To be published in Computer Modeling and Simulation in Science Education;
edited by Nancy Roberts, Wallace Feurzeig, and Beverly Hunter; Springer-Verlag; early 1999.

Multisensory Immersion as a Modeling Environment for Learning Complex Scientific Concepts

Chris Dede
Graduate School of Education
George Mason University
Fairfax, VA 22030
(703) 993-2019
cdede@gmu.edu

Marilyn C. Salzman
Human Factors and Applied Cognitive Psychology
George Mason University
Fairfax, VA 22030
(703) 352-8375
msalzman@gmu.edu

R. Bowen Loftin
Virtual Environment Technology Lab
University of Houston
Houston, TX 77023
(713) 743-1006
bowen@uh.edu

Debra Sprague
Graduate School of Education
George Mason University
Fairfax, VA 22030
(703) 993-2069
dspragu1@gmu.edu

"The power of technology to change one's intellectual viewpoint is one of its greatest contributions, not merely to knowledge, but to something even more important: understanding... it goes beyond the limits of human perception." Arthur C. Clark [1973].

In every aspect of our knowledge-based society, fluency in understanding complex information spaces is an increasingly crucial skill [Dede & Lewis, 1995]. In research and industry, many processes depend on people utilizing complicated representations of information [Rieber, 1994]. Increasingly, workers must navigate complex information spaces to locate needed data, must find patterns in information for problem solving, and must use sophisticated representations of information to communicate their ideas [Kohn, 1994; Studt, 1995]. Further, to make informed decisions about public policy issues such as global warming and environmental contamination, citizens must comprehend the strengths and limits of scientific models based on multivariate interactions. In many academic areas, students' success now depends upon their ability to envision and manipulate abstract multidimensional information spaces [Gordin & Pea, 1995]. Fields in which students struggle with mastering these types of representations include math, science, engineering, statistics, and finance.

Research on learning scientific concepts yields insights into why understanding complex information spaces is difficult. Many scientific domains deal with abstract and multidimensional phenomena that people have difficulty comprehending. Mastery of abstract scientific concepts requires that students build flexible and runnable mental models [Redish, 1993]. Frequently, these scientific models describe phenomena for which students have no real-life referents [Halloun & Hestenes, 1985a] and incorporate invisible factors and abstractions [Chi, Feltovich, & Glaser, 1991; White, 1993]. Students learning science need to be able to sift through complex information spaces, identifying what is important and what is not, as well as recognizing critical patterns and relationships. Learners may need to translate among reference frames, to describe the dynamics of a model over time in order to predict how changes in one factor influence other factors, and to reason qualitatively about physical processes [McDermott, 1991].

Developing effective pedagogical strategies and simulation technologies for teaching complex science concepts presents a substantial challenge for educational researchers and instructional designers. Despite the utilization of new teaching approaches, tools, and technologies, students struggle with abstractions in science. They not only enter their courses with gaps and inaccuracies in their conceptual understanding of the material, but also often leave with unaltered misconceptions [Halloun & Hestenes, 1985b; Reif & Larkin, 1991]. Students' lack of real-life referents for intangible phenomena, coupled with an inability to reify ("perceptualize") abstract models, is an important aspect of this problem. To aid in comprehending abstract information spaces, finding ways to utilize our biologically innate ability to make sense of physical space and perceptual phenomena seems a promising approach.

Using Models and Simulations to Convey Complex Scientific Concepts

Guided inquiry experiences using scientific models that reveal the shortcomings of learners' current conceptual frameworks can help wean students from erroneous beliefs. Before inculcating the formal representations that scientists use, these models can develop learners' abilities to intuitively understand how the natural world functions. Fostering in students the capability to qualitatively predict the behavior of phenomena under investigation is a valuable foundation for teaching them to manipulate quantitative formulas. Also, students are not empty vessels to be filled with theories; they have firmly held, often erroneous beliefs about how reality operates. Model based instruction can help learners evolve their existing mental models to more accurate conceptions of reality.

To date, uses of information technology to apply these pedagogical principles have centered on creating computational tools and two-dimensional virtual representations that students can manipulate to complement their memory and intelligence in constructing more accurate mental models. Perkins [1991] classifies types of "constructivist" paraphernalia instantiated via information technology: information banks, symbol pads, construction kits, phenomenaria, and task managers. Transitional objects (such as Logo's "turtle") are used to facilitate translating personal experience into abstract symbols [Papert, 1988; Fosnot, 1992]. Thus, technology-enhanced constructivist learning currently focuses on how representations and tools can be used to mediate interactions among learners and natural or social phenomena.

