

A Scaffolding Framework to Support Learning of Emergent Phenomena Using Multi-Agent-Based Simulation Environments

Satabdi Basu · Pratim Sengupta · Gautam Biswas

Published online: 8 August 2014
© Springer Science+Business Media Dordrecht 2014

Abstract Students from middle school to college have difficulties in interpreting and understanding complex systems such as ecological phenomena. Researchers have suggested that students experience difficulties in reconciling the relationships between individuals, populations, and species, as well as the interactions between organisms and their environment in the ecosystem. Multi-agent-based computational models (MABMs) can explicitly capture agents and their interactions by representing individual actors as computational objects with assigned rules. As a result, the collective aggregate-level behavior of the population dynamically emerges from simulations that generate the aggregation of these interactions. Past studies have used a variety of scaffolds to help students learn ecological phenomena. Yet, there is no theoretical framework that supports the systematic design of scaffolds to aid students' learning in MABMs. Our paper addresses this issue by proposing a comprehensive framework for the design, analysis, and evaluation of scaffolding to support students' learning of ecology in a MABM. We present a study in which middle school students used a MABM to investigate and learn about a desert ecosystem. We identify the different types of scaffolds needed to support inquiry learning activities in this simulation environment and use our theoretical framework to demonstrate the effectiveness of our scaffolds in helping students develop a deep understanding of the complex ecological behaviors represented in the simulation..

Keywords Inquiry learning · Simulation-based learning environments · Multi-agent simulations · Scaffolding · Design-based research · Multiple representations · Ecology education

S. Basu · G. Biswas (✉)

Institute for Software Integrated Systems, Department of Electrical Engineering and Computer Science, Vanderbilt University, 1025 16th Avenue South, Nashville, TN 37212, USA
e-mail: gautam.biswas@vanderbilt.edu

S. Basu
e-mail: satabdi.basu@vanderbilt.edu

P. Sengupta
Department of Teaching and Learning, Vanderbilt University, Box 230, 230 Appleton Place, Nashville, TN 37203-5721, USA
e-mail: pratim.sengupta@vanderbilt.edu

Introduction

Complex systems represent a class of phenomena across a wide variety of domains, such as traffic jams, chemical reactions, and ecosystems, that are not regulated through central control but self-organize into coherent global patterns (Holland 1995; Kauffman 1995; Goldstone and Wilensky 2008). In such systems, emergence is the process by which collective, global behaviors arise from the properties of individuals and their interactions with each other, often in non-obvious ways. Therefore, in these systems, some of the observed global behaviors or patterns—known as *emergent* phenomena—cannot always be explained easily by the properties of the individual elements (e.g., Bar-Yam 1997, p. 10; Holland 1998). Emergent phenomena are central to several domains, such as population dynamics, natural selection, and evolution in biology, behavior of markets in economics, chemical reactions and statistical mechanics, thermodynamics, and electromagnetism in physics (Mitchell 2009; Darwin 1871; Smith 1977; Maxwell 1871).

Multi-agent-based computational models (MABMs) have been shown to be effective pedagogical tools in learning about emergent phenomena in multiple domains (Chi 2005; Mataric 1993; Wilensky & Reisman 2006). The term “agent” in the context of MABMs indicates individual actors modeled as computational objects (e.g., fish in a simulation of a pond ecosystem; vehicles in a simulation of traffic patterns), whose behaviors are governed by simple rules assigned by the user. When students interact with or construct a MABM, they initially engage in intuitive “agent-level thinking” (i.e., thinking about the actions and behavior of individual actors in the system; Goldstone & Wilensky 2008). Thereafter, with proper scaffolding, they can build upon their agent-level understanding and overcome the challenges of interpreting population dynamics (e.g., formation of a traffic jam, interpreting the predator-prey population dynamics) that emerge from the aggregation of the agent-level interactions (Dickes & Sengupta 2013; Tan & Biswas 2007; Wilensky & Resiman 2006).

Our research focuses on scaffolding in MABM-based learning environments for the domain of ecology. Ecological phenomena are examples of complex systems that a wide range of learners—from middle school to college students—find challenging to understand (Chi 2005; Chi et al. 1994; Okebukola 1990; White 1997). However, with proper scaffolding, MABMs have been shown to be effective pedagogical tools for learning about ecology. For example, in the studies reported by Tan & Biswas (2007) and Dickes & Sengupta (2013), students conducted controlled experiments, primarily using the predict-observe-explain cycle, in MABM environments. In both studies, the researchers conducting the study scaffolded the students’ interactions with the MABMs. In Wilensky and Reisman’s (2006) study, learners designed MABMs to represent and learn predator-prey dynamics and were provided assistance in terms of programming support and reflection prompts by the interviewer.

In this paper, we establish and evaluate a framework that can support the systematic design and evaluation of scaffolds to support students’ learning of complex ecological phenomena using MABMs. We build on the literature on scaffolding in science education and simulation-based learning, and present a theoretical framework for analyzing scaffolding in MABM-based learning environments. We report results from a study in which 8th grade students used a MABM to investigate and learn about a desert ecosystem. We used semi-clinical interviews to identify the different types of scaffolds needed to support students’ learning of emergent phenomena, such as interdependence and population dynamics, as they conducted inquiry using the MABM. Using our theoretical framework, we investigated the effectiveness of these

scaffolds in helping students learn the ecosystem model underlying the simulation and supporting their multi-level reasoning about the relevant ecological phenomena.

Literature Review

Complex Systems in Education

A number of studies have shown that people experience difficulty in understanding emergent phenomena (Hmelo-Silver & Pfeffer 2004; Jacobson 2001; Wilensky and Resnick 1999; Chi 2005). Researchers suggest that learners require designed educational interventions to understand complex phenomena (Charles and d'Apollonia 2004; Klopfer 2003; Penner et al. 2000; Resnick and Wilensky 1993, 1998; Slotta & Chi 2006; Dicks & Sengupta 2013; Leelawong & Biswas 2008). Specifically, in physics and biology education, researchers have shown that novices find it easier to reason about the actions and behaviors of individual agents, and scaffolding them with multi-agent simulations helps them develop an understanding of the relationships between individual-level behaviors and the aggregate-level outcomes that emerge from the aggregation of the individual-level behaviors and interactions (Blikstein & Wilensky 2009; Klopfer 2003; Dicks & Sengupta 2013; Danish et al. 2011).

Levy & Wilensky (2008) have pointed out several approaches that organize and articulate the relationships between agent-based and aggregate reasoning. The first approach is based on a notion of “difference in scale” (Lemke 2000). According to this approach, differences in scale are typically organized in terms of size or temporality. The fundamental processes are invariant, but they act at different scales of magnitude or time. In this view, emergent phenomena do not involve an essentially different mode of reasoning. Instead, they require the ability to reason with multiple scales and transition between the scales. Another approach is based on a notion of “incommensurability” or an ontological divide (Chi 2000, 2005). Chi described two schemas for processes: (1) *direct processes*, such as blood circulation, and (2) *emergent processes*, such as diffusion. A direct process is event-like and often has a clear goal, i.e., it is distinct, sequential, and bounded and has a clear beginning and end. In contrast, an emergent process is one that is uniform, simultaneous, and ongoing with no defined beginning or end. Based on this argument, Chi et al. have advocated that it is only by getting students to discard the naïve ontology of direct processes, and replacing it with the expert ontology of emergent process, that one can engender expert-like thinking in novices. They have argued for fostering such radical conceptual change through direct instruction focused on “ontology training,” i.e., teaching the “process based” or “emergent ontology” (Chi et al. 1994; Slotta & Chi 2006).

A third perspective is based on a notion of a *developmental trajectory*, starting with agent-based reasoning, and moving to aggregate reasoning (Danish et al. 2011; Dicks & Sengupta 2013; Sengupta & Wilensky 2009; Tan & Biswas 2007; Wilensky & Stroup 2003; Wilensky and Reisman 2006). In contrast to Chi et al.'s proposal of ontological discontinuity between agent-level and aggregate-level reasoning, this approach is based on recruiting, rather than discarding students' productive intuitions about agent-level attributes and behaviors. When multi-agent-based models are used, using this approach, learners are introduced to the agent-level behaviors and rules, and gradually develop an understanding of aggregate-level or emergent outcomes by interacting with multiple complementary representations of the putative phenomena: agent-level rules and variables, dynamic visualizations that simultaneously displays actions of all the agents in the microworld, and graphs that show aggregate-level patterns (such as population change in ecosystems). Our work is based on this approach, which

necessitates students taking the perspective of both the individual agent as well as the overall system as a collection of interacting agents.

Naive Biological Knowledge and Learning Ecology Using MABMs

From a cognitive perspective, research in children's early biological knowledge suggests that children draw analogically on their knowledge about humans to interpret unfamiliar biological phenomena (Piaget 1929; Carey 1985; Ingaki & Hatano 2002). For example, they may apply knowledge about their own needs in comparison to plants and other animals. In spite of pre- and elementary school children using a repertoire of biological knowledge to construct inferences about biological phenomena, it is difficult for learners, even at the secondary level, to develop causal explanations for more complex phenomena such as population growth, reproduction, death, decline, and inheritance of traits, which are essential epistemic components for understanding several important ecology concepts (Hendrix 1981; Bernstein & Cowan 1975; Ingaki & Hatano 2002). Chi et al. further argued (see "[Naive Biological Knowledge and Learning Ecology Using MABMs](#)") that novices' alternative ideas about some ecological phenomena are caused by an ontological miscategorization of the phenomena as a *direct* or an *event-like process* rather than an *equilibration process* (Ferrari and Chi 1998; Chi 2005).

However, several scholars have argued to move beyond a *deficit* perspective that primarily focuses on what children *cannot* do and have instead encouraged the field to focus on what children can accomplish with adequate support (Metz 1995; Lehrer & Schauble 2008). Lehrer and Schauble (2004, 2006, 2008), for example, showed that children's intuitive knowledge about the biological world as well as their representational competencies can be productively bootstrapped through the careful design of pedagogical activities that integrate the development of mathematical practices with development of biological concepts. This necessitates that the design of instructional supports and activities, rather than children's pre-instructional knowledge, becomes a central focus. That is, children's difficulties and successes may not be completely due to their pre-instructional knowledge; rather the design of instructional activities also plays a significant role in their conceptual development.

From an instructional perspective, several scholars have described how aggregate-level formalisms are used to teach population dynamics in science classrooms. An example is the Lotka-Volterra differential equation, which depicts how populations of different species in a predator-prey ecosystem evolve over time. This equation still forms the cornerstone of classroom instruction in population dynamics at the high school level and beyond (Wilensky & Reisman 2006). While mathematically sophisticated, these formalisms do not make explicit the underlying agent-level attributes and interactions in the system and therefore remain out of reach of most elementary, middle, and even high school students. On the other hand, in the context of understanding emergent phenomena, it is known that agent-based reasoning is developmentally prior to aggregate reasoning (Levy & Wilensky 2008; Sengupta & Wilensky 2009, 2011; Blikstein & Wilensky 2009). This is because agent-level reasoning leverages children's intuitive knowledge about their own bodies, perceptions, decisions, and actions (Papert 1980; Levy & Wilensky 2008). This body of research suggests that it is the lack of connection between students' natural, embodied, agent-based reasoning, and the aggregate forms of reasoning and representations they encounter in school that creates a barrier to their understanding of emergent phenomena.

