful for the students. In the first study, the research projects were sharable public entities for students, which was an important feature of the constructionist curriculum. As students iteratively engaged in conducting and sharing such research projects, they collectively constructed knowledge about the emergent phenomena. Specific design features of the ESMs facilitated the shaping of research practices of the classroom learning community. I explain these features later in the chapter when I discuss the conclusions of Study 1.

Power and learnability properties of agent-based restructurations make the ESM approach effective in designing and using such learning environments for students to learn about emergent complex phenomena. Power properties of agent-based restructurations make it feasible for designers and educators to develop ESM that models emergent phenomena authentically. These power properties are demonstrated in Chapters 3, 5, and 7, which discuss examples of restructurations of phenomena related to gene regulation, natural selection, and habitat preference behavior. Cognitive properties of agent-based restructurations in an ESM allow students to engage in scientific practices to investigate emergent properties of a system that are otherwise difficult to learn. Social properties of restructurations allow students to share their research designs and finding easily, thus facilitating sharing of ideas and collective knowledge construction. Affective properties of restructurations make student participation in such learning activities *fun* and *playful*. Later in this chapter, I present findings of Study 1 and 2 about how cognitive, social, and affective properties of restructurations in ESMs supported student learning.

The third study of my dissertation was about a co-design partnership with a teacher to create ESM-based curricula. This study demonstrated that the ESM design approach supported teacher engagement in co-designing curricula that were aimed for students to engage in specific epistemic practices. ESM approach was useful for integrating Computational Thinking (CT) practices related to data, modeling and simulation, systems thinking, computational visualization, programming, and algorithm in a science curriculum. This study also demonstrated how teacher contributions in the co-design process and richness of co-designed curricula were enhanced because of the cognitive and social properties of restructurations. The findings of this study underscore the effectiveness of using the ESM approach for co-designing CT-integrated science curricula.

OVERVIEW OF THREE STUDIES

In Study 1, using Engeström's theoretical framework of expansive learning (Engeström, 2001), I conceptualized students' *epistemically expansive learning* and *epistemic expansion* of a classroom activity system, which was about positioning students in epistemically agentic roles in the process of knowledge construction. I analyzed how cognitive, affective, and social properties of restructurations (Wilensky, 2020; Wilensky & Papert, 2010) instantiated through design features of an ESM mediated an epistemic expansion of a classroom activity system. Study 2 built on the findings of Study 1. In Study 2, I investigated student learning in a more scaffolded ESM-based curriculum focusing on their epistemic activities using Collins and Ferguson's theoretical framework of epistemic forms and epistemic games (Collins & Ferguson, 1993). The ESM-based curriculum in Study 2 was designed for students to participate in epistemic games to generate a specific epistemic form. This pedagogical epistemic form was aligned with NGSS recommended science practices. In this study, I conceptualized an idea of *epistemic connections* for characterizing student engagement in scientific inquiry practices to investigate and learn about disciplinary ideas. When students used a practice to learn about a disciplinary idea, I considered it as an epistemic connection between the practice and the idea. Typically, in an

ESM-based learning activity in the curriculum, students used a few practices to investigate specific disciplinary ideas, thus making epistemic connections among those practices and ideas. Using Epistemic Network Analysis (ENA) (Shaffer et al., 2009), I studied how students' epistemic connections changed as they participated in *epistemic games* using an ESM as an experimental model system. These epistemic games were regarding devising and using specific strategies for constructing knowledge about the modeled phenomenon in the ESM. The findings of this study contribute to understanding how an ESM can be used to design a curriculum for students to participate in specific epistemic games to generate desired *epistemic forms*. Study 3 of this study takes this thread of designing ESM-based curricula even further. In this study, I analyzed a co-design partnership between a teacher and me that created ESM-based curricula to engage students in specific practices. This study was focused on understanding the use of the ESM design approach for co-designing science curricula that are integrated with Computational Thinking (CT) practices. In this longitudinal study of three years, I investigated how cognitive and social properties of restructurations were at play for (a) increasing richness of CT-integration, (b) supporting participation of a teacher in the co-design process, and (c) mediating shifts in her pedagogical practices in the classroom.

CONCLUSIONS OF STUDY 1

In Study 1, I investigated student learning an ESM-based curriculum, GenEvo, to answer the following research question: '*How do design features of an ESM support epistemically expansive learning in a science classroom?'* This study focused on how cognitive, social, and affective properties of restructuration in an ESM-based curriculum mediated students' epistemically expansive learning in a science classroom. To shift student roles from *receivers of* *facts* to *doers of sciences,* it is important to design learning environments that consider and attempt the problem of practice (Miller et al., 2018; Russ & Berland, 2019). In this Study, I conceptualized the *epistemic expansiveness* of a learning environment is in terms of providing students opportunities and ways to shape practices to investigate a phenomenon. Using micro-ethnographic analysis guided by Cultural Historical Activity Theory (CHAT) (Engeström, 2001), I identified design features of the ESM-based curriculum that mediated the transformation of a classroom activity system in terms of its shift towards the social construction of disciplinary knowledge. These features embodied agent-based restructurations and constructionist principles. I analyzed how cognitive, social, and affective properties of restructuration instantiated through these features mediated the epistemic expansion. In many traditional biology classrooms, when students are positioned as *receivers of facts*, the intended and enacted object for students is to listen to disciplinary ideas explained by a teacher using static representations. The ESM-based curriculum mediated a transformation of this activity system to position students as *doers of science* by engaging them in the social construction of disciplinary knowledge and evolving epistemic practices using an interactive experimental system that included agent-based restructurations.

The findings of this study show how three features of the ESM, (1) randomness in agent behaviors, (2) representational features of computational objects, and (3) visualization across levels, mediated students' engagement in intermediate learning goals (objects of the activity system) as they progressed with constructing knowledge of disciplinary ideas and evolving scientific practices. The three *objects* (goals) of intermediate activity systems in the ESM-based learning environment were - (1) careful evaluation of evidence to establish a pattern, (2) development of shared vocabulary to construct knowledge, and (3) reasoning about complex emergent patterns. Student enactment towards achieving these objects, in turn, mediated the transformation of the classroom activity system through epistemic expansion. The object of the classroom activity system became the social construction of knowledge about complex disciplinary ideas.

The shifts in the other components of the activity system – rules, community, and division of labor, were also salient in the ESM-mediated epistemic expansion discussed in this paper. Students and the teacher were part of the classroom community. Different features of the ESM-based learning environment and ESM-enabled pedagogical moves shaped interactions of the community as different rules and divisions of labor evolved. For example, in the vignette discussed in Chapter 4, inherent randomness in agent behaviors required students to conduct multiple trials, collect and analyze data systematically to establish a pattern. In the classroom community, student groups performed investigations separately – dividing the labor, and then collectively built the knowledge of disciplinary ideas. The ESM-enabled pedagogical moves of the teacher, such as supporting the use of emerged colloquial words for identifying and addressing computational objects (such pink triangles or potato-shaped proteins) and encouraging sharing and evaluation of evidence facilitated students' active participation in shaping practices of knowledge construction and sharing.

The cognitive properties of agent-based restructuration in an ESM reduced perceptual limitations to reason about an emergent phenomenon that arise because of "level-slippage" (Wilensky & Resnick, 1999). By providing visual access to agent behaviors and emergent patterns, agent-based restructurations allow students to investigate how interactions at the 'agent-level' result in emergent patterns at the 'system-level' (Goldstone & Wilensky, 2008; Wilensky, 2003; Wilensky & Reisman, 2006). In the GenEvo curriculum, learners manipulated agent-level behaviors of proteins and parts of DNA to investigate the effects of those manipulations at the system-level regarding gene regulation, which allowed learners to overcome the confusion of levels (Wilensky & Resnick, 1999). Using the agent-based models in the ESM students could easily ask and investigate their questions by changing agent behaviors and investigating the effects of those changes. Additionally, as a classroom community of learners, students collectively shaped practices to establish their findings as valid knowledge products. Vidya, Samir, and others in the classroom learned about the fundamental aspects of gene regulation by starting with simple experiments and observations using the ESM and then evolving more sophisticated epistemic practices. A vignette in Chapter 4 illustrates how certain rules evolved in the community regarding what counts as evidence (one trial vs multiple trials) as well as how to present and evaluate the evidence. This shift in terms of establishing rules regarding evidence-gathering practices took place without the teacher directly telling students how they should collect and evaluate evidence. As witnessed in their interview responses and the analysis of students' classroom participation, they displayed positive engagement in the social construction of knowledge using the GenEvo ESM. The playful nature of agent-based restructurations and ease of sharing, testing, and incorporating each other's ideas were important aspects of social and affective properties of ESM that supported student learning.

Implications of Study 1

To support students' epistemic agency in a science classroom, it is important to design an environment that can provide them with opportunities to engage in epistemically expansive learning. Such epistemically expansive learning would allow students to shape and evolve practices that they engage in to construct knowledge about a phenomenon. The ESM-based GenEvo learning environment facilitated students' engagement in epistemically expansive learning experiences by socially constructing knowledge of disciplinary ideas about gene regulation and evolution. The evidence presented in Chapter 4 demonstrates that the students not only learned disciplinary ideas about genetics and evolution but also perceived shifts in their epistemic agency in the classroom. This study contributes to the theory of restructuration by explaining the effectiveness of cognitive, social, and affective properties of agent-based restructurations instantiated through specific design features of the ESM. The analysis presented in Chapter 4 highlights how these design features acted as mediational tools to support students' epistemic expansive learning. The identified design features and the complementing pedagogical practices serve as guiding principles to design for and support students' epistemically expansive learning.

CONCLUSIONS OF STUDY 2

While in Study 1, I was a lead designer and a lead teacher of an ESM-based curriculum taught in extra-school programs. In Study 2, a partnering teacher taught an ESM-based curriculum in her regular high school biology classroom. In this study, I investigated how an ESM-based curriculum supported students' epistemic connection-making among science practices and disciplinary ideas as they constructed knowledge about an emergent phenomenon. The research question that guided this study was: *How does the design of an ESM-based curricular unit support student connection-making among scientific inquiry practices and disciplinary ideas?*

The ESM-based curriculum used in this study was designed in a more structured way as compared to the curriculum in Study 1 to specifically engage students in scientific inquiry practices that are recommended by Next Generation Science Standards (NGSS Lead States, 2013). The curricular activities were structured for students to engage in specific practices to learn about specific aspects of disciplinary ideas related to the evolution of a population due to natural selection. The results regarding student learning at an aggregate level suggest that the ESM-based curriculum supported students' engagement in making these connections between practices and ideas, which I call *epistemic connections*, in a sequential and integrated manner. The macro-level analysis illustrates a temporal progression of practices as students moved from the practice of *asking questions* to the practice of *constructing explanations* about emergent patterns, such as changes in a population because of the introduction of a mutant phenotype. In addition, our analysis suggests that there was a progression of disciplinary ideas from genotypic and phenotypic properties of individuals to heredity and change in a population. As new ideas and practices became prominent nodes in student networks, they also contained and were connected to earlier nodes, which demonstrates that students learned these ideas in an integrated manner.

NGSS also recommends designing learning environments that provide meaningful and authentic learning opportunities for students. However, such engagement in authentic scientific inquiry in classroom contexts is instructionally challenging (Chinn & Malhotra, 2002), mainly because designing for such authentic learning experiences requires creating an experimental system that students can use to construct and validate explanations regarding a disciplinary phenomenon by engaging in science practices. This also requires making experimental systems and anchoring phenomena cognitively accessible to students. To design for cognitive accessibility of the experimental system of rock-pocket-mice, I incorporated agent-based restructurations in the ESM design. Because of their cognitive properties, agent-based restructurations make it easy for a user to observe, manipulate, and interpret the properties and behaviors of agents and reason about the emergence of system-level aggregate patterns (Wilensky & Papert, 2010). These cognitive affordances of agent-based restructurations in the ESM supported students in engaging in science practices and learning deeper aspects of disciplinary ideas related to the emergent properties of the system under investigation.

The micro-level analysis revealed that a student iteratively refined his research questions three times in the curriculum. This refinement was mediated by his cognitive engagement with the agents (mice population), their properties and behaviors, and their surroundings. His questions became more specific and included more aspects of DCIs related to natural selection over the lessons. His final question considered two very critical aspects of natural selection, heritability and environment, that he could investigate using the ESM. Similarly, the interactive constructionist ESM provided an experimental system for another student to test her predictions regarding the effects of agent properties (phenotypes) and interactions (reproduction, Mendelian inheritance) on emergent patterns by performing mini-experiments. Analysis of her responses demonstrated how the agent-based restructurations in the ESM allowed her to visualize and interpret the experimental results. This led to the correction of her understanding of Mendelian principles of inheritance. Analysis of responses of a third student revealed that she and her group engaged systematically in scientific inquiry practices in the final lesson, from asking a question to planning investigations to using a model to analyzing data and constructing explanations.

These cases illustrate how the students of the ESM-based curriculum participated in science inquiry practices and learned about various aspects of disciplinary ideas related to natural selection in a scaffolded yet self-driven manner.

