Rethinking Science Learning Through Digital Games and Simulations: Genres, Examples, and Evidence

Douglas Clark[1], Brian Nelson[2], Pratim Sengupta[1], & Cynthia D'Angelo[2]
Vanderbilt University[1] & Arizona State University[2]

Introduction

Science education in the classroom has traditionally focused on facts and rote learning. This is largely a legacy of behavioristic approaches to teaching and models of instruction that focus on the atomistic components and "building blocks" of a discipline rather than engaging students in the actual practices or processes of the discipline (i.e., harnessing those building blocks in service of a larger goal or purpose) with the assumption that these atomistic building blocks must be mastered before proceeding to overarching processes. Students in classrooms traditionally memorize equations and the names of chemicals and bones in the absence of using that knowledge to explore natural phenomena or engage in the processes of science. This view of science learning has been reinforced and entrenched by the behavioristic orientation of the assessments generally employed to assess students' abilities and learning. These assessments have persisted due in part to the absence of other forms of assessment that match their economic and pragmatic ease of implementation.

These traditional definitions of learning, teaching, and assessment, however, do not align with the national standards for science education (AAAS, 1993; NRC, 1996) and the broader 21st century skills recognized as critical for all citizens (NRC, in press). The NRC report, Taking Science to School (Duschl, Schweingruber, & Shouse (Eds.), 2007, pp. 36-41), synthesizes current perspectives on goals for science learning into four strands.

> "Students who are proficient in science:
> 1. know, use, and interpret scientific explanations of the natural world;
> 2. generate and evaluate scientific evidence and explanations;
> 3. understand the nature and development of scientific knowledge; and
> 4. participate productively in scientific practices and discourse."

Essentially, the first strand focuses on integrated understanding of science concepts and the accompanying content knowledge (which we will subsequently refer to as “conceptual understanding” for brevity). The second strand focuses on processes and skills for gathering, creating, and processing that knowledge (which we will refer to as “process skills”). The third strand focuses on understanding the epistemological nature of that knowledge and how it is developed (“epistemological understanding”). The fourth strand focuses on students' attitudes, identities, self-perceptions, and habits of mind relevant to their participation and engagement in scientific practices (which we will refer to as “attitudes and identity”). Hereafter, we often refer to these collectively as the “TSTS science proficiency strands” or “TSTS 1-4” for brevity. This summary is cursory, and readers should consult the Taking Science to School report for complete descriptions, but this summary clearly underscores the degree to which our current understanding of science proficiency has evolved beyond traditional classroom goals for science learning.

While we will organize goals for science learning in this paper primarily in terms of the NRC TSTS proficiency standards, it is important to point out that these standards align closely with other perspectives on learning that we consider critical. In particular, the Preparation for Future Learning (PFL) approach (e.g., Bransford & Schwartz, 1999; Schwartz, Bransford, & Sears, 2005) focuses on learning in terms of how well that learning supports students as they engage in future tasks and learning. The goals of this perspective align well across the TSTS strands and can inform goals within each strand in terms of desirable organization of students’ knowledge and skills, views on the nature of knowledge and knowledge development, and identities and stances as active learners and inquirers. Similarly, Hammer, Elby, Scherr, and Redish (2005) provide a framework for supporting students’ abilities to apply their understanding to new situations in terms of the activation of knowledge elements and the context surrounding students’ understandings of concepts. Thus, while the NRC TSTS proficiency strands provide our primary framework of goals for science learning, these goals align well with, and should be informed by, other research into the preparation of students for future science decisions, issues, and challenges.

In addition to redefining goals and assessments, we also need to rethink traditional approaches to supporting science learning. Traditional approaches, with their focus on explicit formalized knowledge structures, seldom connect to or build upon people's tacit intuitive understandings. Well-designed digital games and simulations, however, are exceptionally successful at helping learners build accurate intuitive understandings of the concepts and processes embedded in the games due to the situated and enacted nature of good game play (e.g., Gee, 2003). Most commercial games fall short as platforms for learning, however, because they do not help people articulate and connect their evolving intuitive understandings to more explicit formalized structures that would support transfer of knowledge to other contexts. Games and simulations hold the potential to support people in integrating people's tacit spontaneous concepts with instructed concepts, thus preparing people for future learning through a flexible and powerful conceptual foundation of conceptual understanding and skills.

We thus need to rethink traditional classroom approaches to science learning in terms of (1) goals, (2) approaches, and (3) assessment. Digital simulations and games hold much promise in support of this shift in both formal and informal settings. This paper explores the value of simulations and games for science learning by providing (1) overviews, explanations, and working definitions within the context of science learning, (2) theoretical discussions of the potential affordances for science learning in formal and informal settings, (3) detailed examples of simulation and game titles from some of the most promising genres for science learning, and (4) overviews of the current evidence for science learning from research in simulations and games organized in terms of the science proficiency strands of Taking Science To School as well as an additional category focused on design structures. After exploring simulations and games independently, the paper then synthesizes the discussions in terms of future directions for research and development.

What are the “best” tools for the job?

Games and simulations can be thought of as potential tools for learning. Just as there are many genres of tools, there are many genres of games and simulations, each with many exemplars and sub-genres. Different tools are more or less appropriate for certain tasks. The list of genres of tools that we would consider valuable for home construction would differ from the list of genres of tools that we would consider valuable for cooking, gardening, or automotive work.

Within this metaphor, it therefore seems more useful to think about the genres of games and simulations that hold the most promise for supporting science learning than to argue about a finite list of “best” titles for science learning. Ultimately, we want to understand the valuable types of tools to have in our toolbox rather than debate the best brands of screwdrivers or whether a hammer is better than a drill because the answers to such questions are extremely context dependent.

This is particularly relevant for simulations and games given how quickly individual titles become outdated by technological progress and how quickly new iterations and advances in each genre evolve. As a result, the following sections explore several genres of simulations and games that seem valuable to keep in our “toolbox” for science learning. We provide a detailed example of a title within each genre, and list other exemplars from that genre, but these choices should not be taken as proclaiming these titles the “best” above all others. The point in each case involves exploring the potential affordances of each genre for science learning.

Digital Simulations and Science Learning

This paper defines digital simulations as computational models of real or hypothesized situations or phenomena that allow users to explore the implications of manipulating or modifying parameters within the models. Following Schwarz and White (2005), we use the phrase ‘scientific modeling’ to mean a combination of the following processes including (a) embodying key aspects of theory and data into a model — frequently a computer model, (b) evaluating that model using criteria such as accuracy and consistency, and (c) revising that model to accommodate new theoretical ideas or empirical findings.

Theoretical Affordances of Simulations

Unlike laboratory-based experimental setups, as Holland (1998) points out, a theoretical scientific model need not bear any obvious resemblance to the thing being modeled. For example, Newton's equations are symbols confined to a sheet of paper, and do not look like planetary orbits; yet they mechanistically model this physical reality better than any physical model of solar systems. Likewise, the core component of a scientific computer model is that it should model (i.e., represent) the "mechanism(s)" underlying a scientific phenomenon. There are various ways of representing mechanisms. While some computer models are based on graphical representations of equations and qualitative relationships between variables (Jackson, Krajick & Soloway, 2000; Jackson et al., 1996; Shecker, 1993), others allow users to create and/or manipulate objects and/or interactions in the model and dynamically display the results in real time, and/or, in the form of inscriptions such as graphs (e.g., Edelson, Gordin, & Pea, 1999; Adams et al., 2008a, 2008b; diSessa, Hammer, Sherin, & Kolpakowski, 1991; Wilensky & Reisman, 2006; Sengupta & Wilensky, 2009; Frederiksen, White, & Gutwill, 1999; Keller, Finkelstein, Perkins and Pollack, 2006), whereas some make the users themselves parts of the model (Wilensky & Stroup, 1999b; Colella, 2000; Klopfer, Yoon and Rivas, 2005), and some have users learn by teaching an intelligent agent (Biswas, Jeong, Roscoe, & Sulcer, 2009; Schwartz et al., 2007).

Simulations provide leverage in terms of harnessing a user's spatial learning and perceptual systems in ways that text and verbal interactions do not (Lindgren & Schwartz, 2009). Simulations can furthermore be started, stopped, examined, and restarted under new conditions in ways that are sometimes impossible in real situations (Holland, 1998) allowing learners to explore the mechanisms underlying scientific phenomena that they experience in everyday lives (such as hitting a ball, projectile motion, etc.) as well as phenomena otherwise inaccessible in their everyday life (such as microscopic properties of matter, electrical conduction, cell biology, etc.).

Dimensions and Genres for Promising Simulation for Science Learning

Simulations for science learning vary along a number of dimensions including (a) the degree of user control, (b) the extent and nature of the surrounding guiding framework in which the simulations are embedded, (c) how information is represented, and (d) the nature of what is being modeled. This list of dimensions is not exhaustive, but provides insights into the range of productive simulations available.

One dimension involves the degree of user control provided. Simulations can provide a large range of user control, from “glass box” models with full user control and programmability to targeted simulations that focus user control on specific variables. Targeted simulations (e.g., many stand-alone Physlets, PhET, and TEAL simulations as well as many of the simulations embedded in digital inquiry environments such as WISE or Pedagogica) provide the user with a specific set of choices to focus attention on key dynamics of interest. This approach provides powerful affordances in terms of implementation and integration. Targeted simulations (1) minimize training time for effective use by students and teachers, (2) support effective exploration and inquiry in short periods of curricular time, (3) focus users on the specific phenomena and interactions of interest, and (4) provide high levels of flexibility for integration into existing and new curricula.

An intermediate level of user control is available in “sandbox” simulations that do not allow the user to modify the programming, but that provide a wide range of controls and modifiability to support open-ended exploration (e.g., SimEarth, SimCity, SimAnt, SimFarm, Interactive Physics, Geode Initiative). Sandbox simulations require more training time for users than targeted simulations and more curricular time for implementation, but allow greater flexibility for conducting open-ended inquiry.

"Glass box" models can provide interfaces for manipulating specific variables, but also allow users to modify the underlying code that generates the model behaviors (e.g., NetLogo (Wilensky, 1999) and StarLogo). An affordance of this genre is that learners can develop more advanced models by modifying the existing code after starting out with some simpler pre-built simulations, or, build new models from scratch using intuitively designed Logo-based programming languages. This involves tradeoffs, however, in terms of significantly higher training times for learners as well as requiring more time within the curriculum for productive implementation.

A fourth genre of user control involves networked participatory simulations where control is spread across multiple connected users (e.g., HubNet (Wilensky & Stroup, 1999a), Live Long and Prosper, and ARMS). As noted by Roschelle (2003), participatory simulations provide separate devices for each student (or each small group of students) and facilitate data exchanges among devices. Overall patterns emerge from local decisions and information exchanges (Roschelle, 2003). These involve levels of control for individual users similar to targeted simulations, but spreads overall control across the group. Participatory simulations have been used in the classroom to enable students to model and learn about the many decentralized scientific phenomena such as swarming ants, epidemics, traffic jams, and flocking birds.

A second dimension for simulations focuses on the extent and nature of the surrounding guiding framework in which the simulations are embedded. Some simulations are relatively stand-alone, allowing users relatively direct access to the simulation with minimal curricular support or constraint. Many Physlets, TEAL, and PhET simulations fall into this category. These simulations allow the instructor to freely integrate them into any other curriculum including hands-on experimentation (see for example, the PhET Electricity simulation, TEAL Electromagnetism simulations, etc.).

Other simulations are embedded in larger contextual frameworks or platforms to guide the user's progress, inquiry, and reflection through one or more simulations. These provide more curricular support for the users in exploring the embedded simulations but are less flexible for integration into other curricula and require more curricular time than standalone simulations. These are typically curricular and/or technological*platforms* (e.g., TELS and Pedagogica) in which simulations (computer models) and/or suites of simulations (or computer models) can be integrated with other tools such as journaling, discussion, brainstorming, probeware data collection, sharing, drawing, and concept mapping activities.

Platforms that come with their own programming environments (such as NetLogo, StarLogo, Molecular Workbench, HubNet) can also be used to program their own surrounding platform for simulations.

A third dimension for simulations involves the variety of alphanumeric, graphed, abstract iconic, and representative iconic representations of information. Most simulations of scientific phenomena involve more than one of these types of representations, but often focus heavily on a subset of these representations. Tradeoffs among these formats are numerous and ultimately depend on the goals of the designers and the nature of the phenomenon being modeled.

The fourth dimension for simulations involves what is actually being modeled and how. This dimension is conceptually the most complex and can be subdivided into four genres: (1) behavior-based, (2) emergent, (3) aggregate models, and (4) composite models of skills and processes.

Behavior-based models typically involve objects and interactions between objects (that are part of the simulation), which users can manipulate and interact with through assigning and/or modifying behaviors or systemic constraints. For example, using the sandbox Interactive Physics simulation environment for physics, learners can create objects of their choice and add behaviors (e.g., movement) and constraints (e.g., gravity and other forces), and then observe the results and conduct further investigation of the phenomenon being modeled. The difference between these simulations and other object-based simulations (such as the ones described in the following two categories) is that they are usually a black box to the learners. The advantage is that the simplified intermediate model that simulations of this type can help students create and integrate may be more accessible to the learners than more detailed but more complicated mechanisms (Lewis, Stern, & Linn, 1993; Linn & Hsi, 2000; White, 1993a, 1993b; White & Fredericksen, 1998).

Emergent, Multi-agent Based Models and Models-based Curricular Units typically model complex *emergent* systems. Emergence is the process by which collective behavior arises out of individuals' properties and interactions, often in non-obvious ways. Such systems in which a coherent, higher-level (i.e., aggregate-level) phenomenon arises out of simple and *de-centralized* interactions between many individual agents or actors are known as *emergent* systems, and the models that represent emergent phenomena are called *emergent* models.

Aggregate modeling (also known as systems dynamics modeling) typically models aggregate-level behavior of complex systems using various forms of representations such as the semiotics of systems dynamics (stocks and flows), and/or graphical representations of qualitative models of aggregate-level behavior. An example is Stella, which is a platform for system dynamics modeling. The "primitives" in Stella are computational objects (e.g., stocks, spigots, circles, and arrows) that can be combined to develop fairly sophisticated models. The Stella model consists of a networked diagrammatic representation of links between these different objects, which is then converted into the mathematical equation represented by the diagram by the software itself. Running the model results in graphs and data tables of the systemic behavior. In addition to mathematical relationships in the form of equations, one can also input qualitative relationships between variables.

Composite models of processes and skills allow users to train for complex tasks in simulated environments. These environments were originally developed for military training, but have spread to medical and general educational training settings as well, such as conducting the launch of a NASA mission (e.g., the ARMS example discussed later in this paper), a chemistry experiment (e.g., ChemLab), or dissecting a frog (e.g., Froguts). They allow the user to engage in the practices associated with the task or practice through an underlying set of interconnected models.