When children construct or use multi-agent-based models (MABMs) to learn complex phenomena, this divide can be bridged (Dillenbourg 1999; Jacobson & Wilensky 2006; Tisue & Wilensky 2004; Sengupta and Wilensky 2011; Goldstone and Wilensky 2008; Dicks & Sengupta 2013). Specifically in the domain of ecology, the following studies show that

MABMs can indeed act as productive instructional tools for elementary and middle school students. Tan & Biswas (2007) conducted a study where 6th grade students used a multi-agent-based simulation environment to conduct science experiments related to a fish tank ecosystem. Students' interactions with the MABM introduced them to agent-level thinking, which in turn allowed them to use their intuitive biological knowledge to develop explanations of the emergent phenomena. The experimental study showed that students who used the simulation showed significantly higher pre-posttest learning gains as compared to a control group that were provided with the results of the simulations but did not have opportunities to explore in the multi-agent simulation environment. In another study, Wilensky and Reisman (2006) showed that high school students were able to develop a deep understanding of population dynamics in a predator-prey ecosystem by building multi-agent-based models of wolf-sheep predation. Dicke & Sengupta (2013) showed that students as young as 4th graders can develop multi-level explanations of population-level dynamics in a predator-prey ecosystem after interacting with a MABM of a bird-butterfly ecosystem through scaffolded learning activities.

Scaffolding in Computer-Based Learning Environments

Types of Scaffolding

Scaffolding has been defined as the process by which a teacher or more knowledgeable peer provides assistance that enables novice learners to solve problems, carry out tasks, or achieve goals, which would, otherwise, be beyond their unassisted efforts (Wood et al. 1976). The origin of this work, as Sherin et al. (2004) and Wertsch (1985) have argued, can be traced back to Vygotsky (1978), who adopted Engel's notion of instrumental mediation as it applies to "technical tools" and extended this notion to cover psychological tools or "signs." Furthermore, higher mental functions were seen by Vygotsky (1978) as first arising in the inter-psychological plane and then on the intra-psychological plane. It is the mediation of signs that makes this possible. The inclusion of computer-based learning environments requires this definition to be expanded to encompass assistance provided by *software tools* that can configure the environment, and provide functionality and resources that enhance the learning experience (Guzdial 1994; de Jong & Njoo 1992; de Jong and Van Joolingen 1998; Sherin et al. 2004). Guzdial (1994) outlined three ways in which software can provide scaffolding: (1) *communicating* processes to learners, (2) *coaching* learners with hints and reminders about their work, and (3) *eliciting* articulations from learners to encourage reflection and self-regulation.

Specifically, in the domain of science education, de Jong & Njoo (1992) have identified two types of process involved in inquiry learning: (1) *transformative processes* that directly yield knowledge such as problem definition, hypothesis formulation, designing an experiment, data collection and verification, etc. and (2) *regulative processes* that are necessary to manage the transformative processes, e.g., planning, verifying, and monitoring. de Jong and Van Joolingen (1998), in an extensive review of the field, concluded that learners face challenges in each of these aspects in the absence of suitable scaffolding. Reid et al. (2003) argued that learners need three kinds of support: (1) *interpretive support* that helps learners in structuring their knowledge of the domain, (2) *experimental support* that helps learners in setting up and interpreting experiments, and (3) *reflective support* which assists learners in reflections on the learning process. Examples of interpretative support are prompting to achieve timely activation of prior knowledge (Reid et al. 2003), modeling tools (Penner 2001), or a concept-mapping tool for organizing the learnt knowledge (Slotta 2004). These tools help learners generate representations of the domain and may also provide valuable feedback to the learners for interpreting their findings (Leelawong & Biswas 2008; Moreno & Mayer 2007). Experimental support

may consist of support for control of variables (Chen & Khlar 1999) or providing guiding questions to help students find relevant data (Reiser et al. 2001). Examples of reflective support include eliciting articulations from learners to encourage reflection and self-regulation using virtual agents or multiple representations (e.g., Azevedo et al. 2011; Biswas et al. 2010).

Building on the work of de Jong, Van Joolingen and several other researchers cited in the previous paragraph, Quintana et al. (2004) describe a systematic set of scaffolding guidelines and strategies for software-based scaffolding of inquiry learning in science, organized around three primary science inquiry components of (1) *sense making*, (2) *process management*, and (3) *articulation and reflection*. The complex “sense-making” processes require scaffolds for identifying relevant variables, designing empirical tests of hypotheses, collecting data, constructing and revising explanations based on evaluation of the data, and communicating arguments. “Process management” scaffolds help learners to generate plans for completing an inquiry task and monitor the planned steps systematically to manage the complex and ill-structured nature of inquiry. Finally, “reflection and articulation” scaffolds use reminders and specific pointers to facilitate productive planning, monitoring, and articulation during sense making. Scaffolds supporting these different inquiry components include malleable and multiple representations and language that highlights relevant information and builds on learners’ intuitive ideas about disciplinary knowledge, automatic handling of routine managerial chores, and decomposing complex tasks. When relevant, the scaffolds also provide expert guidance about scientific practices. One of our central goals in this paper is to identify the particular forms of scaffolds that learners require to support their sense-making, process management, and reflection and articulation processes, in the specific context of their interactions with multi-agent-based learning environments.

Degree of Scaffolding and Fading

To what extent should learners be scaffolded during an inquiry task? In their extensive review of computer games and simulations in science learning, Clark et al. (2009) pointed out that the degree of scaffolding taking into account the “immediacy of feedback” is important when designing the scaffolding mechanisms in simulation- and model-based learning environments. Several researchers have also suggested that certain types of relatively open-ended initial learning activities prior to providing scaffolding may lead to longer term overall learning gains in simulation-based learning environments (Bransford & Schwartz 1999; Kapur 2008; Kapur & Kinzer 2009; Schwartz & Martin 2004). For example, Schwartz & Martin (2004) argued for structuring instructional design around well-designed “invention” activities (i.e., activities that ask students to invent original solutions to novel problems). Building on this theory, Pathak et al. (2008) investigated the notion of “Productive Failure” (PF) (Kapur 2008), where students initially worked on problems in electricity with little scaffolding to support their learning. This was followed by a second phase of scaffolded structured problem solving activities using NetLogo simulations of electric current (Sengupta & Wilensky 2008). This research showed that the PF group performed better on posttest assessments on declarative and conceptual understanding than a group that received scaffolding for both phases of learning. Exit interviews indicated that students in this group appreciated that fact that they could freely explore in the learning environment in phase 1, and this better prepared them to learn in latter phases.

Van Joolingen et al. (2007) argued that scaffolds support learning by providing relevant contextual information, either by providing templates of actions or activities, or by constraining the learner’s interaction with the learning environment. However, ideally, scaffolding should empower learners to achieve the target learning goals with progressive “fading”

of the support (Guzdial 1994; Stone 1998; Sherin et al. 2004; etc.). As Guzdial (1994) pointed out, the fading schema should not use an *all-or-nothing approach*; rather, fading needs to be gradual, and it is likely to vary from individual to individual. McNeil et al. (2006) investigated the influence of fading of a particular form of scaffolding—written guidance for conducting scientific inquiry—on 7th grade students, during an 8-week project-based chemistry unit in which the construction and writing of scientific explanations was a key learning goal. Students received one of two treatments in terms of the type of written support: *continuous*, which involved detailed support for every investigation, or *faded* that involved less support over time. Their analyses showed significant learning gains for students for all components of scientific explanation (i.e., claim, evidence, and reasoning). Yet, on posttest items, where neither group received any scaffolds, the faded group wrote better explanations in terms of the reasoning support for the explanations, as compared to the continuous group. Fading written scaffolds better equipped students to write better explanations when they were not provided with support. In this study, we used a type of fading similar to McNeil et al. (2006), and we explain this at the end of “Scaffolding Using Multiple External Representations (MERs).”

Scaffolding Using Multiple External Representations (MERs)

In our learning environment, students are introduced to multiple external representations that present complimentary information to the learners. Ainsworth (2006) argues that MERs often lead to a better understanding of the represented phenomenon as compared to presenting the same phenomenon using a single representation. She presents three reasons in support of this argument. First, MERs can be used to represent complementary forms of information. For example, Scheiter et al. (2009) point out that realistic and schematic visualization may complement each other. While realistic dynamic visualizations may support the recognition of structures in the real world, schematic animations may support deeper understanding of underlying principles and processes (Van Gendt & Verhagen 2001). MERs may also facilitate analogical and abstract reasoning processes (Goldstone & Son 2005; Schwartz 1995). Second, with MERs, one representation may be used to constrain the interpretation of other representations and thereby reduce the potential for misconceptions. For instance, a realistic dynamic visualization may contain multiple interacting elements that are not fixed in space, and a schematic animation may resolve this problem by highlighting the relevant object in question and leaving out the other objects (Scheiter et al. 2009). Third, MERs may support a deeper understanding of the content by offering multiple perspectives on the same phenomena, which, in turn, may help develop a more flexible use of different representational formats.

However, relating different representational formats to each other is not easy for novices, and they may face a number of problems (Ainsworth 2006; Van Someren et al. 1998). For example, Kozma et al. (1996) argued that one of the challenges learners face can be attributed to the excessive demands on their limited capacities of visual perception and attention due to the multiple, linked representation approach. To address this issue, researchers have suggested two solutions: (1) providing symbolical links between the two forms of representations to support learners in relating the objects and relations shown in one representation to the respective entities displayed in the second representation, i.e., *dynalinking* (Ainsworth 2006; Kozma et al. 1996), and (2) presenting the visualizations one after another (and not all together) (Schwarz & Dreyfus 1993).

Our learning environment uses multiple visual representations—e.g., a dynamic simulation showing the behavior of the various elements in the ecosystem, graphs of the total population of each species over time, textual resources, and causal maps enabling students to model interdependence among the ecosystem elements as cause-effect relationships. The use of these

representational systems can act as scaffolds for the students, and the rationale behind the sequencing of these scaffolds is discussed in “[Rationale for Sequencing and Fading Scaffolds Using Multiple Representations](#).” Furthermore, the availability of these scaffolds to the learner was tied to the process of *fading* (McNeil et al. 2006) the scaffolds in our study. For example, during the initial phase of our study, we provided the learner with all of these representational systems; in addition, the interviewer provided verbal prompts, as needed, based on the difficulties experienced by students. These prompts served the following purposes: to remind students about the particular goal of inquiry that they were currently involved in, what they needed to pay attention to, and how they might construct an explanation based on the results of their inquiry. During the final activity, our goal was to see if learners could generate a multi-level explanation of the relevant phenomena in the form of a causal model. When students started constructing their causal map, they did not have access to the simulation. They generated causal maps and used them to answer additional questions about the aggregate behavior that was posed by the researchers. Therefore, our strategy for fading scaffolds during the final phase can be understood in terms of the following actions: (1) a timely removal of two important forms of external representations—the *dynamic visualization* and *graphs*—that acted as a scaffold for learners during earlier activities and (2) not providing the detailed verbal scaffolds on the conduct of experiments (e.g., setting appropriate parameters of the simulation).

Theoretical Framework for Analyzing Scaffolding in MABMs

The Δ -Shift Framework

Sherin et al. (2004) proposed a scaffolding framework for interpreting and analyzing the effectiveness of sets of scaffolds by comparing students’ performances in a base environment (where the student performs a learning activity without scaffolding) and in a scaffolded environment (an environment that assists students with their learning and problem solving activities). More precisely, analysis of scaffolding involves a comparison of two systems (S), i.e., the base system, S_{base} , which is the unassisted (or baseline) learning environment, and the scaffolded system, S_{scaf} , which is the baseline learning environment with additional scaffolds. Thus, $\Delta s = S_{\text{scaf}} - S_{\text{base}}$ represents the difference between the two systems, and Sherin et al. make an assumption that the situations S_{base} and S_{scaf} are relatively invariant; therefore, Δs is invariant. We adopt this framework, and like Sherin, we define change in learning performance (P) due to scaffolding as $\Delta p = P_{\text{scaf}} - P_{\text{base}}$. Typical students would score P_{base} in the base situation but achieve performance level P_{scaf} in the scaffolded situation S_{scaf} . Similarly, a factor P_{target} is defined as an idealized goal, i.e., the maximum one is expected to learn in the chosen problem domain with the given MABM and simulation. In an ideal situation, P_{scaf} will equal P_{target} . For example, this could represent all of the correct relations between pairs of entities in a MABM environment. In the next section, we discuss how this framework can be adapted for MABMs in general and, more specifically, for a MABM in the domain of ecology.

