Metadata of the article that will be visualized in Online

Please note: Image will appear in color online but will be printed in black and white.

1	Article Title	**Mr. Vetro: A collective simulation for teaching health science**	
2	Article Sub- Title		
3	Article Copyright - Year	**International Society of the Learning Sciences, Inc.; Springer Science+Business Media, LLC 2010 (This will be the copyright line in the final PDF)**	
4	Journal Name	International Journal of Computer-Supported Collaborative Learning	
5	Corresponding Author	Family Name	**Ioannidou**
6		Particle	
7		Given Name	**Andri**
8		Suffix	
9		Organization	AgentSheets, Inc.
10		Division	
11		Address	6560 Gunpark Dr., Boulder 80301, CO, USA
12		e-mail	andri@agentsheets.com
13	Author	Family Name	**Repenning**
14		Particle	
15		Given Name	**Alexander**
16		Suffix	
17		Organization	AgentSheets, Inc.
18		Division	
19		Address	6560 Gunpark Dr., Boulder 80301, CO, USA
20		e-mail	alexander@agentsheets.com
21	Author	Family Name	**Webb**
22		Particle	
23		Given Name	**David**
24		Suffix	
25		Organization	University of Colorado
26		Division	School of Education
27		Address	Boulder 80309-0249, CO, USA
28		e-mail	dcwebb@colorado.edu
29	Author	Family Name	**Keyser**
30		Particle	
31		Given Name	**Diane**
32		Suffix	
33		Organization	University of Colorado

34 Division: School of Education
35 Address: Boulder 80309-0249, CO, USA
36 e-mail: diane.keyser@colorado.edu

Author
37 Family Name: **Luhn**
38 Particle:
39 Given Name: **Lisa**
40 Suffix:
41 Organization: St. Helena High School
42 Division:
43 Address: 1401 Grayson, St. Helena 94574, CA, USA
44 e-mail: ljluhn@msn.com

Author
45 Family Name: **Daetwyler**
46 Particle:
47 Given Name: **Christof**
48 Suffix:
49 Organization: Drexel University
50 Division: College of Medicine
51 Address: 2900 Queen Lane, Philadelphia 19129, PA, USA
52 e-mail: Christof.Daetwyler@DrexelMed.edu

Schedule
53 Received: 19 October 2007
54 Revised:
55 Accepted: 3 March 2010

56 Abstract

Why has technology become prevalent in science education without fundamentally improving test scores or student attitudes? We claim that the core of the problem is *how* technology is being used. Technologies such as simulations are currently not used to their full potential. For instance, physiology simulations often follow textbooks by sequentially exposing individual systems such as the circulatory and respiratory systems one at a time, leaving out essential comprehension of system interactions. Moreover, the standard computer lab hides students behind large monitors and ignores the social aspect of learning. We have created a new kind of infrastructure, called *Collective Simulations* to provide engaging inquiry-based science learning modules that uniquely combine social learning pedagogies with distributed simulation technology. This infrastructure creates immersive learning experiences based on wirelessly connected computers and enables radically different classroom learning experiences that engage students and teachers simultaneously. Collective Simulations allow students to learn about the intricacies of interdependent complex systems by engaging in discourse with other students and teachers. As part of our Mr. Vetro collective simulation, students learn about physiology through technology-enhanced role-play. Each group controls physiological variables of a single organ on their computer. A central simulation gathers all the data and projects the composite view of a human. In an example activity, the heart and lung teams collaborate to adjust parameters and reach homeostasis. Results from formal evaluation studies demonstrate a positive impact on scientific inquiry, student learning, and students' interest in personal health issues. This article describes Mr. Vetro and its underlying architecture and presents the evaluation results.

57 Keywords separated by ' - ': Collective simulations - Distributed simulations - Social learning pedagogies - Interdependent complex systems - Meaningful learning

58 Foot note information:

Computer-Supported Collaborative Learning
DOI 10.1007/s11412-010-9082-8

Mr. Vetro: A collective simulation for teaching health science

Andri Ioannidou · Alexander Repenning · David Webb · Diane Keyser · Lisa Luhn · Christof Daetwyler

Received: 19 October 2007 / Accepted: 3 March 2010

Abstract Why has technology become prevalent in science education without fundamentally improving test scores or student attitudes? We claim that the core of the problem is *how* technology is being used. Technologies such as simulations are currently not used to their full potential. For instance, physiology simulations often follow textbooks by sequentially exposing individual systems such as the circulatory and respiratory systems one at a time, leaving out essential comprehension of system interactions. Moreover, the standard computer lab hides students behind large monitors and ignores the social aspect of learning. We have created a new kind of infrastructure, called *Collective Simulations* to provide engaging inquiry-based science learning modules that uniquely combine social learning pedagogies with distributed simulation technology. This infrastructure creates immersive learning experiences based on wirelessly connected computers and enables radically different classroom learning experiences that engage students and teachers simultaneously. Collective Simulations allow students to learn about the intricacies of

A. Ioannidou (✉) · A. Repenning
AgentSheets, Inc., 6560 Gunpark Dr., Boulder, CO 80301, USA
e-mail: andri@agentsheets.com

A. Repenning
e-mail: alexander@agentsheets.com

D. Webb · D. Keyser
School of Education, University of Colorado, Boulder, CO 80309-0249, USA

D. Webb
e-mail: dcwebb@colorado.edu

D. Keyser
e-mail: diane.keyser@colorado.edu

L. Luhn
St. Helena High School, 1401 Grayson, St. Helena, CA 94574, USA
e-mail: ljluhn@msn.com

C. Daetwyler
College of Medicine, Drexel University, 2900 Queen Lane, Philadelphia, PA 19129, USA
e-mail: Christof.Daetwyler@DrexelMed.edu

interdependent complex systems by engaging in discourse with other students and teachers. As part of our Mr. Vetro collective simulation, students learn about physiology through technology-enhanced role-play. Each group controls physiological variables of a single organ on their computer. A central simulation gathers all the data and projects the composite view of a human. In an example activity, the heart and lung teams collaborate to adjust parameters and reach homeostasis. Results from formal evaluation studies demonstrate a positive impact on scientific inquiry, student learning, and students' interest in personal health issues. This article describes Mr. Vetro and its underlying architecture and presents the evaluation results.

Keywords Collective simulations · Distributed simulations · Social learning pedagogies · Interdependent complex systems · Meaningful learning

Problem: Science illiteracy Q1

Science, Technology, Engineering, and Mathematics (STEM) education in the United States is in dire need for improvement (Committee on Prospering in the Global Economy of the 21st Century 2005). In spite of significant technological investment, the US is still lacking in these fields. In the 2003 PISA study assessing scholastic performance, the USA scored 24th of the 29 countries compared (OECD 2003). In the 2006 PISA survey, the USA had still not managed to get into the top 20 of mathematics, science, or reading.

Particularly concerning is the lack of scientific understanding that goes beyond academics and is essential to everyday life. Low health literacy, that is, the struggle of people obtaining, processing, and understanding basic health information, is a fundamental problem for the entire society. A lack of understanding regarding smoking, substance abuse, and nutrition, for instance, can significantly lower the quality of life.

The urgency of addressing science education challenges—and particularly health science education, both at the academic and pragmatic levels—is commonly understood. However, it is much less clear what the underlying components of science illiteracy and apathy are and what kind of approaches could be employed to increase both scientific knowledge and interest in science. One line of thinking is focused on information access. It is argued that people simply make poor decisions regarding nutrition and drug abuse because they do not have the facts. Various approaches are proposed, such as providing Web sites or podcasts that reach out to the public to inform them about the facts. Some successes have been documented as a result of these campaigns, but by and large the emergence of the Web has surprisingly little impact on public science understanding (National Science Foundation, Division of Science Resources Statistics 2008). These are important indicators that science education must move beyond the simple model of transmitting facts and provide engaging STEM education approaches.

Only addressing motivational concerns is not sufficient. Video games appear to work where other approaches clearly fail (Prensky 2006). Seemingly attention-challenged students spend huge amounts of time solving intricate game puzzles. While the sheer size of the game industry is a clear sign for the cultural relevance of games, it is not obvious at all how educational games can strike a meaningful balance between education and engagement. Educational games typically do not succeed if they simply try to spice up bland facts through supposedly fun activities.

Using technology to teach science, including games, has the potential to address both educational and motivational issues of science education. But why, in spite of the $70 billion investment in technology during the 1990s (Oppenheimer 2003), has K-12 science education failed to excite students? Why has technology become prevalent in science education without fundamentally improving test scores (U.S. Department of Education 2003) or student attitudes? We believe that our educational institutions need a new combination of pedagogy with tools that helps K-12 students understand highly complex and interacting systems and genuinely excites them about science.

Solutions: Collective simulations

Our *Collective Simulations* approach is a way to make science learning and teaching more meaningful, engaging, and effective. The *Collective Simulations* approach is a conceptual framework that integrates *social learning pedagogical models* with *distributed simulation technical frameworks*. This conceptual framework both allows and actively encourages meaningful learning by supporting discovery-oriented social learning processes. In a Collective Simulation, an entire group of students and their teacher—acting as a discourse facilitator—participate in a simulation-enhanced role-play that unfolds in real time on a set of networked computers running distributed simulations. These computers are smart artifacts or tools that augment discourse and support the discovery-oriented scientific process.

Collective Simulations address cognitive, emotional/affective, and social aspects of learning.