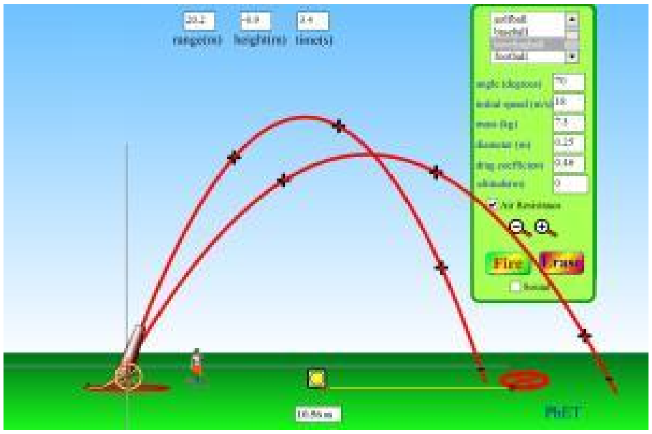
Examples of Productive Simulations for Science Learning

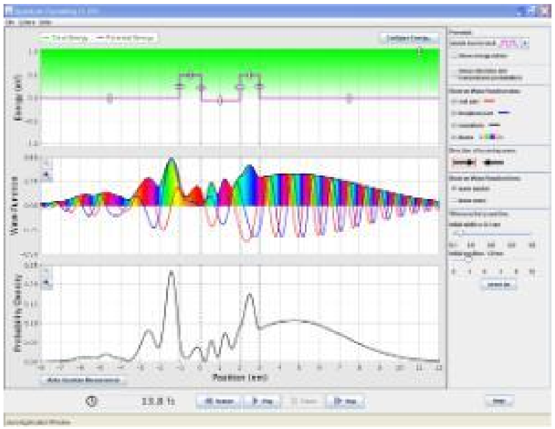
The following sections present detailed examples of the dimensions and genres of productive simulations for science learning described above. As discussed earlier, the choices of examples are intended to exemplify the affordances of the genres rather than to claim that these examples are "the best." Obviously a simulation activity designed for use as a recreational title will benefit from a different combination of choices along each dimension than a title intended to supplement instruction in a classroom or a title intended to provide stand-alone instruction over an extended timeframe. Thus, there is no one "best" choice but rather careful choices to match the characteristics of the simulation to the user context intended. URLs to two-minute video overviews on YouTube.com are provided for several of the examples at:

http://sites.google.com/site/nrcsciencegamessims/

Example. PhET Interactive Simulations

Overview. PhET (http://phet.colorado.edu) provides extensive suites of targeted stand-alone simulations for physics, chemistry, biology, earth science, and math. PhET teams of scientists, software engineers, and science educators use a research-based approach to create simulations that support student engagement with and understanding of scientific concepts. PhET simulations animate what is invisible to the eye through the use of graphics and intuitive controls such as click-and-drag manipulation, sliders, and radio buttons. The simulations also offer measurement instruments such as rulers, stop-watches, voltmeters, and thermometers to encourage inquiry.

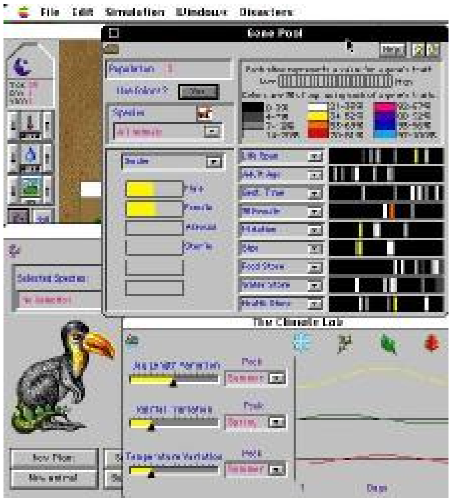


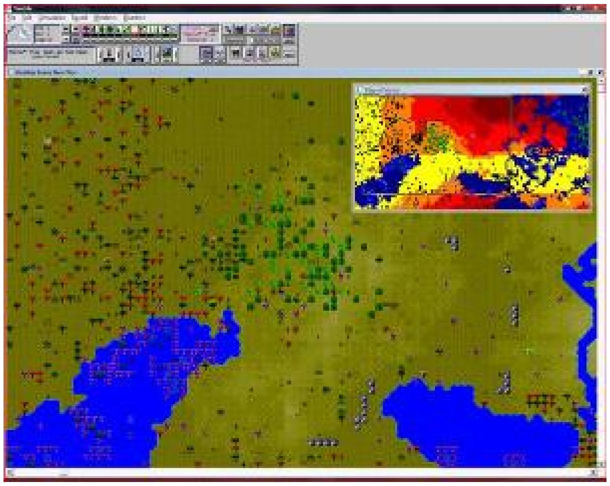


Value for Science Learning. PhET provides one of the largest online libraries of simulations for use in science instruction. These simulations are designed to be targeted and stand-alone to facilitate incorporation into existing curricula across grade levels with minimal prior training for students or teachers. They can thus supplement existing curricula or form the core of new inquiry projects. PhET simulations can engage students in inquiry-based learning activities as well as foster scientific discussion among student-peers in a classroom setting.

Example. Maxis SimLife

Overview. SimLife: The Genetic Playground is a sandbox, stand-alone, multiagent-based commercial “edutainment” title from the early ‘90s that has subsequently been re-released as a virtual console game on various game platforms. SimLife simulates an ecosystem. Players can create and modify the genetics of the plants and animals and design and modify the environments and ecosystems. Players experiment and try to create self-sustaining ecosystems. Players thus can explore the relationships between genetics and the ecosystems, evolution, and what types of traits might help creatures thrive.

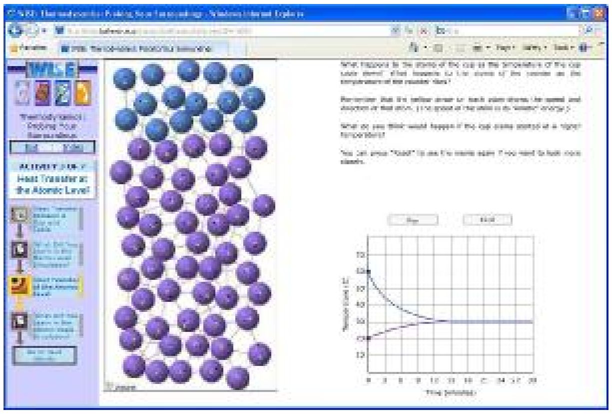


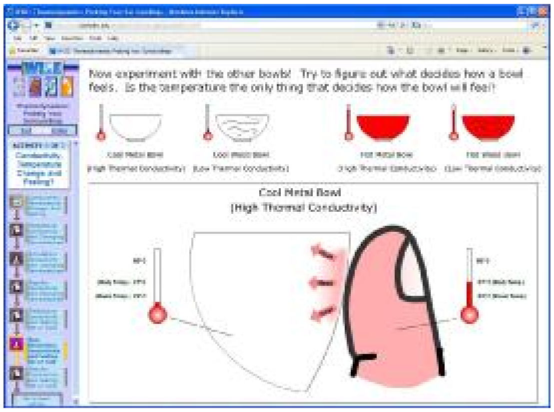


Value for Science Learning. SimLife and other related commercial entertainment titles, such as virtual aquariums, SimEarth, SimAnt, SimCity, and other related titles allow users to explore core scientific, socio-scientific, and engineering issues in such an engaging manner that users are willing to purchase the simulations and play them recreationally.

Example. The Web-Based Inquiry Science Environment (WISE)

Overview. The WISE environment (http://wise.berkeley.edu) is a framework and toolset for engaging students in the intentional process of diagnosing problems, critiquing experiments, distinguishing alternatives, and planning investigations with simulations, discussion tools, journaling and note taking tools, drawing tools, sharing tools, and several other tools. In the WISE: Probing your Surroundings project, for example, students collect real time data about the temperatures of objects and explore interactive simulations dealing with heat transfer, thermal conductivity, and thermal sensation. Students then work in online forums toward a consensus in explaining the patterns observed in the empirical data.



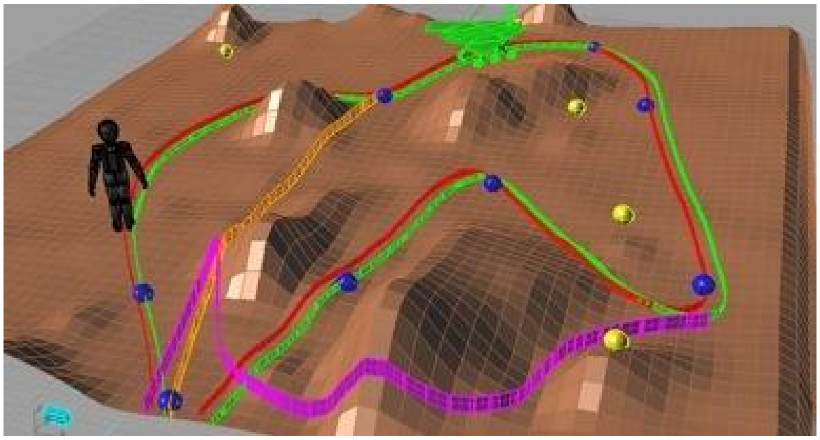


Now experiment with the other bowls! Try to figure out what decides how a bowl feels. Is the temperature the only thing that decides how the bowl will feel?

Value for Science Learning. Partnerships between researchers and teachers use the authoring environment to create these inquiry projects with a specific focus on supporting students engaging with simulations over the past several years. Frameworks like WISE provide scaffolds to students and teachers engaging across all four TSTS science proficiency strands from content knowledge to identity and attitudes. Similar excellent frameworks have been developed by Concord Consortium (http://concord.org).

Example. Astronaut Robot Mission Simulator

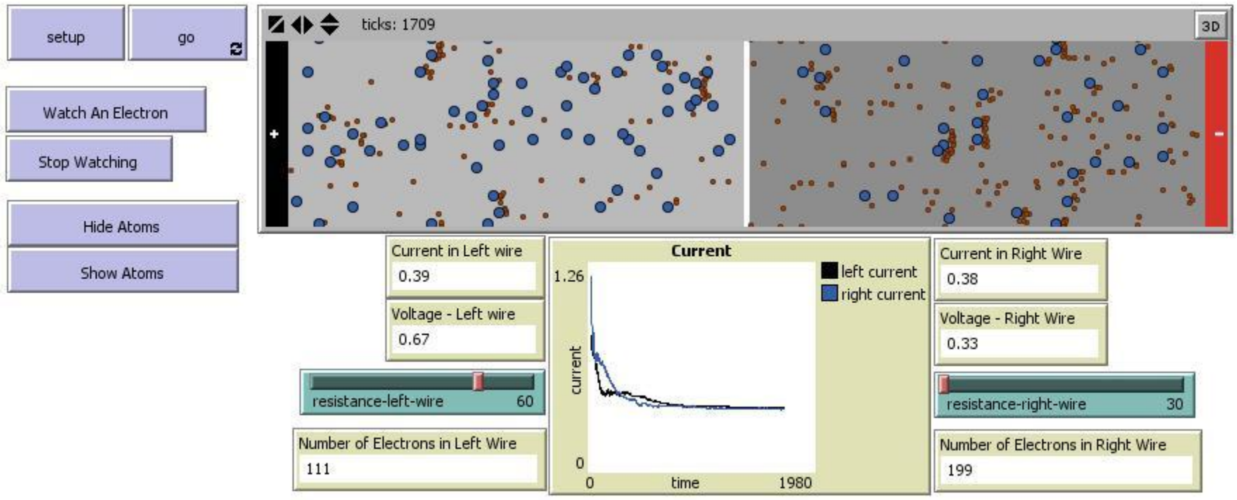
Overview. Astronaut Robot Mission Simulator (ARMS) developed by the Motivational Environments research group at Arizona State University (http://ame2.asu.edu/projects/intrinsic/) in collaboration with the School of Earth and Space Exploration simulates actual planetary exploration missions. The multiplayer participatory simulation provides personally tailored interfaces for each of several mission roles (e.g., robotic engineer, astronaut, physician, geologist, biologist, psychologist, mission control, etc.). The simulation engages all phases and aspects of extra-vehicular activity. Explorer-learners work as a team to safely and optimally return scientific understanding from the Moon and Mars (for comparison, they can conduct missions to diverse Earth biomes, too), learning first-hand about the transdisciplinary creative processes required to advance science, engineering, and exploration.



Value for Science Learning. This multiplayer simulation provides an excellent opportunity for students to engage in roles and science that would otherwise be impossible to experience. Single and multiplayer training simulations of this type allow participants to engage in otherwise dangerous, lengthy, or expensive investigations. Such simulations have also been created by other research groups, companies, and government organizations for medical procedures (e.g. surgical simulators such as http://www.public.asu.edu/~kkahol/kanavDESIGNDEV.html), chemistry labs (e.g., http://modelscience.com), frog dissection (e.g., http://froguts.com, http://digitalfrog.com), and other topics mirroring pioneering simulation work by the military. Simulations such as these engage the participant more directly as an actor participating in the skills and processes of a profession or role and thus support participants across all four TSTS science proficiency strands.

Example. NIELS (NetLogo Investigations in Electromagnetism)

Overview. NIELS (http://ccl.northwestern.edu/NIELS) is a suite of *low-threshold* and *high-ceiling* emergent embedded models built on the NetLogo modeling environment that can be used by students across different ages (fifth grade through undergraduate). These models represent electric current and resistance in simple linear circuits as phenomena that emerge from simple rules of interaction (push, pull, collisions, or bouncing) between thousands of individual level agents (such as electrons and atoms) inside the conducting material. NIELS models are based on Drude's free electron theory, and enable learners to develop a qualitative sense-of-mechanism of the aggregate-level formal representations (such as equations and graphs representing Ohm's Law).



ticks: 1709

Current in Left wire 0.39

Voltage - Left wire 0.67

resistance-left-wire 60

Number of Electrons in Left Wire 111

Current in Right Wire 0.38

Voltage - Right Wire 0.33

resistance-right-wire 30

Number of Electrons in Right Wire 199

Value for Science Learning. Although electricity has been regarded as "notoriously difficult" for students at all levels (middle-school through college) to learn, NIELS enables students as young as fifth grade to understand and explain the concepts of electric current and resistance, their relationships, as well as the behavior of novel electrical circuits.

Example: HubNet

Overview. HubNet is a computer architecture that enables using NetLogo to run *participatory simulations* in the classroom. In a participatory simulation, a whole class takes part in enacting the behavior of a system as each student controls a part of the system by using an individual device, such as a networked computer or Texas Instruments graphing calculator. For example, in the Gridlock simulation, each student controls a traffic light in a simulated city, while the class as a whole tries to make traffic flow efficiently through the city. As the simulation runs, data is collected which can afterwards be analyzed on a computer or calculator. Students engaged in participatory simulations act out the roles of individual system elements and then see how the behavior of the system as a whole can emerge from these individual behaviors.

Value for Science Learning. Such distributed learning environments can make very difficult ideas around ‘distributed systems’ and ‘emergent behavior’ more accessible to students. Researchers have shown that such ideas are often very difficult for students to understand without specialized instruction. Students have rich conceptual resources for reasoning about and thoughtfully acting in playful spaces, and thus can more easily become highly engaged in the subject matter.

Evidence about Simulations and Learning

As discussed earlier in this paper, the NRC’s Taking Science to School report (2007) describes proficiency in science in terms of four strands. We now present evidence for simulations in science learning organized around these strands. The four categories organize evidence in terms of (1) conceptual and process skills learning (TSTS 1 & 2), (2) epistemological understanding (TSTS 3), (3) attitudes, identity, and motivation (TSTS 4), and (4) an additional category focusing on optimal structuring of simulations for learning.

Conceptual and Process Skills Learning (TSTS 1 & 2)

Strong evidence suggests that various types of simulations used in conjunction with appropriate curricula and instruction can foster various aspects of scientific expertise such as model-based reasoning, systems-thinking, construction of scientific explanations, and other conceptual skills and understanding (e.g., Mandinach & Cline, 1993; Raghavan & Glaser, 1995; Richards, Barowy, & Levin, 1992; Edelson, Salierno, Matese, Pitts, & Sherin, 2002; Sandoval et al., 2002; Sandoval, 2003; Schwarz & White, 2005; White & Frederiksen, 1998). Furthermore, many models-based computational learning environments have been shown to be successful in engaging K-12 students in deep reasoning and sophisticated analysis (e.g., Adams, et al., 2008a, 2008b; Edelson, Gordin, & Pea, 1999; Frederiksen, White, & Gutwill, 1999; Goldman-Segall, 1996; Harel & Papert, 1991; Kafai & Harel, 1991; Klopfer, Yoon and Rivas, 2005; Papert, 1980; Roschelle & Teasley, 1995; Rothberg, Sandberg, & Awerbuch, 1994; Sengupta & Wilensky, 2008a, 2009; Tabak & Reiser, 1997; White, 1993b; Wieman, Adams, and Perkins, 2008). Another area of value for simulations is their demonstrated value for helping students create and integrate intermediate models that are more accessible to the learners than a more detailed but more complicated mechanism (Lewis, Stern, & Linn, 1993; Linn & Hsi, 2000; White, 1993a, 1993b; White & Frederiksen, 1998).