Extending the Δ -Shift Framework for MABMs

Understanding complex phenomena using MABMs, as we have discussed earlier, requires the ability to generate multi-level explanations, in which the aggregate-level phenomena, such as population dynamics, can be explained by aggregations of interactions between several individual agents. Previous studies (see “[Literature Review](#)”) have established that students can bring in with them several productive pre-instructional ideas at the agent level in the

domains of ecology and biology. Some of these ideas can be productive in developing deep conceptual understanding, whereas others can hinder the process of learning. Thus, in order to support students' learning of emergent phenomena in ecology using MABMs, we posit that scaffolds need to be designed to accomplish the following objectives:

1. *Encourage use of productive pre-instructional ideas to develop multi-level explanations:* bootstrap intuitive, pre-instructional ideas that are productive for developing deep understanding of the species and their interactions (e.g., correct intuitions like “Birds eat seeds”);
2. *De-emphasize incorrect pre-instructional ideas:* identify incorrect pre-instructional ideas both at the agent level (e.g., “Birds do not eat other birds”) and at the aggregate level (e.g., “In an ecosystem in equilibrium the population of any particular species stays constant over time”), and design a situation that helps the student realize how using these ideas can lead to incorrect explanations or conclusions;
3. *Scaffold the learner by helping her/him identify the corresponding correct idea:* design situations that will help students identify correct alternatives, thus helping them build the correct explanations. Often, (2) and (3) can be accomplished together.

Because our goal is to foster a deep conceptual understanding in students, our theoretical framework is based on hypothetical, idealized representations of students' *existing* and *target* mental models, as shown in Fig. 1. However, we note that what comprises a mental model, i.e., the cognitive structure of conceptual and perceptual elements that comprise novice and expert mental models, is currently a topic of debate in the Learning Sciences (Gupta et al. 2009; Chi et al. 1994; Chi 2005; diSessa 1993; Clark et al. 2009). We consider this debate outside the scope of our paper and, instead, identify the constituent elements of students' mental models in terms of *relationships* that are identified by students as they interact with the simulation. These relationships indicate how agents interact with other agents and their environment. Relationships, therefore, include individual agent-level actions as well as interactions between agents of different species. The lightly shaded oval in Fig. 1 represents the student's initial mental model at the start of the learning activity and it lists their correct and incorrect relations. The

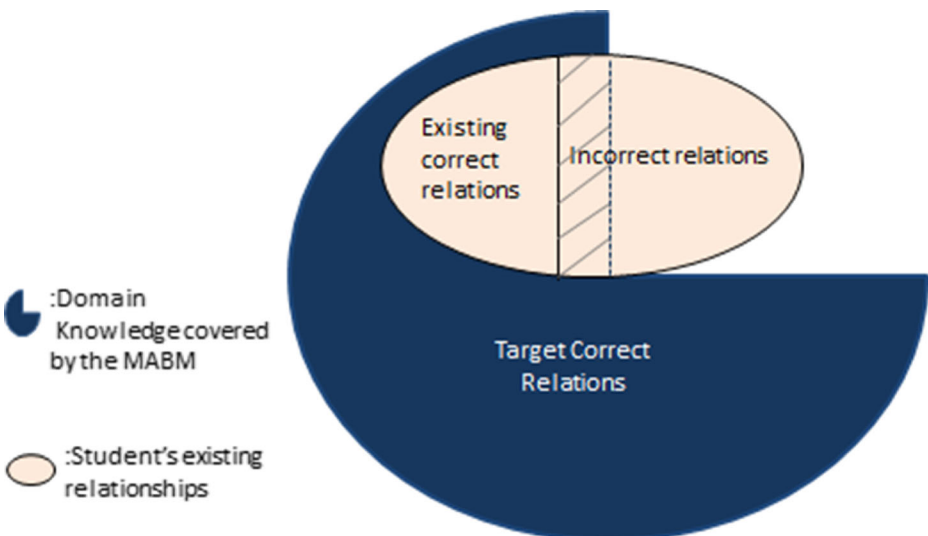


Fig. 1 Students' existing and target mental models of the domain

dark portion represents the correct domain knowledge or target mental model that we would like the student to develop after they conduct their learning activities.

Ideally, as students engage in the learning activities, we expect to see an increase in the number of correct relationships that students can identify in the target domain. At the same time, effective scaffolding should lead to the reduction in the number of incorrect relationships that students may have identified earlier, and both these processes may be inter-dependent. For example, some incorrect relationships might mask or hinder the learning of certain correct relationships. The effects of the scaffolds for removing those incorrect relationships are thus twofold: (1) they support learning by removing the incorrect relations and, in the process, (2) help in learning some correct relation(s) as well. This set of incorrect relations is represented in Fig. 1 by the shaded overlapping region between the students' existing incorrect relations and the target domain relations.

The twofold effects of the scaffolds in the shaded region can each be considered to be a unit Δ -shift according to Sherin et al.'s (2004) scaffolding analysis framework. However, doing so creates a dilemma: to achieve the most effective learning, these two effects often need to happen simultaneously. Identifying and learning a correct relation, in many cases, happens only when the student realizes that a relation s/he had previously assumed to be correct is actually incorrect. Thus, performing a linear sum of the effects of the scaffolds on individual performance measures will account for some effects more than once. In defining our measure for quantifying the effectiveness of scaffolds for our extended Δ -shift framework, we sum the increase in correct relations learned and incorrect relations discarded but reduce this number by the number of linked correct and incorrect relations so as not to double count the effect of a scaffold that simultaneously helps the students overcome incorrect conceptualization while learning the correct ideas.

In this study, our goal as to help students learn a set of pre-defined *relations* in a Saguaro ecosystem in the Sonora Desert in southern Arizona.¹ The learning goals include the set of relations that capture relevant interactions between entities of interest, such as *doves eat seeds* of the cacti and *doves pollinate cacti seeds*. The target performance (P_{target}) is defined as a set of relations that describe a food web, along with relations that describe the equilibration process that sustains the ecosystem. The Saguaro desert ecosystem model is described in more detail in the next section.

We illustrate our extended Δ -shift framework by studying the relations learnt by a student, John,² during the study. Scaffolding provided during the study reduced the number of incorrect relations in John's explanations from 3 to 1. One of these incorrect relations was "Birds do not eat other birds." This is most likely an example of anthropomorphic reasoning, where students liken the behavior of birds to that of humans. At the same time, the number of correct relations in his explanations increased from 2 to 5. One of the new relations he learnt was "Hawks prey on doves." Initially, we noticed that John was reaching the correct conclusion but discarding it citing his existing unproductive idea: "Birds do not eat other birds." Thus, we note that in this case, the students' realization that some birds do eat other birds helped him realize that hawks prey on doves, a relation he could not determine in the early phase of the study. This was also seen in the explanations of some other students who still believed at the end of the intervention that birds do not eat other birds and, therefore, did not recognize that hawks were preying on doves in the simulation. Using our extended Δ -shift framework, the net effects of the scaffolds on John's explanations = $\Delta p = \#(\text{correct relations learned}) + \#(\text{incorrect relations unlearned}) - \#(\text{interlinked correct and incorrect relations}) = (5 - 2) + (3 - 1) - 1 = 4$.

¹ http://www.fossweb.com/CA/modules3-6/Environments/activities/saguaro/content_old.html

² Not the student's actual name

The Simulation-Based Learning Environment

The Saguaran Desert Ecosystem Simulation

We conducted our research study using a NetLogo agent-based simulation (Wilensky 1999) of a Saguaran desert ecosystem (Basu & Biswas 2011; Basu et al. 2011). The desert ecosystem model includes five species modeled as agents: two plants (*ironwood trees* and *cacti*) and three animals (*rats*, *doves*, and *hawks*). A primary component of the desert ecosystem model is the set of food chain relations between the five species. Fruits (pods) and seeds associated with the ironwood trees and cacti, respectively, are additional agents that are represented in the simulation environment.

A review of the literature showed that novice learners bring in a plethora of intuitive agent-level knowledge. Therefore, student learning is best supported by activities that direct learners' attention to the agent-level entities (agents), their attributes, and their interactions in the simulation. As discussed earlier, it is these interactions that give rise to emergent, aggregate-level phenomena—such as behaviors that can be modeled using the Lotka-Volterra equation—that novices typically find counter-intuitive and difficult to learn.

In our MABM, each of the five species behaviors is characterized by a set of simple rules that model the species characteristics and their interactions with the other species. For example, the animals (in our simulation) need energy to survive and they die if their energy falls below pre-defined levels. The rules that describe the energy gain for each species are defined by their place in the food chain. Interactions between pairs of species on the food chain model a predator-prey relationship. The energy gain for rats is, for example, specified by a rule that says rats gain a specific amount of energy by feeding on the seeds of the cacti or the pods from the ironwood trees. Similarly, doves gain energy by eating the seeds of the cacti, while hawks prey on rats and doves to gain energy. A second category of rules specifies that the animals lose energy through locomotion (flying or ground movement) and reproduction. The reproduction rules, defined as stochastic processes, apply only to animals with energy levels above pre-specified values. Pods are produced at fixed intervals in simulation time by the ironwood trees. New ironwood trees are produced from the pods that are not eaten up by the rats. Seeds are produced by cactus plants when pollinated by the doves and then grow into new cactus trees. All animals and plants die when their age exceeds pre-specified age limits.

A screenshot of the simulation environment is shown in Fig. 2. Learners interact with the environment by setting the initial number of species of each type in the simulation run using the sliders on the left side of the user interface. They can start or stop a simulation run at any time using the start/stop buttons on the upper left side of the screen to take a more detailed look at the current state of their simulation. They can also regulate the speed of the simulation using the speed slider located at the top of the screen.

As Fig. 2 shows, the user has access to two types of displays as the simulation runs: (1) an animated pictorial depiction of the interactions between individual species or agents in the simulation window on the right and (2) a set of graphs that displays how the aggregate population for each species changes with time. This feature provides learners with access to individual and population behaviors simultaneously. The first graph in Fig. 2 is setup to plot the populations for all the species in the ecosystem. The second graph allows the user to choose one or more species populations that they want to observe selectively.