- *Cognitive dimensions*. Beyond the ability to learn and retain certain facts, Collective Simulations afford inquiry-based science and enable students through self-directed discoveries to combine facts into higher-level reasoning.
- *Affective dimensions*. Given the non-cognitive components of decision processes concerning health-related issues, such as smoking or substance abuse, it would stand to reason that one should explore how to complement the cognitive part of health science education with influence on student affect for pro-social outcomes. Collective Simulations do that with audiovisual stimuli that immerse students into the scenarios they are exploring.
- *Social dimensions*. Using technology per se cannot improve teaching without understanding the social processes and the context in which technology such as Collective Simulations will be used. The pedagogical dimensions of social processes are key indicators of the usefulness of a simulation in a classroom context. The distributed nature of Collective Simulations stimulates discussions, which in turn lead to meaningful learning (Michael 2001).

Fundamentally, a Collective Simulation is a cyberlearning (National Science Foundation, Task force on Cyberlearning 2008) supported role-playing activity in which a group of people learns about complex systems, such as the human body, by discussing relationships, making decisions, and experiencing interactions through a real-time simulation process. For instance, in human physiology, each person "plays" the role of an individual organ or system such as the heart or the lungs. The technological part of role-play is a handheld or laptop computer simulating and embodying the organ. This cyberlearning infrastructure is responsible for at least two functions. Each individual simulation, for example, the simulation of the lung, is wirelessly connected to a central simulation hub integrating all the individual simulations. Typically, this central simulation is also running on a computer with a display visible to the entire group. In a human physiology Collective Simulation, this computer presents not only all the physiological variables but also a visible and animated manifestation of the entire human being. The Collective Simulation essentially takes place on two levels. At the individual level, local decisions are being made such as what should the heart rate be to match the physiological needs of some physical exercise. At the collective level, the simulation visualizes the consequence of all the local decisions.

Collective Simulations are highly related to participatory simulations (Colella 2000; Wilensky & Stroup 1999). In both kinds of simulations, there are multiple levels of representations capturing the intricate interdependencies of complex systems for educational purposes. Both kinds of simulations involve some role-play supported by networked technology. However, collectivism strongly prioritizes group goals over individual goals, suggesting a focus on a collective function. With Mr. Vetro, the collective function is that of a working, balanced human being able to achieve tasks such as walking, running, or biking. In other words, while both kinds of simulations are instruments to explore emergent, bottom-up phenomena, Collective Simulations must also include top–down perspectives to guide investigations with specific collective goals. Specifically, with Mr. Vetro, collectivism suggests that we would be more interested in defining a group function and then adjusting all the interconnected organs accordingly to achieve said group function over first adjusting organs and then exploring an emergent group function.

The following sections present the technological, medical, and educational approaches for using Collective Simulations in health science education with the Mr. Vetro application.

Technological approach: The Mr. Vetro collective simulation

Collective Simulations are based on the C^5 conceptual framework that we have been developing with NIH support, which integrates distributed simulation technical frameworks with social learning pedagogical models. C^5 is an Information Technology infrastructure that integrates the principles of ***C***ollective, ***C***ompact, ***C***onnected, ***C***ontinuous, and ***C***ustomizable simulations (Repenning & Ioannidou 2005). This framework both allows and actively encourages meaningful learning by supporting discovery-oriented social learning processes. In a Collective Simulation, an entire group of students and their teacher—acting as a discourse facilitator—participate in a collaborative simulation-based activity that unfolds in real time on a set of networked computers running distributed simulations.

We have implemented a prototype of a Collective Simulation, called Mr. Vetro, and have evaluated it in local high schools. In the Mr. Vetro[1] Collective Simulation (Fig. 1), different human systems or organs are simulated on wirelessly connected computers, while a central simulation aggregates parameters from the organs and computes Mr. Vetro's vital signs. Using the distributed simulations, students control Mr. Vetro's organs. The lung group can vary parameters such as breathing rate and tidal volume in response to changing conditions such as exercise or smoking. The heart group can vary parameters such as heart rate and stroke volume. The role-play aspect of Collective Simulations makes students collaborate, as organs or systems, to deal with specific goals. For instance, a group of students can make Mr. Vetro engage in intense exercise—this forces his body and, therefore, the Collective Simulation participants to cope with the need to provide more oxygen to the muscle tissue and deal with increased levels of CO_2. This leads to scientific discourse between the "organs" that includes justifications and results in the adjustment of parameters such as the

[1] An interactive flyer can be found at: http://agentsheets.com/research/c5/documents/interactive%20flier/c5-flier.html.

Fig. 1 Mr. Vetro Collective Simulation: Students control organs (heart, lungs) distributed on wirelessly connected computers. The central simulation projects the entire human being, calculated vital signs, and a blood-centric representation of the body. Mr. Vetro and his doctor communicate the situation (e.g., hyperventilation) via messages. The teacher acts as discourse facilitator

heart rate and stroke volume for the heart or the breathing rate and tidal volume for the lungs to deal with a certain situation.

Mr. Vetro is based on a wireless network connecting computers and integrating the simulated organs into a composite higher-level representation of a human being with calculated physiological variables such as oxygen saturation, partial pressure of carbon dioxide, oxygen needed, and oxygen delivered. Mr. Vetro and his physiological variables are represented in the central simulation and projected to the entire group in order to visualize the consequences of local decisions. The Collective Simulation does not automatically balance Mr. Vetro's life signs. As Mr. Vetro begins to jog, his need for oxygen goes up. The heart and lung teams have to collaborate to keep Mr. Vetro in a healthy state. Changes to physiological variables such as the breathing rate immediately update the entire Collective Simulation (visually and audibly). The immersive nature of Mr. Vetro goes beyond just displaying numerical values—it employs sophisticated real-time visualizations using our unique real-time visualization engine (Agent Warp Engine; Repenning & Ioannidou 2008), and even auditory stimuli providing immersive experiences to learners.

The *Agent Warp Engine* (AWE; Repenning & Ioannidou 2008) is a new end-user framework that is capable of creating sophisticated, interactive, networked visualizations. This framework, which goes beyond regular animations to create medical visualizations and networked simulations such as our Mr. Vetro, is based on spreadsheet ideas in a way that provides rich end-user visualizations and networking. AWE is implemented as a thin layer on top of our Open Agent Engine[2] (OAE) which itself is the open source part of the AgentCubes 3D game and simulation-authoring environment (Repenning & Ioannidou 2006). The OAE implements a simple 3D agent-based simulation engine based on four main components:

- *OpenGL: 2D/3D Graphics*. A highly optimized 2D/3D API[3] for rendering large numbers of agents with 3D shapes efficiently.
- *QuickTime*: An API available for multiple platforms, which provides access to image, movie, and sound files.
- *XMLisp*: An API mapping XML (Extensible Markup Language[4]) expressions to object-oriented language constructs (Repenning & Ioannidou 2007).
- *3D Agents*. Autonomous objects that have a 3D position, orientation, size, velocity, and acceleration. Agents can be composed into scenes, animated, displayed, and user selected.

AWE adds the end-user visualization and end-user networking components to the architecture (Fig. 2).

- *End-User Visualizations* are end-user created custom visualizations using techniques such as shape warping. They are end-user accessible, rich, and efficient. With AWE, end users can make their own intricate realistic visualizations. An important goal of our overall Collective Simulations work is to offer refined kinds of visualizations necessary to communicate complex dynamic processes. In Mr. Vetro, we need to illustrate the function of the heart, the lung, and the human skeleton. All three systems mechanically interact with each other in complex ways. For instance, inhaling air will change the shape of the lung, which in turn will influence the skeleton. Ribs expand and, in the case of deep breathing, even the position of the shoulders and arms can be influenced. AWE offers a number of visualizations, but the most sophisticated one (called "morph") is specifically designed to implement complex visualization based on shape warping (Repenning & Ioannidou 2008). Shape warping is a kind of image warping (Wolberg

[2] http://www.agentsheets.com/lisp/OpenGL.html
[3] http://www.opengl.org
[4] http://www.w3.org/XML/

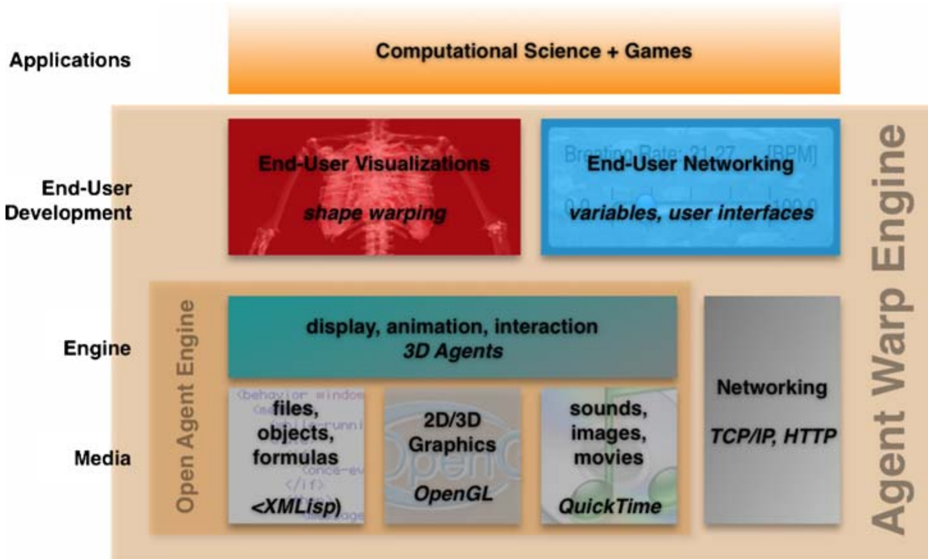

Fig. 2 The agent warp engine is layered on top of the open agent engine

1994, 1996). The media richness, both visual and audio, immerses learners into the interaction. For instance, a set of variables representing a hyperventilating human being should not just be a list of numbers, but should show a human being with heart and lung functioning in a state of hyperventilation, even with sounds that provoke emotional responses.