In a recent meta-analysis of the effectiveness of simulations and models-based learning environments in science education, Chang, Chiu, McElhaney, and Linn (in preparation) show that dynamic visualizations can support virtual experimentation especially for topics that cannot be investigated in the classroom or using physical objects. Chang, et al. also show that dynamic visualizations can successfully link multiple representations such as those at observable, sub-microscopic, and symbolic levels. In the paragraphs that follow, we provide expanded examples of some of the research studies that support the value of simulations in terms of conceptual and process skills learning.

Keller, Finkelstein, Perkins and Pollack (2006), for example, showed that the PhET Circuit Construction Kit (CCK) simulation, which models the behavior of electric circuits, can be used as an effective tool for engaging students in productive discussions about the modeled phenomena. They found that students who were shown the simulation during the lecture showed a comparatively much higher and statistically significant gain in conceptual understanding after discussing the modeled phenomenon with their peers, compared to students who were shown a physical demonstration or who were provided with an equivalent verbal explanation to discuss with their peers. Other studies corroborate the effectiveness of PhET simulations as instructional tools in undergraduate physics courses (Wieman, Adams, & Perkins, 2008; Adams et al., 2008a, 2008b).

Meir, Perry, Stal, Maruca, and Klopfer (2005) showed that students benefit from the ability to explore diffusion and osmosis at the molecular level. They developed OsmoBeaker to allow students to perform inquiry-based experiments at the molecular level. Their results showed 13 of 15 students in the diffusion lab improved from pretest (mean 4.2/10) to posttest (mean 6.7/10) for a mean 60% improvement ($p < .001$). In the osmosis lab, 23 of 31 students improved. The difference in pretest (10/18) and posttest means (12/18) was significant ($p < .001$).

Clark and Jorde (2004) explored the efficacy of helping students re-explain their intuitive ideas related to thermal equilibrium with a simulation that focused on the role of conductivity in how hot or cold an object feels. Results show that students in the experimental group significantly outperformed control group students on the posttests and delayed posttests (approximately $p < .001$ across measures) of their understanding of thermal equilibrium (a topic that is often confusing for students due to their intuitive ideas about how hot or cold different materials feel). This superior performance not only included their tactile understandings, for which the experimental group received the augmented visualization, but also their understanding of thermal equilibrium, for which the experimental and control group's received the same visualization models. These findings were supported through parallel research looking at overarching research of approximately 3,000 students in the CLP curriculum (Clark & Linn, 2003) as well as through detailed microgenetic longitudinal analysis of four students (Clark, 2006).

Logo-based platforms (including Boxer, NetLogo, and StarLogo) have been used in numerous studies to explore how simulations might help students learn scientific concepts and process skills. diSessa and colleagues showed that novice learners (e.g., middle school students) were able to develop a deep understanding of Newtonian mechanics, as well as meta-representational competence, through using Logo-based modeling platforms such as Boxer (diSessa, 2000; diSessa, Hammer, Sherin, & Kolpakowski, 1991). Wilensky (2003) provides an example in which emergent multi-agent models designed in NetLogo (Wilensky, 1999) enabled middle school students to successfully learn statistical mechanics, a topic that is traditionally taught using equation-based representations in advanced physics courses (college level and beyond).

Similar studies have also been conducted in other domains of science, which show that students can make sense of more advanced content at a younger age using such agent-based forms. For example, Sengupta and Wilensky (2006, 2008a, 2009), in 5th, 7th, and 12th grades as well as in freshmen undergraduate classes show that NIELS enables students as young as 5th graders to (a) alleviate commonly held misconceptions in introductory electricity (for a review of misconceptions, see Reiner, Slotta, Chi, & Resnick, 2000), and (b) construct scientifically correct explanations of the behavior of electric current as represented in traditional circuit diagrams, as well as physical setups of commonly used electrical circuits (Sengupta & Wilensky, 2008a, 2008b, 2009). Another example can be found in Blikstein and Wilensky (2005) in the domain of materials science, and Wilensky and Reisman's (1998, 2006) research in biology, where the authors show how NetLogo-based representations can enable novices to engender an expert-like understanding of key concepts in these domains. Even earlier examples include Papert and colleagues' research on 5th graders learning fractions using Logo (e.g., Harel & Papert, 1991).

Further evidence for the effectiveness of emergent models is provided by Klopfer, Yoon, and Um (2005). These authors conducted a study in which they showed that by constructing and exploring models using another multi-agent based learning environment (StarLogo), fifth and seventh grade students can gain an understanding of the importance of several scientific practices such as repetitive hypothesis revision and testing, develop insights into some key systems concepts such as random variation, and become better able to classify models as a result of their participation. These learning outcomes, the authors argued, "address both the need for school science programs to adopt a more authentic scientific approach to investigation, as well as the call for inclusion of complex systems concepts in school science curricula", and that "this finding shows that many of the principles of complex systems (such as random variation) are not too complicated to be integrated into the classroom, even at the elementary school level" (Klopfer, Yoon, & Um, 2005, p. 175).

In the realm of participatory simulations, some prototypical scenarios for modeling include epidemiology in terms of spread of disease (Colella, 2000; Klopfer, Yoon, and Rivas, 2005), genetics and Mendellian inheritance (Klopfer, Yoon, & Rivas, 2005), traffic systems (Wilensky & Stroup, 2000), statistics (Abrahamson & Wilensky, 2004), and other topics. Colella (2000) showed that through interacting with participatory simulations, students were able to identify problems and construct hypotheses, and develop a keen sense of possible outcomes of the phenomena being modeled. Klopfer, Yoon, and Rivas (2005) implemented two participatory simulations in two Boston area high schools (one public school, N = 71; and one private school, N = 117). They found that after interacting with the participatory simulations, the pooled data of students' self-assessments of learning showed that students highly rated their learning about content (genetics or epidemics) (M = 3.64 / 5), technology (M = 3.72 / 5), and experimental design (M = 3.64 / 5). They also expressed strong agreement with the statement that the technology used positively impacted their learning (M = 3.95 / 5) (Klopfer, Yoon, and Rivas, 2005). Klopfer, Yoon, and Perry (2005) conducted five case studies of teachers at various stages in their careers working with participatory simulations of complex systems using ubiquitous and accessible mobile computing devices. Evidence across their data sources showed enhanced motivation, engagement, and self-directed learning by students as well as a large degree of ease of adaptation of the simulations in terms of subject-matter content knowledge and curricular integration, the simulations' facility in attending to teacher-individualized goals; and shifts by the teachers toward adopting learner-centered strategies.

Epistemological Understanding (TSTS 3)

While many simulation studies have focused on conceptual and process skills learning, fewer studies of simulations have focused specifically on epistemological understanding, although some have included epistemological understanding as a component of their goals and findings. Schwarz and White

(2005), for example, explored these issues in the Model-Enhanced ThinkerTools (METT) Curriculum. METT is an inquiry-oriented physics curriculum for middle school students in which they learn about the nature of scientific models and engage in the process of modeling. The METT Curriculum enabled students to create, evaluate, and discuss computer models of their ideas about force and motion. Results from four science classes in an urban middle school show that the students using the METT Curriculum showed significant improvements in inquiry skills and physics knowledge and formed better conclusions because "they had a better idea of the form that a scientific model should take and of the criteria a good model should meet" (Schwarz & White, 2005, p. 75). Students also performed better on some of the far-transfer problems on the physics test. The authors compared their results from this study with previous studies of students using the ThinkerTools curriculum, and concluded that students' engagement in activities about evaluating and discussing models led to the observed gains.

Attitudes, Identity, and Motivation (TSTS 4)

Research has long supported the claim that simulations-based learning environments can (a) produce high levels of intrinsic motivation; (b) encourage self-directed, learner-controlled exploration of an intellectually rich and diverse goal space; and (c) provide the learner with immediate, clear, and informative feedback (e.g., Lawler, 1982; Lepper & Chabay, 1985; Papert, 1980). Building on this research, recent research on constructivist technology-based learning environment design has argued for the motivational importance of authentic, interesting tasks and contexts (e.g., Cognition and Technology Group at Vanderbilt, 1992; Edelson, Gordin, & Pea, 1999; Adams, et al., 2008a, 2008b). Edelson, Gordin, and Pea (1999), for example, showed that scientific visualization (i.e., techniques used to display and analyze scientific data visually through the systematic variation of color, shape, orientation, and position) in the domain of climatology and earth sciences bootstrap students' interests and motivate them to conduct authentic scientific inquiry using the WorldWatcher curriculum.

Similarly, in the domain of physics, based on a corpus of over 250 interviews with undergraduate physics students, Adams et al. (2008a, 2008b) identified "overwhelming evidence that simulations that suitably incorporate interactivity, animation, and context can provide a powerful learning environment where the students productively engage with and master physics content" (Adams et al., 2008a, p. 31). However, it is important to note that both Edelson, Gordin, and Pea (1999) and Adams et al. (2008a, 2008b) concur that bootstrapping students' interest and fostering engagement and motivation is a very challenging task for the instructional designer.

Optimal Structuring of Simulations for Learning

Several studies focusing on learning through simulations have found important guiding principles for developing productive simulations for students. Research from the 1990s showed that while many visualizations showed promise, not all visualizations proved effective (e.g., reviews by Park and Hopkins, 1993, and Rieber, 1990). Research from that timeframe suggested that visualizations should make normally tacit behavior visible (Norman, 1990; Merrill & Reiser, 1993, 1994) and should clearly explain causality (Faraday & Sutcliffe, 1997). Researchers of cognitive load theory (Sweller, 1993, Chandler & Sweller, 1991, Ward & Sweller, 1990) demonstrated that eliminating unnecessary cognitive tasks improves learning. Other research suggested that visualizations should be integrated within the curriculum in a manner that focuses students on the connections and ideas within the visualization (Raghavan & Glaser, 1995; Snir, Smith, & Grosslight, 1993; White, 1993a) and that providing learners with explicit learning objectives and structure could increase active engagement of the learner, leading to higher motivation and better integration and retention of content (Naps, 1996; Hansen, Scrimpsher, & Narayanan, 1998). Another interesting finding is that students learn better when derivational linkages between dynamic models are made explicit, and/or they are scaffolded with "reflection prompts" to mentally construct such derivational linkage(s). Studies conducted by White, Frederiksen, and their colleagues have shown that both these strategies can act as useful principles for designing and sequencing models and simulations in the domain of electromagnetism (Frederiksen, White, & Gutwill, 1999; White and Frederiksen, 1998).

Research in the past decade has built upon and redefined some of these suggestions and perspectives for structuring and leveraging simulations for science learning. Recently, Lindgren and Schwartz (2009) wrote a comprehensive review of simulations for science education exploring how much research has framed simulations in terms of information processing theories and how instead research might be framed from the perspective of perception and spatial learning. By switching frames in this manner, they illustrate four learning effects that can help clarify design decisions and structuring for productive simulations in terms of (1) the picture superiority effect, (2) the noticing effect, (3) the structuring effect, and (4) the tuning effect. Lindgren and Schwartz provide a series of guidelines for considering the positive and negative performance of simulations through these lenses. Lindgren and Schwartz interpret many prior simulation research studies in light of these effects.

Another important aspect of simulation activity design is the concreteness and/or perceptual salience of the features displayed in the simulations. For example, in a simulation about simple harmonic motion Parnafes (2007) noted that students typically tended to attend to the perceptually salient features of the simulation as opposed to conceptually important features (features an expert would attend to). Therefore, when designing simulations, it is important that the salient features of the simulation are ones that will be most productive for the student.

Son and Goldstone (2009) conducted a series of three experiments to examine the influence of different descriptions and perceptual instantiations of the scientific principle of competitive specialization. One of their experiments compared the role and effectiveness of intuitive descriptions to concrete ones. Their study demonstrated that intuitive descriptions led to enhanced domain-specific learning but also deterred transfer. Another experiment demonstrated that “idealized graphics are more effective than concrete graphics even when unintuitive descriptions are applied to them. When graphics are concrete, learning and transfer largely depend on the particular description” (Son & Goldstone, 2009, p. 1).

The level of scaffolding, including the immediacy of feedback, is an important design component of simulations-based or models-based learning environments. Jacobson, Kim, Pathak, and Zhang (2009) builds upon theory and research suggesting that certain types of relatively unstructured initial learning activities might lead to longer term overall learning gains (Bransford & Schwartz, 1999; Kapur, 2008; Kapur & Kinzer, 2009; Schwartz & Martin, 2004). For example, Schwartz and 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). Jacobson et al. investigated 10th grade students in Singapore as they learned concepts about electricity using four NIELS models under two conditions. The first was “productive failure” in which the students worked on a problem for each of the models followed by structured problem activities specified on worksheets. In contrast, the “non-productive failure” condition involved structured worksheet activities for both the initial and second problem solving activity for each of the models. The research showed that the “productive failure” group performed better on posttest assessments of declarative and conceptual understanding.

In addition to research on simulations, research on other computational learning environments can inform the design of simulation environments to explicitly scaffold epistemic practices of scientific inquiry such as making scientific observations, formulate hypothesis, collaboration and critique of peers, and construction of scientific explanations. BGuILE, for example, is a computational learning environment for high school biology (Reiser at al., 2001; Tabak et al., 1995) that engages students in scientific investigations in which they can explore models explaining patterns in data in evolution and ecology. In BGuILE, scaffolding is provided to encourage students to compare “competing hypotheses, articulate predictions, and record interpretations according to specific task models of biological inquiry” (Tabak et al., 1995, p. 1). Furthermore, it also provides a context for collaboration in which the biological task model is used to drive the content of students' discussions. Sandoval (2003) and Sandoval & Reiser (2004) showed that when used in conjunction with Explanation Constructor (v2.0), a computer-based learning environment specifically designed to support and scaffold students’ construction of scientific explanations, students can develop deep and correct scientific explanations of the phenomena represented in BGuILE.

Similarly, "Teachable Agents" (Biswas, Jeong, Roscoe, & Sulcer, 2009; Schwartz et al., 2007) provide another novel way of structuring simulation-based computational learning environments (e.g., Tan & Biswas, 2007; Tan, Skirvin, Biswas, & Catley, 2007). These authors have designed a teachable agent system called Betty's Brain, where students teach an intelligent computer agent using a well-structured visual representation (Biswas, Leelawong, Schwartz, & Vye, 2005; Leelawong & Biswas, 2008). Using the agent's (i.e., Betty's) performance as a motivation, students themselves learn so that they can remediate the agent's knowledge. Betty's Brain is designed for "knowledge domains where qualitative causal chains are a useful structural abstraction (e.g., the life sciences)" (Schwartz et al., 2007, p. 342). Studies show that Teachable Agents can not only foster deep content understanding but also help students develop meta-cognitive strategies (Biswas, Jeong, Roscoe, & Sulcer, 2009; Schwartz et al., 2007).

Overall, on a more general level, we would like to point out that an emergent theme from our discussion and review of the literature is that the "best tools for the job" involve simulations that provide *meaningful* opportunities for learners to interact with the represented phenomena. Each of the types of structuring of simulations discussed in this section enable learners to *explore* and *alter* the represented phenomena in a manner that fosters and supports *epistemic* aspects of scientific inquiry such as formulating and investigating what "causes" the phenomena, developing and/or verifying hypotheses, generating inscriptions, constructing scientific explanations, modifying the existing model and/or developing new scientific models.