The Causal Modeling Tool

We hypothesized that providing students with opportunities to build a visual, qualitative causal model (Leelawong & Biswas 2008) would support students' conceptualization of and

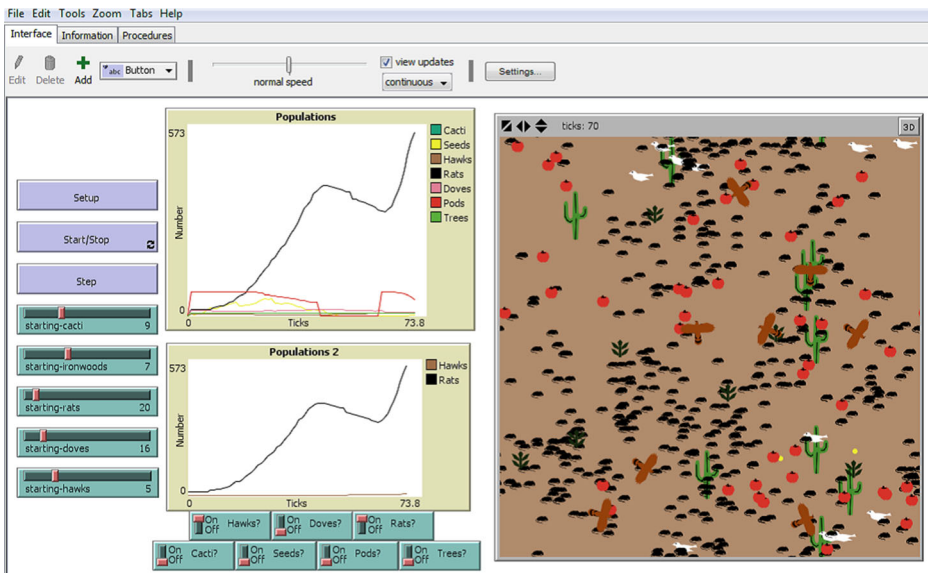


Fig. 2 The user interface (UI) of the Saguaran desert ecosystem simulation environment

reasoning about aggregate-level effects. As a result, students would more successfully construct explanations of aggregate-level phenomena from agent-level relationships. This, in turn, would help students formulate a complementary understanding of aggregate-level phenomena, such as interdependence between species and ecological balance. Figure 3 provides a screenshot of the causal modeling editor that students used to build their maps. A partial map drawn by a student appears in the map window. Students used the controls on the editor (the menu on the left hand side of the screen) to add concepts (the “Teach Concept” button in the menu) and then specified the type of relationships between pairs of concepts by adding annotated links between them (by clicking the “Teach Link” button and dragging the mouse from the “source/cause” concept box to the “target/effect” concept box). An “increase” link (denoted as “+”) means that an increase in the source concept causes an increase in the target concept, while a “decrease” link (denoted as “−”) means that an increase in the source concept causes a decrease in the target concept.

The “Resources” tab in Fig. 3 provided students with a searchable set of hypermedia pages (text and diagrams) with definitions and descriptions of entities and processes in the ecological domain. For example, there were pages corresponding to “Pollination,” “Predator-prey relationship,” “Food chain hierarchy,” and “Dynamic equilibrium.” The resources were available to the students throughout the intervention, i.e., the inquiry activities using the simulation and while they constructed their causal models.

Rationale for Sequencing and Fading Scaffolds Using Multiple Representations

When multiple representations of the same phenomenon are used for scaffolding, the sequencing of these representations plays an important role in ensuring effective learning (Schwarz & Dreyfus 1993). In our learning environment, the animation of the NetLogo agents, graphs that show how agent variables change over time and causal concept maps that capture aggregate relations in the system are all complementary visual representations that capture the interdependence among agents and the population dynamics in the ecosystem. Prior research on

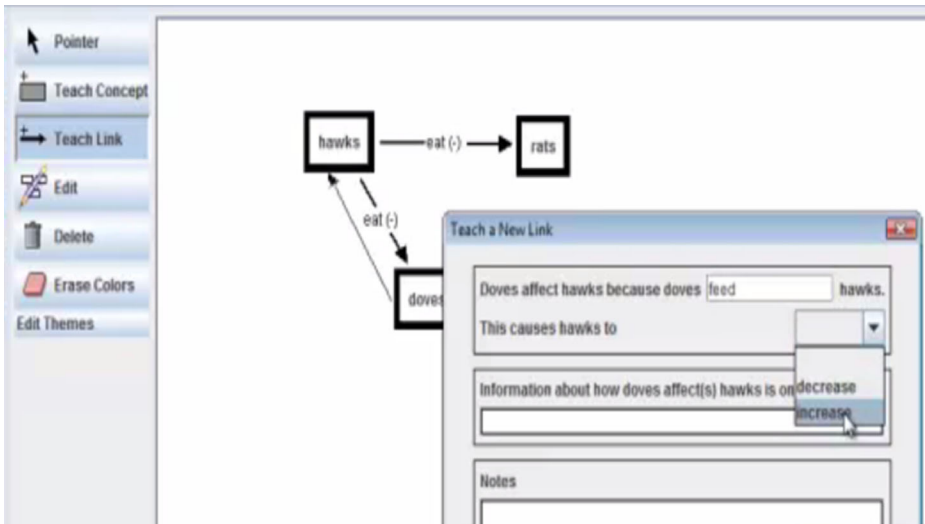


Fig 3 The causal modeling editor and a partial map created by a student

learning using MABMs as well as research on the development of biological knowledge (Goldstone and Wilensky 2008; Dicks & Sengupta 2013; Wilensky and Reisman 2006; Sengupta & Wilensky 2009) provided us with useful guidelines for sequencing these representations to support student learning.

In our system, students begin by interacting with the NetLogo simulation by observing the behavior of individual agents and groups of agents in the simulation window. They engage in deeper interactions with the simulation by changing parameters (variables or sliders in the NetLogo simulation) that affect agent-level behaviors directly. By running the simulation under different conditions, they begin to develop understandings of how the aggregate-level behaviors—e.g., populations of the different species—arise from interactions among the different agents. After this, students are asked to reason about individual graphs and the relationships between the visualizations in the simulation window and the graphs generated by the simulation. This provides students with opportunities to develop agent-aggregate complementary perspectives—i.e., allows them to reason at *both* the individual and aggregate levels of the ecological phenomenon. Students may take notes during this phase and use them for building the causal map of the ecological system later. In the last step, students interact with the causal modeling tool and are not allowed to run experiments with the simulation during this phase. By removing access to the simulation, we constrain students' interactions so that they cannot conduct further experiments to study plots and/or establish and verify any new hypotheses. In essence, we can consider this approach to represent a form of fading of scaffolds. Another interpretation of our approach is that the scaffold fading framework makes students go through an inquiry and model building process in individual discrete steps and keeps the cognitive load of dealing with multiple representations simultaneously low. Students work with the simulation model and study population-level graphs, make notes during this study, and then use their generated notes to build and reason with the causal model.

However, as we argued previously, it is important to note that fading does not imply total disappearance of scaffolds. As Leelawong & Biswas (2008) showed,

causal modeling tools such as the one we used in this study can provide students with an opportunity to learn about indirect effects and how the sum of these effects leads to interdependence and balance in an ecosystem. We demonstrate that this approach to scaffolding and fading of scaffolds leads to effective learning of aggregate-level behaviors in the following sections.

Method

Learning and Research Goals

The learning goals of the study included the following:

1. Learning and reasoning with the relations identified in Table 1;
2. Constructing a causal model of the ecosystem using the relations learnt; and
3. Generating aggregate-level explanations to reason about ecological phenomena.

We established the following research goals for this study:

1. *Identifying the initial state of students' relevant knowledge about the desert ecosystem displayed in the simulation:* This entailed collecting information on students' initial ideas (correct and incorrect) about ecological processes and the species present in the ecosystem and their interactions.
2. *Scaffolding analysis:* This involved (1) identifying the different challenges students faced while interacting with the simulation, constructing the causal model and reasoning causally about the ecosystem, and the types and timing of the scaffolds provided to help overcome each of these difficulties and (2) using our extended Δ -shift framework to assess the effectiveness of the scaffolds in helping students identify and better understand the inter-species relations R1–R6 and eliminate students' incorrect relations.
3. *Measuring the effectiveness of the overall intervention* in helping students develop multi-level explanations of complex ecological processes such as dynamic equilibrium, interdependence, food chain hierarchies, and pollination, as measured by students' pre- to posttest gains.
4. *Determining whether building and reasoning with a causal map representation helps students better understand aggregate-level behaviors*, leading to learning gains in terms of their understanding of aggregate-level phenomena.

Table 1 The relations expected to be learnt from the Saguaran ecosystem simulation

Relations	
R1	Doves eat seeds of the cacti
R2	Rats eat pods (fruits) of the ironwood trees
R3	Rats eat seeds of the cacti
R4	Hawks prey on rats
R5	Hawks prey on doves
R6	Doves help pollinate the seeds

Setting and Procedure

We conducted semi-clinical interviews with 20 8th graders (average age 13–14 years) in a public middle school in an ethnically diverse metropolitan school district in the Southeastern USA. The classroom teacher selected 10 high-performing and 10 low-performing students for this study, based on the students' academic performance in standardized tests and general performance in the science classes. All 20 students gave their written consent to participate in our research study. During the study, each student separately and individually worked with the simulation (MABM) as well as the concept-mapping tool, in a separate room, in the same building, in the presence of a researcher. Throughout their interactions with the software, the researcher interviewed them. During the time we ran our study, the other 8th grade students participated in a different research study conducted by a different set of researchers in their classroom. Participation in our studies did not affect students' class grades in any way.

The study was conducted in two sessions (1 h per day for 2 days, for each student) and was book-ended by written pre- and posttests. The pre- and posttests comprised three multiple choice questions (MCQs), seven short-answer questions (SQs), and six causal reasoning questions (CRQs). The MCQs and SQs were used to gauge students' understanding of aggregate-level phenomena such as the bidirectional predator-prey relationships; effects of big changes to a particular species population on other species of the ecosystem; interdependence in an ecosystem as a result of the food chains, producers and consumers; and the process of pollination. An example MCQ was as follows: "If an ecosystem is in equilibrium/balance, which of the following is true about such an ecosystem?," while an example SQ was as follows: "List, step by step, how the process of pollination helps in creation of new plants." The causal questions (CRQs) checked students' abilities to reason in causal chains using a given causal map representation. For example, students were provided a causal map with concepts A, B, C, D, and E, and asked to reason how a decrease in A would affect C and E. Some of the causal questions tested single-link causal reasoning while others tested multi-link causal reasoning abilities.

During the interviews, the two researchers who conducted the study played dual roles. In the first role as an *interviewer*, they conducted semi-clinical interviews by periodically asking students to provide mechanistic explanations of their observations and predictions. In their second role as a *scaffolder*, they guided the students' inquiry processes. When students made mistakes or when the researchers thought they needed help, the students were provided help using a verbal dialog. The scaffolding was directed to help students conduct relevant experiments in the simulation environment, interpret the results and derive correct ecological relations, construct causal models using the relations, and reason causally about the ecosystem using the causal map. On average, each interview lasted about two 1-h sessions (1 session per day). All the interviewer-student conversations, as well as the students' on-screen actions and movements, were recorded using the Camtasia³ software package and transcribed later by the researchers. The data for this study was derived from the transcriptions. The details of the semi-clinical interview process are described next.

On day 1, each interview consisted of three phases:

1. *Introduction*—Before students started working in the simulation environment, the interviewer asked the students about their prior experience in working with simulations and

³ TechSmith's Camtasia studio is Screen Recording & Video Editing Software. See <http://www.techsmith.com/camtasia/>

- then went over an introductory tutorial to help students learn about the user interface and the use of the controls (e.g., buttons, sliders, plots) provided in the NetLogo environment.
2. *Eliciting initial ideas*—After the introduction, students were asked to define an ecosystem, explain how an ecosystem functions, and what it means for an ecosystem to be in balance (i.e., equilibrium). The students were then asked to run the desert ecosystem simulation a few times and explain what they had observed in the simulation window and the plots, without being provided any goal or context for explanation. Students were also asked questions like “If the hawks were removed from this ecosystem, what would happen to the other species?” to see if they could reason in causal chains. Students’ responses to questions in the initial ideas phase revealed that they had several correct, incorrect, and missing relations in describing the interactions among the ecosystem species. The missing and incorrect relations provided the triggers for using specific scaffolds in the next phase.
 3. *Scaffolding to support learning*—In this phase, when needed, the interviewer used verbal dialog to help students perform guided experiments and systematically study the relations between the pairs of species (agents) in the ecosystem simulation. Students went through multiple iterations of prediction, running the simulation, and observing and explaining the results. The initial scaffolds helped students set the simulation controls to study only a pair of species at a time, understand and explain the ensuing plot(s) that the simulation run generated, and make notes when they discovered a new relation or the lack of a direct relation between two species. They could refer to these notes later when they built the causal model of the ecosystem on day 2. Students also had access to the Resource pages, which provided them with information about important ideas and phenomena like pollination, food chain, equilibrium, and predator-prey relations. If the students were unable to state the correct relation after the initial scaffolding, they received directed scaffolding on experiments they could perform to study inter-species relations. Sometimes, the suggested experiments produced results that contradicted the student’s earlier incorrect predictions. In such cases, the scaffolding helped students see and reason about these direct contrasts, and this may have helped students realize and overcome their previous errors.