These student case studies illustrate how agent-based restructurations provided cognitive ease for students in making *epistemic connections* between practices and ideas. In order for students to engage meaningfully in epistemic activities, there should be alignment between an epistemic form that a learning environment is designed for and the epistemic games that students can play using the learning environment (Wilkerson et al., 2018). The first two student cases illustrate how the alignment between epistemic forms and games created learning opportunities for the students to refine and connect their practices and disciplinary ideas. For example, the first student, Alejandro, refined his question twice, first after he explored the ESM as an experimental system and then after he practiced data collection and analysis with the ESM. As he refined his questions, he operationalized his initial curiosity to ask more specific questions addressing the disciplinary ideas. This created an alignment of his epistemic game with the epistemic form of the curriculum because he could meaningfully investigate those aspects by engaging in scientific inquiry practices. The second student, Jane, refined her disciplinary ideas regarding heredity after attempting to construct explanations based on the data that she collected using the ESM. The third student, Emma, and her group’s epistemic game was about manipulating a system by adding agents with specific properties and investigating the effects of those manipulations. Their experimental investigation focused on change in a mice population in a particular environmental setting.

The results of this study demonstrated how the ESM-based curriculum aligned student epistemic games with the epistemic forms and supported student learning of practices and ideas in a sequential yet integrated manner through iterative refinement of the practices and deeper understanding of various aspects of disciplinary ideas. The ability to engage in more refined and sophisticated aspects of a practice is an important part of the epistemic game that the students were scaffolded to participate in. Alejandro's refinement of a practice, Jane's refinement of disciplinary aspects, and Emma's engagement in practices and ideas in the final lesson all demonstrate how the curriculum leverages disciplinary context and practices in a reciprocal manner to support the learning of each other.

IMPLICATIONS OF STUDY 2

The operationalization of the notions of epistemic forms and games (Collins & Ferguson, 1993) to design for and assess learning of practices and disciplinary ideas provides a structure for developing NGSS-aligned curricula for engaging students in authentic science practices. In this study, I demonstrated how an ESM-based curriculum effectively created learning opportunities for students to engage them in creating an epistemic form aligned with NGSS recommended practices. This form of the canonical science practice is much more complex than the form of tic-tac-toe. It requires sequential and iterative engagement in the practices in the context of the epistemic game that a student chooses to play. In an ESM-based curriculum, students can engage in different variations of an epistemic game to generate the target epistemic form. This way, they engage in knowledge construction of different aspects of disciplinary ideas modeled in the ESM. As witnessed in Emma's case, the variations of the epistemic game are possible because of the structure of an ESM.

The ESM-based curriculum design approach provides ways to iteratively engage students in learning and refining practices and disciplinary ideas. The epistemic form of an ESM-based curriculum and ways to scaffold possible epistemic games provide a framework for increasing alignment between the games that students participate in and the form that the curriculum intends them to generate. Computational models designed as ESMs have agent-based representations that create cognitive ease for students to engage in an epistemic game that uses practices to construct disciplinary ideas. The work also provides guidelines for designing ESM-based curricula that align students' epistemic games with the epistemic form that the curriculum is designed for.

The ESM-based curriculum used in Study 2 is based on an epistemic form aligned with NGSS-recommended science practices. However, it is important to acknowledge that engagement in these practices is only one way to construct knowledge about the world. As I noted in Chapter 4, it is important to consider pluralistic orientations of students' and practitioners' from epistemological perspectives for embracing diversity regarding the ways of knowing (M Bang & Medin, 2010; Warren et al., 2020) and learning with computational learning environments (Turkle & Papert, 1990). In this study, I investigate learning of NGSS recommended practices. These practices are aligned with Western modern scientific ways. In this work, I do not explicitly or implicitly consider these Western modern scientific ways of knowing better than the other ways. Nevertheless, I want to emphasize the importance of learning the practice of constructing and validating knowledge as specified by NGSS by defining a set of Science and Engineering Practices as simply *a* practice of knowledge construction, not *the* practice of knowledge construction.

Expanding the goal of science education from *knowing about science* to *practicing science* requires the creation of learning opportunities that effectively support students in doing both. Designing ESMs and ESM-based curricula for different disciplinary contexts is a way of achieving this goal. Analyzing students' epistemic connections can be an effective way to assess student learning and the effectiveness of restructured curricula that support student learning of scientific practices in disciplinary contexts.

Conclusions of Study 3

In this study, I analyzed a co-design partnership to answer the following research question: *How does restructuration through ESM facilitate the co-design process for integrating Computational Thinking (CT) into science units and its outcomes?* Integrating CT into the K-12 science curriculum can be an effective way to engage students in learning contemporary science practices. However, this requires designing effective CT-integrated curricula and supporting teachers in adopting new teaching practices. Based on the findings of this study, I argue for involving teachers as partners for co-designing such a curriculum to increase their ownership and engagement in crafting effective pedagogical practices. I present a qualitative analysis of the involvement of a science teacher in the co-design process who participated in a Design-Based Implementation Research project aimed at creating CT-integrated curricular materials. I discuss how the underlying design approach of using ESM and ESM-based activities mediated an increase in teacher involvement, the outcomes of teacher involvement in the form of co-designed units, and changes in the classroom practice. Findings yield implications for how best to support teachers in curricular CT-integration.

In Study 3, I demonstrate how agent-based representations and constructionist design features of computational models designed as ESMs make them effective to design CT-integrated science curricula. The analysis presented in this study illustrates how the ESM design approach mediated the co-design process and its outcomes, and the co-design approach reciprocally helped the development of new ESMs and increased the richness of newly designed CT-integrated curricular units. Over three years, the involvement of a partnering teacher changed from providing suggestions for minor text-related changes in a lesson (Year 1) to choosing a topic for CT-integration and co-designing curricular activities (Year 2), to choosing a topic, co-designing models and curricular activities (Year 3). Interview responses of the teacher demonstrate how agent-based representations in ESMs helped the teacher to appreciate the power of such representations to engage in learning disciplinary ideas deeply. This is likely to have helped her in identifying curricular topics for CT-integration with ESMs and co-designing those ESMs and ESM-based curricula. Since ESMs are designed as microworlds, based on the constructionist design approach (Papert, 1980), they are easy to manipulate to engage in *computational data* and *problem solving* practices. These constructionist design features helped the co-design team to create activities for students that were meaningful to engage in CT practices in biological contexts. Over the three years, The CT-integrated curricula became richer in terms of the depth of CT-integration and adequate coverage of all different categories of CT practices included in the taxonomy (Weintrop et al., 2016). Starting from the focus on *modeling and simulation* and *data* practices in the first year, the curricula became more thorough in terms of integration of all the four CT practices, with a specific focus on *systems thinking* and *computational problem-solving practices* in the third year.

These shifts in teacher participation in the co-design process and unit designs were mediated by the ESM design framework. In the first year, the partnering teacher, Tracy, realized the power of agent-based restructurations (Wilensky & Papert, 2010) in ESMs as she saw that her students learn deeply about disciplinary ideas using the ESMs through constructionist investigations of the microworlds. In the second and third years, Tracy chose topics for CT-integration and co-designed even more advanced self-directed learning activities using an ESM. The learning activities in Year 2 and Year 3 Curricula prompted students to engage in *data* practices more deeply and investigate emergent patterns in the system by engaging in *systems thinking* practices.

ESMs are restructurated agent-based representations, which include computational visualization of agent behaviors and the computational code that is used to encode those behaviors. Computational visualizations that allow students to see agent-level interactions and system-level emergent patterns are highly effective to understand an emergent phenomenon (eg, Levy & Wilensky, 2008; Wilensky & Resnick, 1999). Whereas changing, debugging, and creating the code for agent behavior is useful to learn computational problem solving and systems thinking. Tracy's involvement in designing both visualizations and code in the co-design process influenced her design of CT activities. Tracy's co-designed activities included students learning about robust experimental procedures to account for randomness. The agent behavior in an ESM can be specifically coded to have random variations, which still results in robust system-level patterns. Because of randomness in encoded agent behaviors, experiments need to be designed methodically to establish system-level patterns. This requires careful considerations regarding *data practices* such as sample size, multiple trials, etc. Tracy used this feature of ESM to design learning activities for students to learn deeply about *computational data practices* in the context of experimental design.

ESM enables easy manipulation of underlying code as well as changes to interface elements. Tracy and her co-design team used this feature to design an activity in Year 3 Curriculum that involved modifying a computational model to add new factors based on the experiments that students conducted in the real world. This activity involved students engaging in *computational problem-solving practices.* Another activity in the same unit included student coding behaviors of rollypolly agents and visualizing emergent patterns regarding their habitat preference, thus engaging in *systems thinking practices*.

Tracy's pedagogical practices in the classroom also shifted as her participation in the ESM-mediated co-design process increased. Tracy's interactions with students became more specific in terms of facilitating their engagement in and learning from different CT practices. She made more explicit connections between the CT activities and disciplinary ideas. She facilitated student engagement in CT practices by asking questions to help them think through and arrive at solutions. The interaction analysis of Tracy and her student demonstrates how Tracy facilitated engagement in debugging practice of a student while constructing a computational model. Tracy herself had engaged in debugging a model during the co-design process. This example illustrates how Tracy's co-design experiences in the ESM-based co-design approach mediated shifts in her pedagogical practices.

Implications of Study 3

Even though the learning sciences research community is increasingly engaging in designing for integrating CT in disciplinary contexts (Lee et al., 2020), the research community still has a lot to learn about how to teach CT in science contexts, and even less is known about how to support teachers in designing and implementing this integration (Sands, Yadav, & Good, 2018). In this study, I used a novel co-design approach that used ESMs for creating CT-integrated science units. ESM design approach supported CT-integration into biology curricula and mediated a co-design partnership for such CT-integration. Visual and cognitive access to agent-level representations in an ESM facilitated deeper and meaningful participation of the teacher in the co-design process. Shifts in the teacher's participation resulted in the curricular units becoming richer in terms of the depth of the content and CT integration. Increased teacher participation in the co-design process also resulted in shifts in her pedagogical practices in terms of supporting student learning of CT in a biology context. This study demonstrates the effectiveness of our ESM-mediated co-design approach to increase teacher ownership and engagement in designing CT-integrated curricula and crafting effective pedagogical practices. As illustrated in this study, computational visualization plays an important role in engaging learners deeply with investigating and computationally representing a phenomenon. The current version of CT taxonomy (Weintrop et al., 2016) includes Computational Visualization in *Modeling and Simulation*, and *data practices.* The work discussed in this chapter supports the newly proposed version of CT Taxonomy of Practices which has Computational Visualization Practices as a separate category (Peel et al., 2021).

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

Ackermann, E. (2001). Piaget’s constructivism, Papert’s constructionism: What’s the difference. *Future of Learning Group Publication*, *5*(3), 438.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). How genetic switches work. In *Molecular Biology of the Cell. 4th edition*. Garland Publishing, Incorporated.

Arastoopour, G., Chandra, S., Dabholkar, S., Horn, M., & Wilensky, U. (2019). Classifying Emergent Student Learning in a High School Computational Chemistry Unit. *American Education Research Association (AERA) Conference. Toronto, CA*.

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, 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, G., Shaffer, D. W., Swiecki, Z., Ruis, A. R., & Chesler, N. C. (2016). Teaching and assessing engineering design thinking with virtual internships and epistemic network analysis. *International Journal of Engineering Education*, *32*(3), 1492–1501.

Aslan, U., Dabholkar, S., & Wilensky, U. (2018). Developing multi-agent-based thought experiments: A case study on the evolution of gamete dimorphism. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 10798 LNAI*. https://doi.org/10.1007/978-3-319-91587-6_4

Aslan, U., LaGrassa, N., Horn, M., & Wilensky, U. (2020). Phenomenological programming: a novel approach to designing domain specific programming environments for science learning. *Proceedings of the Interaction Design and Children Conference*, 299–310.

Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution*, *41*(2), 203–226.

Axtell, R. L., & Epstein, J. M. (1994). Agent-Based Modeling: Understanding Our Creations. *The Bulletin of the Santa Fe Institute*.

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

Bandini, S., Manzoni, S., & Vizzari, G. (2009). Agent based modeling and simulation: an informatics perspective. *Journal of Artificial Societies and Social Simulation*, *12*(4), 4.

Bang, M, & Medin, D. (2010). Cultural processes in science education: Supporting the navigation of multiple epistemologies. *Science Education*, *94*(6), 1008–1026.

Bang, M, Medin, D., Washinawatok, K., & Chapman, S. (2010). Innovations in culturally based science education through partnerships and community. In *New science of learning* (pp. 569–592). Springer.

Bang, Megan, Warren, B., Rosebery, A. S., & Medin, D. (2012). Desettling Expectations in Science Education. *Human Development*, *55*(5–6), 302–318. https://doi.org/10.1159/000345322

Bar-Yam, Y., McKay, S. R., & Christian, W. (1998). Dynamics of complex systems (Studies in nonlinearity). *Computers in Physics*, *12*(4), 335–336.

Bar-Yam, Y. (2004). A mathematical theory of strong emergence using multiscale variety. *Complexity*, *9*(6), 15–24.