- *End-User Networking* is essentially a formula language featuring variables distributed on different computers and equations connecting them. End-User Networking provides transparent networking where an end user does not have to use network APIs (Advanced Programming Interfaces) or any other complex mechanism to implement variable sharing. This approach enables end users to easily create distributed applications, such as Mr. Vetro. End users (and developers of Collective Simulations) should be able to define, change, and connect these networked variables as easily as they use variables in spreadsheets. For our Collective Simulations, it is essential that these variables can be shared in real time by different clients running on the same computer or separate networked computers. Being able, as a group, to see how individual users change simulation variables in real time and how the system reacts to change helps the perception of causality (Michotte 1962).

AWE is a flexible XML-based architecture that allows for extension of existing and creation of new Collective Simulations. With this modular architecture, activity developers are able to add new systems/organs, new variables (input or output), and new formulae connecting them. Moreover, one can create different interactive activities with any subsets of organs where only some aspects of the systems/organs are visible and controllable by students while other relationships can still be in the background but not necessarily visible. These capabilities provide a plug-and-play system that opens up the possibility of creating multiple educational activities and units without essentially changing the underlying software. Customizations can happen just at the end-user level with end-user programming and customization techniques. AWE was instantiated as the new Mr. Vetro prototype (Fig. 1).

Medical approach: Blood-centric view of human body

Our current version of Mr. Vetro features a blood-centric approach to the human body (Fig. 1 right). Initially, the Mr. Vetro system was used to illustrate the heart as a blood pump and the lung as an air pump, and how the parameters pO_2 and pCO_2 are dependent on how these two pumps work together—which were represented by two teams, each working on Mr. Vetro's lungs or heart. We worked with medical doctors to design the system and make a teaching tool that is more versatile, engaging, and relevant. The questions that the medical team tried to answer first to guide the development of Mr. Vetro included: What model of anatomy/physiology could/should be taught at the high school and middle school level? For which concepts will student learning be most enhanced by the Mr. Vetro simulation technology?

As long as we were looking at the organs from the organ system's viewpoint, it was not easy to find a way to add new systems and show how they interact. We needed to switch the viewpoint from the organs to the blood, which could be described as a very complex fluid organ, to find the solution that we now call "blood-centric" approach. In this approach, we imagine a blood cell that circulates in the blood stream. At any given moment, we can measure what substances (such as oxygen, carbon dioxide, or glucose) are present in its surrounding. We also know at which organ systems the blood can get rid of specific substances, if there is too much of the substance present—or where it can get more of that substance if there is a need for more.

Having a system like Mr. Vetro, where groups of students play the role of organ systems, using the blood-centric approach allows the system to easily engage with all the parameters in a meaningful way. Collective Simulations allow the students to explore the concepts by running the disseminated Mr. Vetro simulations in a way that results not only in deep learning, but creates a fascination for the sciences.

The comprehension of interdependent complex systems (Epstein & Axtell 1996; Senge 1990; Simon 1981) is an enormous challenge to learners. We reviewed physiology teaching materials at middle, high school, and university levels; we included textbooks, interactive CD-ROMs, and Web-based resources and titles offered by major publishers such as Pearson.[5] We found a ubiquitous absence of explanations for how systems *interact*. All the materials we reviewed structured the human body into decoupled subsystems (e.g., respiratory and circulatory) isolated in separate chapters or animations with little, if any, connection. These neglected interactions are central to understanding how the human body works. Indeed, the general need to comprehend interactive systems is part of current educational standards (National Research Council 1996; Thompson & Celva 2005). Realizing the need for more rigorous instruction about the interactions of systems instead of each system in isolation, many medical schools have changed or are in the process of changing their curricula to reflect a more integrated systems approach. However, this kind of change has largely been ignored at the K-12 level because of the inadequacy of traditional teaching media. Collective Simulations can address this problem by creating an immersive learning environment that lets students collaboratively *experience* system interactions. These same interactive system concepts apply not only to physiology, but to other domains such as economics, ecology, and biology.

Educational approach: Inquiry-based, sheltered science teaching, meaningful learning

Working with high school science teachers and educational experts from the University of Colorado School of Education, we developed educational materials for a complete

[5] Pearson/Benjamin Cummings *Interactive Physiology* from http://www.aw-bc.com/info/ip/

cardiopulmonary unit that covered aspects of the normal physiology of the circulatory and respiratory systems, exercise physiology as it relates to the cardiopulmonary system, and scenarios of abnormal states of the system due to situations such as substance abuse. One of our main objectives for this unit and essential question we wanted the students to explore was understanding how smoking and exercise affects the cardiopulmonary system. In activities using the Mr. Vetro Collective Simulation, students compared an average person with a sedentary smoker and an elite athlete. We wanted students completing this unit to be able to explain what it means to "be in shape" in terms of the heart, lungs, and blood cells and to be able to articulate what the physiological difference is between an athlete and a "couch potato." By using Mr. Vetro for these activities, students discovered typical values for the heart rate, breathing rate, stroke volume, and tidal volume of an average person, an elite athlete, and a sedentary smoker. Students also simulated Mr. Vetro exercising and discovered his anaerobic threshold and how to keep him in the aerobic range while exercising. We also integrated math into the activity by asking students to predict and collect data from the simulation for the three Mr. Vetros (average, elite athlete, sedentary smoker) then graph and interpret their data.

The educational materials developed included complete lesson plans, student worksheets, teacher background material, overheads, off-computer activities and associated material, and evaluation material that consisted of pre–and post-assessments to gauge content learning and a survey to assess student attitudes toward science.

Development of these materials was based on our philosophy of *inquiry-based, self-directed science* using topics that are relevant to students' personal lives, using *sheltered instruction* techniques and encouraging social aspects for more *meaningful learning* experiences. The design work involved several iterative design studies (Brown 1992) that involved one of the co-authors trialing prototype lessons with small groups of students in regular classes, after school, or during the summer. In these design and feasibility studies (Repenning et al. 2010), we focused on the development of activities that promoted student engagement and scientific inquiry.

Inquiry science is a student-centered approach to the teaching of science that allows students to use their natural curiosity to learn about the world. In inquiry science, lectures are replaced with investigations, discussions, and problem solving. Inquiry is not just hands-on learning, but it involves thought and discussion centered on classroom activities (Stoddart et al. 2002). Inquiry science is the exploration of scientific phenomena with hands-on, process-oriented experiments. In the context of inquiry, students are allowed to be actively engaged in the process of science (Hampton & Rodriguez 2001). Students solve problems with the use of hands-on activities and discussion of their thinking with other students. New understandings of content are made possible by this type of classroom interaction. The Mr. Vetro simulation and lessons developed around it allow for true inquiry. Not only are students actively learning how the simulation works, they are making predictions, testing their predictions, and incorporating new understandings into their mental model of how the human body works. The Mr. Vetro activities require students to synthesize information and draw connections between old and new information. These higher-order thinking skills are difficult to achieve and assess through traditional methodologies such as lectures, textbook work, and basic lab experiences. With even further developments, the Mr. Vetro Collective Simulation could easily replace more of these activities with a dynamic, experiential, collaborative, and inquiry-based classroom activity.

The lessons were developed for use in a mainstream classroom to allow for the integration of English language learners with native English speakers. Sheltered instruction

JrnlID 11412_ArtID 9082_Proof# 1 - 09/03/2010

is content teaching to English language learners (ELL) in strategic ways that make the concepts comprehensible while promoting the students' academic language development (Short and Echevarria 2004). The goal of sheltered instruction is to achieve full and equitable education for ELLs. To do this, the curriculum must be adapted to allow content to be accessible while the English proficiency of ELLs is increased. If this kind of curriculum is to be adopted in a mainstream classroom, the rigor must not be compromised for the non-ELL students in the process. Stoddart et al. (2002) believe that inquiry-based science is a natural and effective tool for the integration of the usually separate fields of language and science and that a synergistic union of language development and content instruction is possible. The Mr. Vetro activities encourage the use of language—it is essential to make the system operate effectively. Students must read the message and decode the numerical values while communicating the changes they are making to the system with the group. Mr. Vetro is the foundation for the synergism between science and language. The use of sheltered curriculum has been described as the best way to integrate the development of the English language and the teaching of content for students who are English language learners (Echevarria et al. 2004; Gibbons 2002; Hill and Flynn 2006).

It would be difficult to effectively teach science with technology without understanding the social processes and the context in which educational simulations will be used. Joel Michael's notion of *meaningful learning* is based on ideas of a social process that includes active externalization of ideas and engagement in discourse:

> One of the most "active" ways to help students to challenge with their own mental models is to get them to "talk physiology." Discussing, justifying, or explaining their answers to questions with one another, or with the instructor, is a powerful way to encourage meaningful learning. (Michael 2001)

Active externalization, reflection, and discourse as learning processes are highly consistent with Piaget's notion of constructivism (Bruner 1990; Piaget 1972, 1990; Yager 1995). According to him, learning is the process of adjusting our mental models to accommodate new experiences. Simulation-based learning processes are also consistent with Papert's notion of constructionism, in which learning is the result of a externalization-as-construction process of physical and conceptual artifacts (Papert 1980; Papert 1993). In the medical realm, Miller's Pyramid is similar in spirit. Miller's Pyramid is a framework for assessing clinical competence (Norcini 2003). At the lowest level of the framework's pyramid is knowledge (knows), followed by competence (knows how), performance (shows how), and action (does). This categorization is consistent with constructionist approaches in the educational realm. Our Collective Simulations framework combines the action/construction part of these frameworks with another important dimension: *communication*. The social aspect of Michael's notion of meaningful learning is implied by his focus on communication. Vygotzky's theory of social construction of knowledge (Vygotsky 1980) indicated that a learning process is substantially aided by active collaboration that includes discourse. His notion of "Zone of Proximal Development" implies that a child's development is determined by the social interaction and collaborative problem solving. Hutchins' theories of distributed cognition go one step further by adding the notion of tools and artifacts. He points out that human cognitions are not just a framework confined to individuals; they are distributed across individuals, tools, and artifacts in an environment (Hutchins 1990, 1995). Discovery-oriented educational software can enormously benefit from this distributed framework. As educators and technology builders, we chose to depart from a viewpoint in which a classroom is simply a container of isolated technology pieces used by individuals and moved toward a socio-technical cooperative environment that actively facilitates a social learning process. The intent of the lessons that we developed was to scaffold students into discussions of physiology. Minimal whole-class, direct instruction was needed. The teacher was intended to function as a guide, leading students through activities and providing support as necessary. In this way, the activities can be challenging all students to think in new ways about the connections between physiological systems.