The interviewer also asked students questions about predator-prey relationships, sustained equilibrium, food chains, and pollination. Some of the questions directed the students’ attention to reasoning in causal chains to discover indirect relations about the food web (e.g., through a chain of predator-prey relations) and other ecological processes (e.g., pollination). For example, when a student correctly identified the relationship between a predator (hawks) and its prey (rats), they were asked further follow-up questions, such as the following: (1) “What will the population graphs look like if you keep running this simulation?” (2) “If the hawks are eating the rats, why aren’t the rats dying out?” (3) “Are these two species in balance?” and (4) “How can they be in balance if their populations keep changing?” To check if they could reason in causal chains, to answer “what-if” questions, they were asked questions like “What will happen to the ecosystem if you remove hawks?”

The primary components of the day 2 intervention were as follows:

1. *Model building*—Students were introduced to causal modeling and the causal representation tool, and then asked to build a causal model of the ecosystem that captured all of the relations they had learned the previous day. They could refer to their notes from the previous day but were not allowed access to the multi-agent simulation window or the graphs they generated while running experiments. Figure 4 represents a correct and complete causal model of the desert ecosystem. Each of the food chain relationships

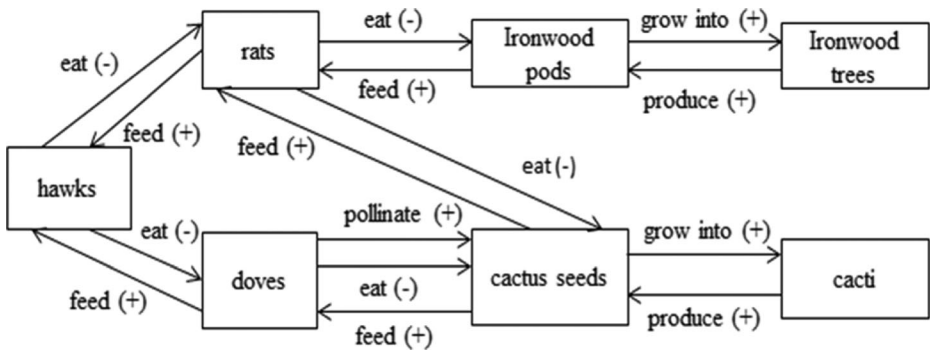


Fig. 4 Expected causal map of the desert ecosystem (all but the pollination link between doves and cactus seed refer to bidirectional predator-prey relations)

R1–R5 correspond to bidirectional “eat” and “feed” links on the causal map. R6 corresponds to a unidirectional “pollinate” link between “doves” and “cactus seeds.” Students used their notes from day 1, and scaffold S6 was used, when needed, to remind students to add all of the relations they had discovered the previous day to complete their model. Scaffolding was also used to make students aware of the bidirectional nature of the food chain relationships;

2. *Reasoning about ecosystem scenarios*—Students were asked three questions about the ecosystem and asked to use their map to generate answers to the questions. An example question asked was “Imagine that a disease killed more than half the doves in the desert ecosystem, how would this affect the rest of the ecosystem?” Students were scaffolded as they used causal explanations derived from their map to answer the questions by encouraging them to verify their predictions by running relevant simulations. To answer our fourth research question (*the link between causal map building and understanding aggregate-level behaviors*), we measured students’ reasoning capabilities for the three questions we asked them before providing them scaffolding or encouraging them to try running relevant simulations to verify their reasoning.

Rationale and Plan for Analysis

In order to answer research questions 1 and 2, the transcribed videotaped responses from the students were used to derive the following information: (a) students’ initial state of knowledge, (b) categorization of the types of scaffolds provided, and (c) a measure of the effectiveness of the scaffolds using the extended Δ -shift framework with the two performance measures: p_1 —the number of correct relationships learnt—and p_2 —the number of incorrect relationships. The goal was to help each student achieve the scores, $p_{1\text{scaf}}=p_{1\text{target}}=6$, where $p_{1\text{scaf}}$ represents the number of correct relations learnt after the scaffolding intervention and $p_{1\text{target}}$ represents all of the relations needed to build a correct model of the desert ecosystem (in this case $p_{1\text{target}}=6$). Similarly, we defined the number of incorrect relations the student had after scaffolding as $p_{2\text{scaf}}$ and ideally, we would like this number to be 0, i.e., $p_{2\text{scaf}}=p_{2\text{target}}=0$. The set of incorrect relations identified during the study are described in “Results.”

For studying research question 3, the pre and posttests were graded and scored, and the overall learning gains as well as the learning gains for each category for questions were tested for significance using paired *t* tests. To answer research question 4, students' answers to the questions asked on day 2 (reasoning about ecosystem scenarios) were scored and used as a measure of their abilities to reason effectively with the new causal map representation. Answering the questions correctly required reasoning through multiple links in the causal model as well as understanding that an ecosystem in equilibrium self-adjusts to small changes in populations. Students were awarded points for each correct link they used to answer a question, and when they identified and combined results from multiple paths to generate a correct answer. They received an additional point if they could explain the nature of dynamic equilibrium in an ecosystem correctly. Our analysis focused on how the students' scores correlated with their responses in the pre- and posttests.

Results

Our findings show that the intervention was effective and produced significant pre- to posttest gains. The scaffolds provided were beneficial and helped students to correct several of their existing incorrect relations. As the intervention progressed, students showed a better understanding of the causal relations that governed the desert ecosystem model and the ecological processes underlying the simulation. Overall, students became proficient in reasoning about complex ecological processes. For example, at the end of the scaffolding phase on day 1, most students came to the realization that “for an ecosystem in balance, the number of each species may keep changing with time, but no species dies off even if the simulation is run for a long time.” They also gained a better understanding about the bidirectional nature of the predator-prey relations between species at different levels of the food chain. Initial responses to questions were mere mentions of outcomes, but as the study progressed, their explanations became more nuanced and mechanistic.

Initial State of Students' Knowledge

At the start of the study, 19 of the 20 students had no prior experience of working with simulation environments, and none of the students had any experience with NetLogo. Thus, all the students started out with more or less similar experience levels, and none of them were at any added advantage based on prior exposure to NetLogo or other simulation environments. Our analysis shows that students' initial attempts to explain the simulation results (initial ideas phase) were primarily guided by their prior knowledge about the species in the desert ecosystem and about ecosystems in general. Most students described an ecosystem as a habitat where animals and plants ate, reproduced, and died. Their explanation of an ecosystem in balance emphasized equal numbers of each of the species and that these numbers did not change with time. Also, when students initially identified a predator-prey relation and were asked to describe the relation, their answers generally reflected a one-directional “eat” relation (as opposed to a bidirectional predator-prey relationship).

Student's initial descriptions contained an average of 1.45 correct relations (see Table 6 in “Measuring the Effectiveness of the Scaffolds”). The initial ideas also revealed several incorrect relations (IR1–IR7), which are summarized in Table 2 along with the number of students who stated that incorrect relation. Initially, the

Table 2 Incorrect relations and their frequency of occurrence

Incorrect relations		Number of students
IR1	Birds do not eat other birds	3
IR2	Species prey on members of its own population	2
IR3	Equal number of species of each kind needed to maintain equilibrium	4
IR4	The population of any particular species stays constant over time in an ecosystem in equilibrium	9
IR5	Dispersal of seeds increases the total number of seeds in an ecosystem	6
IR6	Bird waste provides water for seeds to grow	6
IR7	Plants are not dependent on animals	3

average number of incorrect relations per student was 1.6, and this number was fairly consistent across achievement profiles, as seen in Table 6. The incorrect relations were linked to students' prior knowledge and required repeated scaffolding to overcome.

Some of the difficulties faced by the students in identifying the correct relations could be attributed to conflicts with their prior knowledge. Some of the pre-existing incorrect relations impeded students' abilities to learn the correct target relations R1–R6. For example, students who exhibited IR1 (“Birds do not eat other birds”) found it difficult to learn and remember R5 (“Hawks prey on doves”). Also, IR5 “Dispersal of seeds increases the total number of seeds in an ecosystem” and IR6 “Bird waste provides water for seeds to grow” got in the way of correctly understanding relationship R6 “Doves help pollinate seeds.” Other difficulties in identifying the correct relations originated in students' choice of inappropriate simulation parameters like initial population values. For example, students often deduced a conclusion from a single simulation run without making any attempts to study the generality of the conclusion. Sometimes, the initial population values chosen for a simulation run produced results, which made it hard for students to identify the correct relations. Trying to run the simulation multiple times with different initial population values would have helped students in these cases. Also, most students were not aware of how to control variables and study the interactions between pairs of species at a time. Also, they were not always able to delineate individual relationships from the overall simulation plots that represented the interactions among the five species.

Once students had provided their initial descriptions about relations between the different species in the ecosystem, the interviewer asked questions whose answers required the students to reason in causal chains. Some questions required reasoning in forward chains, e.g., “If hawks increase, the rat population will decrease, and this will increase the amount of uneaten ironwood pods.” Other questions, such as “Knowing that hawks eat rats and rats eat pods, what would happen if hawks were removed from the ecosystem?” required students to be aware of the bidirectional nature of relations in the food chain. When asked such questions, all students initially identified only the unidirectional nature of the relationships, e.g., “If there were no hawks to eat the rats, rats would increase. So, pods would decrease and disappear soon.” The students did not reason further about the consequences of the lack of pods on the population level of rats. We will see in “[Effectiveness of Causal Modeling and Reasoning Using Another Representation](#)” that this is in stark contrast to students' responses at the end of the intervention when they take into account bidirectional relations while answering the same or similar questions.

Scaffolding Analysis

Categories of Scaffolds

Based on the interviewer's field notes and a post hoc analysis of the interview transcripts, we categorized the scaffolds provided according to their purpose and the conditions under which they were provided. We identified seven categories of scaffolds (S1–S7) on day 1 (see Table 3) when students interacted with the simulation and learnt how species in the desert ecosystem were related to each other. Most students required a combination of the categories of scaffolds before they were able to identify all of the relationships.

Our categorization of scaffolds for inquiry-based learning in science using a MABM bears similarities with the frameworks provided by Quintana et al. (2004) and Guzdial (1994). For example, scaffolds S1 and S2 are similar to Quintana's "sense-making" scaffolds; S3, S6, and S7 can be categorized as "process management" scaffolds; and S4 and S5 are scaffolds for "articulation and reflection." In Table 3, we have linked each of our scaffolds S1–S7 to similar or related scaffolds mentioned in the literature. In the experimental study, S3, S4, S6, and S7 were provided as initial scaffolds to help students understand the approach for learning from the simulation environment. Each scaffold was then used, based on the triggering conditions identified in Table 3, in order to help the students work through their learning tasks in the simulation environment.

In some cases, we had to provide multiple scaffolds to address difficulties pertaining to a single idea. This was necessary for two reasons: (1) Conceptual challenges pertaining to students' pre-instructional knowledge, in some cases, required asking them to design or set up specific experimental conditions in the simulation as well as asking them to explain what they were noticing in the simulation and (2) to accommodate the variation in students' thinking, in terms of their responses to the same scaffold. This means that although triggering conditions were indeed theoretically determined, because of individual differences, the triggering conditions varied across students. We believe that addressing this issue makes teaching using a MABM in classrooms a complex endeavor. Table 4 presents excerpts from an exchange between an experimenter and a student during the scaffolded learning phase on day 1. It depicts how a single incorrect relationship [IR1]—"Birds do not eat birds" necessitates the use of different scaffolds: S3 and S4 at time T2, S1 and S4 at T3, S4 and S5 at T4, and S2 at T5. We found that several students used IR1 initially, and we scaffolded the student by asking them to set up a simulation run with hawks and doves only and then explain the observed results (S3, S4). If this did not help students overcome the incorrect relationship, they were asked to slow down the simulation and make another attempt to explain the resultant graphs (S1, S4). A further scaffold reminded students of a similar simulation involving hawks and rats, which they had interpreted correctly earlier (S4, S5). If this additional level of scaffolding did not work, the interviewer helped the student interpret the graph correctly (S2).