Simulations or models that afford these types of meaningful interactions stand in stark contrast with traditional educational multi-media "toolkits" as Rogers and Scaife (1998) point out. The latter mostly limit learners to "point and click" type interactions that simply allow users to start and stop animations or "efficient page-turning and channel hopping capabilities" (Rogers & Scaife, 1998, p. 3). In contrast, in the context of PhET simulations, Wieman, Adams, and Perkins (2008) argued that students are not able to learn from the simulations *just* from watching; rather they should be engaged in active interactions with the simulation. They wrote: “most of the learning occurs when the student is asking herself questions that guide her explanation of the simulation and her discovery of the answers.” (Wieman, et al., 2008, p. 683). Designers of models-based learning environments therefore need to pay very careful attention to the activities that models and any surrounding learning environments afford. Animations that only allow learners to start and stop otherwise uninteractive “movies” will likely not foster as deep learning experiences as those afforded by simulations integrating true interactivity.

Digital Games and Science Learning

Overview of Digital Games.

Games are more challenging to define than simulations, and there are currently a number of compelling definitions (e.g., Dickey, 2005; Huizinga, 1980; Juul 2003; Klopfer, Osterweil, Grof, & Haas, 2009; Klopfer, Osterweil, & Salen, 2009; Newman, 2004; Provenzo, 1991; Salen and Zimmerman, 2003, 2004; Wittgenstein, 1958, 1972). Many of these definitions focus on how games incorporate some combination of rules, choices, play, and systems for tracking progress or success. For the purposes of this paper, which focuses on digital games and simulations for science learning, we will define digital games specifically in terms of their relationship to simulations.

Digital games share many core characteristics with digital simulations in the sense that digital games typically involve interactive models that allow players to make choices that change the states of those models. Some games build on models very similar to the models intrinsic to traditional simulations (e.g., the commercial *World of Warcraft* massively-multiplayer fantasy role-playing game) while the models in other games may stretch a bit further away from models found in traditional simulations (e.g., the models inherent in collectible card based strategy games like the digital version of *Magic the Gathering*).

One potential difference between digital games and simulations is that games are often typically defined as engendering certain levels of play, engagement, and enjoyment as core characteristics while these elements are not generally defined as core requisite characteristics for simulations. But such factors are very subjective in terms of individual users' tastes and interests, and furthermore, many simulations engender play, engagement, and enjoyment for individual users. Thus these characteristics may prove less useful for distinguishing between simulations and games.

The biggest difference between digital simulations and games as defined in this paper is that games incorporate rules and explicit goals for players to achieve or progress through, often with accompanying scoring or reward systems to track a player's progress. The borders between simulations and games become fuzzy, however, because players can also add their own explicit rules and goals to simulations, effectively transforming them into games. *The Sims* by Maxis (one of the biggest selling entertainment title franchises), for example, is arguably a simulation and not a game, but players frequently create goals and challenges within *The Sims* that transform it into a game by our definition (e.g., *The Sims* community defined a detailed challenge to successfully raise a family as an uneducated single mother).

In summary, we define digital games for the purposes of this paper as involving (a) digital models that allow users to make choices that affect the states of those models, (b) an overarching set of explicit goals with accompanying systems for measuring progress, and (c) subjective opportunities for play and engagement.

Theoretical Affordances of Games.

While our definition for digital games is relatively simple, creating digital games that simultaneously provide excellent opportunities for play and engagement (i.e., a good game) AND excellent opportunities to learn proves challenging. Despite this challenge, investigation into the use of games for learning has grown from a small niche area to a major focus of research over the past decade (Gee, 2003; 2007), and support for research on gaming for learning has simultaneously increased. In 2006, the Federation of American Scientists issued a widely publicized report stating their belief that games offer a powerful new tool to support education and encouraging governmental and private organizational support for expanded funded research into the application of complex gaming environments for learning. In 2009, a special issue of Science (Hines, Jasny, & Mervis, 2009) echoes and expands this call.

One example of this kind of increased focus on games for learning can be seen in the "Digital Media and Learning" initiative, an on-going $50 million project supported by the MacArthur Foundation that investigates how games and other digital media impact young people. Another is the Games for Learning Institute (G4LI), a $3 million research effort funded by Microsoft and New York University. G4LI brings together researchers from multiple universities and from Microsoft to investigate the features of computer games that best support engagement and learning (http://research.microsoft.com/en-us/collaboration/institutes/gamesinstitute.aspx). The stakes and potential are high according to a report in *Science* that

> in the 2000-to-2005 time frame, ~450,000 students graduated annually in the United States with a bachelor's degree in STEM. These numbers pale in comparison to the reach of a single computer video game. World of Warcraft, a fantasy game, has over 10 million current subscribers, with ~2.5 million in North America. Food Force, the U.N.-produced game on the mechanics of food aid distribution, saw 1 million players in its first 6 weeks and 4 million players in its first year. Additionally, in the realm of K-to-12 science and math education, the virtual world Whyville, with its game-based activities, now sports 4 million subscribers (90% North American), with the dominant demographic being 8- to 14-year-old girls. Although traditional education institutions pride themselves on educating citizens, they do so at a relatively small scale compared with the media now available. Is it possible to greatly expand the reach of STEM education with the use of video games as the medium? (Mayo, 2009, p 79)

Overview of Promising Game Genres for Science Learning

As with simulations, productive games for science learning can be categorized along a number of dimensions. In fact, games comprise a broader and more heterogeneous range of titles and genres than simulations and there are therefore likely many more dimensions of import to consider. To focus our discussion, however, we will limit ourselves to only three dimensions: (1) the nature of science learning connected to the game, (2) the duration and nature of the game participation, and (3) the intended purpose of the game along an entertainment/curricular spectrum.

The first dimension categorizes the nature of the science learning supported by the game (note that some games may encompass more than one genre along this dimension): (1) inquiry/argumentation as the primary goal within the game, (2) simulation-based science content and processes learning within the game, (3) inquiry / argumentation / design / engineering learning among members of a community outside the game, (4) familiarity with other discipline specific representations, tools, and processes, and (5) science content knowledge (learning of which generally is also integrated into most games with a focus on science learning even if it is not the primary goal).

The second dimension categorizes the duration and nature of game participation, mirroring a distinction in the commercial gaming world between short-term "casual games" and longer, often narrative-based, experiences. In categorizing games for science learning, we see: (1) short interaction casual games, (2) longer duration finite games organized with specific start and stop time, and (3) on-going participation games in which players become members of a persistent ongoing community in and/or around the game.

The third dimension categorizes the intended purpose of the game: (1) fully recreational games, typically commercial, that are designed for entertainment purposes, (2) serious game for informal contexts that maintain many design elements of recreational games but with a more purposeful curricular focus, (3) serious games designed for formal instructional contexts that are designed primarily as curricular interventions for use in classroom settings, and (4) assessment games that are designed primarily as a vehicle for assessing existing knowledge/understanding rather than as a learning platform.

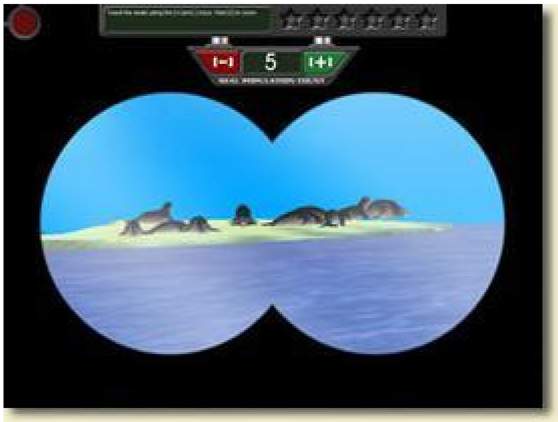
The dimensions described here are not mutually exclusive nor are they exhaustive. Any given game or genre for science learning may contain elements from multiple dimensions, while weighting toward one in particular. For example, virtual world-based games for science developed to date have tended to focus primarily on inquiry and argumentation skills, but have also contained elements of simulation-based content and process learning embedded within them. In the following sections, we introduce a wide range of digital games for science learning that offer exemplars of recent games that fall within and across these dimensions. URLs to two-minute video overviews on YouTube.com are provided for several of the examples at:

http://sites.google.com/site/nrcsciencegamessims/

Before introducing the examples, though, we should note that while many recreational commercial off the shelf games support learning of 21st century skills that are relevant to science learning (e,g,, Gee, 2003; Klopfer, Osterweil, & Salen, 2009; Steinkuhler & Duncan, 2008) and the development of important digital programming and design skills (e.g., Hayes, 2008; Hayes & King, 2009; Squire, Devane, & Durga, in press; Squire & Durga, in press), few commercial recreational titles currently focus in their gameplay on discipline-specific science skills and ideas in as productive a manner as the examples presented below (which are largely developed for educational not-for-profit or research purposes). While the nominally evolution-focused game Spore, for example, might seem on the surface promising, gameplay considerations outweigh scientific accuracy to such a degree that Spore does not hold the same discipline-specific educational power for science learning as antecedent titles by the same parent company (e.g., SimEarth, SimAnt, SimLife, and SimFarm). Some commercial recreational games offer excellent base material to build upon, as for example SURGE (presented below) tries to build on the gameplay dynamics of popular physics-based games, but few recreational commercial games currently bring discipline-specific science ideas and skills to the fore as a central focus. There are some shining exceptions, such as Portal, which focuses players on important ideas around kinematics and gravity, or the Civilization franchise, which focuses on socio-scientific issues in addition to its primary focus on socio-historical and socio-political issues, but we would love to see further merging of discipline-specific science skills and ideas directly into the gameplay a larger number of commercial recreational titles, perhaps by further refining some of the approaches outlined in the examples below to the point that aspects of those approaches might become increasingly attractive for commercial projects.

Example. Operation: Resilient Planet

Overview. Operation: Resilient Planet (http://jason.org) is an immersive 3D world intended for medium length single-player engagement focusing on inquiry into marine science phenomena. It is a recreational title developed by a commercial game company (http://filamentgames.com) in collaboration with National Geographic. Students pilot a remote operated vehicle through 3D underwater settings gathering data to solve mysteries such as the causes for dramatic shifts in shark and monk seal populations in Hawaii.



Value for Science Learning. This format of medium length science inquiry projects in 3D immersive worlds is shared by a number of single and multiplayer science games developed for formal and informal contexts such as Crystal Island (http://intellimedia.ncsu.edu), Quest Atlantis (http://atlantis.crlt.indiana.edu/), Wolf Quest (http://www.wolfquest.org), River City (http://muve.gse.harvard.edu/rivercityproject/) and EcoMUVE (http://www.ecomuve.org/) (see images below). These games allow the students to engage in the processes of inquiry and scientific argumentation while also learning the accompanying content (TSTS Proficiency Strands 1 and 2) in an active role rather than simply a bystander. These games thus also offer great opportunities in terms of TSTS Proficiency Strands 3 and 4 in terms of students understanding of how scientific knowledge is developed as well as their own identities and interest in engaging in science inquiry.


THE TIMES
ISLAND DISCOVERED


Ranger Station

Example. Whyville

Overview. Whyville (http://whyville.net)is a popular web-based 2D massively multi-player online game (MMOG) for pre-teens and teens with a predominately female player base. Visitors to the cartoon-like 2D Whyville world can take part in games and activities that feature a mixture of entertainment and educational purposes. Whyville features a number of in-world games with goals related to science learning, however, these games are only part of a larger game world. Players earn game currency by playing games (often with science focus), use the currency to refine and enhance the appearance of their "avatar" (their representation or character within the game) and their personal space within the game, start businesses, write for the newspaper, and participate in formal and informal events and socializing. The core of the game is its persistent strong community nature.

Value for Science Learning. Whyville is a persistent online virtual community that explicitly encourages science simulation learning and exploration of science ideas as part of a larger engaging community experience. The simulation-focused games encourage not only direct participation by the participants, but also spur players to create elaborate "cheat" websites and discussions where these simulation-focused science games are discussed and analyzed. Larger community events, like the "Whypox" epidemic, engage the entire community in exploring and discussing important science phenomena. Often, middle school is a time that that turns many students, particularly female students, away from science, but Whyville engages more than 4 million players in their free time exploring science-related concepts and discourse.

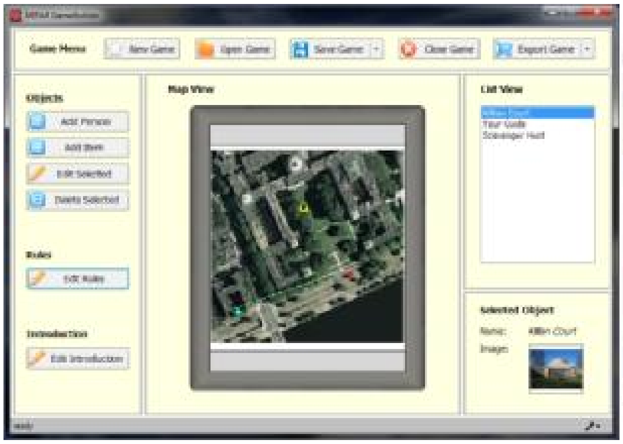
Example. World of Warcraft

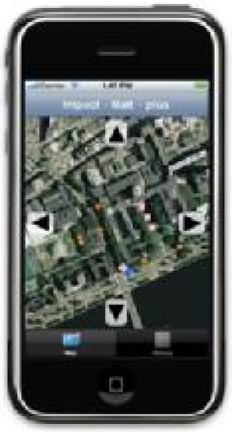
Overview. World of Warcraft (http://worldofwarcraft.com) is a massively multiplayer online role-playing game with a current subscriber base of more than 11 million players worldwide. Players control avatars in an online persistent 3D virtual fantasy world with thousands of other players sharing the same world at any given time on any given game server. Players pursue a number of challenging goals alone, in small teams, and in huge groups (known as “raids”) as they battle monsters, complete quests, engage in crafting and commerce, and socialize in a variety of venues from informal gatherings to highly structured communities (known as “guilds”).

Value for Science Learning. While the content has little to do with science learning, much STEM-related learning can occur around persistent recreational titles like World of Warcraft in terms of argumentation discourse in community forums, development of 21st century skills and literacies through immersion in group dynamics, and design and programming of digital content, modifications, and extensions for the core game software. This type of learning happens across many genres of commercial recreational titles, such as the best-selling title The Sims, which has a primarily female player base. While the STEM-related learning is not distributed evenly across the player base, players and their communities spend massive amounts of free time and effort in these worlds and provide valuable and engaging opportunities for informal STEM-related learning. Commercial recreational games such as these also have the potential to support situated content learning (such as the excellent Civilization game series for history and social studies learning) but few current science based recreational games remain true enough to the content and concepts in this manner.

Example. MITAR Augmented Reality Games

Overview. The MIT Scheller Teacher Education Program has been developing Augmented Reality software called MITAR. MITAR games and simulations use location-aware mobile devices to add a digital layer of information on a real-world context. As users navigate the physical space, they work collaboratively to explore and solve complex problems. Current research also includes investigations of the educational potential of MITAR toolkits which allow non-programmers to author their own AR games.

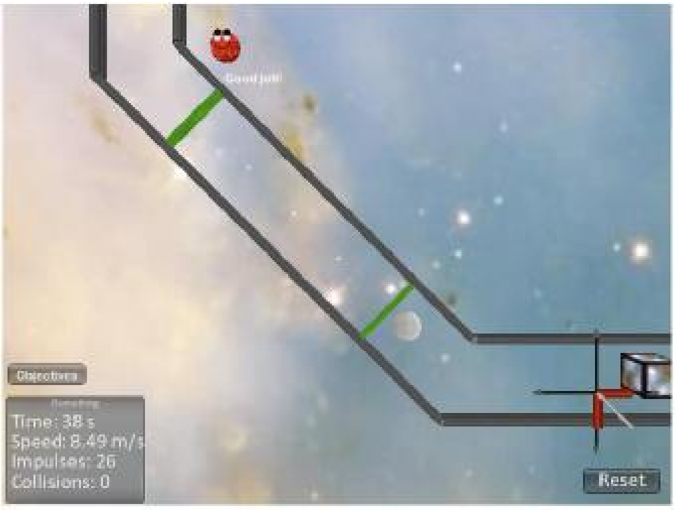




Value for Science Learning. Augmented reality games, such as Environmental Detectives, Outbreak @ The Institute, and Savannah, equip players with handheld computers as they engage in a game in “real” space. In Savannah, for example, players take on the roles of lions as they prowl the savannah in real space while the handheld computer adds information to the game and “augments” reality with other input. Similarly, Outbreak @ The Institute allows players to participate in a simulated outbreak of an epidemic on the MIT campus. Such games allow users to participate in simulations in real space rather than simply on a computer screen.