Basu, S., Biswas, G., Sengupta, P., Dickes, A., Kinnebrew, J. S., & Clark, D. (2016). Identifying middle school students' challenges in computational thinking-based science learning. *Research and Practice in Technology Enhanced Learning*, *11*(1), 13. https://doi.org/10.1186/s41039-016-0036-2

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

Billari, F. C., & Prskawetz, A. (2012). *Agent-based computational demography: Using simulation to improve our understanding of demographic behaviour*. Springer Science &

Business Media. https://doi.org/10.1007/978-3-7908-2715-6

Blikstein, P., & Wilensky, U. (2005). Less is more: Agent-based simulation as a powerful learning tool in materials science. *IV International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2005)*. http://ccl.sesp.northwestern.edu/2005/2005-blikstein-wilensky.pdf

Blikstein, P., & Wilensky, U. (2007). Bifocal modeling: a framework for combining computer modeling, robotics and real-world sensing. *Annual Meeting of the American Educational Research Association (AERA 2007), Chicago, USA*.

Blikstein, Paulo, & Wilensky, U. (2010). MaterialSim: A Constructionist Agent-Based Modeling Approach to Engineering Education. In M. J. Jacobson & P. Reimann (Eds.), *Designs for Learning Environments of the Future* (pp. 17–60). Springer US. https://doi.org/10.1007/978-0-387-88279-6_2

Blomkamp, E. (2018). The promise of co-design for public policy. *Australian Journal of Public Administration*, *77*(4), 729–743.

Brady, C., Holbert, N., Soylu, F., Novak, M., & Wilensky, U. (2015). Sandboxes for Model-Based Inquiry. *Journal of Science Education and Technology*, *24*, 265–286. https://doi.org/10.1007/s10956-014-9506-8

Brandes, A., & Wilensky, U. (1991). Treasureworld: A computer environment for the study and exploration of feedback. *Constructionism*, 102–124. http://ccl.northwestern.edu/papers/treasureworld/treasure.html

Brewer, M. S., & Gardner, G. E. (2013). Teaching evolution through the Hardy-Weinberg principle: a real-time, active-learning exercise using classroom response devices. *The American Biology Teacher*, *75*(7), 476–479.

Buechley, L., Eisenberg, M., & Elumeze, N. (2007). Towards a curriculum for electronic textiles in the high school classroom. *Proceedings of the 12th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education*, 28–32.

Chi, M. T. H. (2005). Commonsense Conceptions of Emergent Processes: Why Some Misconceptions Are Robust. *Journal of the Learning Sciences*, *14*(2), 161–199. https://doi.org/10.1207/s15327809jls1402_1

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically Authentic Inquiry in Schools: A Theoretical Framework for Evaluating Inquiry Tasks. *Science Education*, *86*(2), 175–218. https://doi.org/10.1002/sce.10001

Clark, D., Nelson, B., Sengupta, P., & D'Angelo, C. (2009). Rethinking science learning through digital games and simulations: Genres, examples, and evidence. *Learning Science: Computer Games, Simulations, and Education Workshop Sponsored by the National Academy of Sciences, Washington, DC*. https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_080068.pdf

Cole, M. (1998). *Cultural psychology: A once and future discipline*. Harvard University Press.

Collins, A., & Ferguson, W. (1993). Epistemic forms and epistemic games: Structures and strategies to guide inquiry. *Educational Psychologist*, *28*(1), 25–42.

Dabholkar, S., Anton, G., & Wilensky, U. (2018). Genevo-an emergent systems microworld for model-based scientific inquiry in the context of genetics and evolution. *Proceedings of International Conference of the Learning Sciences, ICLS*.

Dabholkar, S., Arastoopour, G., Horn, M., & Wilensky, U. (2020). Students' Epistemic Connections Between Science Inquiry Practices and Disciplinary Ideas in a Computational Science Unit. *Proceedings of International Conference of the Learning Sciences, ICLS*.

Dabholkar, S., Bain, C., & Wilensky, U. (2016). *NetLogo GenEvo 1 Genetic Switch*. Center for Connected Learning and Computer-Based Modeling, Northwestern University.

Dabholkar, S., Granito, T., Horn, M., & Wilensky, U. (2019). *Evolution of Populations to Speciation (Advanced)*. https://ct-stem.northwestern.edu/curriculum/preview/961/

Dabholkar, S., Granito, T., & Wilensky, U. (2020). *Habitat Preference Behavior*. Center for Connected Learning and Computer-Based Modeling, Northwestern University.

Dabholkar, S., Hall, K., Woods, P., Bain, C., Horn, M., & Wilensky, U. (2018). *From Ecosystems to Speciation*. https://ct-stem.northwestern.edu/curriculum/preview/199/

Dabholkar, S., Swanson, H., & Wilensky, U. (2019). Epistemic considerations for modeling: Understanding the usefulness and limitations of models with Emergent Systems Microworlds. *The Annual Meeting of the National Association of Research in Science Teaching (NARST)*.

Dabholkar, S., & Wilensky, U. (2016a). *GenEvo Systems Biology Curriculum*. Center for Connected Learning and Computer-Based Modeling, Northwestern University.

http://ccl.northwestern.edu/curriculum/genevo/

Dabholkar, S., & Wilensky, U. (2016b). *NetLogo GenEvo 2 Genetic Drift model*. Center for Connected Learning and Computer-Based Modeling, Northwestern University.

Dabholkar, S., & Wilensky, U. (2016c). *NetLogo GenEvo 3 Genetic Drift and Natural Selection model*. Center for Connected Learning and Computer-Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo/models/GenEvo3GeneticDriftandNaturalSelection

Dabholkar, S., & Wilensky, U. (2016d). *NetLogo GenEvo 4 Competition model*. Center for Connected Learning and Computer-Based Modeling, Northwestern University.

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

Dabholkar, S., & Wilensky, U. (2019). Designing for esm-mediated collaborative science learning. *Computer-Supported Collaborative Learning Conference, CSCL*, *2*.

Dabholkar, S., Woods, P., Bain, C., Hall, K., Horn, M., & Wilensky, U. (2018). *Evolution of populations*. https://ct-stem.northwestern.edu/curriculum/preview/681/

Danish, J. A. (2013). Designing for Technology Enhanced Activity to Support Learning. *The Journal of Emerging Learning Design*, *1*(1). https://www.learntechlib.org/p/172810

Danish, J. A. (2014). Applying an Activity Theory Lens to Designing Instruction for Learning About the Structure, Behavior, and Function of a Honeybee System. *Journal of the*

Learning Sciences, *23*(2), 100–148. https://doi.org/10.1080/10508406.2013.856793

Danish, J. A., & Enyedy, N. (2020). Constructing with and through the body. In N. Holbert, M. Berland, & Y. B. Kafai (Eds.), *Designing constructionist futures: The art, theory, and practice of learning designs* (pp. 245–253). MIT Press.

De Jong, T., Linn, M. C., & Zacharia, Z. C. (2013). Physical and virtual laboratories in science and engineering education. *Science*, *340*(6130), 305–308.

DeAngelis, D. L. (2018). *Individual-based models and approaches in ecology: populations, communities and ecosystems*. CRC Press.

DeAngelis, D. L., & Mooij, W. M. (2005). Individual-Based Modeling of Ecological and Evolutionary Processes. *Annual Review of Ecology, Evolution, and Systematics*, *36*(1), 147–168. https://doi.org/10.1146/annurev.ecolsys.36.102003.152644

DeLiema, D., Enyedy, N., & Danish, J. A. (2019). Roles, Rules, and Keys: How Different Play Configurations Shape Collaborative Science Inquiry. *Journal of the Learning Sciences*, *28*(4–5), 513–555. https://doi.org/10.1080/10508406.2019.1675071

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, *32*(1), 5–8.

Dey, S., Dabholkar, S., & Joshi, A. (2006). The effect of migration on metapopulation stability is qualitatively unaffected by demographic and spatial heterogeneity. *Journal of Theoretical Biology*, *238*(1). https://doi.org/10.1016/j.jtbi.2005.05.013

Disessa, A. A. (1987). The third revolution in computers and education. *Journal of Research in*

Science Teaching, *24*(4), 343–367.

DiSessa, A. A. (1982). Unlearning Aristotelian physics: A study of knowledge-based learning. *Cognitive Science*, *6*(1), 37–75.

DiSessa, A. A. (2001). *Changing minds: Computers, learning, and literacy*. Mit Press.

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

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the Challenges of Inquiry-Based Learning Through Technology and Curriculum Design. *Journal of the Learning Sciences*, *8*(3–4), 391–450. https://doi.org/10.1080/10508406.1999.9672075

Edwards, L. D. (1995). Microworlds as Representations. In A. A. DiSessa, C. Hoyles, R. Noss, & L. D. Edwards (Eds.), *Computers and Exploratory Learning* (pp. 127–154). Springer. https://doi.org/10.1007/978-3-642-57799-4_8

Edwards, L. D. (1997). Exploring the territory before proof: Student‘s generalizations in a computer microworld for transformation geometry. *International Journal of Computers for Mathematical Learning*, *2*(3), 187–215.

Elby, A., & Hammer, D. (2010). Epistemological resources and framing: A cognitive framework for helping teachers interpret and respond to their students’ epistemologies. *Personal Epistemology in the Classroom: Theory, Research, and Implications for Practice*, *4*(1), 409–434.

Engeström, Y. (1991). Non scolae sed vitae discimus: Toward overcoming the encapsulation of school learning. *Learning and Instruction*, *1*(3), 243–259.

Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, & R. Punamaki (Eds.), *Perspectives on activity theory* (Vol. 19, Issue 38, pp. 19–30). Cambridge University Press.

Engeström, Y. (2001). Expansive Learning at Work: toward an activity theoretical reconceptualization. *Journal of Education and Work*, *14*(1), 133–156. https://doi.org/10.1080/13639080123238

Epstein, J. M., & Axtell, R. (1996). *Growing artificial societies: social science from the bottom up*. Brookings Institution Press.

Erickson, F. (1986). Qualitative methods in research on teaching. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 119–161). Macmillan.

Erickson, F. (2006). Studying Side by Side: Collaborative Action Ethnography in Educational Research. In G. Spindler & L. Hammond (Eds.), *Innovations in educational ethnography: Theory, methods and results* (pp. 235–257). Erlbaum.

Erickson, F. (2010). Classroom ethnography. In *International encyclopedia of education* (pp. 320–325). Elsevier.

Ferrari, M., & Chi, M. T. H. (1998). The nature of naive explanations of natural selection. *International Journal of Science Education*, *20*(10), 1231–1256.

Feurzeig, W. (1986). Algebra slaves and agents in a Logo-based mathematics curriculum.

Instructional Science, *14*(3–4), 229–254.

Fields, D. A., Kafai, Y. B., Morales-Navarro, L., & Walker, J. T. (2021). Debugging by design: A constructionist approach to high school students' crafting and coding of electronic textiles as failure artefacts. *British Journal of Educational Technology*, *52*(3), 1078–1092. https://doi.org/10.1111/bjet.13079

Finzer, W. (2016). *Common online data analysis platform (CODAP)*. Concord Consortium. https://codap.concord.org

Glaser, B. G., & Strauss, A. L. (2017). *The Discovery of Grounded Theory*. Routledge. https://doi.org/10.4324/9780203793206

Gobert, J. D., O'Dwyer, L., Horwitz, P., Buckley, B. C., Levy, S. T., & Wilensky, U. (2011). Examining the Relationship Between Students' Understanding of the Nature of Models and Conceptual Learning in Biology, Physics, and Chemistry. *International Journal of Science Education*, *33*(5), 653–684. https://doi.org/10.1080/09500691003720671

Goldman, S. R., Ko, M.-L. M., Greenleaf, C., & Brown, W. (2018). Domain-specificity in the practices of explanation, modeling, and argument in the sciences. In F. Fischer, C. Chinn, K. Engelmann, & J. Osborne (Eds.), *Interplay of Domain- Specific and Domain-General Aspects of Scientific Reasoning and Argumentation Skills* (pp. 121–141). Taylor Francis.

Goldstein, J. (1999). Emergence as a Construct: History and Issues. *Emergence*, *1*(1), 49–72. https://doi.org/10.1207/s15327000em0101_4

Goldstone, R. L., & Wilensky, U. (2008). Promoting transfer by grounding complex systems principles. In *Journal of the Learning Sciences* (Vol. 17, Issue 4). https://doi.org/10.1080/10508400802394898

Granito, T., Dabholkar, S., Horn, M., & Wilensky, U. (2020). *Animal behavior lab (isopods)*. https://ct-stem.northwestern.edu/curriculum/preview/1523/

Grimm, V., & Railsback, S. F. (2005). Individual-based Modeling and Ecology. In *Individual-based Modeling and Ecology* (Vol. 8). Princeton University Press. https://doi.org/10.1515/9781400850624

Grimm, V., & Railsback, S. F. (2013). *Individual-based Modeling and Ecology*. Princeton University Press. https://doi.org/doi:10.1515/9781400850624

Groff, J., & Mouza, C. (2008). A Framework for Addressing Challenges to Classroom Technology Use. *AACE Review (Formerly AACE Journal)*, *16*(1), 21–46. https://www.learntechlib.org/p/24421

Grover, S. (2019). Thinking about Computational Thinking: Lessons from Education Research. *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, 1283.