On day 2, we identified four types of scaffolds provided (S8–S11) as students built causal models of the desert ecosystem and then reasoned causally using their models (see Table 5). S8 and S9 helped with construction of the causal model, while S10 and S11 helped with causal reasoning using the model so that students could reason in multi-link chains and understand the notions of interdependence and balance in the ecosystem.

Measuring the Effectiveness of the Scaffolds

The effectiveness of the scaffolds was measured in terms of the increase in the number of correct relations (R1–R6) learnt and the number of incorrect relations (IR1–IR7) that were

Table 3 Categories of scaffolds provided on day 1 and their triggering conditions

Scaffold number	Description of scaffold	Similar scaffolds in the literature	Triggering condition
S1	Setting up a simulation run with the correct parameters	Experimental support (Reid et al. 2003); process management scaffold (Quintana et al. 2004)	(a) Students cannot interpret a simulation run because the simulation speed was too fast to recognize how the changes were occurring or some of the species died off too soon; (b) Students initial choice of population values produce non-standard results; (c) Students do not observe the simulation long enough to observe the effect of interactions among species
S2	Interpreting results of a simulation run	Interpretive support (Reid et al. 2003); sense-making scaffold (Quintana et al. 2004)	(a) Students draw conclusions after a single simulation run with a set of initial values but do not repeat the experiment with different values to confirm the trend; (b) Students cannot correlate individual agent-level behaviors with the aggregate population graphs
S3	Controlling variables systematically to construct the model underlying the simulation	Experimental support (Reid et al. 2003); process management scaffold (Quintana et al. 2004)	(a) Students cannot figure out how to study pairwise inter-species relationships; (b) Students forget which pairs of species they have already studied and what conclusions they have drawn
S4	Self-explanations to support making predictions and what-if analyses	Reflective support (Reid et al. 2003); reflection and articulation scaffold (Quintana et al. 2004)	(a) Students draw an incorrect conclusion or give a correct answer without explanations—using self-explanation prompts to get them to reflect on simulation results; (b) Ask students to make predictions before they run a simulation, giving them opportunities to reflect, and compare their ideas with observed simulation results
S5	Creating cognitive conflict	Reflective support (Reid et al. 2003); reflection and articulation scaffold (Quintana et al. 2004)	Students draw incorrect conclusions about inter-species relationships—they are then reminded them about previous contradictory findings, and simulations are re-run with different parameters to help create conflict
S6	Scaffolding to encourage making notes and using them to monitor progress	Reflective support (Reid et al. 2003); reflection and articulation scaffold (Quintana et al. 2004)	Students do not keep track of their progress or of previous similar or contradictory findings they made on day 1

Table 3 (continued)

Scaffold number	Description of scaffold	Similar scaffolds in the literature	Triggering condition
S7	Scaffolding by providing resources	Interpretive support (Reid et al. 2003); support for domain knowledge (De Jong & Van Joolingen 1998); sense-making scaffold (Quintana et al. 2004)	Students have limited knowledge about certain aspects of the ecosystem and need to read more detailed information about them

Table 4 Experimenter (E) scaffolding a student’s (S) use of an incorrect relationship

Time T1	Initial idea (phase 2)	Student (S): I guess the hawks are eating the rats or maybe the seeds because they can’t be eating the doves Experimenter (E): and why can’t they? S: I mean a bird can’t eat another bird
Time T2	Later, when simulation is run with hawks and doves only	E: What do you think, what’s going on here? S: I don’t think there’s any relationship here. Both the hawks and the doves have nothing to eat and so they die off
Time T3	Scaffold: let’s slow down the simulation	E: Let’s reduce the speed and observe this more closely S: Okay E: Let’s stop it here. See, the hawks stay constant here while the doves go down. Why do you think this is happening? Student 1: Oh, I see. I guess the hawk is actually eating the dove and so... Student 2: I don’t know. I really don’t know
Time T4	For student 2, scaffold: cause cognitive conflict	E: Okay, let’s run the simulation for hawks and rats once again S2: Okay E: You said there was a relationship here, right? S2: Ya, the hawks eat the rats E: and what makes you think so? S2: The rats go down and then when all the rats are gone, the hawks start dying out too E: But, couldn’t it happen that both hawks and rats were dying off due to lack in food? S2: No. Initially the hawks stay constant E: So, getting back to the hawks and dove graph, why do you think the hawks were staying constant there? Effect 1—S2: Okay, okay, I see...I was getting it wrong. The hawks do eat the doves Effect 2—S2: It’s weird. I really don’t understand. Hawks won’t eat doves, so they shouldn’t stay constant. But, I don’t get it
Time T5	If effect 2	E interprets the graph for S2 and provides the correct relationship

overcome. The results show that the scaffolds provided were highly effective. Table 6 reports the average values of the two performance measures p_1 and p_2 , i.e., the number of correct relations learned and incorrect relations unlearned, respectively, for the high and low achieving students as well as for all students. These values were computed at three points: (1) initial ideas phase ($p_{1\text{base}}$ and $p_{2\text{base}}$), (2) after the general scaffolds phase, i.e., when only S3, S4, S6, and S7 were administered ($p_{1\text{intermediate}}$ and $p_{2\text{intermediate}}$), and (3) at the end of the scaffolding phase ($p_{1\text{scaf}}$ and $p_{2\text{scaf}}$). Figures 5 and 6 depict how the number of students who possessed each correct and incorrect relation varied in the different phases.

The performance measures p_1 and p_2 in Table 6 show that the average number of correct relationships contained in students’ responses increased from 1.45 (s.d.=0.83) in the “Initial Ideas” phase to 3.85 (s.d.=1.04) in the “General scaffolds” phase and to 5.8 (s.d.=0.52) at the end of the “Scaffolding” phase ($\text{gain}=4.35$). Simultaneously, the average number of incorrect relationships identified by each student remained constant at 1.6 (s.d.=0.94) from the “Initial

Table 5 Categories of scaffolds introduced on day 2 and their triggering conditions

Scaffold number	Type of scaffold	Triggering condition
S8	Scaffolds to remind students to refer to their notes while building the causal model and model all relations noted down	Students do not consider all of the relationships they have learnt while building the causal model on day 2
S9	Scaffolds reminding students to include the bidirectional nature of food chain relationships	Students model either “predator eats prey” or “prey feeds predator,” whichever way they had noted the relationship previously
S10	Scaffolding by explaining how to reason in chains of links by reasoning through successive single links	Students can only reason with single links, but cannot propagate the effects through a sequence of adjacent links in the chain
S11	Scaffolding by asking students to verify their answers to the reasoning questions by running appropriate simulations	Students predict outcomes of different scenarios in the ecosystem by reasoning with the causal map they have built, but they are not sure whether their predictions are correct or not

Ideas” phase to the “General Scaffolds” phase but then dropped to 0.25 (s.d.=0.44) at the end of the “Scaffolding” phase (*loss*=1.35).

Table 6 shows that the scaffolds had similar effects across achievement profiles but were slightly more beneficial for high-performing students compared to the low-performing students ($t=0.71, p=0.49$ in case of p_1 , and $t=0.76, p=0.47$ in case of p_2). It is interesting to observe that when calculated individually, increase in correct relations and decrease in incorrect relations were slightly higher for high-performing students compared to low-performing students (4.4 versus 4.3, and 1.5 versus 1.2, respectively), but the combined effect (net gain as described in “[Extending the Δ-Shift Framework for MABMs](#)”) was 5.0 and exactly the same for both groups of students. Thus, using the results of our extended Δ-shift framework provides the strong result: *The scaffolds helped high- and low-performing students produce equally strong learning gains.*

Effectiveness of the Intervention in Helping Students Explain Aggregate-Level Phenomena

The effect of the intervention on students’ understanding of aggregate-level phenomena was measured through students’ performances on the pre- and posttests (detailed descriptions of the

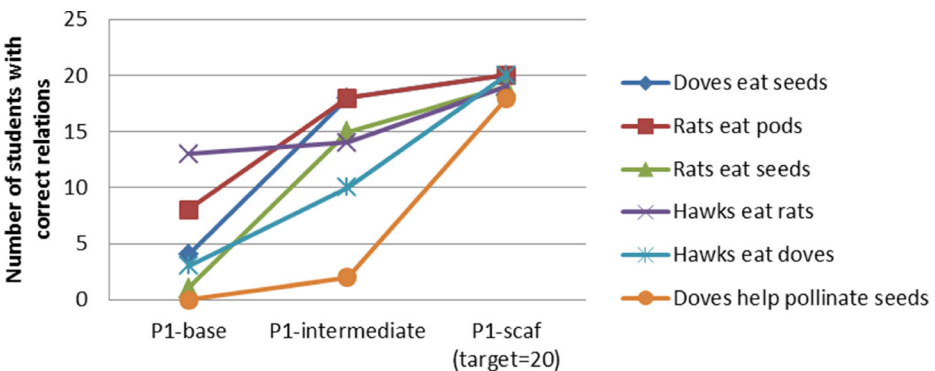


Fig. 5 Effectiveness of scaffolds in helping students learn correct relations

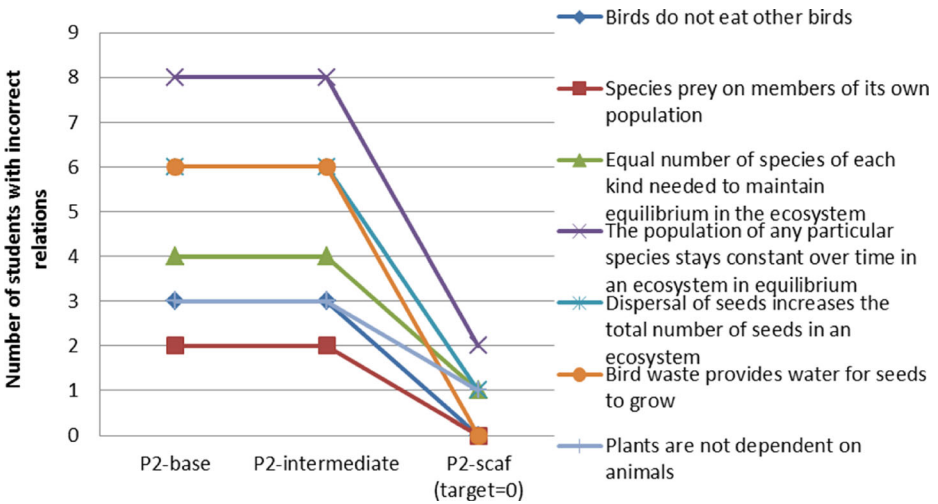


Fig. 6 Effects of scaffolds in helping students eliminate incorrect relations

tests and example questions appear in “Setting and Procedure”). Table 7 shows significant pre- to posttest gains (paired *t* tests with $p < 0.0001$) and large effect sizes (effect sizes > 1.7) for all categories of questions.