Example. SURGE (Scaffolding Understanding by Redesigning Games for Education)

Overview. SURGE (http://surgeuniverse.com)builds on the popular game mechanics of casual, short duration, physics-based games (e.g., Switchball, Marble Madness, Orbz, etc.) to help players connect intuitive understanding with formal concepts and representations. SURGE overlays and integrates formal physics representations directly into game play. Levels are designed and sequenced to simultaneously scaffold students' mastery of the game and physics concepts. Storyline elements incorporate key physics ideas. Players must think carefully about navigation decisions to manage their limited resources, avoid collisions, maximize fuel, and minimize travel time. Play supports all players in a gentle learning curve that entices replay and strategy refinement.



Value for Science Learning. Schools traditionally focus on powerful but abstract formalisms without building connections to student's tacit intuitive ideas. This results in brittle fragmented learning and minimal transfer. For three decades, many popular video games have organized their play around core science concepts (e.g., Incredible Machine, many golf games, Lunar Lander, etc.). While playing these popular video games, players can develop tacit intuitive understandings relevant to kinematics and Newton's laws, but aren't supported in articulating or extending these ideas. Well-designed digital games can build, bridge, and connect students' intuitive understandings with formal physics concepts and representations through popular game mechanics. Games like SURGE, Supercharged, London Museum's Launchball (http://sciencemuseum.org.uk/launchpad/launchball/), ImmuneAttack (http://fas.org/immuneattack/) and MIT's Weatherlings are examples that build gameplay around core science concepts and support students in exploring and articulating these ideas.

Example. Adventure Lab

Overview. Rather than playing digital games, another possibility is to support student learning by having them design video games. Adventure Lab (http://NCGRADUATE.com) supports students in building video games. Students use the environment to create video games that will teach science to the players using an environment that makes scripting the game mechanics much more accessible while retaining the polish and power of immersive gaming environments. The overall approach therefore supports the learning of the author and all of the other participants. Through a progression of cross-disciplinary steps, authors engage in science content by developing rich narratives that focus on STEM careers and 21st century skills.

Value for Science Learning. The general axiom that the act of teaching can support excellent learning is at the core of this approach. This engages students deeply in the content as they think about how they might teach other students using a medium they understand and find compelling. In addition, the motivation of creating something for a real audience, rather than writing an artificial report that no one will read, provides another strength to this approach. Several other environments exist to support students in creating games, such Scratch (http://scratch.mit.edu), AgentSheets (http://agentsheets.com) Game Maker (http://yoyogames.com/gamemaker/) and Alice (http://alice.org). Related to these scripting languages and environments is an actual game about designing games called Gamestar Mechanic that doesn't focus on science content, but does provide an engaging way to introduce students to programming and design of games (http://gamestarmechanic.com)

Example. Save Science

Overview. Situated Assessments using Virtual Environments for Science Content and Inquiry *(SAVE Science)* implements a series of game-based assessment 'quests' designed to evaluate science content and inquiry skills among middle school students. These short-duration, game-based modules are based on an assessment rubric designed to capture evolving patterns of scientific understanding of content and inquiry simultaneously among participating students. Quests in SAVE Science evaluate student learning of content and skills by employing situated scientific inquiry (data gathering, analysis, and hypothesis formation) in an immersive 3D game context. For example, in the module pictured here, players interact with virtual farmers and sheep as they conduct inquiry to demonstrate their understanding of how organisms adapt to a given environments over time.

Value for Science Learning. The value for science learning of SAVE Science project is that its research into the development of game-based assessments may lead to a new model of assessment that can provide rich, meaningful data for teachers, parents, and stakeholders about how and to what extent students are building knowledge that leads them toward goals for learning as described by Taking Science to School, while increasing engagement and self-efficacy among science learners. The VirtualAssessment project (http://virtualassessment.org) has developed a similar approach to authentic assessment in a multiuser environment (see image below).

Evidence about Games and Learning.

We now present and categorize the evidence for the potential of games for science learning in terms of (1) conceptual and process skills learning (TSTS 1 & 2), (2) epistemological understanding (TSTS 3), (3) attitudes, identity, and motivation (TSTS 4), and (4) optimal structuring of games for learning. We would like to acknowledge the excellent review papers by Fletcher and Tobias (2006, 2008), de Freitas (2006), Frohberg, Goth, and Schwabe (2009), and Dieterle (in press). These review papers helped us identify many studies that we might not otherwise have discovered in light of the broad dispersal of games research across so many disciplines and communities.

Evidence: Conceptual Understanding (TSTS 1) and Process Skills (TSTS 2)

Many games for science learning focus on a combination of the first two TSTS proficiency strands. The first proficiency strand of TSTS centers on students' knowledge and ability to use and interpret scientific explanations of the natural world. This includes conceptual understanding and the encompassed content. The second strand focuses on generating and evaluating scientific evidence and explanations, in terms of students' understanding and ability to apply processes and skills to operate on content and conceptual understanding. Many games are predicated on the premise that collaborative virtual environment-based games are uniquely suited to allow learners to accomplish goals associated with the two strands in unison.

A series of studies by Moreno and Mayer (2000; 2004) investigated the impact of design principles applied to computer games on student retention of science content and on problem-solving transfer questions. In two of these studies, undergraduate university students played a computer game about environmental science (Lester's "Design-a-plant" game) that included personalized (1st and 2nd person language) instructional content, delivered as narrated speech by a pedagogical agent. Students who heard personalized content outperformed students who received neutral content on problem-solving transfer questions (*p* < .0001, effect size 1.55) and post-implementation retention questions asking them to write down things they saw in the lesson (*p* < .005, effect size .83). In a follow-up study using the same game, but with personalized content delivered via text (not voice), they found similar results. Students who saw personalized content outperformed students who received neutral content on problem-solving transfer questions (*p* < .0001, effect size 1.58) and post-implementation retention questions asking them to write down things they saw in the lesson (*p* < .05, effect size .57). Moreno and Mayer (2004) continued these studies and added in an immersion dimension that involved wearing a head-mounted display. They found that this version of head-mounted immersion did not impact learning. Note that this use of the term immersion is different than that generally employed in the science gaming world (c.f., Dede, 2009), which does not require virtual reality equipment, but focuses more on creating an experience for the player of being embedded in a virtual world through the nature of the game environment rather than through equipment.

Another series of studies was conducted in the Supercharged game (Barnett, Squire, Higginbotham, & Grant, 2004; Jenkins, Squire, & Tan, 2004). Supercharged is a 3D game in which players utilize and explore the properties of charged particles and field lines to navigate their ship through space. The spaceship is moved through the game world by taking advantage of the properties of charged particles in the space. Three middle school classes participated in a mixed methods pilot study comparing learning outcomes in kids playing Supercharged (n = 35) and those using a guided inquiry in-class curriculum (n = 61). Average post-test scores were significantly higher ($p < .05$) for the students who played Supercharged. The test included 12 questions on electromagnetism combined with pre-post interviews of a random sub-set of the students that were then transformed into additional quantitative data (Squire, Barnett, Grant, Higginbotham, 2004).

Anderson and Barnett (in press) continued the investigation of Supercharged with pre-service elementary teachers. The control group in their study learned through a series of guided inquiry methods while the experimental group played Supercharged during the lab sessions of the course. Supercharged stuents significantly outperformed the inquiry control students in terms of pre-post assessment gains ($F(2,134) = 4.8$, $p < 0.05$, $\eta^2=0.59$), but Supercharged students rated their knowledge of the topic lower than the inquiry group (M post-control = 3.0, M post-experiment = 2.7).

Klopfer, Scheintaub, Huang, Wendel, & Roque (2009) describe a series of pilot studies using StarLogo TNG to combine games, simulations, engineering, and science for students. Many of the pilot studies involve qualitative descriptions of student learning of content and programming skills. Klopfer et al. also detail the findings in a pilot study involving a simulation-based game called The Planet Game with 47 students. A comparison of pretests and posttests suggested that a number of the students addressed important misconceptions through the short activity.

Hickey, Ingram-Goble, and Jameson (2009) studied the impact of the 15-hour Taiga ecological sciences curriculum in Quest Atlantis. Gains in individual understanding of science content and socioscientific inquiry were assessed with a performance assessment that presented related problems in a related context. Gains in aggregated achievement were measured using randomly sampled released achievement test items that were aligned to targeted content standards but independent of the Taiga curriculum. The first study involved a sixth grade teacher using the curriculum with two classes and obtaining larger gains in understanding and achievement (0.3 and 0.2 SD, respectively) than two of his other classes that used expository text on the same concepts and skills. After two new types of virtual formative feedback were developed for the in-game Quest submissions, the same teacher used the Taiga curriculum in all four of his classes the subsequent year, resulting in substantially larger gains in understanding and achievement (1.1 SD and 0.4 SD). In another Taiga study, Barab, Scott, Siyahhan, Goldstone, Ingram-Goble, Zuiker, and Warrant (2009) compared learning outcomes among 51 undergraduate participants in 4 conditions: electronic book-based content, 'simplistic framing' of content presented as a web-based 2D curriculum, immersive world-based (pair-based), and single-player immersive world-based. They found that learners in either of the virtual world-based conditions significantly outperformed learners in the electronic book group ($p=.01$), and outperformed book and 'simple-framing' groups on a transfer test ($p=.01$). Anderson and Barnett (2008) found similarly positive outcomes ($p < 0.01$) on standardized learning measures in their study of a Quest Atlantis implementation with 26 elementary school students.

Work over several studies in River City, another 3D massively multiplayer environment, investigated students' engagement in inquiry, content learning, and motivation. Students in River City work collaboratively to solve a simulated 19th century city's problems with illness by interacting with each other, digital artifacts in the game, computer agents in the game, and various data collection tools in the world. Ketelhut, Dede, Clarke, & Nelson (2006) describe the results from the three implementations in 2004 with approximately 2000 adolescent students. Their results show that students learned biology content, that students and teachers were highly engaged, that student attendance improved, that disruptive behavior dropped, that students were building 21st century skills in virtual communication and expression, and importantly, that using this type of technology in the classroom can facilitate good inquiry learning.

Further work on River City by Dieterle (2009b) showed that a set of 574 students who completed the curriculum showed significant differences from pretest to posttest in terms of Science Content

Understanding (CONTENT) and Self-Efficacy in Scientific Inquiry (SEISI) ($p < .01$). His work further showed that (a) students who preferred creating and sharing artifacts through the Internet were well-suited for learning about disease transmission and scientific problem solving skills in the curriculum and that (b) students who felt highly connected with the media, tools, and people they used for communication, expression, and problem solving in the curriculum were more likely to believe they are able to successfully complete the activities that a scientist might engage in.

Neulight, Kafai, Kao, Foley, and Galas (2007) investigated two sixth-grade classes' (46 students) understanding of a virtual infectious disease from a participatory simulation in a game-like world (Whypox in Whyville) in relation to their understanding of natural infectious diseases. They found that there was a significant shift in students' responses between pre and post from pre-biological to biological explanations ($t(44) = -3.5$, $p = 0.001$; $t(44) = -3.496$, $p = .001$) demonstrating that twice as many students reasoned about natural infectious disease with biological reasoning by the end of the curriculum.

Kafai, Quintero, and Feldon (in press) further studied students during Whypox epidemics in Whyville. Part of their study focused on students' use of two simulators that allowed players to run small-scale and fast simulations of the epidemics. The simulations allowed the players to make predictions and compare their predictions to the simulation results. During Whypox outbreaks, simulation usage peaked with more than 1400 simulations performed by 171 players in their sample. Kafai and colleagues found that 68% of the players conducted some form of systematic investigation by running the simulations 3 or more times, 49% of those players demonstrated significant improvements in the accuracy of their predictions, and that 70% of players pursued engineering type goals in the process rather than scientific strategies as indicated by the relationship between the independent variables and the accuracy of users' predictions.

Holbert (2009) coded observational data of talk and gesture collected during ethnographic observations of and individual clinical interviews with children playing popular video games (Mario Kart Wii and Burnout Paradise), Holbert identified that children's intuitive schema of velocity, acceleration, and momentum were at play while they were playing these games. These schemas have been previously identified as registrations (Roschelle, 1991) and phenomenological primitives (p-prims) (diSessa, 1993), and have been shown to play productive roles in the development of understanding of physics.

Rosenbaum, Klopfer, and Perry (2006) studied 21 urban high school students playing Outbreak @ The Institute, an augmented reality game where players take on the roles of doctors, technicians, and public health experts trying to contain a disease outbreak. Players interact with virtual characters and employ virtual diagnostic tests and medicines while they move across the university campus setting in real life with handheld computers. Rosenbaum, Klopfer, and Perry found that surveys, video, and interviews of the students showed that the students perceived the game as authentic, felt embodied in the game, engaged in the inquiry, and understood the dynamic nature of the model in the game.

Clark, Nelson, D'Angelo, Slack, and Menekse (2009) analyzed pre-post test data from 24 undergraduate and graduate students playing SURGE, a Newtonian mechanics based game as part of an ongoing series of studies (e.g., Clark, Nelson, D'Angelo, Slack, & Menekse, in press). The data reinforce the potential of games to help students learn, but also underscore their potential to reinforce alternative conceptions as well as normative conceptions. The game actually resulted in a significant decrease ($\chi^2 = 4.75$, $p = .029$) on one item by unintentionally focusing students' attention on another physics relationship (not all of the intended functionality had been added to the interface), but the students demonstrated significant ($p = .037$) gains on the rest of the posttest when that first question was excluded. In-game and post-interview data indicated that players made successful (although variable) use of growing tacit understanding of the physics concepts in the game to complete levels of the game (D'Angelo, Clark, Nelson, Slack, and Menekse, 2009).

Annetta, Mangrum, Holmes, Collazo, and Cheng (2009a) studied seventy-four fifth graders playing the game Dr. Friction, a teacher-created Multiplayer Educational Gaming Application (MEGA), in the middle of a unit on simple machines. Using a pre-post test design, students overall did significantly better ($p < .001$, effect size .65) on the post-test (regarding simple machines). The study also looked at gender differences and found no significant difference. In a separate paper, however, Annetta, Mangrum, Holmes, Collazo, and Cheng (2009b) studied 66 students using a teacher-created game about genetics (as a review in class) and found no significant increase on a post-test when compared to 63 students not playing the game. All students had the same general instruction with the same teacher in four high school biology classes.

Miller, Moreno, Estrera, and Lane (2004) studied middle school students playing an episodic adventure game, MedMyst, about infectious diseases and microbes. Gain scores from pre to post tests showed that students retained information from the game (most comparisons significant at $p < .001$). The same game was used with high school students with smaller gains.

In a study related to real-world transfer of skills, Greenfield, Camaioni et al. (1994) had university students in the U.S. and Rome play a video game for 2.5 hours, and then take a test to measure their ability to generalize and apply principles from a few demonstrated examples (in this case, images of components in an electronic circuit). Video game performance was significantly correlated with improvement in scores on the electronic circuit test.

In addition to the studies above that we were able to locate, Mayo (2009) reports that McClean, Saini-Eidukat, Schwert, Slator, and White (2001) studied 238 college students playing Virtual Cell and 273 students playing Geography Explorer. According to Mayo, McClean et al. found that students playing Virtual Cell showed 40% gains over lecture and students playing Geography Explorer showed 15-40% gains over lecture.