Grover, S., & Pea, R. (2013). Computational Thinking in K-12: A Review of the State of the Field. In *Educational Researcher*. https://doi.org/10.3102/0013189X12463051

Guo, Y., & Wilensky, U. (2018). Mind the gap: Teaching high school students about wealth inequality through agent-based participatory simulations. *Constructionism 2018 Conference*, *September*. https://www.researchgate.net/publication/327388505_Mind_the_gap_Teaching_high_schoo l_students_about_wealth_inequality_through_agent-based_participatory_simulations

Gutiérrez, K. D., Jurow, A. S., & Vakil, S. (2020). Social design-based experiments. *Handbook of the Cultural Foundations of Learning*, 76.

Guzdial, M. (1994). Software-Realized Scaffolding to Facilitate Programming for Science Learning. *Interactive Learning Environments*, *4*(1), 1–44. https://doi.org/10.1080/1049482940040101

Hashimov, E. (2015). Qualitative Data Analysis: A Methods Sourcebook and The Coding Manual for Qualitative Researchers. *Technical Communication Quarterly*, *24*(1), 109–112. https://doi.org/10.1080/10572252.2015.975966

Hiebeler, D. (1994). *The swarm simulation system and individual-based modeling*.

Hjorth, A., Head, B., Brady, C., & Wilensky, U. (2020). LevelSpace: A NetLogo Extension for Multi-Level Agent-Based Modeling. *Journal of Artificial Societies & Social Simulation*, *23*(1).

Hjorth, A., & Wilensky, U. (2014). Redesigning your city - a constructionist environment for urban planning education. *Informatics in Education*. https://doi.org/10.15388/infedu.2014.02

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

Hmelo, C. E., Holton, D. L., & Kolodner, J. L. (2000). Designing to Learn About Complex

Systems. *Journal of the Learning Sciences*, *9*(3), 247–298. https://doi.org/10.1207/S15327809JLS0903_2

Holman, B., Berger, W. J., Singer, D. J., Grim, P., & Bramson, A. (2018). Diversity and democracy: Agent-based modeling in political philosophy. *Historical Social Research/Historische Sozialforschung*, *43*(1), 259–284.

Horn, M., Baker, J., & Wilensky, U. (2020). *NetTango Web*. Center for Connected Learning and Computer-Based Modeling, Northwestern University. http://ccl.northwestern.edu/nettangoweb/

Horn, M., & Wilensky, U. (2012). NetTango: A mash-up of NetLogo and Tern. *AERA 2012*.

Hoyles, C., & Noss, R. (1987). Synthesizing mathematical conceptions and their formalization through the construction of a logo-based school mathematics curriculum. *International Journal of Mathematical Education in Science and Technology*. https://doi.org/10.1080/0020739870180411

Hoyles, C., & Noss, R. (1992). A pedagogy for mathematical microworlds. *Educational Studies in Mathematics*. https://doi.org/10.1007/BF00302313

Hutchins, N. M., Biswas, G., Maróti, M., Lédeczi, Á., Grover, S., Wolf, R., Blair, K. P., Chin, D., Conlin, L., & Basu, S. (2020). C2STEM: A system for synergistic learning of physics and computational thinking. *Journal of Science Education and Technology*, *29*(1), 83–100.

Jacobson, M. J., & Wilensky, U. (2006). Complex systems in education: Scientific and educational importance and implications for the learning sciences. In *Journal of the*

Learning Sciences. https://doi.org/10.1207/s15327809jls1501_4

Jeong, H., & Hmelo-Silver, C. E. (2016). Seven affordances of computer-supported collaborative learning: How to support collaborative learning? How can technologies help? *Educational Psychologist*, *51*(2), 247–265.

Kandel, E. R., Dudai, Y., & Mayford, M. R. (2014). The molecular and systems biology of memory. *Cell*, *157*(1), 163–186.

Kitano, H. (2002). Computational systems biology. *Nature*, *420*(6912), 206–210. https://doi.org/10.1038/nature01254

Kitano, H. (2017). Biological Complexity and the Need for Computational Approaches. In S. Green (Ed.), *Philosophy of Systems Biology* (pp. 169–180). Springer, Cham. https://doi.org/10.1007/978-3-319-47000-9_16

Klopfer, E. (2003). Technologies to support the creation of complex systems models—using StarLogo software with students. *Biosystems*, *71*(1), 111–122. https://doi.org/https://doi.org/10.1016/S0303-2647(03)00115-1

Klopfer, E., Scheintaub, H., Huang, W., Wendel, D., & Roque, R. (2009). The simulation cycle: Combining games, simulations, engineering and science using StarLogo TNG. *E-Learning and Digital Media*, *6*(1), 71–96.

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

Konold, C., Finzer, W., & Kreetong, K. (2017). Modeling as a Core Component of Structuring Data. *Statistics Education Research Journal*, *16*(2).

Krist, C., Schwarz, C. V, & Reiser, B. J. (2019). Identifying essential epistemic heuristics for guiding mechanistic reasoning in science learning. *Journal of the Learning Sciences*, *28*(2), 160–205.

Kuhn, T. S. (2012). *The structure of scientific revolutions*. University of Chicago press.

Kwok, R. (2019). AI empowers conservation biology. *Nature*, *567*(7746), 133–135.

Kwon, S. M., Wardrip, P. S., & Gomez, L. M. (2014). Co-design of interdisciplinary projects as a mechanism for school capacity growth. *Improving Schools*, *17*(1), 54–71.

Kynigos, C. (2015). Constructionism: Theory of Learning or Theory of Design? In S. J. Cho (Ed.), *Selected Regular Lectures from the 12th International Congress on Mathematical Education* (pp. 417–438). Springer International Publishing. https://doi.org/10.1007/978-3-319-17187-6_24

Kyza, E. A., & Georgiou, Y. (2014). Developing In-Service Science Teachers' Ownership of the Profiles Pedagogical Framework through a Technology-Supported Participatory Design Approach to Professional Development. *Science Education International*, *25*(2), 57–77.

Lab 8. (2001). In *AP(R) Biology Lab Manual for Students*. College Entrance Examination Board.

Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. In *Situated learning: Legitimate peripheral participation.* Cambridge University Press.

https://doi.org/10.1017/CBO9780511815355

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., & Wilensky, U. (2008). Inventing a "Mid Level" to Make Ends Meet: Reasoning between the Levels of Complexity. *Cognition and Instruction*, *26*(1), 1–47. https://doi.org/10.1080/07370000701798479

Levy, S., & Wilensky, U. (2009). Students' learning with the connected chemistry (CC1) curriculum: Navigating the complexities of the particulate world. *Journal of Science Education and Technology*. https://doi.org/10.1007/s10956-009-9145-7

Levy, S., & Wilensky, U. (2011). Mining students' inquiry actions for understanding of complex systems. *Computers & Education*, *56*(3), 556–573.

Linn, M. C., & Hsi, S. (2000). *Computers, teachers, peers: Science learning partners*. Routledge.

Loscalzo, J., & Barabasi, A. (2011). Systems biology and the future of medicine. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, *3*(6), 619–627.

Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K., & Balan, G. (2005). Mason: A multiagent simulation environment. *Simulation*, *81*(7), 517–527.

Maroulis, S., Guimera, R., Petry, H., Stringer, M. J., Gomez, L. M., Amaral, L. A. N., & Wilensky, U. (2010). Complex systems view of educational policy research. *Science*, *330*(6000), 38–39.

Miller, E., Manz, E., Russ, R., Stroupe, D., & Berland, L. (2018). Addressing the epistemic elephant in the room: Epistemic agency and the next generation science standards. *Journal of Research in Science Teaching*, *55*(7), 1053–1075. https://doi.org/10.1002/tea.21459

Mitchell, V., Ross, T., May, A., Sims, R., & Parker, C. (2016). Empirical investigation of the impact of using co-design methods when generating proposals for sustainable travel solutions. *CoDesign*, *12*(4), 205–220.

Morales-Doyle, D. (2017). Justice-centered science pedagogy: A catalyst for academic achievement and social transformation. *Science Education*, *101*(6), 1034–1060.

Müller-Hill, B. (1996). *The lac Operon: A short history a Genetic Paradigm*. Walter de Gruyter.

Nachman, M. W., Hoekstra, H. E., & D'Agostino, S. L. (2003). The genetic basis of adaptive melanism in pocket mice. *Proceedings of the National Academy of Sciences*, *100*(9), 5268–5273.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. The

National Academies Press.

North, M., Collier, N., & Vos, J. (2006). Experiences creating three implementations of the repast agent modeling toolkit. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, *16*(1), 1–25.

Noss, R., & Hoyles, C. (2017). Constructionism and microworlds. In *Technology Enhanced Learning* (pp. 29–35). Springer.

Novak, M., & Wilensky, U. (2015). *NetLogo Bacteria Food Hunt model*. Center for Connected Learning and Computer-Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo/models/BacteriaFoodHunt.

Papert, S. (1972). Teaching Children Thinking. *Programmed Learning and Educational Technology*, *9*(5), 245–255. https://doi.org/10.1080/1355800720090503

Papert, S. (1980). Mindstorms: Children , Computers , and Powerful Ideas and Powerful Ideas. In *NY: Basic Books*. Basic books. https://dl.acm.org/citation.cfm?id=1095592

Papert, S. (1993). *The children's machine*. BasicBooks.

Papert, S., & Harel, I. (1991). Situating constructionism. *Constructionism*, *36*(2), 1–11.

Peel, A., Dabholkar, S., Anton, G., Wu, S., Wilensky, U., & Horn, M. (2020). *A case study of teacher professional growth through co-design and implementation of computationally enriched biology units*.

Peel, A., Dabholkar, S., Wu, S., Horn, M., & Wilensky, U. (2021). An Evolving Definition of Computational Thinking in Science and Mathematics Classrooms. *Computational Thinking*

Education.

Peppler, K. (2013). STEAM-powered computing education: Using e-textiles to integrate the arts and STEM. *Computer*, *46*(09), 38–43.

Peters, V. L., & Slotta, J. D. (2009). Co-Designing Curricula to Promote Collaborative Knowledge Construction in Secondary School Science. *Proceedings of the 9th International Conference on Computer Supported Collaborative Learning - Volume 1*, 204–213.

Piaget, J. (1970). *Science of education and the psychology of the child.* Viking Press.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., Kyza, E., Edelson, D., & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences*, *13*(3), 337–386.

Railsback, S. F., & Grimm, V. (2019). *Agent-based and individual-based modeling: a practical introduction*. Princeton university press.

Resnick, M. (1997). *Turtles, termites, and traffic jams: Explorations in massively parallel microworlds*. MIT Press.

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. (2009). Scratch: Programming for All. *Commun. ACM*, *52*(11), 60–67. https://doi.org/10.1145/1592761.1592779

Resnick, M., & Wilensky, U. (1998). Diving into complexity: Developing probabilistic decentralized thinking through role-playing activities. *The Journal of the Learning Sciences*, *7*(2), 153–172.

Rogoff, B. (2008). Observing sociocultural activity on three planes: Participatory appropriation, guided participation, and apprenticeship. *Pedagogy and Practice: Culture and Identities*, 58–74.

Roschelle, J. M. (1991). *Students' construction of qualitative physics knowledge: Learning about velocity and acceleration in a computer microworld*. University of California, Berkeley.

Rosengren, K. S., Brem, S. K., Evans, E. M., & Sinatra, G. M. (2012). *Evolution challenges: Integrating research and practice in teaching and learning about evolution*. Oxford University Press.

Russ, R. S., & Berland, L. K. (2019). Invented science: A framework for discussing a persistent problem of practice. *Journal of the Learning Sciences*, *28*(3), 279–301.

Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education*, *92*(3). https://doi.org/10.1002/sce.20264

Sabelli, N. H. (2006). Complexity, technology, science, and education. *The Journal of the Learning Sciences*, *15*(1), 5–9.

Scardamalia, M., & Bereiter, C. (1991). Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. *The Journal of the Learning Sciences*, *1*(1), 37–68.

Schwartz, R. S., Lederman, N. G., & Crawford, B. A. (2004). Developing views of nature of science in an authentic context: An explicit approach to bridging the gap between nature of science and scientific inquiry. *Science Education*, *88*(4), 610–645.

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.

Schwarz, C. V, & White, B. Y. (2005). Metamodeling Knowledge: Developing Students’ Understanding of Scientific Modeling. *Cognition and Instruction*, *23*(2), 165–205. https://doi.org/10.1207/s1532690xci2302_1

Sengupta, P., & Wilensky, U. (2009). Learning electricity with NIELS: Thinking with electrons and thinking in levels. *International Journal of Computers for Mathematical Learning*, *14*(1). https://doi.org/10.1007/s10758-009-9144-z

Severance, S., Penuel, W. R., Sumner, T., & Leary, H. (2016). Organizing for teacher agency in curricular co-design. *Journal of the Learning Sciences*, *25*(4), 531–564.