The resulting curriculum does not strive to replace traditional instruction, but complements it. It gives students the opportunity to explore the relationships between physiological systems in an integrative matter that extends the reach of traditional approaches. Lessons scaffold students into discussions of physiology and specialized subject areas. The teacher functions as a guide and leads students through activities and provides support as necessary. In this way, activities can be differentiated, in the sense that they challenge all students to think in new ways about the connections between systems.

Related work: Physiology simulations and multi-user simulation architectures

In the process of creating the Collective Simulations infrastructure and specifically Mr. Vetro, we examined physiology software and participatory simulation systems.

The JavaMan Project[6] includes a Web-based computer model of human physiology. It lets users explore the respiratory and cardiovascular systems of the human body. The model includes a very large set of variables, which can make it unmanageable. An abstract, mostly textual interface, does not include a visualization of a human being, and does not allow the simulation to be distributed.

SimBioSys[7] uses simulated Emergency Room situations as training context. It exposes students to real-time situations based on patient case studies. Without the proper treatment, a patient is likely to die. SimBioSys does not deal very directly with physiological variables; it is more concerned with the process of quickly determining which sensors to use and what kind of medication to administer.

The Logal Science Explorer Series by Riverdeep[8] is much more K-12 oriented. It provides a set of activities in a science curriculum that covers key concepts in biology, chemistry, and physics. Through what Riverdeep calls "computer simulation activities," students can discover and understand scientific concepts. We have specifically explored the Riverdeep units on the cardiovascular system and the respiratory system. Logal activities provide good content on individual systems. Students interact with simulations through worksheets that encourage them to predict and estimate feasible value ranges of physiological variables. However, these explorations are focused on variables of the same system, for example, the lung volume over time, and do not support interactions *between* different systems.

Multiple projects at Stanford University Medical Media and Information Technologies (SUMMIT)[9] include interactive technology for medical education. VirtualLabs and simPHYSIO (Huang 2004) are Web-delivered interactive visual and simulation-based teaching modules for physiology. These modules are designed to be engaging and to promote active learning. The learners can see multiple-linked representations, make

[6] http://www.health.adelaide.edu.au/paed-anaes/javaman/
[7] http://www.critcon.com/ccipublic/products/clinics.php3
[8] http://www.riverdeep.net/
[9] http://virtuallabs.stanford.edu/

predictions, learn problem-solving skills with a “hands-on” approach, apply and test scientific methodologies, and see immediate outcomes from their manipulations. Most modules are also focused on a single system instead of the interactions between systems. Perhaps the most closely related systems are the multiplayer “serious games” that SUMMIT has created, for instance, the Virtual 3D World for Emergency Medical Team Training, a distributed simulation where avatars controlled by a geographically dispersed networked team are supposed to help patients. This system used to train medical students is similar in spirit with our unique Collective Simulations framework and is an approach that has shown promise in medical education. In usability and learning achievement trials with senior medical students and residents, it demonstrated significant improvement in their learning.

Participatory simulations (Diehl 1990) involve people using or building simulations collaboratively. Desktop, handheld, and wearable computing devices can be connected in various ways to support participatory simulations. In HubNet, collaborating students explore math functions (Wilensky & Stroup 1999) by providing individual data points to a continuous mathematical function. A central server collects these data points and represents them as points on a plot visible to the entire class. In the Science Theater project (Cherry et al. 1999), groups of elementary school children learn about ecosystems by building simulations that include designing and exchanging animal agents over the Internet through the AgentSheets Behavior Exchange (Repenning & Ambach 1997; Repenning et al. 1999; Repenning et al. 1998). In Savannah (Benford et al. 2005), collaborative ecosystem simulations go a step further, incorporating location-aware technology that enables students to role play animal behavior and explore location-based consequences. Using probeware and handhelds has been explored as a viable science teaching method by others as well, for instance, Metcalf and Tinker (2004). In the MIT Thinking Tag Virus simulation game, each player wears a small computational tag that can communicate with nearby tags through infrared links (Colella 2000). This hardware is used to create a role-playing game in which a group explores the mechanism of virus transfer. Similarly, the Geney project (Danesh et al. 2001; Mandryk et al. 2001) uses Palm Pilot handheld devices to simulate the breeding of a species across a distributed population. The Cooties system simulates the spread of viral disease through Palm Pilots (Morehead 2001). EDC (Arias et al. 1999) and CarettaKids (Deguchi et al. 2006) developed collaborative simulations for city planning simulations. CarettaKids also integrates personal learning by allowing students to develop individual plans before projecting them to the group. Multi-User SimCity is a networked version of the popular Maxis SimCity[10] simulation game that lets a group of environmental design students play city planning games.

This list is by no means comprehensive. Earlier multi-user simulation systems had to make difficult trade-offs between pedagogy and technology. Multi-user simulations in education were either powerful or running on handhelds. Collective Simulations uniquely join regular desktop and powerful handheld computers into sophisticated, yet economically feasible classroom simulations. The resulting combination of social learning pedagogies with distributed simulation technology is not just a means of more engaging and efficient learning; it enables new approaches of learning not previously possible without information technology.

Assessment of collective simulations

Q1

An evaluation team from the University of Colorado, School of Education, designed and carried out an extensive evaluation study of the Mr. Vetro technology and curriculum in

[10] http://en.wikipedia.org/wiki/SimCity

local high schools. The team consisted of assessment experts with experience in conducting studies of classroom practice and professional development of teachers in K-12 settings (Webb et al. 2005). The evaluation team addressed both formative and summative evaluation needs of the design team in terms of teacher implementation, classroom practice, student motivation, and learning outcomes. They were an integral part of shaping the design of educational activities built in the Mr. Vetro system and the technology itself. They worked closely with the design team and the lead teacher to continually align the evaluation with educational material and the technology.

The evaluation team designed the evaluation study using established assessment material and processes. These included the *Reformed Teacher Observation Protocol* (RTOP; Adamson et al. 2003; Piburn et al. 2000) for observing classroom implementations, the *Colorado Learning Attitudes about Science Survey*[11] (CLASS; Adams et al. 2004; Perkins et al. 2006), material from *standardized tests* (SAT, ACT, PISA, and AP Biology tests), and a framework for assessing student understanding adapted from Shafer and Foster (1997)

Based on feedback from pilot implementations and teacher training sessions, we made the necessary alterations of both the educational activities and the assessment material to make them usable in a real classroom environment. When the instruments were ready for classroom use, the evaluation team carried out the evaluation study, which included classroom observation during implementation.

We conducted the experiment designed by the evaluation team in three different high schools in a local school district. Five science teachers participated: two teaching only using Mr. Vetro (treatment); two teaching only in their conventional way (comparison), and one teaching both using Mr. Vetro and in the conventional way. A total of fifteen science classes (Biology, AP Biology, Anatomy and Physiology, and IB Biology) participated. About 400 high school students, ages 14–18, were part of those classes.

Mr. Vetro teachers were trained to use the technology and curriculum by acting as students and working through the lessons. Teachers were encouraged to allow the natural process of student problem solving and sense making, rather than take a lecture-oriented approach to disseminating information and concepts. The three comparison teachers were given the content objectives of each lesson in the unit and were encouraged to teach these topics in the way that they felt most comfortable. They were also given the pre–and post–assessments but were encouraged not to “teach to the test.” They could develop their own lesson plans for the cardiopulmonary system or utilize suggested activities that paralleled the Mr. Vetro simulation. Both groups implemented the unit over 5–7 days, in multiple science classes during the same semester.

The evaluation study results are presented below grouped in four categories:

1) *teaching outcomes* that focus on the impact of the Mr. Vetro Collective Simulation on teaching practices;
2) *student learning outcomes* that focus on the impact of the simulation on knowledge acquisition and performance on tests;
3) *student motivation outcomes* that focus on the effects of the simulation on student attitudes and beliefs toward science;
4) *real-life consequences* that focus on the impact of the Mr. Vetro experience on student lives.

[11] Available at http://www.colorado.edu/sei/class/

Teaching outcomes Q1

In this section, we present the impact of the Mr. Vetro Collective Simulation and associated curriculum on teaching practices. These results demonstrate how the use of the simulation increased the level of student inquiry in the classroom by motivating the use of higher-order thinking skills, student-to-student discourse, hypothesis formation, and testing.

One goal of the Mr. Vetro study was to determine if the use of Collective Simulations and associated lesson plans increased the level of teacher performance, specifically indicated by decreased time of direct instruction, use of sheltering strategies, and increased performance on the RTOP. We hypothesized that the Mr. Vetro simulation and instructional materials would provide teachers opportunities to support investigations that incorporated elements of inquiry-based science teaching, more so than what might be observed in similar lessons not supported by such an interactive simulation. The lesson design incorporated the use of hypothetical scenarios, prediction, and testing of results to support student discussions of cardiopulmonary physiology. By design, the instructional materials and activities assumed the teacher would function as a guide, leading students through activities and providing support as necessary, but not explicitly solving the problems or demonstrating related solution strategies to address questions posed in the scenarios.