Finally, not all science learning needs to happen inside the games themselves. Steinkuehler and Duncan (2008) study scientific habits of mind demonstrated in the discussion forums around the commercial massively multiplayer online role-playing game World of Warcraft. The game itself focuses on fantasy themes, but Steinkuehler and Duncan analyzed 1,984 posts by users in 85 different discussion threads and found that 86% of the posts involved social knowledge construction, more than 50% of the posts evidenced systems-based reasoning, roughly 10% evidenced model-based reasoning, and 65%

displayed evaluative epistemologies supportive of argumentation as a means for knowledge construction. Steinkuehler and Duncan argue that this is evidence that even popular commercial titles without a direct connection to science can support scientific thinking processes.

Evidence: Epistemological Understanding (TSTS 3)

As described previously the third strand of the Taking Science To School proficiency standards focuses on students' epistemological understandings of the nature and development of scientific knowledge. The immersive virtual contexts (Dede & Ketelhut, 2003) of digital games, combined with the situated embodiment they engender (Barab, Zuiker, Warren, Hickey, Ingram-Goble, Kwon, Kouper, & Herring, 2007) seem to support learners in better understanding the complex and sometimes messy nature and development of scientific knowledge in the real world. As Ketelhut states (personal communication, 2009), scientific inquiry is about learning the ways in which scientists develop new knowledge. Multi-player virtual environment-based games such as Quest Atlantis, River City, and Whyville, which enable students to learn and practice authentic inquiry skills collaboratively, also help them better understand the nature and development of scientific knowledge. For example, in their Whyville study, Neulight, Kafai, Kao, Foley, and Galas (2007) investigated the degree to which participation in the Whyville game might bolster students' understanding of the causes of disease. Qualitative analysis of student chat in Whyville and in answers to surveys showed improved accuracy in participants' understanding the spread of infectious disease. Nelson (2007) conducted a River City study in which he explored the impact of embedded guidance messages on student understanding of real-world science inquiry processes and knowledge, as measured by pre and post-implementation survey questions. The study found that increased viewing of guidance messages was associated with significantly higher ($p < .05$) score gains on questions related to scientific inquiry and disease transmission.

In addition, participants in these kinds of games report changes in their understanding of what science means and of how they view themselves in relation to science. Participants in a 2004 River City study of more than 1,000 students reported feeling like real scientists for the first time (Clark & Dede, 2005).

A new theme in science learning research focuses on augmented reality games. Squire (in press), for example, presents a case study investigating one enactment by 55 students of an augmented reality game-based curriculum called Sick at South Beach. The case study describes (a) how fictional elements of the augmented reality game situated the learning experience and encouraged academic practices, (b) how student-created inscriptions influenced the students' emerging understandings, and (c) how the game-based curriculum's design enhanced students' conceptual understandings, and (d) how learning through a technology-enhanced curriculum encourage students' identities as independent problem solvers.

Squire and Jan (2007) present a cross-case comparison of three cases involving approximately twenty-eight students engaged participating in the place-based augmented reality game, Mad City Mystery as the students learned about environmental science. Squire and Jan demonstrate that Mad City Mystery engages students in meaningful scientific argumentation as they develop narrative accounts of scientific explanation.

Finally, Squire and Klopfer's (2007) case study research on the augmented reality game *Environmental Detectives* demonstrates how students can (a) be supported in negotiating complex problem spaces that demand the integration of multiple information data sources and (b) develop a narrative of science on which they can build deeper understandings in the future of both the conceptual content and the socially situated nature of scientific practice.

Evidence: Attitudes and Identity (TSTS 4)

The fourth strand of the Taking Science to School proficiency standards focuses on students' attitudes, identity, and habits of mind in terms of their willingness to engage and participate productively in scientific practices and discourse. Many studies have focused on the value of immersive game environments as platforms for situated curricula that can motivate and engage students to learn and apply science content and inquiry skills. Early pilot implementations of the River City curriculum were found to be strongly motivating for participants as an alternative to classroom-based science curricula, especially for students with lower academic backgrounds (Dede, Ketelhut, & Ruess, 2002). Ketelhut (2007) investigated sources of student engagement in River City, finding that students reported the ability to conduct inquiry as a key motivating element, along with the ability to use virtual tools such as bug catchers and microscopes to aid in their inquiry.

Tuzan (2004) investigated the motivational elements that supported student participation in Quest Atlantis, identifying a large number of elements centered on identity, play, immersion, and social relationships. Barab, Arici, and Jackson (2005) reported on their iterative design process in creating and modifying Quest Atlantis to support engagement, finding (among other things) the need for a strong narrative backstory. In a study investigating the use of a science inquiry curriculum called Whypox built into Whyville, Galas (2006) found that the curriculum was engaging for students and supported realistic opportunities to conduct collaborative scientific inquiry. The Whypox curriculum featured a virtual disease unleashed on students in Whyville that affected their virtual avatars.

In addition to focusing on motivation and engagement, a number of studies have focused on issues relevant to the fourth TSTS strand in terms of how game-based curricula in multi-player virtual environments can support and promote authentic scientific practices and use of science-centered discourse. For example, Barab and his colleagues have conducted a number of studies in this area around their Taiga curriculum in the Quest Atlantis virtual environment (Barab, Zuiker, Warren, Hickey, Ingram-Goble, Kwon, Kouper, & Herring, 2007; Barab, Sadler, Heiselt, Hickey, Zuiker, 2007; Barab, Scott, Siyahhan, Goldstone, Ingram-Goble, Zuiker, & Warrant, 2009). In one such study, they report on a mixed methods study into the power of the Taiga curriculum to support the kinds of productive inquiry practices and scientific discourse described in the NRC report (Barab, Sadler, Heiselt, Hickey, Zuiker, 2007). In the design experiment with 28 fourth grade students in a gifted class, Quest Atlantis researchers found that all participants were actively engaged in discourse related to the inquiry tasks of the curriculum, and that they participated actively and productively in inquiry practices (data gathering, negotiation, data interpretation, etc.). Anderson (2009; in press) examined the impact of peer-peer dialog and embedded scaffolding on science learning in Quest Atlantis. In one study with nine 5th grade students, it was found that scaffolds in the game were supportive of players in helping guide their discussions about inquiry with other students (Anderson, 2009). In a second case study with two fifth grade Quest Atlantis players, Anderson found that the game seemed to support the acquisition of science content and skills, and to support the students' ability to express that knowledge to peers.

In her Whyville implementation, Galas (2006) found that the Whypox curriculum supported realistic opportunities to conduct collaborative scientific inquiry. Middle school students in the study worked together to track the spread of the disease. Study participants also visited Whyville's "Center for Disease Control" to gather and share information about the disease outbreak, and used an embedded simulation that modeled the ways in which diseases spread through a population. Dede and Ketelhut (2003) report on the impact that participation in authentic science practices and discourse in the River City game had on student self-efficacy, finding significantly higher ($p < .05$) levels of 'global science self-efficacy' among River City participants than among their peers in a project-based classroom curriculum.

Research has also focused on attitudes, identity, and motivation in other domains. At a general level, for example, Sanford et al. (2006) conducted a survey of what kids and teachers think about commercial games. Sanford et al. found that (a) kids love games and want to play them and (b) teachers don't play them but think that they might be useful for learning because they are motivating and engaging to kids. At a more specific level, de Freitas (2006) reports that Galloway (2006) found that integrating learning activities at a college in the Neverwinter Nights game engine increases course completion and grades for specific skills between 30% to 100%.

In addition to work on how games can support identity, attitudes, and self efficacy, research by McQuiggan, Mott, and Lester (2008) has investigated how intelligent tutoring systems and games might assess student's sense of self-efficacy using a combination of question, physiological, and behaviors in the environment to correctly classify approximately 85% of instances. This carries important implications for games. To further complement this work, McQuiggan and Lester (2007) explore models for modeling and evaluating empathy in embodied companion agents, which could translate in games to support learning dramatically. In addition, McQuiggan, Rowe, and Lester's research (2008) determined that empathetic characters had a significant effect on measurements of students' overall presence, involvement and control, and naturalism of the experience. This all ties in well with Lester's group's work on Crystal Island, a 3D virtual environment where players engage in inquiry solving a health mystery.

Evidence: Optimal Structuring of Games for Science Learning

Not many studies exist that specifically look at how changes in the structure and design of games impact science learning. In many studies, however, researchers have been able to make some statements about how the design or structure influenced how students interacted with each other and with the game itself. In research on Savannah, an augmented reality game where players take on the roles of lions in a pride, Facer et al. (2004) describe their design choices and how the students were rewarded in the game for making certain decisions that fit with how the game designers wanted them to act. In some cases the students did not act as they were intended to (for instance, they stuck with one strategy that worked instead of trying new things) due to a simplification of the design. While the students learned the rules of the game quickly, these rules were not always sufficient to help the students learn the science ideas at the appropriate level. "The main challenge to designers is to develop sufficiently sophisticated games rules, and sufficiently focused challenges in order to encourage the children to attempt different strategies to overcome these problems" (Facer et al., 2004, p. 407). It is one thing to create a fun and engaging game that students will want to play. It is another to create one that will also teach them the intended concepts and ideas.

Moreno and Mayer (2005) looked at the role of guidance (explaining the reasons for a correct answer) and reflection (having students explain their answer) in Lester's Design a Plant game. The study consisted of 105 undergraduate students in 4 groups (guidance/reflection, guidance/no reflection, no guidance/reflection, no guidance/no reflection) playing the game. They were given retention, transfer, and program rating tests. MANOVA analysis saw significant differences on transfer measures between guidance and no guidance groups, but no difference between reflection/no reflection groups. On other measures, there were marginal but non-significant differences between guidance/no guidance groups.

Mayer, Mautone, and Prothero (2002) in a study with 105 college students found that providing pre-training in the Profile Game before playing the game by showing players pictures of possible geological features that would need to be identified through the game, led to significantly better performance on identifying those geographical features in the game.

Jones, Minogue, Tretter, Negishi, & Taylor (2006) investigated the impact of haptic (sense of touch) augmentation of a science inquiry simulation/game (Mystery of the Sick Puppy) on 36 middle school and high school students' learning about viruses and nanoscale science. They compared use of a sophisticated haptic desktop device, a haptic gaming joystick), and a mouse (no haptic feedback). Results showed that the addition of haptic feedback from the haptic-gaming joystick and the sophisticated haptic desktop device provided a more immersive learning environment that made the instruction more engaging (roughly $p < .001$ across engagement measures) and influenced the way in which the students constructed analogies in their understandings about the abstract science concepts (roughly $p<.05$ across measures). Thus, providing the students with increasingly haptic feedback greatly increased the efficacy of the environment.

Squire, Giovanetto, Devane, and Durga (2005) report in their qualitative study of high school students playing Civilization III, a game that focuses on socio-scientific issues, that "we can build better game-based learning environments by starting with developing game mechanisms where players literally perform the kinds of understandings we want them to have" (p. 40). Games can be structured in many ways and the learning outcomes can change based on the ways in which the game is structured. Some learning goals may be fact-based knowledge (e.g. What is an aqueduct?), while other goals may require a deeper understanding (e.g. How does an aqueduct affect a nearby city?). Squire et al. (2005) found that collaborative competitive games could support this deeper understanding. They also found that it was important to offer multiple ways and choices for students to engage in the gaming experience.

Squire and Durga (in press), Squire, DeVane, and Durga (in press), and DeVane, Durga, and Squire (in press) build on this research through design-based research studies of organizing learning communities around Civilization for disadvantaged students. The studies demonstrate through qualitative analysis the connection between gameplay and "Modding" (the use of software tools provided by the game to program or create extensions or variants of the game), demonstrate that students in these after school learning communities can learn to "mod" and design games themselves, and demonstrate that students in these communities develop important academic skills, systems thinking skills, and problem-solving strategies.

In recent study that looked at the issue of incentives and competition, Hickey, Filsecker, & Kwon (2009) contrasted two versions of the Taiga curriculum in Quest Atlantis. Students in two classes that used a "Public Recognition" condition were given badges to place on their game avatar and invited to move a paper version of their avatar to illustrate both their progress and status in the game. Students in two matched classes were not given incentives and instead were encouraged to engage in the curriculum for more intrinsic reasons. Students in the public recognition condition were shown to use more of the core science concepts in the quest assignments, and to use them more correctly, and showed significantly larger gains in understanding, and larger gains in achievement. Arguing against the concern that incentives may lower intrinsic motivation, the students in the public recognition condition also showed slightly higher intrinsic motivation during the game and slightly larger gains in interest toward solving scientific and socioscientific problems.

Discussion

We began this paper by outlining our belief that games and simulations are tools that can support science learning when appropriately designed and implemented. As we have shown, there are many simulation and game genres of value to science learning with differing goals, aimed at diverse audiences, and designed for implementation in a wide spectrum of settings.

We have also shown that, in spite of the recent spike of interest in these tools, there is a long history stretching back at least 30 years into their use for learning. The lessons of this long history offer valuable insights for researchers today exploring the use of games and simulations for science learning. The first lesson is that the tool metaphor itself is one to use with caution. History is littered with technological tools for learning that ultimately made little ultimate impact learning but that enjoyed levels of enthusiasm and research interest similar to that now afforded simulations and games (Cuban, 2001). Thomas Edison famously predicted that motion pictures would completely revolutionize education, replacing textbooks and traditional classroom instruction. But when studies were conducted on the impact of film on learning, results were mixed (Oppenheimer, 2003). In a similar sense, it simply isn’t true that all games and simulations used for science learning are better tools than traditional classroom instruction. This is due to the wide range of possible learning experiences that one can design in the form of a game-based or simulation-based learning environment. However, in this paper, we have tried to identify evidence pertaining to specific examples of “what works” along several dimensions.

Tools are just tools until they are applied to some end. From motion pictures to educational television to games and simulations, all technological tools are ultimately used to achieve goals that themselves arise from needs, and those needs reflect beliefs about learning. Application of games and simulations as learning tools will not only vary based on setting, audience, and outcome goals but also as a reflection of the theories of learning held by those who wield the tools. Researchers of games and simulations for learning have based their questions and inquiry on a multitude of theoretical viewpoints including behaviorism, cognitive science, constructionism, constructivism, situated cognition, socioconstructivism, and more. The theoretical perspectives underlying the studies on visual response times differ from the theoretical perspectives underlying the studies on immersion and presence in games, which differ from the theoretical perspectives underlying studies on socio-scientific collaboration. Theoretical views behind the design and implementation of simulations and games for STEM learning have evolved, and will continue to evolve, as surely, if not as rapidly, as the technology of the tools.

One lesson to take from this diversity of views and approaches is to strive for agnosticism in applying simulations and games for science learning. Apply and adapt the tools to the tasks, recognizing the theoretical framework from which a given tool, learning perspective, or analytical methodology has sprung while carefully adapting the tools to new goals and settings. Likewise, the researchers should take a pragmatic approach in applying findings from research done in the past and/or conducted under a different theoretical banner from their own situations, audiences, and settings. Our own organizational framework, examining past and current research as it can be applied to the goals defined in the NRC Taking Science To School report is an example of this kind of pragmatic approach. The report's goals reflect the current thinking on what it means to "know science." Much of the work described in this paper was conducted under different assumptions, but still provides guidance for current and future work under prevailing views of learning, such as preparation for future learning.

Another impact of the diverse views of games and simulation may be seen in the highly individualistic approaches to research reflected in the studies we have described. We have touched upon studies conducted by individuals or small teams from higher education, K12, the military, medicine, government, and commercial enterprise. Much has been learned from these studies that can be built upon, but this approach presents two problems as we move ahead: (1) individual 'silo' studies can lead to high levels of replication of effort, and (2) widely divergent and isolated research efforts negatively impact the ability to share and build upon findings, slowing down the field as a whole. We agree with Chris Dede's position in his paper for this NRC workshop that larger-scale research efforts are critical to advancing the field. This doesn’t mean, from our perspective, that the disparate views on games and simulations for science learning somehow need to coalesce or homogenize. This is neither possible nor desirable.