Table 6 Effectiveness of the scaffolds by achievement profiles

		Low-performing students mean (s.d.)	High-performing students mean (s.d.)	All students mean (s.d.)
Average number of correct relationships identified	p_{1base}	1.4 (0.84)	1.5 (0.85)	1.45 (0.83)
	$p_{1intermediate}$	3.6 (0.84)	4.1 (1.19)	3.85 (1.04)
	p_{1scaf}^a ($p_{1target=6}$)	5.7 (0.67)	5.9 (0.32)	5.8 (0.52)
	Number of relations learned	4.3	4.4	4.35
	<i>t</i> value	12.65	15.32	19.86
	<i>P</i> value (2-tailed)	<0.0001	<0.0001	<0.0001
	Effect size	5.66	6.85	6.28
Average number of incorrect relationships	p_{2base}	1.7 (1.06)	1.5 (0.85)	1.6 (0.94)
	$p_{2intermediate}$	1.7 (1.06)	1.5 (0.85)	1.6 (0.94)
	p_{2scaf}^b ($p_{2target=0}$)	0.5 (0.53)	0.00 (0.00)	0.25 (0.44)
	Number of incorrect relations unlearned	1.2	1.5	1.35
	<i>t</i> value	3.2	5.58	5.82
	<i>P</i> value (2-tailed)	<0.005	<0.0001	<0.0001
	Effect size	1.43	2.49	1.84
Overall performance gain	5.0	5.0	5.0	

^a The goal is for students to learn all the relations

^b The corresponding goal is for students to have no incorrect relations

Table 7 Paired *t* test results and effect sizes for pre- and posttest scores ($n=20$, $df=19$)

Category	Pre (S.D.)	Post (S.D.)	<i>t</i> statistic	Sig (2-tailed)	Effect size
MCQ (max score=3)	0.95 (0.83)	2.40 (0.5)	8.54	<0.0001	2.12
SQ (max score=40)	18.6 (6.10)	28.10 (4.92)	8.02	<0.0001	1.71
Causal (max score=21)	6.30 (4.38)	15.35 (5.75)	6.24	<0.0001	1.77
Total (max score=64)	25.85 (8.85)	45.85 (9.47)	8.84	<0.0001	2.18

We also computed the pre-to-post gains separately for the high and low-performing students. Figure 7 shows that both high- and low-performing students gained significantly from the intervention. However, the low-performing students gained at a slightly higher rate (average gain of 22 points versus a gain of 18 points for high-performing students). The high-performing students started with a much higher average pretest score compared to the low-performing students (30.4 versus 21.3) but ended with a mean posttest score only slightly higher than the low-performing students (48.4 versus 43.3). This indicates that the intervention was not only beneficial for both groups of students but also helped to narrow the gap between their scores.

Effectiveness of Causal Modeling and Reasoning Using Another Representation

Results indicate that the students scored an average of 34.95 out of a maximum possible score of 38 on the three scenario questions involving causal reasoning. As described in “[Setting and Procedure](#),” an example reasoning question asked was “Imagine that a disease killed more than half the doves in the desert ecosystem, how would this affect the rest of the ecosystem?” At this stage of the intervention, students were able to reason causally taking into account the bidirectional nature of the food chain relations, unlike their initial answers that we described in “[Initial State of Students’ Knowledge](#)”. Table 8 lists the Pearson correlations between normalized pre-post gains for each category of questions and the causal reasoning scores. To avoid ceiling effects, we calculated normalized gains (*NGains*) for each student by dividing the student’s pre-post gain by the maximum amount the student could gain depending on the pretest scores. The results indicate that creating an explicit causal model helped students develop a deep conceptual understanding of the relevant ecological phenomena, and students

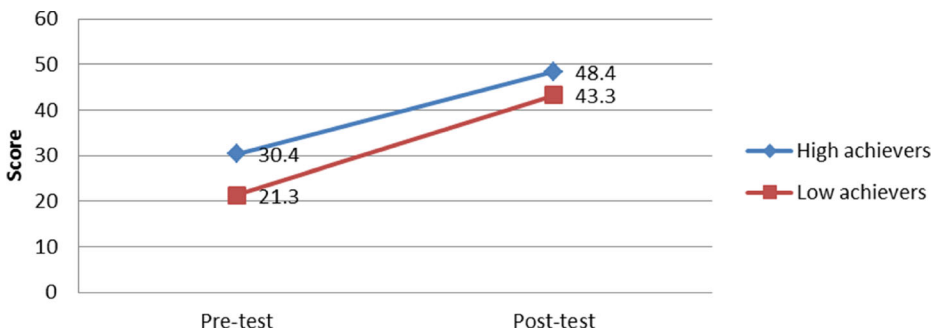
**Fig. 7** Pre and post test scores for high- and low-performing students respectively

Table 8 Correlations between normalized pre-post gains and causal model reasoning scores

	Pearson correlation	Sig (2-tailed)
MCQ NGain	0.41	0.07
SQ NGain	0.79	<0.0001
CRQ NGain	0.71	0.0004
TOTAL NGain	0.95	<0.0001

who reasoned more effectively with the causal map representation showed higher pre- to posttest gains [highly significant ($p < 0.0001$) positive correlation ($r = 0.947$) between causal map score and pre-post test gains].

Discussion and Conclusions

In this paper, we presented a theoretical framework for designing and evaluating scaffolds in a multi-agent simulation-based learning environment for middle school students in the domain of ecology. In order to validate our theoretical framework, we also presented a study we conducted with 20 8th grade middle school students. The study helped us identify the different types of scaffolds required for helping students to learn the appropriate agent-level relationships represented in a desert ecosystem simulation and use them to construct explanations of important aggregate-level ecological phenomena like interdependence and balance among the species in the ecosystem.

During the learning activities, students interacted with several forms of external representations: dynamic visualizations (multi-agent simulation), graphs, and causal maps. They conducted guided experiments using the simulation and developed causal explanations for the observed relations between entities in the ecosystem. Finally, they integrated the individual pieces of causal knowledge into a causal map structure that reflected their understanding of the desert ecosystem. To further test their understanding, the students were asked to provide causal explanations for different scenarios in the ecosystem using the causal model, which included aggregate-level relations between species (e.g., how would a big change in the population of a single species affect the rest of the ecosystem). For this task, students were not allowed to go back and interact with the simulation or the graphs. We found that the students were able to develop agent-aggregate complementary explanations, i.e., their aggregate-level explanations combined use of the causal map as well as individual-level relations that they had verified during the inquiry learning process. This provides evidence for the effectiveness of the scaffolds that were provided for mapping between the agent- and aggregate-level relations, and for helping students identify and establish connections between the causal map representation that captured the aggregate-level relations into a system-level structure, and the MABM simulation that helped them transition from individual agent-level to aggregate-level relations.

More generally, we have presented a theoretical framework for scaffolding analysis that connects the nature of the scaffolds required by students during their interaction with MABMs to the broader literature on scaffolding and simulations in science education. Our paper makes two significant contributions to this end. First, Tables 3 and 5 report the scaffolds we used in our study in response to students' challenges during their learning activities in the MABM simulations. We have listed how each of the scaffolds used during interaction with the MABM corresponds to scaffolds identified previously by other researchers. We have also extended

Sherin et al.'s (2004) Δ -shift framework in two important ways. First, we adopted the framework specifically for MABMs. As Simon (1969) said, the goal of any designed artifact is to bring about an alignment between an existing state and a desired state, by moving the former toward the latter. In the case of learning, the existing state corresponds to students' existing mental models that they enter the instructional setting with, and the desired state corresponds to a target mental model that is more expert-like. More specifically, in the case of learning complex phenomena with MABMs, as we have pointed out in the Literature Review, development of expert-like mental models requires students to develop multi-level explanations, i.e., they need to identify agent-level relationships and extend them to aggregate-level explanations starting from the agent-level relationships. We have argued that Δ -shifts, which are defined as the smallest units of pedagogical action indicating an individual scaffold in Sherin et al.'s framework, should correspond to scaffolds provided for learning both these forms of relationships. Second, we have also argued that different performance measures for evaluating the effectiveness of a scaffold or set of scaffolds may not be independent of one another. So, we proposed a model that enables our extended Δ -shift framework to deal with such multiple, inter-dependent, performance measures simultaneously, rather than treating them as independent performance measures. The net effects of the scaffolds is computed, not as a linear sum of the individual effects, but as the joint effect of the scaffolds on learning correct relations and overcoming related incorrect ones. Based on this extended Δ -shift framework, our analysis reveals that the scaffolds were able to produce significant learning gains and that their combined effects were similar across achievement profiles. The pre-post scores also reveal significant learning gains, and causal reasoning using the causal map representation is shown to have a strong positive influence on these gains.

The design and use of agent-based simulations hold great promise for science education in classroom settings. With this paper, we propose to shift the conversation around agent-based simulations in science education research. Most studies using MABMs thus far have focused on "making the case" for the effectiveness of multi-agent computer models as effective pedagogical tools for teaching and learning complex scientific processes in K-12 settings (Dicke & Sengupta 2013; Tan & Biswas 2007; Goldstone and Wilensky 2008; Wilensky and Reisman 2006; Danish et al. 2011). The emphasis of these studies is on highlighting the *fit* between the computational representational infrastructure of multi-agent-based models, and children's pre-instructional knowledge and representational competence. That is, these studies show that the representational infrastructure of multi-agent-based models makes complex phenomena accessible to young learners without introducing them to equation-based formalisms. We have taken a different position in this paper. As computer scientists and cognitive scientists, we believe that there is now ample evidence that agent-based computational representations can indeed make complex phenomena accessible for young learners. However, we believe that this powerful form of a relatively new pedagogical approach may never be adopted for classroom use in a wide scale if we do not focus on the issue of scaffolding—i.e., identify the forms of scaffolds that students may need and teachers may need to provide as they interact with these simulations. As our analysis shows, scaffolding students in their inquiry using MABMs is a complex endeavor, as this may involve using a combination of multiple scaffolds to address students' conceptual challenges pertaining to one specific learning goal. This is partly due to the nature of the difficulty itself and partly due to the variability of student thinking in a classroom. We believe that our results presented here will guide educational designers of agent-based systems to develop scaffolds that will help bridge the missing link between development of these simulations in the research lab and students productively using these simulations in classrooms in a well-scaffolded manner.

Acknowledgments The authors' acknowledgment support from NSF Cyberlearning grant # 1237350 for this work.