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

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

Sinatra, G. M., Brem, S. K., & Evans, E. M. (2008). Changing minds? Implications of conceptual change for teaching and learning about biological evolution. *Evolution: Education and Outreach*, *1*(2), 189–195.

Small, M. L. (2011). How to conduct a mixed methods study: Recent trends in a rapidly growing literature. *Annual Review of Sociology*, *37*.

Smith III, J. P., DiSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, *3*(2), 115–163.

Smith, M. U., & Baldwin, J. T. (2015). Making sense of Hardy-Weinberg equilibrium. *The American Biology Teacher*, *77*(8), 577–582.

Stieff, M., & Wilensky, U. (2003). Connected Chemistry—Incorporating Interactive Simulations into the Chemistry Classroom. *Journal of Science Education and Technology*, *12*(3). https://doi.org/10.1023/A:1025085023936

Stroupe, D. (2014). Examining classroom science practice communities: How teachers and students negotiate epistemic agency and learn science-as-practice. *Science Education*, *98*(3), 487–516.

Swanson, H., Anton, G., Bain, C., Horn, M., & Wilensky, U. (2019). Introducing and Assessing Computational Thinking in the Secondary Science Classroom. In *Computational Thinking Education*. https://doi.org/10.1007/978-981-13-6528-7_7

Tesfatsion, L. (2002). Agent-based computational economics: Growing economies from the bottom up. *Artificial Life*, *8*(1), 55–82.

Trischler, J., Dietrich, T., & Rundle-Thiele, S. (2019). Co-design: from expert-to user-driven ideas in public service design. *Public Management Review*, *21*(11), 1595–1619.

Turkle, S. (2011). Introduction: The Things That Matter. In S. Turkle (Ed.), *Evocative objects: Things we think with*. MIT press.

Turkle, S., & Papert, S. (1990). Epistemological pluralism: Styles and voices within the computer culture. *Signs: Journal of Women in Culture and Society*, *16*(1), 128–157.

Turkle, S., & Papert, S. (1992). Epistemological pluralism and the revaluation of the concrete. *Journal of Mathematical Behavior*, *11*(1), 3–33.

VanLehn, K. (2013). Model construction as a learning activity: a design space and review. *Interactive Learning Environments*, *21*(4), 371–413. https://doi.org/10.1080/10494820.2013.803125

Vogelsberger, M., Genel, S., Springel, V., Torrey, P., Sijacki, D., Xu, D., Snyder, G., Bird, S., Nelson, D., & Hernquist, L. (2014). Properties of galaxies reproduced by a hydrodynamic simulation. *Nature*, *509*(7499), 177–182.

Voogt, J., Laferrière, T., Breuleux, A., Itow, R. C., Hickey, D. T., & McKenney, S. (2015). Collaborative design as a form of professional development. *Instructional Science*, *43*(2), 259–282. https://doi.org/10.1007/s11251-014-9340-7

Vygotsky, L. S. (1978). *Mind in Society: the development of higher psychological processes*. Cambridge University Press.

Wagh, A., Cook-Whitt, K., & Wilensky, U. (2017). Bridging inquiry-based science and constructionism: Exploring the alignment between students tinkering with code of computational models and goals of inquiry. *Journal of Research in Science Teaching*, *54*(5). https://doi.org/10.1002/tea.21379

Wagh, A., & Wilensky, U. (2018). EvoBuild: A quickstart toolkit for programming agent-based models of evolutionary processes. *Journal of Science Education and Technology*, *27*(2), 131–146.

Warren, B., Vossoughi, S., Rosebery, A., Bang, M., & Taylor, E. (2020). Multiple Ways of Knowing*: Re-Imagining Disciplinary Learning. In *Handbook of the cultural foundations of learning* (pp. 277–294). Routledge.

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). https://doi.org/10.1007/s10956-015-9581-5

Wilensky, U. (1993). *Connected mathematics--Building concrete relationships with mathematical knowledge*. Massachusetts Institute of Technology.

Wilensky, U. (1997a). *NetLogo Wolf Sheep Predation model*. Center for Connected Learning and Computer-Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo/models/WolfSheepPredation

Wilensky, U. (1997b). *StarLogoT*. Center for Connected Learning and Computer-Based Modeling, Northwestern University.

Wilensky, U. (1999a). GasLab—An extensible modeling toolkit for connecting micro-and macro-properties of gases. In *Modeling and simulation in science and mathematics education* (pp. 151–178). Springer.

Wilensky, U. (1999b). *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. (2005). *Connected Chemistry 7 Ideal Gas Law model*. Center for Connected Learning and Computer-Based Modeling, Northwestern University.

Wilensky, U. (2020). Restructuration Theory and Agent-based Modeling: Reformulating Knowledge Domains Through Computational Representations. In N. Holbert, M. Berland, & Y. B. Kafai (Eds.), *Designing Constructionist Futures: The Art, Theory, and Practice of Learning Designs* (pp. 287–300). MIT Press.

Wilensky, U., Brady, C. E., & Horn, M. S. (2014). Fostering computational literacy in science classrooms. *Communications of the ACM*, *57*(8). https://doi.org/10.1145/2633031

Wilensky, U., & Jacobson, M. J. (2014). Complex systems and the learning sciences. In *The Cambridge Handbook of the Learning Sciences, Second Edition* (pp. 319–338). Cambridge

University Press.

Wilensky, U., Levy, S., & Novak, M. (2004). *Connected Chemistry Curriculum*. Center for Connected Learning and Computer-Based Modeling, Northwestern University. http://ccl.northwestern.edu/curriculum/chemistry/

Wilensky, U., & Novak, M. (2010). Teaching and Learning Evolution as an Emergent Process: The BEAGLE project. In R. Taylor & M. Ferrari (Eds.), *Epistemology and Science Education: Understanding the Evolution vs. Intelligent Design Controversy*. Routledge.

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

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

Wilensky, U., & Reisman, K. (2006). Thinking like a wolf, a sheep, or a firefly: Learning biology through constructing and testing computational theories - An embodied modeling approach. In *Cognition and Instruction* (Vol. 24, Issue 2). https://doi.org/10.1207/s1532690xci2402_1

Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, *8*(1), 3–19.

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

Wilkerson-Jerde, Michelle, Wagh, A., & Wilensky, U. (2015). Balancing Curricular and

Pedagogical Needs in Computational Construction Kits: Lessons From the DeltaTick Project. *Science Education*, *99*(3), 465–499. https://doi.org/10.1002/sce.21157

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.

Williams, K. R., Wasson, S. R., Barrett, A., Greenall, R. F., Jones, S. R., & Bailey, E. G. (2021). Teaching Hardy-Weinberg Equilibrium using Population-Level Punnett Squares: Facilitating Calculation for Students with Math Anxiety. *CBE—Life Sciences Education*, *20*(2), ar22.

Windschitl, M., Thompson, J., Braaten, M., & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education*, *96*(5), 878–903.

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33–35.

Xiang, L., & Passmore, C. (2015). A framework for model-based inquiry through agent-based programming. *Journal of Science Education and Technology*, *24*(2–3), 311–329.

Yin, R. K. (2012). Case study methods. In *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological.* (pp. 141–155). American Psychological Association. https://doi.org/10.1037/13620-009

Yoon, S. A., Goh, S.-E., & Park, M. (2017). Teaching and Learning About Complex Systems in K–12 Science Education: A Review of Empirical Studies 1995–2015. *Review of Educational Research*, *88*(2), 285–325. https://doi.org/10.3102/0034654317746090

Yoon, S. A., & Hmelo-Silver, C. (2017). Introduction to special issue: models and tools for systems learning and instruction. In *Instructional Science* (Vol. 45, Issue 1). https://doi.org/10.1007/s11251-017-9404-6

APPENDIX

APPENDIX 1: CURRICULAR ACTIVITIES OF A GENEVO COURSE

Introductory activity

Be a Bacterium!

Question 0.1: Imagine that you are a bacterium. How would your typical day be? List at least 5 things that you would do throughout the day.

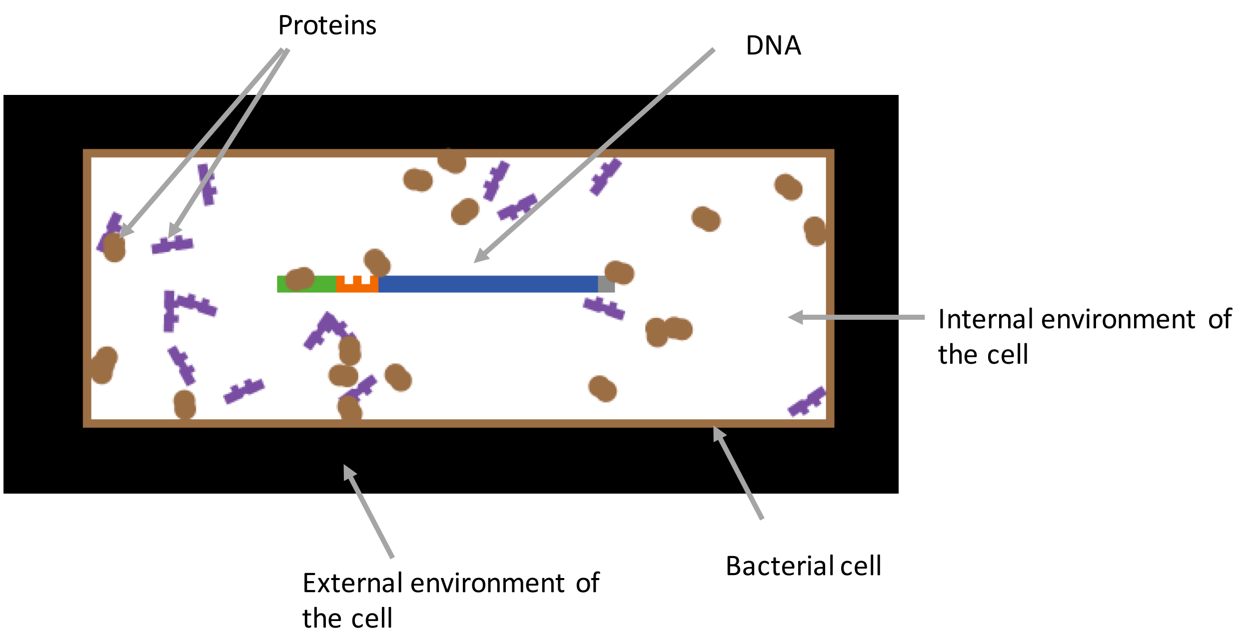
Question 0.2: What information about the world and about yourself that you would need to live successfully as a bacterium?
List at least 5 questions.

Stage 1: Genetic Switch

Overview: Getting to know the model

We are going to be real scientists to figure out if bacteria can make smart decisions. We are going to use a computational model to perform our research investigations.
Let's get to know the model first!

Components of the model:



How to run the model:

1. Click 'SETUP' to set the initial state for the bacterial cell.

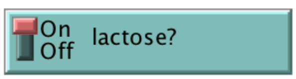
This step is to setup the initial positions of the violet and brown proteins inside the cell. If you click 'SETUP' again, the positions of the violet and brown proteins change, whereas the position of the DNA stays the same.

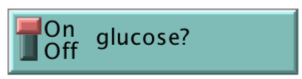
2. Click 'Go' to run the model.

This model is a computational simulation of the external and internal environments of a bacterial cell. When you click 'Go', you can see the protein molecules move around inside the cell. They do not go outside of the cell. Some of them interact with DNA. Observe their interactions with the DNA. DNA and proteins are molecular machines. *Smart* decisions that cells make are because of interactions between genes and proteins.

3. Sugar control:

We are going to investigate how bacteria cells smartly make decisions to eat different sugars. In their natural environments, bacteria use different food sources to produce energy. They need energy to survive and reproduce. If they don't get enough energy they die.
In our experiments, we can control which sugar is available to bacteria by turning ON or OFF the following switches:

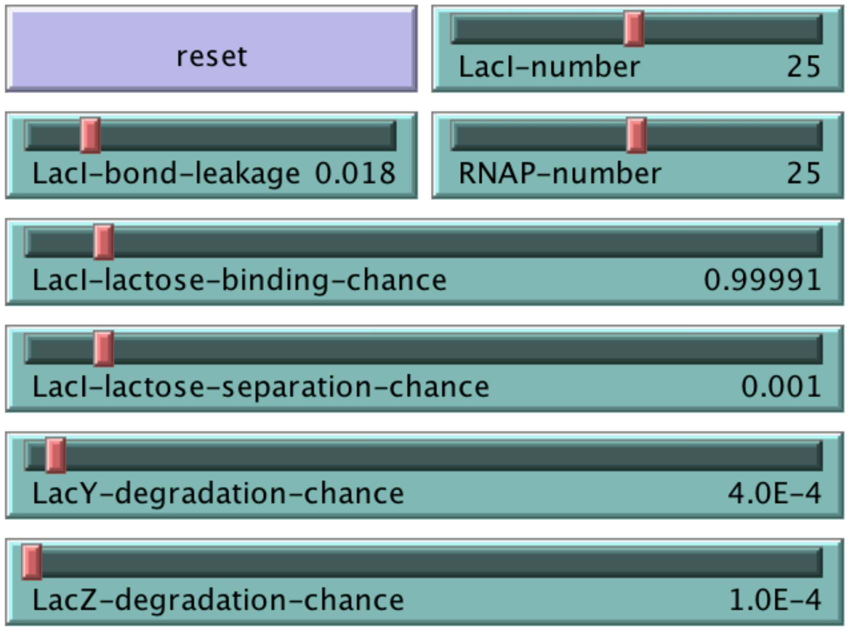




Glucose and lactose are two different types of sugars. Using these switches, we can have different combinations of these two sugars available to bacteria.
For example, keeping these two switches ON means both these sugars are available to bacteria.