Rather, we believe that, in order to maximize impact, the field needs to move beyond the current extended exploratory stage. Much of the research in the field to date has focused on proofs of concept. O’Neil, Wainess, and Baker (2005), for example, used Kirkpatrick’s levels of evaluation and the CRESST model of learning as a way to assess learning in games in a meta-review of 15 year’s worth of articles on games and simulations. Of the 1,000s of articles they found, they say that only 19 studies meet their standards for empirical research (meaning they contain data of some kind). We now need to focus on more careful studies that measure what is actually learned and what design principles best support this learning.

This should not be interpreted as a call only for randomized trials because we see important roles for rigorously conducted qualitative and quantitative research. Rather, the field now needs to focus on applying rigorous qualitative and quantitative methods in the conduct of studies that are data-focused and are submitted to and published in peer-reviewed journals such that the rigor of the methods and data can be sanctioned by the field in our pursuit to better understand what and how games and simulations can help people learn science. This suggestion applies more to games research than to simulation research, because simulations research has reached a more mature phase of study, but research in both fields would benefit from a heavier emphasis on rigorous analysis of qualitative and quantitative data of what exactly is being learned, by whom, and how.

Finally, to support these enhanced research efforts, new approaches, research networks, databases, and clearinghouses are needed to help coordinate research efforts by facilitating connections between and among researchers and across communities and contexts. Whereas keeping abreast of current research is relatively simple in many fields, tracking or finding research on games and simulations for learning is not. Games and simulation research is conducted across an almost unparalleled breadth of fields and contexts, spanning a wide range of academic disciplines, commercial enterprises, and government and military organizations. Further supports to connect researchers and developers across disciplines and contexts in terms of compiling and sharing pragmatic design principles, research, and lessons learned will greatly facilitate advances in our field. The Digiplay searchable games research database (digiplay.info), the Games, Learning, and Society conferences (GLS) (www.glsconference.org), and Digital Games Research Association (DIGRA) (www.digra.org) provide excellent examples of such initiatives, but further infrastructure development would significantly catalyze games and simulations research toward its true potential.

References

Abrahamson, D., & Wilensky, U. (2004). *SAMPLER: Collaborative interactive computer-based statistics learning environment*. Paper presented at the 10th International Congress on Mathematical Education, Copenhagen, July 4 - 11.

Adams, W. K., Reid, S., LeMaster, R., McKagan, S. B., Perkins, K. K., Dubson, M., & Wieman. C. E. (2008a) A study of educational simulations part I - Engagement and learning. *Journal of Interactive Learning Research, 19*(3), 397-419.

Adams, W. K., Reid, S., LeMaster, R., McKagan, S. B., Perkins, K. K., Dubson, M., & Wieman, C. E. (2008b). A study of educational simulations part II - Interface design. *Journal of Interactive Learning Research, 19*(4), 551- 577.

American Association for the Advancement of Science (1993). *Benchmarks for scientific literacy*. New York: Oxford University Press.

Anderson, J. (2009). *Real Conversations in Virtual Worlds: The impact of student conversations on understanding science knowledge in elementary classrooms*. Paper presented at the American Educational Research Association Annual Meeting. April 13-17, 2009. San Diego, California.

Anderson, J., & Barnett, G. M. (in press). Using Video Games to Support Pre-service Elementary Teachers Learning of Basic Physics Principles. To appear in *Journal of Science Education and*

Technology.

Annetta, L. A., Minogue, J., Holmes, S. Y., & Cheng, M.-T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers and Education, 53*(1), 74-85.

Annetta, L., Mangrum, J., Holmes, S., Collazo, K., & Cheng, M.-T. (2009). Bridging reality to virtual reality: Investigating gender effect and student engagement on learning through video game play in an elementary school classroom. *International Journal of Science Education*, *31*(8), 1091-1113.

Barab, S. A., Arici, A., & Jackson, C. (2005). Eat your vegetables and do your homework: a design based investigation of enjoyment and meaning in learning. *Educational Technology, 45*(1), 15-20.

Barab, S. A., Sadler, T., Heiselt, C., Hickey, D., Zuiker, S. (2007). Relating Narrative, Inquiry, and Inscriptions: A Framework for Socio-Scientific Inquiry. *Journal of Science Education and Technology, 16*(1), 59-82.

Barab, S. A., Scott, B., Siyahhan, S., Goldstone, R., Ingram-Goble, A., Zuiker, S., & Warrant, S. (2009). Transformational play as a curricular scaffold: Using videogames to support science education. *Journal of Science Education and Technology 18*, 305-320.

Barab, S. A., Zuiker, S., Warren, S., Hickey, D., Ingram-Goble, A., Kwon, E-J., Kouper, I., & Herring, S. C. (2007). *Situationally Embodied Curriculum: Relating Formalisms and Contexts. Science Education, 91*(5), 750-782.

Barnett, M., Squire, K., Higginbotham, T. & Grant, J. (2004). Electromagnetism Supercharged! In Proceedings of the 2004 International Conference of the Learning Sciences. Los Angeles: UCAL Press.

Biswas, G., Jeong, H., Roscoe, R., & Sulcer, B. (2009). *Promoting Motivation and Self-regulated Learning Skills Through Social Interactions in Agent-based Learning Environments.* 2009 AAAI Fall Symposium Series, Memphis, TN.

Biswas, G., Leelawong, K., Schwartz, D., & Vye, N. (2005). Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence, 19,* 363-392.

Blikstein, P. & Wilensky, U. (2005) *Less is more: Agent-based simulation as a powerful learning tool in materials science.* Paper presented at the 4th International Joint Conference on Autonomous Agents and Multiagent Systems. Utrecht, Netherlands.

Bransford, J. D. & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education, 24,* 61-100.

Chandler, P. & Sweller, J. (1991). Cognitive load theory and the format of instruction *Cognition and Instruction, 8*(4), 293-332.

Chang, H.-Y., Chiu, J., McElhaney, K, & Linn, M. C. (in preparation). Can dynamic visualizations improve science learning: A synthesis study.

Clark, D. B. (2006). Longitudinal conceptual change in students' understanding of thermal equilibrium: An examination of the process of conceptual restructuring. *Cognition and Instruction, 24*(4), 467–563.

Clark, D. B., & Jorde, D. (2004). Helping students revise disruptive experientially-supported ideas about thermodynamics: Computer visualizations and tactile models *Journal of Research in Science Teaching 41*(1), 1-23.

Clark, D. B., & Linn, M. C. (2003). Scaffolding knowledge integration through curricular depth *Journal of Learning Sciences, 12*(4), 451-494.

Clark, D. B., Nelson, B., D'Angelo, C. M., Slack, K. & Menekse, M., (in press). *Connecting students' intuitive understandings about kinematics and Newtonian mechanics into explicit formalized frameworks.* Paper to be presented at the American Association for the Advancement of Science (AAAS) Conference 2010. San Diego, California.

Clark, D. B., Nelson, B., D'Angelo, C. M., Slack, K., & Menekse, M., (2009). *Intended and unintended Learning in SURGE.* Results presented at the Technology Enhanced Learning in Science 2009 meeting. Minneapolis, MN.

Clarke, J., & Dede, C. (2005). *Making learning meaningful: An exploratory study of using multi-user environments (MUVEs) in middle school science.* Paper presented at the American Educational Research Association Conference, Montreal, Canada.

Cognition and Technology Group at Vanderbilt. (1990). Anchored instruction and its relationship to situated cognition. *Educational Researcher, 19*, 2–10.

Colella, V. (2000). Participatory simulations: Building collaborative understanding through immersive dynamic modeling. *Journal of the Learning Sciences, 9*(4), 471–500.

Cuban, L. (2001). Oversold and Underused: Computers in the Classroom. London, Harvard University Press.

D’Angelo, C. M., Clark, D. B., Nelson, B. C., Slack, K., & Menekse, M. (2009). *The effect of vector representations on students’ understanding of motion.* Poster presented at the Physics Education Research Conference (PERC)/American Association of Physics Teachers (AAPT) 2009 meeting. Ann Arbor, Michigan.

de Freitas, S. (2006). Using games and simulations for supporting learning. In C. Martin & L. Murray (Eds), Learning, Media and Technology: Special Issue on Gaming.

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science, 323*(5910), 66–69.

Dede, C., & Ketelhut, D. J. (2003). *Designing for motivation and usability in a museum-based multi-user virtual environment.* Paper presented at the American Educational Research Association Conference, Chicago, IL.

Dede, C., Ketelhut, D. J., & Ruess, K. (2002). Motivation, usability, and learning outcomes in a prototype museum-based multi-user virtual environment. *Proceedings of the Fifth International Conference of the Learning Sciences*, 406-408.

DeVane, B., Durga, S., Squire, K. (in press). “Economists who think like ecologists”: Re-Framing systems thinking in games for learning. To appear in *E-Learning*.

Dickey, M. D. (2005). Engaging by design: How engagement strategies in popular computer and video games can inform instructional design. *Education Training Research and Development, 53*(2), 67-83.

Dieterle, E. (2009). Neomillennial learning styles and River City. *Children, Youth and Environments, 19*(1), 245–278.

Dieterle, E. (in press). Games for science education. In A. Hirumi (Ed.), Playing [video] games in schools: Engaging learners through interactive entertainment. Eugene, OR: International Society for Technology in Education.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, *10*(2/3), 105-225.

diSessa, A. A. (2000). *Changing minds: Computers, learning and literacy*. Cambridge, MA: MIT Press.

diSessa, A. A., Hammer, D. M., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Children's meta-representational expertise. *The Journal of Mathematical Behavior*, *10*(2), 117.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). Taking science to school: Learning and teaching science in grades K-8. Washington, D.C.: The National Academics Press.

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the Learning Sciences*, *8*(3/4), 391-450.

Edelson, D. C., Salierno, C., Matese, G., Pitts, V., & Sherin, B. (2002). Learning-for-Use in Earth science: Kids as climate modelers. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, New Orleans, LA, April 2002.

Facer, K. Joiner, R. Stanton, D. Reid, J. Hull, R. & Kirk, D. (2004). Savannah: Mobile gaming and learning? *Journal of Computer-Assisted Learning, 20*, 399-409.

Faraday, P. M. & Sutcliffe, A. G. (1997). Designing effective multimedia presentations. Proceedings of CHI '97 ACM (pp. 272-278). Atlanta, GA.

Fletcher, J. D., & Tobias, S., (2006). *Using Computer Games and Simulations for Instruction: A Research Review.* Paper in Proceedings of the Society for Applied Learning Technology Meeting, New Learning Technologies, Orlando, FL, February 2006.

Fletcher, J. D., & Tobias, S., (2008). *What Do We Know About The Learning Effectiveness of Computer Games?* Paper presented at the American Educational Research Association 2008 conference.

Frederiksen, J. R., White, B. Y., & Gutwill, J. (1999). Dynamic mental models in learning science: The importance of constructing derivational linkages among models. *Journal of Research in Science Teaching, 36*(7), 806-836.

Frohberg, D., Goth, C., & Schwabe, G. (2009). Mobile Learning projects - a critical analysis of the state of the art. *Journal of computer assisted learning, 25*(4), 307-331.

Galas, C. (2006). Why Whyville? *Learning and Leading with Technology, 34*(6), 30-33.

Galloway, J. (2006). A great balancing act. Times Educational Supplement. April.

Gee, J. P. (2003/2007). What video games have to teach us about learning and literacy. New York: Palgrave Macmillan.

Goldman-Segall, R. (1996). Challenges facing researchers using multimedia tools. *Computer Graphics Quarterly, 28*(1), 48–52.

Greenfield, P. M., Camaioni, L., Ercolani, P., Weiss, L., & Lauber, B.A. (1994). Cognitive socialization by computer games in two cultures: Inductive discovery or mastery of an inconic code. *Journal of Applied Developmental Psychology, 15,* 59-85.

Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. In J. P. Mestre (Ed.), *Transfer of learning from a multidisciplinary perspective* (pp. 89-119). Greenwich, Connecticut: Information Age Publishing.

Hansen, S., Scrimpsher, D., & Narayanan, N. H. (1998). *From algorithm animations to animation-embedded hypermedia visualizations*. Technical Report CSE98-05, Dept. of Computer Science and Engineering, Auburn University.

Harel, I., & Papert, S. (1991). "Software Design as a Learning Environment." In I. Harel and S. Papert (Eds.), *Constructionism*. Norwood, NJ: Ablex.

Hayes, E. R. (2008). Game content creation and it proficiency: An exploratory study*Computers & Education, 51*(1), 97-108.

Hayes, E. R., & King, E. M. (2009). Not just a dollhouse: What The Sims2 can teach us about women's IT learning. *On The Horizon, 17*(1), 60-69.

Hickey, D. T., Filsecker, M. K. & Kwon, E. J. (2009).*Situative Considerations of Incentives and Competition in Educational Videogames*. Paper presented at the Educational Technology Special Interest Group Invited Symposium at the Biennial Meeting of the European Association for Research on Learning and Instruction, Amsterdam, August 2009.

Hickey, D., Ingram-Goble, & Jameson, E. (2009). Designing Assessments and Assessing Designs in Virtual Educational Environments.*Journal of Science Education and Technology, 18*(2),187–208.

Hines, P. J., Jasny, B. R., & Merris, J. (2009). Adding a T to the three R's*Science, 323*, 53.

Holbert, N. (2009). Learning Newton while crashing cars. Poster presented at GLS 2009, Madison, WI, June, 10-12.

Holland, J. (1998). *Emergence: From Chaos to Order*. Addison-Wesley, CA.

Huizinga, J. (1980). *Homo Ludens: A study of the play element in culture*. London: Routledge and Kegan.

Jackson, S. L., Stratford, S. J., Krajcik, J., & Soloway, E. (1996). A Learner-Centered tool for students building models.*Commununications of the ACM, 39*(4), 48-49.

Jackson, S., Krajcik, J., & Soloway, E. (2000). Model-It: A design retrospective. In M. J. Jacobson & R. B. Kozma (Eds.), Innovations in science and mathematics education: Advanced designs for technologies of learning (pp. 77–115). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Jacobson, M., Kim, B., Pathak, S., & Zhang, B. (2009). *Learning the physics of electricity with agent-based models: Fail first and structure later?* Paper presented at the Annual Meeting of American Educational Researchers Association. San Diego, CA: April 5, 2009.

Jenkins, H., Squire, K. and Tan, P. (2004). You can't bring that game to school! Designing Supercharged!

In B. Laurel (Ed.) *Design research*. Cambridge, Massachusetts: MIT Press.

Jones, M. G., Minogue, J., Tretter, T., Negishi, A., & Taylor, R. (2006). Haptic augmentation of science instruction: Does touch matter? *Science Education, 90*(1), 111-123.

Juul, J. (2003). The game, the player, the world: Looking for a heart of gameness. M. Copier and J. Raessens (Eds), *Proceedings of Level-Up: Digital games research conference* Utrecht: University of Utrecht, 30-45.

Kafai, Y. B., & Harel, I. (1991). Children's learning through consulting: When mathematical ideas, programming knowledge, instructional design, and playful discourse are intertwined. In I. Harel & S. Papert (Eds.), *Constructionism* (pp. 85-110). Norwood, NJ: Ablex.

Kafai, Y. B., Quintero, M., & Feldon, D. (in press). Investigating the 'Why' in Whypox: Casual and Systematic Explorations of a Virtual Epidemic. To appear in *Games and Culture*, January 2010.

Kapur, M. (2008). Productive failure. *Cognition and Instruction, 26*(3), 379 - 424.

Kapur, M., & Kinzer, C. (2009). Productive failure in CSCL groups. *International Journal of Computer-Supported Collaborative Learning, 4*, 21-46.

Keller, C. J., Finkelstein, N. D., Perkins, K. K., & Pollock, S. J. (2006). *Assessing the effectiveness of a computer simulation in introductory undergraduate dnvironments.* PERC Proceedings, 2006.

Ketelhut, D. J. (2007). The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in River City, a multi-user virtual environment *The Journal of Science Education and Technology, 16*(1), 99–111.