The *Reformed Teacher Observation Protocol* (RTOP; Adamson et al. 2003; Piburn et al. 2000) was used for classroom observation in both the Mr. Vetro and the non-Mr.-Vetro comparison classrooms. The RTOP is a tool designed to document characteristics of classroom practice that define reformed teaching. It is designed to measure five related constructs of inquiry-oriented teaching (Lesson Design and Implementation, Propositional Knowledge, Procedural Knowledge, Communicative Interactions, Student-Teacher Relationships).

As noted in its Reference Manual (Piburn et al. 2000), the RTOP "provides an operational definition of what is meant by 'reformed teaching.'" The items arise from a rich research-based literature that describes inquiry-oriented standards-based teaching practices in mathematics and science" (p. 41). RTOP scores "predict improved student learning in mathematics and science classrooms at all levels. Analysis of the RTOP suggests that it is largely a unifactorial instrument that taps a single construct of inquiry" (p. 24). The RTOP includes five items for each category that are rated on a scale from 0 to 4.

Mr. Vetro (n=39 observations) and comparison (n=22 observations) classes were observed using the RTOP. In addition, parts of class sessions were videotaped and notes were taken by observers.

Two members of the evaluation team conducted the observations for participating teachers. Several weeks prior to conducting observations, the research team watched video clips of math and science classrooms included in the RTOP training materials and rated the observed practices using the protocol. Differences between the scores of the research team and the suggested ratings in the manual were deliberated. We continued to observe videos and compare scores until the research team reached a sufficient threshold of inter-rater agreement (overall score within 5 points; no more than a 1 point difference on any item). During preliminary trials of the Mr. Vetro materials, we also conducted paired observations in which the two researchers would observe a real-time lesson and compare RTOP scores. Each teacher was evaluated in every class using the RTOP. Individual teacher scores were compiled. The average rating in each category for all comparison teachers and all Mr. Vetro teachers was compared.

Moreover, a *teacher interview protocol* was used to provide qualitative, anecdotal evidence of teachers' opinions regarding the degree to which the simulation promotes scientific inquiry, on the role of the Mr. Vetro simulation as an instructional tool, and teachers' perceptions of the impact of Mr. Vetro on student motivation and learning. The semi-structured teacher interview was used after the Mr. Vetro simulation. Teacher interviews were audiotaped and transcribed. Four key constructs were explored in the design of the interview protocol and analysis:

1) teachers' conceptions of scientific inquiry;
2) teachers' conceptions of student learning;
3) teachers' observation of relative student engagement; and
4) any lingering effects that students experience with the Mr. Vetro simulation.

Teachers' perceptions of the impact of Mr. Vetro on student motivation and learning offer an important perspective on the relative difference in student motivation or engagement in response to the experience, compared to regular science instruction. Instructional activities that are dependent on technology can present new and unexpected challenges to teachers, even those in science who include lab experiments and practicals as part of instruction. Also, group interaction that is required to complete the scenarios can present a different set of challenges to teachers who have not established norms for student-to-student interaction.

The comparison of means for Mr. Vetro and comparison groups are displayed in Fig. 3. Means for all five RTOP categories were greater for Mr. Vetro classes than means for the comparison classes by 4 to 5 points on a 20-point scale. All differences were statistically significant at the $p<0.0001$ level. The total of mean values for each category shows an overall difference of over 20 points on a scale of 100, suggesting that the Mr. Vetro simulation fostered far greater opportunities for student inquiry.

The data, reported in more detail in (Luhn 2010), shows that RTOP performance was increased in all categories in the Mr. Vetro classes compared to the comparison classes. In general, teachers moved around the classroom, were engaged one-on-one with students,

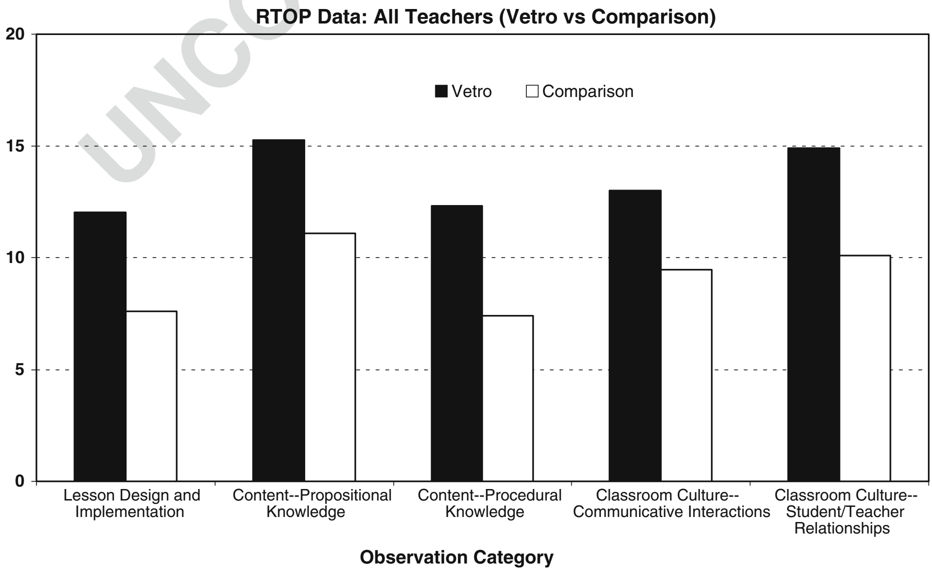

Fig. 3 Graph of RTOP Data comparing Mr. Vetro treatment with comparison groups

answered group questions as needed, and spent limited time in didactic information transfer. Mr. Vetro allowed for true inquiry. Students were actively learning how the simulation worked, making predictions, testing their predictions, and incorporating new understandings into their mental model of how the human body works. The Mr. Vetro activities required students to synthesize information and draw connections between old and new information. Finally, the activities not only encouraged the use of language, but they were essential to make the system operate effectively. Students read the messages and decoded the numerical values while communicating with the group the changes they were making to the system. Mr. Vetro classrooms exemplified the synergism between science and language.

Based on teacher interviews, across the three teachers who used Mr. Vetro in their classes, there was strong agreement that Mr. Vetro *promoted scientific inquiry* among their students. As one teacher in her first year of full-time teaching described:

> Most labs you do are pretty minds off. You know, follow these instructions. What do you see? Does that support your hypothesis? So scientific inquiry to me means that students are creating their own experiments…. [With Mr. Vetro], students were really given an opportunity to think and test their hypotheses…. The way lessons were set up with predicting, and then putting your predictions in and seeing if you get the results that you are trying to simulate. That piece is the inquiry piece.

A veteran science teacher of 26 years agreed, but offered a slightly different perspective:

> It's really kind of a process of coming up with questions. And questions are based on anything that can happen in your environment that makes you curious or seems interesting. And then coming up with a strategy that is a plausible approach, that is something you can calculate or measure. You have an idea of how things might be, and that leads to some predictions. And then you test those predictions out. We saw that with Mr. Vetro. You'd say, what happens if I do this.

It is important to note that although teachers recognize the importance of promoting scientific inquiry, they also articulated some of the inherent challenges in providing those experiences to students on an ongoing basis. Mr. Vetro offered a unique learning experience that readily drew students into sustained problem solving and working with the simulation to investigate the proposed scenarios.

With respect to student learning, all three teachers agreed that *students learned more about the cardiopulmonary system as a system through Mr. Vetro than they have in years past*. As one teacher described:

> My hope is that they wouldn't learn that the systems work in isolation, which is frequently the way we teach it. And there is nothing right or wrong with that, because it simplifies things for students. To say we are just dealing with the cardiovascular system now. They do appreciate it when you bracket things off, and isolate them. It makes their initial learning easier. But it really misses a very important aspect of physiology, which is interrelationships. We could have made this really weird and pulled in the nervous system, and the endocrine system. And we did a little bit, because something has to change those parameters. But it did give them a chance to explore two systems juxtaposed in real time. So as one thing changes the other thing has to.

Typically, the systems of the human body are taught in isolation to familiarize students with specific functions, terminology, and in some cases, relationships to other systems. The synergistic aspects of the Mr. Vetro simulation combined with the collaboration of students who modeled the interaction and interdependency of these two systems reinforced the way in which blood gases are transported by these two systems through various contexts.

One teacher who used both the Mr. Vetro unit and a comparison unit, described that some features of the cardiopulmonary system were very difficult to teach without the use of the simulation.

> The hardest thing for me to teach [the comparison class], without the Mr. Vetro, was tidal volume and stroke volume…. You can do it using direct instruction… but I think that they get general relationship between exercise and the strength of their heart. The Vetro group really seemed to get it all. They really absorbed most of it. I think they, as an overarching whole, really got tidal volume and stroke volume and how that relates to the cardiopulmonary systems.

Lastly, all of the teachers agreed that the Mr. Vetro simulation *evoked a level of sustained engagement that they do not typically observe among their high school science students*. The teachers affirmed that several important instructional principles such as use of relevant contexts, student-to-student collaboration, and use of more challenging activities contributed to a level of student engagement that, in one teachers' opinion, changed his students' overt expressions of passive engagement.

> It had a game-like quality to it that I think sucked them in. They knew when things weren't going right, and so they had to intervene. There is no question that there was a higher level of engagement than I usually get with some kind of a simulation…. Students were into it. For multiple days they were working for longer stretches of time without needing bathroom breaks, and getting up to walk around and throw trash away. They were definitely more focused then they usually would be.

Overall, participating teachers' experiences with Mr. Vetro were positive. In the interviews, they were able to offer several examples of how the experience enhanced student engagement, increased opportunities for scientific inquiry, and contributed to students' deeper understanding of the functions of the cardiovascular and pulmonary systems and their interrelationship.