4. Genetic Control:

We have several sliders available to control the genetic properties of the bacterial cell.



We will investigate what each of these sliders do during the course of our investigation.
Use the RESET button to set the values to default.
Molecular biologists and synthetic biologists, which are special types of scientists, make such changes in real cells. We will make these changes to our computational cell!

Exploration

Explore the Genetic Switch Model. Write down the observations that you find interesting.

Observations:

Activity 1.1: Energy and Cell Division
Guiding Questions:

Question 1.1.1: What are the effects of the presence or absence of sugar/s in the environment on the energy of the cell?

[Explain how you found the answer. What did you change in the model? What were your observations? How did you arrive at the answer using your observations?]

Question 1.1.2: What are the changes you observe in the energy graph as time changes? What makes the energy of a cell increase and what makes it decrease? What happens when the energy of the cell becomes twice as much as its initial energy?

[Explain how you found the answers. What did you change in the model? What were your observations? How did you arrive at the answers using your observations?]

Question 1.1.3: What are the effects of cell division on the energy of a cell? What other factors are affected by cell division?

[Explain how you found the answers. What did you change in the model? What were your observations? How did you arrive at the answers using your observations?]

Activity 1.2: DNA-Protein Interactions

Guiding Questions:
[For each question explain how you found the answer. What did you change in the model? What were your observations? How did you arrive at the answer using your observations?]

Questions 1.2.1: In this model, all the molecules that **only** appear inside the cell are proteins. How many different types of proteins are there in this model?

Hint: Change the sugar settings and see when certain types of proteins appear and disappear.

Question 1.2.2: Notice that certain proteins seem to perform certain functions. Observe and write down the functions performed by different proteins. Justify your observations based on evidence.

Hint: Certain functions might be dependent on the external environment. Change the settings and carefully establish the functions.

Question 1.2.3: Some proteins seem to interact with DNA. Describe your observations about these interactions. Do you think these interactions are important? What is the importance of these interactions?

Activity 1.3: Genetic Regulation

Guiding Questions:

[For each question explain how you found the answer. What did you change in the model? What were your observations? How did you arrive at the answer using your observations?]

Question 1.3.1: Some proteins are always present, whereas some proteins appear and disappear based on certain conditions, like the presence or absence of certain sugars. What are the conditions that make certain proteins appear and some proteins disappear?

Question 1.3.2: Why do you think the appearance and disappearance of proteins is important?

Question 1.3.3: When there are changes in the external environment in terms of presence or absence of certain sugars, different proteins are produced. This is called ‘genetic regulation’ of protein production. How is this achieved in the cell?
How is it related to the energy changes of the cell?

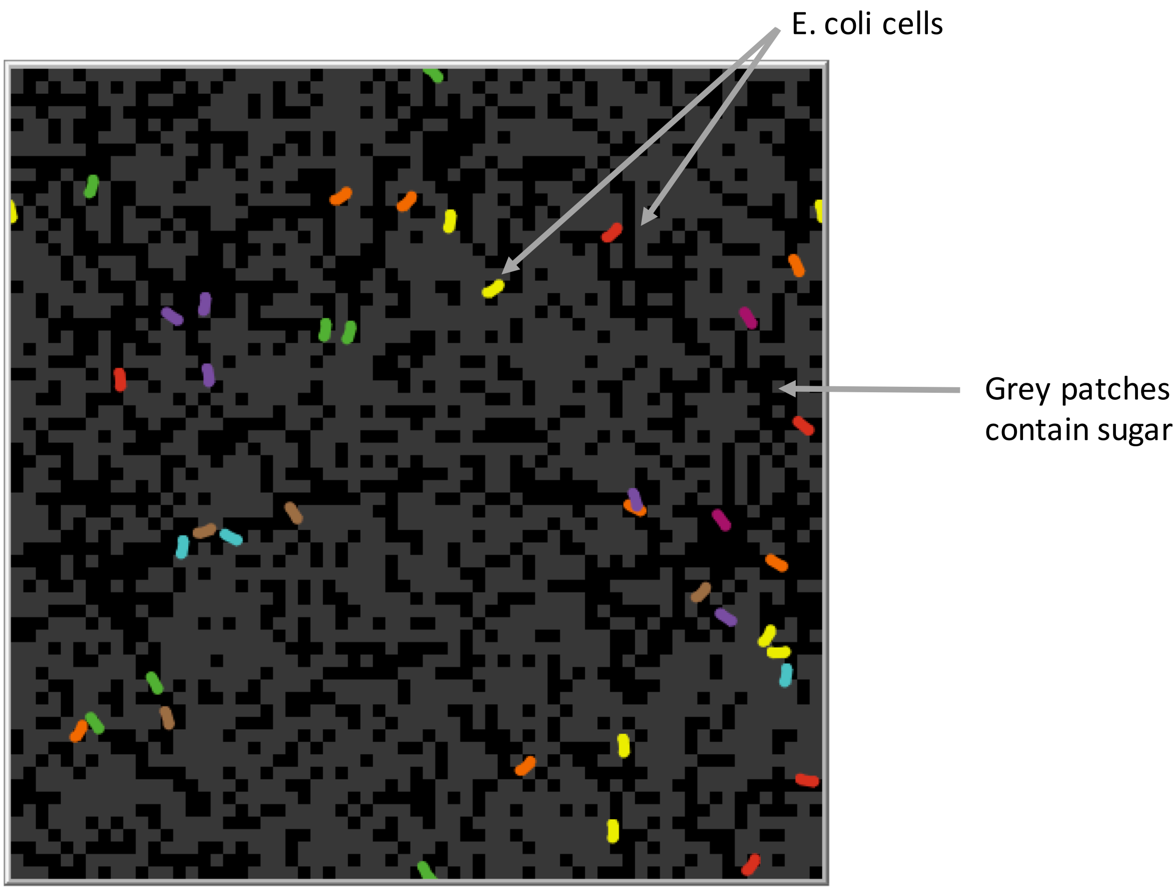
Question 1.3.4: We have observed how a cell responds differently in terms of protein production to the presence or absence of lactose or glucose in the environment. Molecular biologists refer to this as a Genetic Switch. Can you think of a reason why? Can you explain why it is a genetic switch?

Stage 2.1: Genetic Drift

Overview: Getting to know the model
This is a model of a population of bacterial cells, *E. coli*.

The model starts with different colored *E. coli* cells, randomly distributed across the world. The *E. coli* cells move around the world and eat sugar if it’s available to them where they are present. Grey patches (in the image below) contain sugar. Eating sugar increases the energy of an *E. coli* cell, whereas movement and basic metabolic processes decrease its energy. When the energy of a cell doubles, it reproduces to form two daughter cells of its type (of the same color). If the energy of an *E.coli* cell reduces to zero, the cell dies.
Different colored cells do not have any ‘advantage’ over other cells in terms of growth rate or sugar consumption.

Components of the model:

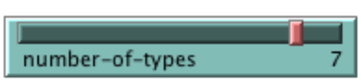


How to run the model:

1. Click ‘SETUP’ to set the initial population of the bacterial cells.
2. Click ‘Go’ to run the model.

This model simulates the growth of a bacterial population. As the model progresses the cells move around. If they are at a patch that has sugar, they eat it.

3. Number of types:



Use this slider to set the initial number of types (colors) of bacteria in the world.

4. Maximum initial population:

Use this slider to set the maximum number of bacteria of all colors in the initial population in the world.

5. Carrying capacity:

Use this slider to set the carrying capacity of the world. Carrying capacity is the maximum population that can be sustained in the world. This slider changes the availability of sugar in the world and thus controls the maximum population.

Activity 2.1.1: Population dynamics basics

1. This is a model simulating the growth of a bacterial population in an environment containing sugar. Bacteria eat sugar and divide. Thus, the population of bacteria grow. Start the simulation with **<u>one type</u>** of bacteria. Record your observations. How does the population change?

2. Change the ‘carrying capacity’ of the environment. How does the carrying capacity affect the growth of the population?

Activity 2.1.2: Prediction - Non-selective process of microevolutionary changes
Predication time!
Set the carrying capacity to medium. Set the number of types of bacteria to ‘two’. Set maximum initial population to 10. Do NOT run the model, yet.

What do you expect to happen after a few thousand ticks (5000 ticks)?
Will bacteria of both the colors survive or will one color win the evolutionary *race* if you run it for a really long time?

Run the model. Explain your observations.

Activity 2.1.3: Effects of the changes in carrying capacity on genetic drift

1. Another prediction!

Increase the number of types of bacteria to 6 or 7. How do you think the results will be different than when you had 2 types? Make a prediction. Do NOT run the model yet.

Run the model. Explain your observations.

2. Let's investigate the effects of carrying capacity on this process of genetic drift.

Genetic drift is the process of "one color" surviving without having any selective advantage. How would the process of genetic drift differ at high and low carrying capacities? Make a prediction.

Run the model. Explain your observations.

Stage 2.2: Genetic Drift and Natural Selection

Activity 2.2.1: Understanding faster reproduction

There is something called %-advantage in this model. We want to understand if it increases the rate of reproduction for E. coli.

Start the simulation with the following conditions –

- number-of-types = 1
- carrying-capacity = "high"
- ecoli-with-selective-advantage = "red"
- natural-selection? ON
- max-initial-population 1

a. Run the simulation with %-advantage = 0 for exactly 250 ticks. Record the number of bacteria in the population.

b. Run the simulation with %-advantage = 1 for exactly 250 ticks. Record the number of bacteria in the population.

Is there any difference? Run the simulation for each conditions a and b multiple times and see if these differences are consistent.

What do your observations tell you about %-advantage and rate of reproduction?

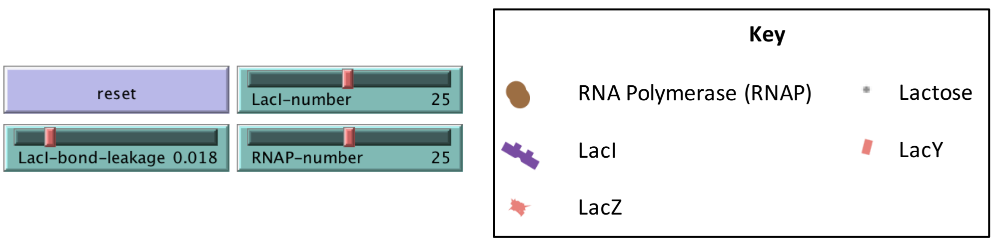
Activity 2.2.2: Exploring natural selection as a mechanism of microevolution

Design an experiment to see if **%-advantage** helps a color to *win* in case of natural selection. Describe your experiment and results.

Do all values of %-advantage help for a color to win? Why? Explain your answer.

Stage 3: Designing Genetic Circuits

Activity 3.1: Understand the parameters
Open the Genetic Switch model.



Let's focus on these three parameters.
Explain how each parameter affects the model. What changes do you see in the model when you change the values of these parameters? Describe the changes in terms of the number of protein molecules, or DNA-protein interaction.

1. LacI-number

2. RNAP – number

3. LacI-bond-leakage

Activity 3.2: Understand how the parameters affect the behavior

Change one or more of the three parameters shown above in the model.
Make a prediction about how it will influence the behavior of the genetic switch. Design a test to see if your cell does better or worse in terms of responding to the change in the external environment and fast cell division rate.
Write down your changes and predictions before you run the model.

Run the model and explain your observations.

Activity 3.3: Design your genetic circuit

Question 3.3.1: What would be a beneficial behavior to get a cell to reproduce faster in a changing external environment in terms of the availability of sugar?
[In Evolutionary Biology lingo, this is called '*Phenotype* that has higher *fitness*'.]

Question 3.3.2: Each of the teams are going to design 'genetic circuits' now. Work with your group. Design a 'genetic circuit' in your cell, so that the cell will have higher fitness. List the changes you made and explain why you made those changes.

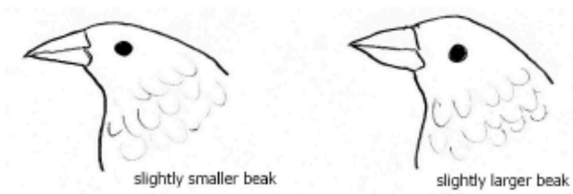
Activity 3.4: Design your genetic circuit – Part 2

Modify your 'genetic circuit'. Based on the performance of your cell in the evolutionary competition, modify the design of its genetic circuit. List the changes you made. Explain how those changes will modify the behavior of the cell.