Ketelhut, D. J., Dede, C., Clarke J., & Nelson, B. (2006). A multi-user virtual environment for building higher order inquiry skills in science. Paper presented at the 2006 AERA Annual Meeting, San Francisco, CA, 7 to 11 April 2006; available at http://muve.gse.harvard.edu/rivercityproject/documents/rivercitysympinq1.pdf

Klopfer, E., Osterweil, & Salen (2009). *Moving Learning Games Forward*. The Education Arcade. Massachusetts Instiute of Technology. Retrieved on September 1, 2009 from http://www.educationarcade.org/

Klopfer, E., Osterweil, S., Groff, J., & Haas, J. (2009). Using the Technology of Today in the Classroom Today. The Education Arcade. Massachusetts Instiute of Technology. Retrieved on September 1, 2009 from http://www.educationarcade.org/

Klopfer, E., Scheintaub, H., Huang, W., Wendal, D. & Roque, R. (2009). The Simulation Cycle: combining games, simulations, engineering and science using StarLogo TNG. *E-Learning, 6*(1), 71-96.

Klopfer, E., Yoon, S. Um, T. (2005) Teaching complex dynamic systems to young students with StarLogo. Journal of Computers in Mathematics and Science Teaching [Online]. Available: http://dl.aace.org/16982. 24(2): 157-178.

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

Klopfer, E., Yoon, S., & Rivas, L. (2004). Comparative Analysis of Palm and Wearable Computers for Participatory Simulations. *Journal of Computer Assisted Learning, 20*, 347-359.

Lawler, R. (1982). Designing computer-based microworlds. *Byte, 8*, 138-155.

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.

Lepper, M. R. & Chabay, R. W. (1985). Intrinsic motivation and instruction: Conflicting views on the role of motivational processes in computer-based education. *Educational Psychologist, 20*(4), 217-231.

Lewis, E. L., Stern, J., & Linn, M. C. (1993). The effect of computer simulations on introductory thermodynamics understanding. *Educational Technology, 33*(1), 45-58.

Lindgren, R. & Schwartz, D. L. (2009). Spatial learning and computer simulations in science. *International Journal of Science Education, 31*(3), 419-438.

Linn, M. C., & Hsi, S. (2000). *Computers, teachers, peers: Science learning partners*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mandinach, E., & Cline, H. (1993). Systems, science and schools. *System Dynamics Review, 9*(2), 195-206.

Mayer, R. E., Mautone, P., & Prothero, W. (2002). Pictorial Aids for Learning by Doing in a Multimedia Geology Simulation Game. *Journal of Educational Psychology*, *94*, 171–185.

Mayo, M. J. (2009). Video games: A route to large-scale STEM education? *Science, 323*, 79-82.

McClean, P., Saini-Eidukat, B., Schwert, D., Slator, B., & White, A. (2001). In J. A. Chambers (Ed.), Selected Papers from the 12th International Conference on College Teaching and Learning (pp. 111–118). Jacksonville, FL: Center for the Advancement of Teaching and Learning.

McQuiggan, S. & Lester, J. (2007). Modeling and Evaluating Empathy in Embodied Companion Agents. *International Journal of Human-Computer Studies, 65*(4), 348-360.

McQuiggan, S., Mott, B., & Lester, J. (2008). Modeling Self-Efficacy in Intelligent Tutoring Systems: An Inductive Approach. *User Modeling and User-Adapted Interaction, 18*(1-2), 81-123.

McQuiggan, S., Rowe, J., & Lester, J. (2008). The Effects of Empathetic Virtual Characters on Presence in Narrative-Centered Learning Environments. In Proceedings of the 2008 SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy, pp. 1511-1520.

Meir, E., Perry, J., Stal, D., Maruca, S., and Klopfer, E. (2005). How effective are simulated molecular-level experiments for teaching diffusion and osmosis? *Cell Biology Education, 4*, 235-248.

Merrill, D. C. & Reiser, B. J. (1993). Scaffolding the acquisition of complex skills with reasoning-congruent learning environments. In Proceedings of the workshop in graphical representations, reasoning, and communication from the world conference on artificial intelligence in education. Edinburgh, Scotland.

Merrill, D. C. & Reiser, B. J. (1994). Reasoning congruent learning environments: Scaffolding learning by doing in new domains. Technical Report, Institute for Learning Sciences, Northwestern

University, Evanston, IL.

Miller, L. M., Moreno, J., Estrera, V. & Lane, D. (2004). Efficacy of MedMyst: An internet teaching tool for middle school microbiology *Microbiology Education, 5*, 13-20.

Moreno, R., & Mayer, R.E. (2000). Engaging students in active learning: The case for personalized multimedia messages. *Journal of Educational Psycholog, 92*, 724-733.

Moreno, R., & Mayer, R.E. (2004). Personalized messages that promote science learning in virtual environments. *Journal of Educational Psychology*, *96*, 165-173.

Moreno, R., & Mayer, R.E. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology, 97,* 117-128.

Naps, T. L., Chair, Working Group on Visualization (1996). An overview of visualization: Its use and design. In Proceedings of the conference on integrating technology into computer science education (pp. 192-200). Barcelona, Spain.

National Research Council. (1996). *The national science education standards.* Washington, D.C.: The National Academy Press.

National Research Council. (In Press). Hilton, M. (Ed.), Exploring the Intersection of Science Education and 21st Century Skills: A Workshop Summary. Washington, D.C.: National Academy Press.

Nelson, B. (2007). Exploring the use of individualized, reflective guidance in an educational multi-user virtual environment. *Journal of Science Education and Technology, 16*(1), 83-97.

Neulight, N., Kafai, Y. B., Kao, L., Foley, B., and Galas, C. (2007). Children's participation in a virtual epidemic in the science classroom: Making connections to natural infectious diseases. *Journal of Science Education and Technology, 16*(1), 47-58.

Newman, J. (2004). *Videogames*. London, UK: Routledge.

Norman, D. A. (1990). The Design of Everyday Things. Doubleday/Currency. Doubleday, New York.

O'Neil, H. F., Wainess, R., & Baker, E. L. (2005). Classification of learning outcomes: Evidence from the computer games literature. *The Curriculum Journal, 16,* 455-474.

Oppenheimer, Todd. (2003). The Flickering Mind. New York: Random House.

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

Park, O. & Hopkins, R. (1993). Instructional conditions for using dynamic visual displays: A review. *Instructional Science, 21*, 427-449.

Parnafes, O. (2007). What does "fast" mean? Understanding the physical world through computational representations. *Journal of the Learning Sciences, 16*(3), 415-450.

Provenzo, E. F. (1991). *Video kids: Making sense of Nintendo.* Cambridge, MA: Harvard University Press.

Raghavan, K., & Glaser, R. (1995). Model-based analysis and reasoning in science: The MARS curriculum. *Science Education, 79*(1), 37-61.

Reiner, M., Slotta, J. D., Chi, M. T. H., & Resnick, L. B. (2000). Naïve physics reasoning: A commitment to substance-based conceptions. *Cognition and Instruction, 18*(1), 1-34.

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: Lawrence Erlbaum.

Richards, J., Barowy, W., & Levin, D. (1992). Computer simulation in the science classroom. *Journal of Science Education and Technology, 1*(1), 67-79.

Rieber, L. P. (1990). Animation in computer-based instruction. *Educational Theory, Research, and Development, 38*(1), 77-86.

Rogers, Y., & Scaife, M. (1998). How can interactive multimedia facilitate learning? In Lee, J. (Ed.) *Intelligence and Multimodality in Multimedia Interfaces: Research and Applications.* Menlo Park, CA: AAAI. Press.

Roschelle, J. (1991). Students' construction of qualitative physics knowledge: Learning about velocity and acceleration in a computer microworld. Unpublished doctoral dissertation, University of

California, Berkeley.

Roschelle, J. (2003). Unlocking the learning value of wireless mobile devices. *Journal of Computer Assisted Learning, 19*(3), 260-272.

Roschelle, J., & Teasley, S. (1995). The Construction of Shared Knowledge in Collaborative Problem Solving. In C. O'Malley (Ed.), *Computer Supported Collaborative Learning*. Berlin: Springer-Verlag.

Rosenbaum, E., Klopfer, E., and Perry, J. (2006). On location learning: authentic applied science with networked augmented realities. In Press for *Journal of Science Education and Technology.* Rothberg, M.A.,

Rothberg, M.A., Sandberg, S., & Awerbuch, T.E. (1994). Educational software for simulating risk of HIV infection. *Journal of Science Education and Technology, 3*(1), 65-70.

Salen, K., & Zimmerman, E. (2003). Rules of Play: Game Design Fundamentals. Cambridge: MIT Press.

Sandberg, S., & Awerbuch, T.E. (1994). Educational software for simulating risk of HIV infection. *Journal of Science Education and Technology, 3*(1), 65-70.

Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *Journal of the Learning Sciences, 12*(1). 5-51.

Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education, 88*. 345-372.

Sanford, R., Ulicsak, M., Facer, K., & Rudd, T. (2006). *Teaching with games: Using commercial off-the-shelf computer games in formal education*. Bristol: Futurelab.

Schecker, H. (1993). Learning Physics by Building Models. *Physics Education, 28*, 102 - 106.

Schwartz, D. L. & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction, 22*(2), 129-184.

Schwartz, D. L., Bransford, J. D., & Sears, D. (2005). Efficiency and innovation in transfer. In J. P.

Mestre (Ed.), *Transfer of learning from a multidisciplinary perspective* (pp. 1-51). Greenwich, Connecticut: Information Age Publishing.

Schwartz, D. L., Chase, C., Chin, C., Oppezzo, M., Kwong, H., Okita, S., Biswas, G., Roscoe, R. D., Jeong, H., & Wagster, J. D. (2007). Interactive metacognition: Monitoring and regulating a teachable agent. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition in Education*. Mahwah, NJ: Erlbaum.

Schwarz, C., & White, B. (2005). Meta-modeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction, 23*(2), 165-205.

Sengupta, P. & Wilensky, U. (2006) *NIELS: An agent-based modeling environment for learning electromagnetism*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Sengupta, P., & Wilensky, U. (2008a). *Designing Across Ages: On The Low-Threshold-High-Ceiling Nature of NetLogo Based Learning Environments* . Paper presented at Paper presented at the annual meeting of the American Educational Research Association (AERA 2008), New York, NY.

Sengupta, P., & Wilensky, U. (2008b). *On The Learnability of Electricity As A Complex System*. In M. Jacobson (Chair) and R. Noss (Discussant), "Complex Systems & Learning: Empirical Research, Issues & "Seeing" Scientific Knowledge With New Eyes." In Proceedings of the International Conference for the Learning Sciences, 2008 (Manuscript in press).

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.

Snir, J., Smith, C., & Grosslight, L. (1993). Conceptually enhanced simulations: A computer tool for science teaching. *Journal of Science and Technology, 2*(2), 373-388.

Son, J. Y., & Goldstone, R. L. (2009). Fostering General Transfer with Specific Simulations. *Pragmatics and Cognition, 17*, 1-42.

Squire, K. (in press). From information to experience: Place-based augmented reality games as a model for learning in a globally networked society. To appear in *Teacher's College Record*

Squire, K. D. & Jan, M. (2007). Mad City Mystery: Developing scientific argumentation skills with a place-based augmented reality game on handheld computers. *Journal of Science Education and Technology, 16*(1) 5-29.

Squire, K. D., DeVane, B. & Durga, S. (in press). Designing centers of expertise for academic learning through video games. To appear in Theory Into Practice.

Squire, K., & Durga, S. (in press). Productive gaming: The case for historiographic game play. To appear in R. Ferdig (Ed.) The handbook of educational gaming. Hershey, PA: Information Science Reference.

Squire, K., & Klopfer, E. (2007). Augmented reality simulations on handheld computers. Journal of the Learning Sciences, 16(3), 371 - 413.

Squire, K., Barnett, M., Grant, J. M., & Higginbotham, T. (2004). Electromagnetism supercharged!: learning physics with digital simulation games. In Y. B. Kafai, W. A. Sandoval, N. Enyedy, A. S. Nixon, & F. Herrera, (Eds.), Proceedings of the 6th International Conference on Learning Sciences (pp. 513–520). Los Angeles: UCLA Press.

Squire, K., Giovanetto, L., Devane, B. & Durga, S. (2005). *From users to designers: building a self-organizing game-based learning environment.* Paper presented at the annual convention of the American Educational Research Association, Montreal, April, 2005.

Steinkuehler, D., & Duncan, S. (2008). Scientific habits of mind in virtual worlds. *Journal of Science Education and Technology, 17*(6), 530-543.

Stieff, M., & Wilensky, U. (2003). Connected Chemistry - incorporating interactive simulations into the chemistry classroom. *Journal of Science Education and Technology, 12*(3), 285-302.

Sweller, J. (1993). Some cognitive processes and their consequences for the organization and presentation of information. *Australian Journal of Psychology, 45*(1), 1-8.

Tabak, I., & Reiser, B. J. (1997). Complementary roles of software-based scaffolding and teacher-student interactions in inquiry learning. In R. Hall, N. Miyake, & N. Enyedy (Eds.), In Proceedings of the Computer Supported Collaborative Learning Conference (pp. 289-298).

Tabak, I., Sandoval, W. A., Smith, B. K., Agganis, A., Baumgartner, E., & Reiser, B. J. (1995). Supporting Collaborative Guided Inquiry in a Learning Environment for Biology. In J. L. Schnase & E. L. Cunnius (Eds.), *Proceedings of CSCL '95: The First International Conference on Computer Support for Collaborative Learning*, (pp. 362-366). Bloomington, IN: Erlbaum.

Tan, J., Skirvin, N., Biswas, G. & Catley, K. (2007). *Providing Guidance and Opportunities for Self-Assessment and Transfer in a Simulation Environment for Discovery Learning.* The twenty-ninth Annual Meeting of the Cognitive Science Society, Nashville, Tennessee, (pp. 1539).

Tan, J. & Biswas, G. (2007). *Simulation-Based Game Learning Environments: Building and Sustaining a Fish Tank.* The first IEEE International Workshop on Digital Game and Intelligent Toy Enhanced Learning, Jhongli, Taiwan, (pp. 73-80).

Tuzan, H. (2004). *Motivating Learners in Educational Computer Games.* Indiana University, Bloomington.

Ward, M. & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction, 7*(1), 1-39.

White, B. (1993a). Intermediate causal models: A missing link for successful science education? In R. Glaser (Ed.), *Advances in instructional psychology, Volume 4* (pp. 177–252). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

White, B. Y. (1993b). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction, 10*(1), 1–100.

White, B., & Frederiksen, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*(1), 3-118.

Wieman, C. E., Adams, W. K., & Perkins, K. K. (2008). PhET: Simulations that enhance learning.

Science, 322, 682-683.

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

Wilensky, U. (2003). Statistical mechanics for secondary school: The GasLab Modeling Toolkit. *International Journal of Computers for Mathematical Learning*, *8*(1), 1-41.

Wilensky, U., & Reisman, K. (1998). Learning biology through constructing and testing computational theories - An embodied modeling approach. In Y. Bar-Yam (Ed.), *Proceedings of the Second International Conference on Complex Systems*. Nashua, NH: New England Complex Systems Institute.

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 & Instruction, 24*(2), 171-209.

Wilensky, U., & Stroup, W. (1999a). HubNet. http://ccl.northwestern.edu/netlogo/hubnet.html. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL.

Wilensky, U., & Stroup, W. (1999b). *Learning through participatory simulations: Network-based design for systems learning in classrooms.* Proceedings of Computer Supported Collaborative Learning (CSCL'99). Stanford, CA, December 12 - 15.

Wilensky, U., & Stroup, W. (2000). Networked gridlock: Students enacting complex dynamic phenomena with the HubNet architecture. Proceedings of the Fourth Annual International Conference of the LearningSciences, Ann Arbor, MI, June 14 - 17.

Wittgenstein, L. (1972). The blue and brown books: Preliminary studies for the 'Philosophical Investigations'. Oxford: Basil Blackwell.