The Mr. Vetro simulation and associated curriculum demonstrated an impact on teacher practice that supported greater student scientific inquiry. They authentically integrated sheltering strategies into the mainstream classroom. The activities allowed for higher-order thinking skills to be used. Although the research states that higher-order thinking skills are important in high school, in the majority of high school classrooms, more than 80% of classroom activities are at the acquisition level of thought (Echevarria et al. 2004). Higher-order thinking skills are difficult to access through traditional methodologies such as lectures, textbook work, and basic lab experiences; however, they are the essence of rigorous curriculum. Contrasting portraits of teachers in simulation and non-simulation classrooms for similar content suggests that such technology can be more effective in supporting student scientific inquiry, assuming the design principles and affordances of the related curricula support such an approach.

Learning outcomes

Q1

Our objective for evaluating student learning was to design a series of assessments to compare the learning gains for high school students studying the cardiopulmonary system. Assessment design incorporated competencies similar to those used for the Programme for International Student Assessment (PISA) by the Organization for Economic Co-Operation and Development (OECD 2003). Assessments were designed to elicit students' knowledge and their ability to apply that knowledge using three types of tasks: Level I questions were geared toward recalling basic facts; Level II questions involved making connections; and Level III questions required application of knowledge in new situations.

We predicted that students in the Mr. Vetro group would show higher achievement gains than those in the comparison group especially at the higher levels of thinking, because the simulation demonstrates the relationship between the cardiovascular system and pulmonary system that is difficult to demonstrate otherwise and promotes collaboration and meaningful discussions between students as they control Mr. Vetro's organ parameters.

Method and data sources To explore the relative impact on student performance, data from a Mr. Vetro treatment group used the computer-based simulation with related activities. For the comparison group, a parallel tech-free curriculum was used that did not involve use of the software. Both groups had between 5–7 days of instruction on the cardiopulmonary system.

Students completed written assessments just prior to and just after the unit. Level I questions were selected response questions. Students chose the typical values for "at rest" heart and lung parameters and oxygen saturation. Level I questions also consisted of multiple-choice questions that were collected from various conventional assessments, such as the AP Biology Exam, NY State Regents Exam, PISA, Trends in International Mathematics and Science Study (TIMSS), National Assessment of Educational Progress (NAEP), and the Scholastic Achievement Test (SAT). These questions were chosen based on best-fit alignment with the content in the unit.

Level II questions required definitions of vocabulary words used throughout the unit, and were open response. Level III questions, which were only administered in the post-assessment, were open response questions requiring transferring knowledge learned in the classroom to a novel scenario. For example, students were asked to predict and explain the effects of having asthma or coronary artery disease on the cardiopulmonary system.

Level I questions were marked as correct (1 point) or incorrect/no response (0 points). Level II and III questions were graded on a 0–4 point rubric developed by the design team. An inter-rater reliability process was conducted to improve the clarity of the rubric among five raters. A sample of ten papers was independently scored using the rubric and these scores were compared for consistency. Differences in scoring student responses were discussed; points of clarification and exemplars of student performance for various levels were included in subsequent iterations of the rubric. This process was conducted at least twice until additional papers were scored with consistent inter-rater scoring exceeding an acceptable threshold of 85%.

Results Independent two-tailed t-tests and effect sizes were calculated. Only students who took both the pre-assessment and the post-assessment (and for which we had permission to use the data) were included in the analysis; that is, out of the over 400 participating students, 264 were included in this analysis. The treatment group involved three teachers and 169 students in nine classrooms. The comparison group involved three teachers and 95 students in five classes.

Four smaller groups of Mr. Vetro versus comparison classes were selected based on similar class types and the differences between the scores from all types of the pre-assessments were found to not be statistically significant:

Group 1: multiple choice, $t(15)=2.13$, $p=0.75$; typical values, $t(24)=2.06$, $p=0.24$; vocabulary, $t(24)=2.06$, $p=0.86$.

Group 2: multiple choice, t(94)=1.99, p=0.44; typical values, t(99)=1.98, p=0.32; vocabulary, t(97)=1.98, p=0.95.
Group 3: multiple choice, t(57)=2.00, p=0.64; typical values, t(55)=2.00, p=0.41; vocabulary, t(58)=2.00, p=0.31.
Group 4: multiple choice, t(88)=1.99, p=0.48; typical values, t(71)=1.99, p=0.78; vocabulary, t(68)=2.00, p=0.098).

Some classes were placed in more than one group and others were only used in the aggregate study. A description of each group is in Table 1.

Table 2 shows that for each of these four groups, the learning gains for multiple-choice questions were not notable (p<0.05, effect sizes very small to moderate). However, learning gains for typical values and vocabulary were statistically significant in favor of the Mr. Vetro group, and the effect sizes were large to very large.

The comparison of pre–and post-unit responses to the assessments demonstrate that the tech-based Collective Simulation approach, for similar content, results in stronger learning gains for all groups regardless of the level of the course (e.g., regular or honors). Statistically greater gains especially occurred in the results of Level I recall of typical values of the heart and lung parameters and in Level II vocabulary definitions. Differences in the gains in Level I recall of multiple-choice questions from conventional assessments were more moderate, or in some matched groups, not significant. It is also notable that responses to post-assessment Level III questions, which require critical thinking and application, were significantly more complex in the Mr. Vetro groups than those in the matched Comparison groups (Keyser 2010).

Table 1 Groups of students

Group #	Description	Group type	Teacher	School	# of Classes	# of students	Type of Class
1	Same teacher at same school teaching same course	V	A	A	One	17	Regular Biology
		C	A	A	One	12	Regular Biology
2	Two different teachers from different schools teaching same course	V	A	A	One	44	Regular Biology
			B	B	One		
		C	A	A	One	57	Regular Biology
			C	C	Two		
3	Two different teachers from different schools teaching same course	V	D	C	Two	29	Anatomy and Physiology
		C	E	B	Two	38	Anatomy and Physiology
4	Two different teachers from same school teaching advanced students (mixed class type)	V	B	B	Three	85	2 AP Bio, 1 Adv Bio
		C	E	B	Two	38	Anatomy and Physiology

(*V* Mr. Vetro curriculum, *C* Comparison classes)

Table 2 Learning gains, Mr. Vetro vs. comparison groups

Assessment type	Group #	Mean, Mr. Vetro	Mean, comparison	p-value	Effect size
Multiple choice	1	1.82	2.08	$t(25)=2.06$, $p=-0.702$ No statistical diff	-0.15 Very small, Neg
	2	1.70	0.88	$t(93)=1.99$, $p=0.037$	0.58 Moderate
	3	1.38	0.97	$t(43)=2.02$, $p=0.369$ No statistical diff	0.24 Small
	4	1.72	0.97	$t(97)=1.98$, $p=0.010$	0.44 Small
Typical values	1	2.71	0.83	$t(26)=2.06$, $p=0.0031$	1.03 Very large
	2	2.68	0.79	$t(89)=1.99$, $p<0.0000001$	1.28 Very large
	3	1.83	0.89	$t(43)=2.02$, $p=0.014$	0.65 Large
	4	2.44	0.89	$t(92)=1.99$, $p<0.0000001$	1.06 Very large
Vocabulary	1	15.24	11.25	$t(21)=2.08$, $p=0.036$	0.81 Large
	2	15.82	7.14	$t(99)=1.98$, $p<0.0000001$	1.15 Very large
	3	14.62	6.29	$t(52)=2.01$, $p<0.0000001$	1.44 Very large
	4	16.55	6.29	$t(89)=1.99$, $p<0.0000001$	1.60 Very large

Student motivation outcomes Q1

As part of this study, we also documented the impact of the Mr. Vetro Collective Simulation on student attitudes and beliefs towards science. We used the Colorado Learning Attitudes about Science Survey (Adams et al. 2004; Perkins et al. 2006) an adaptation of a survey designed for monitoring undergraduate attitudes toward science. Results indicated that the Mr. Vetro simulation has neither a positive or negative effect on student attitudes toward science whereas the comparison group shows mixed results (Webb 2010). It is worth mentioning that based on historical use of the survey a typical result for the administration of the CLASS is a decline between pre–and post-unit administration. So, results that are apparently "flat" are viewed as relatively successful. A further analysis of the data compared honors versus regular biology students. The fact that attitudes were influenced negatively in the honors group (Webb 2010) was viewed as unexpected by the evaluation team given the strong learning gains demonstrated by the honors students on the achievement measures.

To explore, in part, some of the differences in student attitudes in response to the Mr. Vetro experience, students from all classes were asked if they would volunteer to participate in student focus groups. During the focus group, students also completed a two-part survey. The first part was responding to seven written questions. All questions were free-response. Below are some highlights from the written portion of the survey.

I. Have you ever done work similar to what a scientist might do in this class? What about other science classes?

Most students in all types of classes responded that they have done labs and dissections, which they deemed "similar," presumably because it was not a lecture. A few notable comments are below:

- Many scientific simulations that we have visited on the internet are similar to Mr. Vetro, just not as deep in terms of interactivity or group use. *~AP Biology Student*

- I have done labs and hands-on dissections, but I have never done anything as intense as this. *~AP Biology Student*
- The Mr. Vetro activity was like nothing I have ever done before. It seemed like it should belong in classes that taught students who wanted to be future doctors or nurses. *~Anatomy and Physiology Student*

These responses indicate that Mr. Vetro was viewed by this sample of students as a unique experience. Moreover, several students suggested that the material was more challenging and substantive than other simulations, and perhaps should be used with students interested in medical careers.