References

- Ainsworth, S. E. (2006). DeFT: a conceptual framework for learning with multiple representations. *Learning and Instruction, 16*(3), 183–198.
- Azevedo, R., Johnson, A., Chauncey, A., & Graesser, A. (2011). Use of hypermedia to convey and assess self-regulated learning. In B. Zimmerman & D. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 102–121). New York: Routledge.
- Bar-Yam, Y. (1997). *Dynamics of complex systems* (Vol. 213). Reading, MA: Addison-Wesley.
- Basu, S., & Biswas, G. (2011). *Multiple representations to support learning of complex ecological processes in simulation environments* (In proceedings of the 19th international conference on computers in education). Thailand: Chiang Mai.
- Basu, S., Biswas, G., and Sengupta, P. (2011). Scaffolding to support learning of ecology in simulation environments. In Susan Bull and Gautam Biswas (Eds.), *Proceedings of the 15th International Conference on Artificial Intelligence in Education*. Auckland, New Zealand.
- Bernstein, A. C., & Cowan, P. A. (1975). Children's concepts of how people get babies. *Child Development, 46*, 77–91.
- Biswas, G., Jeong, H., Kinnebrew, J., Sulcer, B., & Roscoe, R. (2010). Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Research and Practice in Technology-Enhanced Learning, 5*(2), 123–152.
- Blikstein, P., & Wilensky, U. (2009). An atom is known by the company it keeps: a constructionist learning environment for materials science using multi-agent simulation. *International Journal of Computers for Mathematical Learning, 14*(1), 81–119.
- Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: a simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education, 24* (pp. 61–101). Washington DC: American Educational Research Association.
- Carey, S. (1985). *Conceptual change in childhood*. Cambridge: MIT Press.
- Chen, Z., & Klahr, D. (1999). All other things being equal: acquisition and transfer of the control of variables strategy. *Child Development, 70*(5), 1098–1120.
- Chi, M. T. H. (2000). Self-explaining expository texts: the dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161–238). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Chi, M. T. H. (2005). Common sense conceptions of emergent processes: why some misconceptions are robust. *Journal of the Learning Sciences, 14*, 161–199.
- Chi, M. T. H., & Ferrari, M. (1998). The nature of naïve explanations in natural selection. *International Journal of Science Education, 20*(10), 1231–1256.
- Chi, M. T. H., Slotta, J. D., & deLeeuw, N. (1994). From things to processes: a theory of conceptual change for learning science concepts. *Learning and Instruction, 4*, 27–43.
- Clark, D., Nelson, B., Sengupta, P., & D'Angelo, C. (2009). *Rethinking science learning through digital games and simulations: genres, examples, and evidence*. Washington DC: National Research Council.
- d'Apollonia, S., Charles, E., & Boyd, G. (2004). Acquisition of complex systemic thinking: mental models of evolution. *Educational Research and Evaluation, 10*(4), 499–521.
- Danish, J. A., Peppler, K., Phelps, D., & Washington, D. (2011). Life in the hive: supporting inquiry into complexity within the zone of proximal development. *Journal of Science Education and Technology, 20*(5), 454–467.
- Darwin, C. (1871). *The descent of man, and selection in relation to sex* (1st ed.). London: John Murray, ISBN 0801420857, retrieved 2009-06-18
- de Jong, T., & Njoo, M. (1992). Learning and instruction with computer simulations: learning processes involved. *Journal of Research in Science Teaching, 30*(8), 821–844.
- de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*, 179–202.
- Dickes, A., & Sengupta, P. (2013). Learning natural selection in 4th grade with multi-agent-based computational models. *Research in Science Education, 43*(3), 921–953.
- Dillenbourg, P. (1999). Collaborative learning: cognitive and computational approaches. *Advances in Learning and Instruction Series*. Elsevier Science, Inc., PO Box 945, Madison Square Station, New York, NY 10160-0757.

- diSessa, A. A. (1993). Towards an epistemology of physics. *Cognition and Instruction*, 10(2–3), 105–225.
- Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *Journal of the Learning Sciences*, 14, 69–114.
- Goldstone, R. L., & Wilensky, U. (2008). Promoting transfer by grounding complex systems principles. *Journal of the Learning Sciences*, 17(4), 465–516.
- Gupta, A., Hammer, D., & Redish, E. F. (2009). *Ontology of physics concepts: variability and local coherences*. AERA 2009 Annual Meeting, San Diego, CA.
- Guzdial, M. (1994). Software-realized scaffolding to facilitate programming for science learning. *Interactive Learning Environments*, 4(1), 1–44.
- Hendrix, J. R., Mertens, T. R., & Baumgartner, R. S. (1981). Individualizing instruction through concept assessment. *American Biology Teaching*, 43(5), 246–253.
- Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science*, 28(1), 127–138.
- Holland, J. (1995). *Hidden order: how adaptation builds complexity*. Cambridge: Perseus.
- Holland, J. (1998). *Emergence: from chaos to order*. Reading, MA: Perseus Books.
- Ingaki, K., & Hatano, G. (2002). Life, liveliness, and living kinds: how young children think about the biological world. Review of “Young Children’s Naive Thinking about the Biological World”. *International Journal of Behavioral Development*.
- Jacobson, M. J. (2001). Problem solving, cognition, and complex systems: differences between experts and novices. *Complexity*, 6(3), 41–49.
- Jacobson, M. J., & Wilensky, U. (2006). Complex systems in education: scientific and education importance and implications for the learning sciences. *The Journal of the Learning Sciences*, 15(1), 11–34.
- Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379–424.
- Kapur, M., & Kinzer, C. K. (2009). Productive failure in CSCL groups. *International Journal of Computer-Supported Learning*, 4(1), 21–46.
- Kauffman, S. (1995). *At home in the universe: the search for laws of self-organization and complexity*. New York: Oxford University Press.
- Klopfer, E. (2003). Technologies to support the creation of complex systems models—using StarLogo software with students. *Biosystems*, 71(1), 111–122.
- Kozma, R. B., Russell, J., Jones, T., Marx, N., & Davis, J. (1996). The use of multiple linked representations to facilitate science understanding. In S. Vosniadou, R. Glaser, E. De Corte, & H. Mandel (Eds.), *International perspective on the psychological foundations of technology-based learning environments* (pp. 41–60). Hillsdale: Erlbaum.
- Leelawong, K., & Biswas, G. (2008). Designing learning by teaching agents: the Betty’s brain system. *International Journal of Artificial Intelligence in Education*, 18(3), 181–208.
- Lehrer, R., & Schauble, L. (2004). Modeling natural variation through distribution. *American Educational Research Journal*, 41(3), 635–679.
- Lehrer, R., & Schauble, L. (2006). Cultivating model-based reasoning in science education. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 371–388). New York: Cambridge University Press.
- Lehrer, R., & Schauble, L. (2008). *A modeling foundation for reasoning about evolution*. Presentation to the Consortium for Policy Research in Education Learning Progression Working Group, Ann Arbor, MI.
- Lemke, J. L. (2000). Across the scales of time: artifacts, activities, and meanings in ecosocial systems. *Mind, Culture, and Activity*, 7(4), 273–290.
- Levy, S. T., & Wilensky, U. (2008). Inventing a “mid-level” to make ends meet: reasoning through the levels of complexity. *Cognition & Instruction*, 1–47.
- Mataric, M. J. (1993). Designing emergent behaviors: from local interactions to collective intelligence. In *Proceedings of the Second International Conference on Simulation of Adaptive Behavior* (pp. 432–441), August.
- Maxwell, J.C. (1871). Remarks on the mathematical classification of physical quantities. *Proceedings of the London Mathematical Society*, s1–3, 224–233.
- McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students’ construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of the Learning Sciences*, 15(2), 153–191.
- Metz, K. E. (1995). Reassessment of developmental constraints on children’s science instruction. *Review of Educational Research*, 65(2), 93–127.
- Mitchell, M. (2009). *Complexity: a guided tour*. New York: Oxford University Press.
- Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, 19(3), 309–326.
- Okebukola, P. A. (1990). Attaining meaningful learning of concepts in genetics and ecology: an examination of the potency of the concept-mapping technique. *Journal of Research in Science Teaching*, 27, 493–504.
- Papert, S. (1980). *Mindstorms: children, computers and powerful ideas*. New York: Basic Books.

- Pathak, S. A., Jacobson, M. J., Kim, B., Zhang, B., & Feng D. (2008). *Learning the physics of electricity with agent based models: paradox of productive failure*. Paper presented at the International Conference in Computers in Education. Oct. 27–31 Taipei.
- Penner, D. E. (2001). Cognition, computers, and synthetic science: building knowledge and meaning through modeling. *Review of Research in Education*, 25, 1–37.
- Penner, R., Steinmetz, E., Soken, N., Nelson, K., & Whitlow, S. (2000). Model-based design automation for mixed-initiative interactions with complex digital control systems. *International Journal of Systems, Man, and Cybernetics*.
- Piaget, J. (1929). *The child's conception of the world*. London: Routledge & Kegan Paul.
- Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., et al. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences*, 13(3), 337–386.
- Reid, D. J., Zhang, J., & Chen, Q. (2003). Supporting scientific discovery learning in a simulation environment. *Journal of Computer Assisted Learning*, 19(1), 9–20.
- Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B. K., Steinmuller, F., & Leone, A. J. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. M. Carver, & D. Klahr (Eds.), *Cognition and instruction: twenty-five years of progress* (pp. 263–305). Mahwah, NJ: Erlbaum.
- 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.
- Resnick, M., & Wilensky, U. (1993). *Beyond the deterministic, centralized mindsets: New thinking for new sciences*. Paper presented at the annual conference of the American Educational Research Association, Atlanta, GA.
- Scheiter, K., Gerjets, P., Vollmann, B., & Catrambone, R. (2009). The impact of learner characteristics on information utilization strategies, cognitive load experienced, and performance in hypermedia learning. *Learning and Instruction*, 19(5), 387–401.
- Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. *The Journal of the Learning Sciences*, 4(3), 321–354.
- Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for learning: the hidden efficiency of original student production in statistics instruction. *Cognition & Instruction*, 22, 129–184.
- Schwarz, B., & Dreyfus, T. (1993). Measuring integration of information in multirepresentational software. *Interactive Learning Environments*, 3(3), 177–198.
- Sengupta, P., & Wilensky, U. (2008). On learning electricity with multi-agent based computational models (NIELS). *Proceedings of the International Conference for the Learning Sciences*.
- 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), 21–50.
- Sengupta, P., & Wilensky, U. (2011). Lowering the learning threshold: multi-agent-based models and learning electricity. In M. S. Khine & I. M. Saleh (Eds.), *Dynamic modeling: cognitive tool for scientific inquiry* (pp. 141–171). New York: Springer.
- Sherin, B., Reiser, B. J., & Edelson, D. (2004). Scaffolding analysis: extending the scaffolding metaphor to learning artifacts. In: *Journal of the Learning Sciences*, 13(3), 387–421.
- Simon, H.A. (1969). *The sciences of the artificial*. [The Karl Taylor Compton lectures.] Cambridge: The MIT Press.
- Slotta, J.D. (2004). The Web-based Inquiry Science Environment (WISE): scaffolding knowledge integration in the science classroom. In M.C. Linn, P. Bell and E. Davis (Eds). *Internet Environments for Science Education*. 203–232. LEA.
- Slotta, J. D. & Chi, M. T. H. (2006). The impact of ontology training on conceptual change: helping students understand the challenging topics in science. *Cognition and Instruction*, 24(2), 261–289.
- Smith, A. (1977). *An inquiry into the nature and causes of the wealth of nations*. University of Chicago Press.
- Stone, C. A. (1998). The metaphor of scaffolding: its utility for the field of learning disabilities. *Journal of the Learning Disabilities*, 31, 344–364.
- Tan, J., & Biswas, G. (2007). Simulation-based game learning environments: building and sustaining a fish tank. In *Proceedings of the First IEEE International Workshop on Digital Game and Intelligent Toy Enhanced Learning* (pp. 73–80). Jhongli, Taiwan.
- Tisue, S., & Wilensky, U. (2004). NetLogo: a simple environment for modeling complexity. In International Conference on Complex Systems (pp. 16–21), May.
- Van Gendt, K. & Verhagen, P. (2001). *Visual testing. Searching for guidelines*. Paper presented at the 24th National Convention of the Association for Educational Communications and Technology, Atlanta, Georgia.
- Van Joolingen, W. R., De Jong, T., & Dimitrakopoulou, A. (2007). Issues in computer supported inquiry learning in science. *Journal of Computer Assisted Learning*, 23(2), 111–119.

- van Someren, M. W., Reimann, P., Boshuizen, H., & de Jong, T. (1998). *Learning with multiple representations. Advances in Learning and Instruction Series*. New York: Elsevier Science, Inc.
- Vygotsky, L. S. (1978). *Mind and Society: The development of higher mental processes*. Cambridge: Harvard University Press.
- Wertsch, J. V. (1985). *Vygotsky and the social formation of mind*. Cambridge: Harvard University Press.
- White, P., A. (1997). Naïve ecology: causal judgments about a simple ecosystem. *The British Journal of Psychology*, 88, 219–233.
- Wilensky, U. (1999). NetLogo (and NetLogo User Manual), Center for Connected Learning and Computer-Based Modeling, Northwestern University.
- 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. *Cognition and Instruction*, 24(2), 171–209.
- 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.
- Wilensky, U., & Stroup, W. (2003). Embedded complementarity of object-based and aggregate reasoning in students developing understanding of dynamic systems. In *annual meeting of the American Educational Research Association, Chicago, IL*.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 17, 89–100.