Additional space to design new experiments, note observations and analyze data

APPENDIX 2: AN EXAMPLE QUESTION SET USED FOR PRE-TEST

1. A species of bird ate many types of seeds, each type coming from a different species of tree. The birds' beaks varied in size, with some individuals having slightly smaller beaks and others having slightly larger beaks.



A few years went by without much rain, and the only species of tree that survived had large seeds. Many generations later, almost all the birds had the slightly larger beaks. How is this best explained?

A. The birds with larger beaks were better at eating the large seeds than those with smaller beaks, so only the birds with larger beaks got enough food to survive, reproduce, and pass the trait of large beaks to the next generation.

B. The birds with smaller beaks had to work harder than those with larger beaks to crack open the large seeds. The more they used their beaks, the larger their beaks became, so they were able to get enough food to survive, reproduce, and pass the traits of large beaks to the next generation.

C. The birds with smaller beaks grew their beaks so that they would be better able to eat the large seeds and get enough food to survive, reproduce, and pass the trait of larger beaks to the next generation.

D. It was a chance occurrence that all the individual birds' beaks in the next generation were larger. They were therefore able to eat the large seeds and get enough food to survive, reproduce, and pass the trait of large beaks to the next generation.

2. Which of the following correctly describes what happens when a population of bacteria becomes resistant to an antibiotic? Note: a bacterium is one individual in a group of bacteria.

A. During treatment with an antibiotic, each individual bacterium tries to become resistant to the antibiotic. Only some are able to willingly become resistant, and these individuals survive to pass this trait to their offspring.

B. During treatment with an antibiotic, all of the bacteria gradually become more resistant to the antibiotic the more they are exposed to it. They all survive and pass this trait to their offspring.

C. During treatment with an antibiotic, a population of bacteria usually dies. Sometimes by chance, all members of the population become resistant at once, survive, and pass their resistance to their offspring.

D. During treatment with an antibiotic, only those individual bacteria that already have a trait that helps them survive the effects of the antibiotic will live. Their offspring in the next generation will also have this trait.

3. Could individuals of a species look different today than individuals of the same species did many generations ago? Why or why not?

    A. Yes, all individuals can change a little and pass those changes on to their offspring.

    B. Yes, some individuals can change a little and pass those changes on to their offspring.

    C. Yes, some individuals with certain traits are more likely to survive and pass those traits on to their offspring.

    D. No, species do not change even after many generations, so individuals of the same species would not look different.

4. Some of the individual members of a species of organism were moved to a new location that had different environmental conditions than their original location. According to the theory of natural selection, what could happen after many generations to the descendants of the organisms that had been relocated?

    A. The descendants would look the same as the original individuals because species do not change.

    B. The descendants would look different from the original individuals in some ways, and they would look the same in some ways.

    C. The descendants would become a completely different species that would have no similarities to the original individuals.

    D. The descendants would look the same as the original individuals because the environment does not affect how species look.

5. A population is a group of individuals of the same species. Can the percent of individuals with certain traits in a population change because the environment changes?

    A. Yes, when the environment changes, individuals in a population can change their inherited traits to better fit the environment, and this changes the percent of individuals with certain traits in that population.

    B. Yes, when the environment changes, individuals with certain inherited traits survive and reproduce and other individuals with different inherited traits die, and this changes the percent of individuals with certain traits in a population.

    C. No, the percent of individuals with certain inherited traits in a population changes randomly from one generation to the next, never as a result of changes to the environment.

    D. No, the percent of individuals with certain inherited traits in a population cannot change because a population is all one species and so will always have the same inherited traits.

6. A species lives in a particular environment. What is TRUE about the environment that the species lives in and about how the species will look over thousands of years?

    A. The environment will stay the same, and the traits of the species will stay the same.

    B. There will be changes to the environment, but the traits of the species will stay the same.

    C. There will be changes to the environment that could lead to changes in the traits of the species.

    D. There will be changes to the environment, and the traits of the species will change, but the changes in the environment could never lead to changes in the traits of species.

7. Which type of molecule contains genetic information that is passed from parents to offspring?

    A. Fat molecules

    B. DNA molecules

    C. Protein molecules

    D. Carbohydrate molecules

8. Which of the following are functions of protein molecules within cells?

    A. Protein molecules speed up chemical reactions in cells and help other molecules get in and out of cells.

    B. Protein molecules speed up chemical reactions in cells but do not help other molecules get in and out of cells.

    C. Protein molecules help other molecules get in and out of cells but do not speed up chemical reactions in cells.

    D. Protein molecules do not speed up chemical reactions in cells or help other molecules get in and out of cells.

9. Many different types of protein molecules are made within cells. Which of the following could be influenced by the actions of those protein molecules?

    A. Both an organism's physical characteristics and its behaviors

    B. An organism's physical characteristics but not its behaviors

    C. An organism's behaviors but not its physical characteristics

    D. Neither an organism's physical characteristics nor its behaviors

10. In some kinds of organisms, reproduction occurs asexually (without the combining of two sex cells). Which of the following statements is TRUE about the information that is passed in DNA from a parent to its offspring in these kinds of organisms?

    A. The information in the DNA of a parent is identical to the information in the DNA of its offspring.

    B. The information in the DNA of a parent is completely different from the information in the DNA of its offspring.

    C. Half of the information in the DNA of a parent and the DNA of its offspring is identical.

    D. The information that is identical in the DNA of a parent and in the DNA of its offspring varies from organism to organism within the same species.

APPENDIX 3: STUDENT INTERVIEW PROTOCOL

Pre-interview Science and science learning questions:

1. Can you mention some of the topics that you learned in your science class in the last year?
2. Pick one topic and explain how you learned it? <Ask probing questions to get explanation about the learning process, especially pertaining to their agency>
3. Recall and describe how your teacher taught the topic in the class.
4. Can you tell me the names of some scientists?
5. Do you know what scientists do in their daily work? <Ask more specific question. Reword based on their previous answer.>
6. Scientists develop knowledge about the world. How do you think scientists do that?

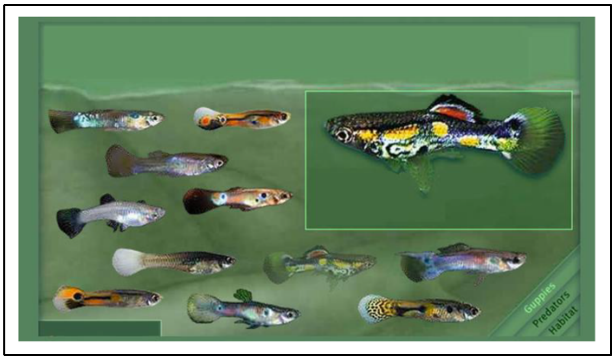
Post-interview Science and science learning questions:

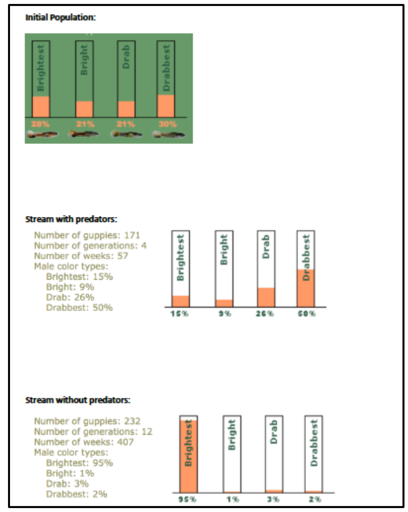
1. Can you mention some of the topics that you learned in this course?
2. Pick one topic and explain how you learned it?
3. Recall and describe how your learned it in the class.
4. Do you know what scientists do as their daily work?
5. Scientists develop knowledge about the world. How do you think scientists do that?
6. Do you think how you learned <add topic that the student mentions> like a scientist would learn about the world? Explain your answer.

Evolution questions:

1. Do you know what guppies are?
   Guppies are small fish that live in ponds and streams. Guppies come in many different colors.

   [Show picture of a group of guppies of many different colors and accompanying results after 100 years]





A scientist placed identical groups of mixed colored guppies in two different parts of a stream that are far away from each other so the guppies can't mix. One part of this stream had predators that ate these guppies. The other part didn't have predators. Guppies typically live to be about 2-3 years of age. When another scientist returned to this stream a hundred years later, she found that there were mostly grey and dull-colored male guppies in the part of the stream with predators, and mostly bright colored guppies where there were no predators.

Why do you think this might be?

If another scientist decided to conduct this same experiment in a similar stream at a different location. So she [repeat procedure]. Do you think she'll find the same results, or would she find something different?

Draw and show what you think her results might be.

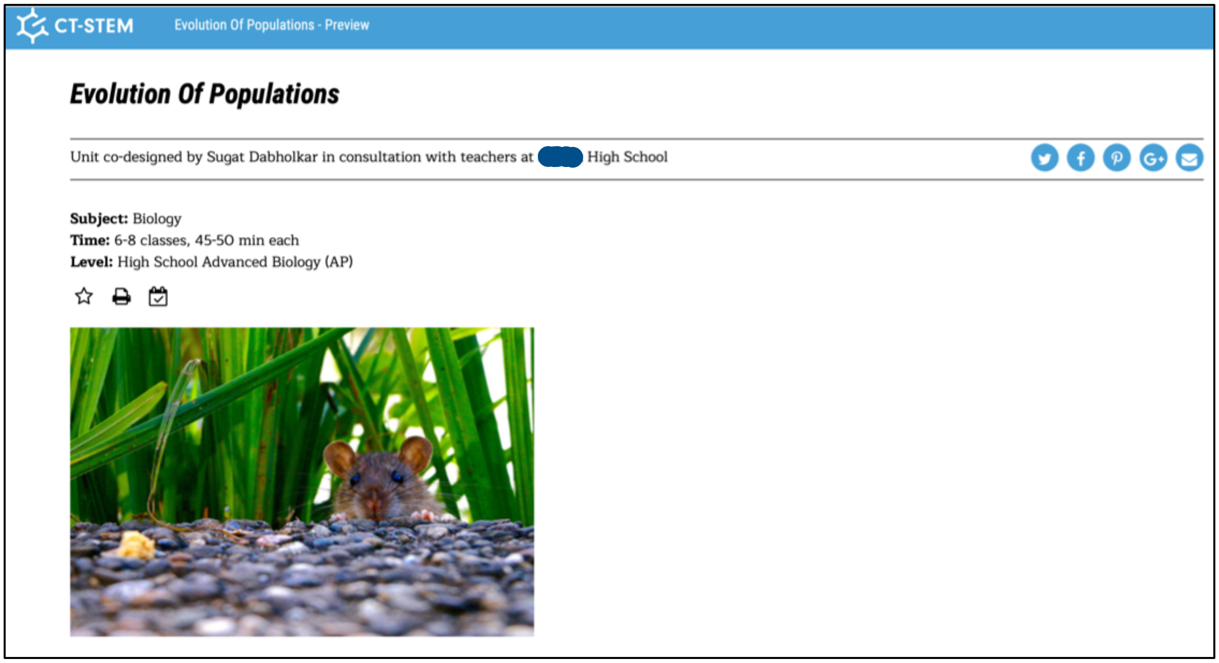
Modeling questions:

Computer models, like ones you will be using in class next week, help people understand things in the world <Give an example from previous question of guppies>.

Some people think that, with the right tools, anyone can make a computer model to help make sense of stuff in the world. Others think that only scientists and trained professionals can make models for others to use.

Who do you agree with? Why?

APPENDIX 4: CT-STEM ROCK POCKET EVOLUTION UNIT
(https://ct-stem.northwestern.edu/curriculum/preview/681/)



Unit Overview

This unit is designed for students to develop an understanding of how populations evolve by studying the case of rock pocket mice. Students will use computational models to understand prey-predator population dynamics and natural selection.

Lessons

- **Intro to Learning with Computational Models**

This is an introductory lesson for using certain types of computational models designed using a software called NetLogo.

In this lesson, students will learn -

- how to computationally study the spread of wildfire

- how to engage in the scientific inquiry practices of constructing knowledge in the context of an Emergent Systems Microworld (ESM)
- how to engage in computational thinking practices in the context of an ESM.

We will focus on four computational thinking practices: data practices, modeling and simulation practices, computational problem solving practices, and systems thinking practices.

- **Using Blocks to Model Ecosystems**

This unit introduces computational thinking practices which include data practices, modeling and simulation practices, computational problem solving practices, and systems thinking practices. These practices are introduced to students in the context of a biology unit about ecology.

- **Using a Model to Make Predictions**

Now that students have built a basic model, we're going to try and use this model to make predictions about the future of Isle Royale and the effects of different potential conservation efforts.

- **Introduction to the Case of Rock Pocket Mice**

In this lesson, students will be introduced to the 'anchoring phenomenon' of rock pocket mice, specifically how the color of the fur coat changed because of the change in the environment where they live. Students will explore a computational model of a population of rock pocket mice and observe changes in the population over time.

- **Natural Selection: Part 1**

In this lesson, students explore the computational model of a population pocket mice further. Specifically, they investigate how inheritance works in this model.

- **Natural Selection: Part 2**

In this lesson, students investigate how natural selection affects the genetic constitution of a population over time. They design and perform an experiment about natural selection in the case of rock pocket mice using a computational model to test their hypotheses.