II. How would you compare the Mr. Vetro activity with other activities you've had in this or other science classes?

There were a variety of responses to this question, although most students favored Mr. Vetro over other science units. In contrast to the aggregate results from AP/IB students on the student attitude survey, the responses from these honors students suggest the experience was seen as important, intriguing, more interactive, and motivating. In fact, for one student "It was the coolest thing [he] ever did in a science class." More mixed or lukewarm responses suggest that reasons for the decline in student attitudes may have involved a mismatch between students' expectations (e.g., preference for lectures, interest in more challenging problems) and the goals of the Mr. Vetro simulation and activities. Considering that this study used the same simulation and instructional activities for all participating classes, these responses indicate that future use of Mr. Vetro in advanced courses should include a different, more challenging set of activities than those used with regular biology students. That is, there should be some consideration of differentiating the curriculum for advanced courses. In contrast, students from regular biology courses indicated that the goals, expectations, and activities included in the Mr. Vetro unit were both accessible and challenged regular biology students. There was no evidence that regular biology students preferred a different approach to learning similar material.

- The Mr. Vetro activity was different because it felt like it was almost reality, and that is the reason why it got my attention. *~AP Biology Student*
- I loved it. I had more fun with Mr. Vetro than with the usual lectures that I have. I also felt more engaged because it was a hands-on experience. *~Advanced Biology Student*
- It was great that we got to figure it out by ourselves, but at some point it was a bit confusing. Most biology classes I've taken, we have just done notes and pictures in this section instead of being able to apply it to a "living human". *~Advanced Biology Student*

III. Compared to other activities you've had in this or other science classes, do you think you learned more, less, or the same from these activities? Explain.

For this question, there were a higher proportion of mixed responses from students in honors classes, as exemplified by the last comment in this list:

- I actually did learn from the activity. I really didn't know the basic vocabulary of the human body and I learned how the human body works much more that I did before. It was more realistic and that's what made it fun. I learn better if I have fun. *~ AP Biology Student*
- Probably a lot more because there was a lot of guidance with Mr. Vetro. I was able to be more interactive in small groups. *~ AP Biology Student*

- I think I learned less from these activities than I normally would because there wasn't much direct learning involved. I think I participate better when we are given a lecture. ~ *AP Biology Student*
- I don't feel like I learned as much as I could. It was just a scramble to keep him alive. We never discussed anything in class so I never knew if I was right. We needed more discussions. I felt like I had no solid foundation. ~ *Advanced Biology Student*

In general, the responses from AP/Honors students who participated in the focus groups suggest that the Mr. Vetro simulation did have a positive impact on the majority of students' attitudes toward biology and the simulation experience. In many cases, these students indicated a preference for the Mr. Vetro simulation when identifying authentic science activity or asked to compare Mr. Vetro to regular classroom activities (e.g., lecture, taking notes). We did notice, however, that there were a higher proportion of mixed or negative responses from AP/Honors students in reaction to the prompt regarding whether they learned more, less, or about the same with Mr. Vetro. For some honors students, there remains a strong preference for more structured and direct instruction. For others, some additional whole class discussion for additional feedback and clarification was needed.

Real-life consequences Q1

Also interesting is anecdotal evidence about knowledge transfer and real-life consequences. We have reports from multiple teachers that outstanding transfer of concepts learned in the Mr. Vetro unit occurred for their students. For instance, after completing the Mr. Vetro unit, in a subsequent unit the same students were studying disease spreading and they analyzed those scenarios using terminology and knowledge of the cardiopulmonary system gained from Mr. Vetro.

> [Students have a better understanding of how the circulatory and respiratory systems work together] because the other day we did an outbreak scenario and the disease was malaria, but the patients had some symptoms and they didn't use the word "hypoxia" [a term used in Mr.Vetro] but the kids did. And they talked about what the heart rate was in the scenario and they were talking about it in terms of Mr. Vetro and what they had learned, which I think was really cool! Somebody learned something and they were applying it to a whole different scenario!~*High School Science Teacher*

Anecdotal evidence from observations of the design and evaluation team indicates that students are very engaged and learn about the intricate connections between the cardiovascular and the respiratory systems in ways that were not observed in comparison classes. Teachers report that students who typically are not engaged in class were very much engaged and actively learning with Mr. Vetro.

Some of the most exciting results are ones that impact students' personal lives. The overwhelming majority of students replied that they learned about their own body from Mr. Vetro. For instance, students report that they now understand their own bodies and physical conditions better. We also have several reports of student athletes talking to their coaches about ways to increase their performance based on what they had experienced with Mr. Vetro exercise scenarios. Some students are trying to quit smoking, while other students indicate that they are more motivated to refrain from smoking or taking other substances because of their experience with Mr. Vetro.

- I have a bunch of lung issues and I never really got what my doctor was telling me, and now, all of a sudden I get it!~*High School Student*

- I am definitely more motivated to avoid those sorts of things; I'm more motivated to avoid substance use and anything of that nature now. ~ *High School Student*
- I did learn a lot about the difference between an average person, an unhealthy person and an elite athlete. It made me want to be more healthy and active. ~ *High School Student*
- I now understand more what's happening in my heart and lungs when I run and compete in sports. ~ *High School Student*
- I have one student who wants to quit smoking. She was asking how to go about doing it, because that [the Mr. Vetro experience] encouraged her. Her mom emailed me and was so excited. Nothing else ever touched her daughter to stop smoking. This is what did it!~*High School Science Teacher*

Conclusions and future work

Q1

The Collective Simulations framework with the Mr. Vetro prototype and curriculum was successful along many dimensions:

- *Technically* it features a flexible, modular architecture that enables us to create and easily customize educational material for different inquiry-based science activities and for different K-12 levels.
- *Educationally* both learners and teachers benefit from Collective Simulations. Teacher practice results provide compelling evidence that the instructional practices and learning experiences provided in the Mr. Vetro class were more conducive to promoting scientific inquiry and student learning of concepts. Learning outcome results suggest that students completing Mr. Vetro do just as well as the comparison group in multiple-choice items and outperform students on items that are more likely to measure their knowledge of typical physiological phenomena and responses of these related systems in realistic scenarios.
- *Motivationally* data shows that less advanced students get more engaged and interested in science than advanced placement students, regardless of the strong learning gains. Academically, it would be interesting to pursue some research focused on exploring the reasons behind this finding.
- *Pragmatically* some of the most exciting results are ones that impact students at a personal level. For us, the fact that there are indications showing that the experiences with Mr. Vetro promote self-awareness and healthier life styles is non-trivial, but instead, some of the most revealing and exciting results of this project.

These positive results are extremely encouraging; but a lot more work is needed to create a technically and educationally viable solution. Technically, we need to ensure that the technology can be effectively and affordably adopted by schools with diverse technological infrastructures; therefore, we need to make the technology work seamlessly on a variety of platforms in order to accommodate the majority of school computer setups. These setups include school computer labs equipped with desktop machines (Macs & PCs), dedicated or shared laptop pools for use in science classes—including low-end laptops (e.g., Intel NetBooks), and handheld devices and multi-user control mechanisms (e.g., clickers and MultiPoint technology). Educationally, we need to provide effective curriculum to span the entire spectrum of K-12 science to ensure systemic adoption by schools. We have created and tested a complete unit but we also need further research in pedagogical issues to develop solutions that provide a complete and viable educational solution for all students. Some of the more interesting mixed results from our evaluation study reported above (i.e., the pattern of mostly positive responses from regular biology students and a greater prevalence of mixed responses from the advanced/AP groups) will also motivate future R&D. We need to engage in additional forward-looking, technical and curricular research in how to more effectively use Mr. Vetro in advanced classes.

We also need to create more content by incorporating more human systems and creating ready-to-use learning activities that cover multiple units of science at all educational levels. Our AWE modular architecture supports the easy addition of new organs or human systems, as well as the input and output parameters and equations linking them. We will utilize this flexibility to expand Collective Simulations activities to cover more science curriculum. The issues of expanding to other human systems no longer involve just technical issues about what kind of abstractions and visualizations to use. They are more conceptual and pedagogical: How do you present the complexity of the human body without overwhelming K-12 students with enormous numbers of parameters and relationships, but instead enabling them to understand sophisticated interactions?

Finally, technology, content, and curriculum will need to be complemented with professional development/teacher training programs to ensure that the commitment and ability of school districts to adopt Collective Simulations are high.

Acknowledgments This work was supported by the National Institutes of Health, National Center for Research Resources (grant numbers 1R43 RR022008-01 and 1R43 RR022008-02) and previously by the National Science Foundation (grant number DMI SBIR 0232669). Opinions expressed are those of the authors and not necessarily those of the National Institutes of Health or the National Science Foundation.

References

Adams, W.K., Perkins, K.K., Dubson, M., Finkelstein, N.D., & Wieman, C.E. (2004). *The design and validation of the colorado learning attitudes about science survey.* Paper Presented at the 2004 Physics Education Research Conference, Boulder, CO.

Adamson, A. E., Banks, D., Burtch, M., Cox, F., III, Judson, E., Turley, J. B., et al. (2003). Reformed undergraduate instruction and its subsequent impact on secondary school teaching practice and student achievement. *Journal of Research in Science Teaching, 40*(10), 939–958.

Arias, E., Eden, H., Fischer, G., Gorman, A., & Scharff, E. (1999). *Beyond access: informed participation and empowerment.* Paper presented at the Conference on Computer Supported Collaborative Learning (CSCL ’99), Stanford, CA, USA.

Benford, S., Rowland, D., Flintham, M., Drozd, A., Hull, R., Reid, J., et al. (2005). *Life on the edge: supporting collaboration in location-based experiences.* Paper presented at the Conference on Human Factors in Computing Systems (CHI) 2005, Portland, OR.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2*(2), 141–178.

Bruner, J. (1990). *Acts of meaning*. Cambridge: Harvard University.

Cherry, G., Ioannidou, A., Rader, C., Brand, C., & Repenning, A. (1999). *Simulations for lifelong learning.* Paper presented at the National Educational Computing Conference (NECC), Atlantic City, NJ.