Appendix 5: Protocol for teacher interview [about teaching a CT-STEM unit]

Tell me about your experience this time with developing and teaching this CT-STEM unit. [optional introductory question]

Students' experience with the unit:

How do you think your students experienced the unit?

What were some of their initial reactions when you started the unit?

Did their reactions change over time?

Can you give a specific example of one student's experience?

Students' learning with the unit:

What do you think your students learned in this unit?

What do you think your students learned about evolution in this unit?

Do you think your students' science learning was different using these computational tools or the same as previous years without these tools?

Can you describe how your students engaged with the NetLogo models?

Were there any advantages or disadvantages for using NetLogo models?

In some parts of this unit, students posed their own questions and designed and performed their experiments using the model. Do you think these activities had an effect on your students' learning?

Do you think there was anything else students learned that we haven't talked about?

Teacher's experience teaching the unit:

What was your experience with teaching this unit?

How was your role different or the same compared to other units that you typically teach?

How did you interact with you students?

What strategies did you use to support student learning in this unit?

Teacher's experience co-designing the unit:

What was your experience with co-designing this unit?

How would you describe your role in the co-design process?

What do you think went well with this collaboration?

Did your involvement in co-designing help you prepare for this unit? How?

If you could change something about the co-designing experience, what would you change?

For this CT-STEM project, we are working with multiple teachers across several schools. Based on your experiences, do you have some suggestions as to how this co-designing process could work in other classrooms and schools?

APPENDIX 6A: A CODEBOOK FOR EPISTEMIC NETWORK ANALYSIS (PRACTICES)

These practices are the overlapping practices between CT-STEM practices (Weintrop et al., 2016) and NGSS science and engineering practices (SEPs) (NGSS Lead States, 2013). In the following list of practices, we have adapted the practices from NGSS SEPs that specifically align with the computational activities in this curricular unit. In this unit, students use Emergent Systems Microworlds (computational tools and models designed in constructionist way) to make sense of natural phenomena in exploratory way. They explore the models, come up with the questions that they can investigate, design experiments to investigate those questions, perform those experiments computationally, collect and analyze data and construct explanations based on their investigations.

Code	sep.asking.questions (Asking Questions and Defining Problems)
Definition	A practice of science is to ask and refine questions that lead to descriptions and explanations of how the natural and designed world works and which can be empirically tested
Sub-practices	Ask questions that arise from examining models or a theory, to clarify and/or seek additional information and relationships; Ask questions to determine relationships, including quantitative relationships, between independent and dependent variables; Ask questions to clarify and refine a model, an explanation, or an engineering problem, Evaluate a question to determine if it is testable and

	relevant; Ask and/or evaluate questions that challenge the premise(s) of an argument, the interpretation of a data set, or the suitability of the design.
Keywords	What (but not if "I had" follows), Why, ?, where, if * what, if* how, when did, how did
Examples	If we added 6 mutants (with dark colorer [color] fur) in a population of light colored mice that are living on a light background environment what would happen to my dark colored mice in the light back ground after 360 generations?

Code	sep.using.models (Developing and Using Models)
Definition	A practice of both science and engineering is to use and construct models as helpful tools for representing ideas and explanations. These tools include diagrams, drawings, physical replicas, mathematical representations, analogies, and computer simulations.
Sub-practices	Develop and/or use multiple types of models to provide mechanistic accounts and/or predict phenomena,and move flexibly between model types based on merits and limitations; Use a model to provide mechanistic accounts of phenomena; Develop and/or use a model (including mathematical and computational) to generate data to support explanations, predict phenomena, analyze systems, and/or solve problems.

Keywords	slide, I/we/you change/turn/raise/left/use/set/lower/decrease/increase/made/out/play/put/move
Examples	When I changed the settings for predation from 0.05 to 0.25 I observe that there were more white mice.

Code	sep.planning.investigations (Planning and Carrying Out Investigations)
Definition	Scientists and engineers plan and carry out investigations in the field or laboratory, working collaboratively as well as individually. Their investigations are systematic and require clarifying what counts as data and identifying variables or parameters.
Sub-practices	Plan an investigation or test a design individually and collaboratively to produce data to serve as the basis for evidence as part of building and revising models, supporting explanations for phenomena, or testing solutions to problems. Consider possible variables or effects and evaluate the confounding investigation's design to ensure variables are controlled. Make directional hypotheses that specify what happens to a dependent variable when an independent variable is manipulated. Manipulate variables and collect data about a complex model of a proposed process or system to identify failure points or improve performance relative to criteria for success or other variables.

Keywords	set, I/we/you will/experiment/would (but not followed by "like to know")/could/are going to/am going to
Examples	we experimented with the light back ground and the light mice where more successful because they blended in with the background more than the dark background. and in the dark background the dark mice population increased and the light mice population decreased.

Code	sep.planning.investigations (Analyzing and Interpreting Data (includes making observations))
Definition	Scientific investigations produce data that must be analyzed in order to derive meaning. Because data patterns and trends are not always obvious, scientists use a range of tools—including tabulation, graphical interpretation, visualization, and statistical analysis—to identify the significant features and patterns in the data. Scientists identify sources of error in the investigations and calculate the degree of certainty in the results. Modern technology makes the collection of large data sets much easier, providing secondary sources for analysis.
Sub-practices	Record information (observations, thoughts, and ideas), Represent data in tables and/or various graphical displays (bar graphs, pictographs, and/or pie charts) to reveal patterns that indicate relationships, Construct, analyze, and/or interpret graphical displays of data and/or large data sets to identify linear and nonlinear

	relationships., Analyze data using tools, technologies, and/or models (e.g., computational, mathematical) in order to make valid and reliable scientific claims,
Keywords	Trial, trail, while the, that the, i see, hypothesis, observation, the way the
Examples	the higher the chance-of-predation the higher chance of the mice getting eaten Trial 1- predation=0.5 Results= 31 white colored 314 dark colored are left and there are way more owls there Trial 2- predation=0.05 Results= 200 white colored mice 361 dark colored mice are left and they are less owls Trial 3- predation=0.35 Results=168 while colored mice 285 dark colored mice and there are a little more owls than before

Code	sep.constructing.explanations (Constructing Explanations (includes making prediction))
Definition	Making a quantitative and/or qualitative claim regarding the relationship between dependent and independent variables. Constructing and revising an explanation based on valid and reliable evidence obtained from a variety of sources
Sub-practices	Make a quantitative and/or qualitative claim regarding the relationship between dependent and independent variables. Construct and revise an explanation based on valid and reliable evidence obtained from a variety of sources (including students' own investigations, models, theories, simulations, peer review) and the assumption

	that theories and laws that describe the natural world operate today as they did in the past and will continue to do so in the future. Apply scientific ideas, principles, and/or evidence to provide an explanation of phenomena and solve design problems, taking into account possible unanticipated effects. Apply scientific reasoning, theory, and/or models to link evidence to the claims to assess the extent to which the reasoning and data support the explanation or conclusion.
Keywords	because, explain, if * then, since, so (but not in the start of the sentence), which causes
Examples	When a mice was able to hide to the same background color as their fur, they are able to reproduce more since the predators can't see them clearly on their corresponding background.

APPENDIX 6B: A CODEBOOK FOR EPISTEMIC NETWORK ANALYSIS (DISCIPLINARY CORE IDEAS)

The codes for Disciplinary Core Ideas (DCIs) are based on the topic of the unit, which is about understanding evolution of populations, specifically focusing on natural selection:

Code	dci.agents
Definition	Referring to organisms, populations of organisms, individual organisms
Keywords	animals, mice, mouse, animal, population
Examples	This occurred; because in the light background, the light colored fur mice survived because that way they can camouflages and their predators would not affect them. (This also occurred to the mice with the dark background) .

Code	dci.phenotype
Definition	Referring to properties, or phenotypic (observable) characteristics or behaviors of agents in the model such as fur coat color or gender
Keywords	homozygous, heterozygous, black/dark, white/light, males, females
Examples	The number of light mice decreased and the number of dark mice increased.

Code	dci.genotype
Definition	Referring to genotype of organisms (AA, Aa, aa)

Keywords	homozygousm heterozygous, AA, Aa, aa
Examples	SO basically I just changed the initial setting which I turn off all the dominate male and female but also I turn off the heterozygous and only had recessive male and female which I put to 100 and they were all white which was (aa).

Code	dci.environments
Definition	Referring to environment of an organism or a population, like the background color (which may affect an orgnisms ability to hide from predators), or presence/absence of predators
Keywords	background, light/dark background, predators
Examples	When a mice was able to hide to the same background color as their fur, they are able to reproduce more since the predators can't see them clearly on their corresponding background.

Code	dci.heritability
Definition	Referring to passing on of a trait or genotype from generations to generations
Keywords	generation, reproduce, pass on, parent, baby, offspring, offsprings

Examples	after a lot of generation the initial mice are generally reproducing and the population is growing, the white mice is what we will see.

Code	dci.survival
Definition	Discussing an organism or a population having advantage or disadvantage for survival under certain conditions
Keywords	survival, advantages, adaptation, selection, successful, fit, fitness
Examples	Light mice in a dark BG as well as when Dark mice in a light BG, resulted in the likely probability of death because mice found in a BG with a disadvantage had a higher chance of dying.

Code	dci.adaptation
Definition	Specifically describing an adaptive mechanism that may have given survival advantage to an organism
Keywords	camouflage, hide, visible, see
Examples	The predators found mice that were not able to hide well a lot easier causing them to die off and the mice that were able to hide, survived and reproduced more mice with the same fur color.

Code	dci.change
Definition	Referring to change in an organism or a population over time
Keywords	increase, decrease, mutation, variation
Examples	The mice have mutations in their color so they're able to camouflage in their new enviroment [environment].

Appendix 7: Protocol for teacher interview [about co-design experience and computational thinking]

This protocol had been developed for the CT-STEM project. It is a protocol that I used for one of the interviews that I conducted during the 3-years of co-design partnership in Study 3. Other interview protocols were similar with minor changes. The interview protocol was designed for understanding teachers' views and perceptions regarding broader goals of the CT-STEM project. I used parts of this interview in which my partnering teacher talked about co-designing and using ESMs and ESM-based curricula.

Design Interview Questions:

Interviewer: Thank you for agreeing to participate in the research project. Do I have permission to audiotape the interview?

RECORDER TURN ON

Again, my name is ________ and this is a CTSI planning interview. For the recording please state your name.

Personal Domain: CT

In this set of questions, I will ask you about your prior experience teaching with computing.

1) Why did you want to participate in the CTSI workshop?
2) How would you define/describe computational thinking?
3) How would you define/describe CT in science and math?

    Is that different than CT/ your first answer? If so, how?

4) Prior to the CTSI, did you use any CT or computing in your science instruction? If so, please describe.

5) What is the value in integrating CT and science/math?

6) Are there aspects of CT you are still uncertain about CT?

Personal Domain: Beliefs

In this first set of questions, I am going to ask you about your beliefs about teaching and learning.

1) What are your overall goals for teaching science to your students?

2) As a teacher, what's your role in the learning process?

3) How do you think students best learn science?

4) What aspects of your goals or beliefs, if any, might potentially be in conflict with the CT-STEM approach?

Personal Domain PCK Set:

In the next set of questions, I will ask you about your prior experience teaching the science content of your curriculum unit.

1) What is the science content/topic in your curriculum unit? Have you taught this topic before?

If yes, describe *how* you taught the unit.

If no, skip to the next section:

<u>EVERYONE:</u>

1) Now that you have co-designed the curriculum unit, describe new instructional strategies/ activities/ representations you will be trying in your unit.

Prompt: what computing tools are you using in your unit? Why did you choose those tools?

2) In the curriculum unit, what do you anticipate will be difficult/challenging for your students? Any potential misconceptions?

3) In the curriculum unit, what do you anticipate will be beneficial for your students?

4) How are you planning to assess student learning in your unit?

a. How is that similar/different from what you've used in the past?

6) When you teach your curriculum unit, what aspects do you think will be difficult for you as a teacher?

Teachers' Perception of their Learning: from Interaction with the External Domain

This next set of questions asks you to reflect on the CTSI workshop.

1) From the workshop, what have you learned?

2) Did you have any "ah-ha" moments? If so please describe.

3) Which resources, in particular, have been the most helpful?

4) From your perspective, what were the most challenging aspects of designing a CT-STEM unit?

5) What feedback do you have for us about CTSI?

Teachers' Perception of their Learning: from Interaction with their co-design partners

In this set of questions, I will be asking you to reflect on the co-design process with your team.

1) Describe your co-design process.

2) What were some benefits from working with your co-design partner?

3) What were some challenges or tensions from working with your co-design partner?

4) How has the curriculum design process at this workshop been different for you?

5) To successfully implement a CT-STEM unit, what do you think you still need to learn?

6) How can we help support you in the future?

Influence of Domain of Practice:

In this next set of questions, I will be asking you to reflect on your school context.

1) In what ways has your school context influenced the design of your curriculum unit?

   Prompts: Influences of:

   a. School administrators?

   b. Other science teachers in your building?

   c. Your professional learning team?

   d. Past and future students?

   e. Your community?

2) One final question: When do you plan to implement your unit?