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

Committee on Prospering in the Global Economy of the 21st Century: An Agenda for American Science and Technology Committee on Science, Engineering, and Public Policy, National Academy of Sciences, National Academy of Engineering, Institute of Medicine of the National Academies. (2005). *Rising above the gathering storm: Energizing and employing America for a brighter economic future.* Washington: The National Academies.

Cummings, P.B. Interactive physiology, from http://www.aw-bc.com/info/ip/

Danesh, A., Inkpen, K. M., Lau, F., Shu, K., & Booth, K. S. (2001). *Geney: Designing a collaborative activity for the palm handheld computer.* Paper presented at the Conference on Human Factors in Computing Systems (CHI), Seattle, USA.

Deguchi, A., Yamaguchi, E., Inagaki, S., Sugimoto, M., & Kusunoki, F. (2006). *CarettaKids: A system for supporting children's face-to-face collaborative learning by integrating personal and shared spaces.* Paper presented at the Interaction Design and Children (IDC) 2006, Tampere, Finland.

Diehl, E. (1990). Participatory simulation software for managers: The design philosophy behind microworlds creator. *European Journal of Operations Research, 59*(1), 203–209.

Echevarria, J., Vogt, M., & Short, D. (2004). *Making content comprehensible for English learners: The SIOP model*. Boston: Allyn and Bacon.

Epstein, J. M., & Axtell, R. (1996). *Growing artificial societies*. Cambridge: MIT.

Gibbons, P. (2002). *Scaffolding language, scaffolding learning teaching second language learners in the mainstream classroom*. Porthsmouth: Heinemann.

Hampton, E., & Rodriguez, R. (2001). Inquiry science in bilingual classrooms. *Bilingual Research Journal, 24*, 417–434.

Hill, J., & Flynn, K. (2006). *Classroom instruction that works with English language learners*. Alexandria: Association for Supervision and Curriculum Development.

Huang, C. (2004). Virtual labs: E-Learning for tomorrow. *PLoS Biol, 2*(6). Available at http://www.plosbiology.org/article/info:doi%2F10.1371%2Fjournal.pbio.0020157.

Hutchins, E. L. (1990). The technology of team navigation. In J. Galegher, R. E. Kraut, & C. Egido (Eds.), *Intellectual teamwork: Social and technological foundations of coorperative work* (pp. 191–220). Hillsdale: Erlbaum.

Hutchins, E. L. (1995). *Cognition in the wild*. Cambridge: MIT.

Keyser, D. (2010). *A comparative analysis of student learning with a collaborative computer simulation of the cardiopulmonary system.* Paper to be presented at the Annual Meeting of the American Educational Research Association (AERA), Denver, Colorado.

Luhn, L. (2010). *The influence of a cardiopulmonary computer simulation on teacher practice and opportunities for student inquiry.* Paper to be presented at the Annual Meeting of the American Educational Research Association (AERA), Denver, Colorado.

Mandryk, R.L., Inkpen, K.M., Bilezikjian, M., Klemmer, S.R., & James A.L. (2001). *Supporting children's collaboration across handheld computers.* Paper presented at the Extended Abstracts of CHI, Conference on Human Factors in Computing Systems, Seattle, USA.

Metcalf, S. J., & Tinker, R. F. (2004). Probeware and handhelds in elementary and middle school science. *Journal of Science Education and Technology, 13*(1), 43–49.

Michael, J. (2001). In pursuit of meaningful learning. *Advances in Physiology Education, 25*(3), 145–158.

Michotte, A. (1962). *The perception of causality.* Andover: Methuen.

Morehead, N. (2001). Handheld computers earn high marks, Federal Computer Week magazine, July 2, 2001. Available at http://fcw.com/articles/2001/07/02/handheld-computers-earn-high-marks.aspx

National Research Council. (1996). *National Science Education Standards*. Washington: National Academy.

National Science Foundation, Division of Science Resources Statistics (2008). Chapter 7. Science and technology: Public attitudes and understanding *science and engineering indicators 2008*.

National Science Foundation, Task Force on Cyberlearning (2008). *Fostering Learning in the Networked World: The Cyberlearning Opportunity and Challenge. A 21st Century Agenda for the National Science Foundation*: National Science Foundation. Available at http://www.nsf.gov/pubs/2008/nsf08204/nsf08204.pdf

Norcini, J. (2003). Clinical review: ABC of learning and teaching in medicine—work based assessment. *BMJ—British Medical Journal, 326*, 753–755.

Oppenheimer, T. (2003). *The flickering mind: The false promise of technology in the classroom and how learning can be saved*. Toronto: Random.

OECD: Organization for Economic Co-Operation and Development, Programme for International Student Assessment (2003). *PISA 2003 technical report*.

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

Papert, S. (1993). *The children's machine*. New York: Basic.

Perkins, K.K., Barbera, J., Adams, W.K., & Wieman, C.E. (2006). *Chemistry vs. physics: A comparison of how biology majors view each discipline.* Paper presented at the 2006 Physics Education Research Conference, Boulder, CO.

Piaget, J. (1972). *The psychology of the child*. New York: Basic.

Piaget, J. (1990). *The child's conception of the world*. New York: Littlefield Adams.

Piburn, M., Sawada, D., Falconer, K., Turley, J., Benford, R., & Bloom, I. (2000). *Reformed Teaching Observation Protocol (RTOP). The RTOP rubric form, training manual and reference manual containing statistical analyses (available from* http://PhysicsEd.BuffaloState.Edu/AZTEC/rtop/RTOP_full/PDF/ *).*

Prensky, M. (2006). *Don't bother me Mom–I'm learning!* St. Paul: Paragon.

Repenning, A., & Ambach, J. (1997). *The agentsheets behavior exchange: Supporting social behavior processing.* Paper presented at the CHI 97, Conference on Human Factors in Computing Systems, Extended Abstracts, Atlanta, Georgia.

Repenning, A., & Ioannidou, A. (2005). *Mr. Vetro: A collective simulation framework.* Paper presented at the ED-Media 2005, World Conference on Educational Multimedia, Hypermedia & Telecommunications, Montreal, Canada.

Repenning, A., & Ioannidou, A. (2006). *AgentCubes: Raising the ceiling of end-user development in education through incremental 3D.* Paper presented at the IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'06), Brighton, United Kingdom.

Repenning, A., & Ioannidou, A. (2007). *X-expressions in XMLisp: S-expressions and extensible markup language unite.* Paper presented at the ACM SIGPLAN International Lisp Conference (ILC 2007), Cambridge, England.

Repenning, A., & Ioannidou, A. (2008). *Agent warp engine: formula based shape warping for networked applications.* Paper presented at the 2008 International Conference on Advanced Visual Interfaces (AVI 2008), Napoli, Italy.

Repenning, A., Ioannidou, A., Rausch, M., & Phillips, J. (1998). *Using agents as a currency of exchange between end-users.* Paper presented at the WebNET 98 World Conference of the WWW, Internet, and Intranet, Orlando, FL.

Repenning, A., Ioannidou, A., & Phillips, J. (1999). *Collaborative use & design of interactive simulations.* Paper presented at the Computer Supported Collaborative Learning Conference at Stanford (CSCL'99).

Repenning, A., Ioannidou, A., Dättwyler, C., Luhn, L., & Repenning, N. (2010). Mr. Vetro: Assessing a Collective Simulation framework. *Journal of Interactive Learning Research (JILR), Association for the Advancement of Computing in Education (AACE), 21*(3).

Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. New York: Doubleday.

Shafer, M. C., & Foster, S. (1997). The changing face of assessment. *Principled Practice in Math and Science Education, 1*(2), 1–8. (retreived Nov 5, 2009 from http://ncisla.wceruw.org/publications/newsletters/fall97.pdf).

Short, D., & Echevarria, J. (2004). Teacher skills to support English language learners. *Educational Leadership, 62*(4), 8–13.

Simon, H. A. (1981). *The sciences of the artificial* (2nd ed.). Cambridge: MIT.

Stoddart, T., Pinal, A., Latzke, M., & Canaday, D. (2002). Integrating inquiry science and language development for english language learners. *Journal of Research in Science Teaching, 39*, 664–687.

Thompson, J., & Celva, E. (2005). *2005-2006 Colorado student assessment program (CSAP) demonstration packet: Science grade 10.* Colorado Department of Education, Unit of Student Assessment. Available at http://www.cde.state.co.us/cdeassess/documents/csap/released_items/sci_demo_g10_2005.pdf

U.S. Department of Education (2003). *The nation's report card: Science 2000*. (NCES 2003-453, by C. Y. O'Sullivan, M. A. Lauko, W. S. Grigg, J. Qian, and J. Zhang). Washington, D.C.: Institute of Education Sciences, National Center for Education Statistics.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University.

Webb, D. C. (2010). *Collective simulations as a context for student scientific inquiry: Effects on student attitudes and beliefs towards science.* Paper to be presented at the Annual Meeting of the American Educational Research Association (AERA), Denver, Colorado.

Webb, D. C., Romberg, T. A., Burrill, J., & Ford, M. J. (2005). Teacher collaboration: Focusing on problems of practice. In T. A. Romberg & T. P. Carpenter (Eds.), *Understanding math and science matters* (pp. 231–251). Mahwah: Lawrence Erlbaum Associates.

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

Wolberg, G. (1994). *Digital image warping*. IEEE Computer Society.

Wolberg, G. (1996). *Recent advances in image morphing*. Paper presented at the 1996 Conference on Computer Graphics International.

Yager, R. (Ed.). (1995). *Constructivism and learning science*. Mahway: Erlbaum.

AUTHOR QUERY

AUTHOR PLEASE ANSWER QUERY.

Q1. Please check capturing of section levels if appropriate.