However, high-performance computing and communications capabilities are creating a new possibility in modeling scientific phenomena [Dede, 1995]. Like Alice walking through the looking glass, the virtual reality interface enables learners to immerse themselves in distributed, synthetic environments. They can become "avatars" (computer-graphics representations that serve as personas of human participants in the virtual world) who collaborate in inquiry-based learning-by-doing and use virtual artifacts to construct knowledge. The key features virtual reality adds to modeling as a means of constructivist learning are:

- *Immersion:* Learners develop the subjective impression that they are participating in a "world" comprehensive and realistic enough to induce the willing suspension of disbelief [Heeter, 1992; Witmer & Singer, 1994]. By engaging students in learning activities, immersion may make important concepts and relationships more salient and memorable, helping learners to build more accurate mental models. Also, inside a head-mounted display, the learner's attention is focused on the virtual environment without the distractions presented in many other types of educational environments.
- *Multiple three-dimensional representations and frames of reference:* Spatial metaphors can enhance the meaningfulness of data and provide qualitative insights [Erickson, 1993]. Enabling students to interact with spatial representations from various frames of reference may deepen learning by providing different and complementary insights [Arthur, Hancock, & Chrysler, 1994].
- *Multisensory cues:* Via high-end VR interfaces, students can interpret visual, auditory and haptic displays to gather information, while using their proprioceptive system to navigate and control objects in the synthetic environment. This potentially deepens learning and recall [Psotka, 1996].
- *Motivation:* Learners are intrigued by interactions with well designed immersive "worlds," inducing them to spend more time and concentration on a task [Bricken & Byrne, 1993].
- *Telepresence:* Geographically remote learners can experience a simultaneous sense of presence in a shared virtual environment [Loftin, 1997].

By using a virtual reality interface, instructional designers can not only display how a model can aid in interpreting a scientific phenomenon, but also can enable learners (1) to experience being part of the phenomenon, and (2) to participate in a shared virtual context within which the meaning of this experience is socially constructed.

The Potential of Multisensory Immersion for Learning Scientific Concepts

The virtual reality interface has the potential to complement existing approaches to science instruction. By themselves becoming part of a phenomenon (e.g., a student becomes a point-mass undergoing collisions in an immersive virtual environment without gravity or friction), learners gain direct experiential intuitions about how the natural world operates. In particular, good instructional design can make those aspects of virtual environments that are useful in understanding scientific principles salient to learners' senses. For example, in two-dimensional Newtonian microworlds students often ignore objects' velocities, instead focusing on position. In our comparable immersive environment, NewtonWorld, learners “inside” a moving object are themselves moving; this three-dimensional, personalized frame of reference centers attention on velocity as a variable. In NewtonWorld, we heightened this saliency by using multisensory cues to convey multiple, simultaneous representations of relative speeds. As another example of the power of “perceptualization,” learners who struggled with the concepts underlying our vector-field-based immersive environment, MaxwellWorld, reported that representations providing redundant data simultaneously through visual, auditory, and haptic stimuli aided their comprehension. Transducing data and abstract concepts (e.g., energy) into mutually reinforcing multisensory representations may be an important means of enhancing understanding of scientific models.

In addition, researchers are documenting that the social construction of knowledge among students in a shared, text-based virtual environment enables innovative, powerful types of collaborative learning [Turkle, 1995; Bruckman & Resnick, 1995]. As discussed later, adding immersive, multisensory representations to these textual “worlds” could potentially increase communicative and educational effectiveness. Overall, we believe that various aspects of multisensory immersion, when applied to scientific models, can provide learners with experiential metaphors and analogies that (1) aid in understanding complex phenomena remote from their everyday experience (e.g., relativity, quantum mechanics) and (2) help in displacing “common sense” misconceptions with alternative, more accurate mental models.

Challenges in Using Virtual Reality for Learning

However, many barriers intrinsic to current virtual reality technology can block students' mastery of scientific concepts. These challenges to educational design include:

- Virtual reality's physical interface is cumbersome [Krueger, 1991]. Head-mounted displays, cables, 3-D mice, and computerized clothing all can interfere with interaction, motivation, and learning.
- Display resolution is inversely proportional to field of view. A corresponding trade-off exists between display complexity and image delay [Piantanida, Boman, & Gille, 1993]. The low resolution of current VR displays limits the fidelity of the synthetic environment and prevents virtual controls from being clearly labeled.
- VR systems have limited tracking ability with delayed responses [Kalawsky, 1993].
- Providing highly localized 3-D auditory cues is challenging, due to the unique configuration of each person's ears. Also, some users have difficulty localizing 3-D sounds [Wenzel, 1992].
- Haptic feedback is extremely limited and expensive. Typically, only a single type of haptic feedback can be provided by computerized clothing; for example, one glove may provide heat as a sensory signal, but cannot simultaneously provide pressure. In addition, using computerized clothing for output can interfere with accurate input on users' motions.

- Virtual environments require users to switch their attention among the different senses for various tasks [Erickson, 1993]. To walk, users must pay attention to their haptic orientation; to fly, users must ignore their haptic sense and focus on visual cues. Also, as Stuart & Thomas [1991] describe, multisensory inputs can result in unintended sensations (e.g., nausea due to simulator sickness) and unanticipated perceptions (e.g., awareness of virtual motion, but feeling stationary in the real world).
- Users often feel lost in VR environments [Bricken & Byrne, 1993]. Accurately perceiving one's location in the virtual context is essential to both usability and learning.
- The magical (unique to the virtual world) and literal (mirroring reality) features of VR can interact, reducing the usability of the interface [Smith, 1987]. Also, some researchers have demonstrated that realism can detract from rather than enhance learning [Wickens, 1992].

As virtual reality technology evolves, some of the challenges to educational design will recede. At present, however, achieving the potential of immersive, synthetic worlds to enhance learning requires transcending these interface barriers through careful attention to usability issues.

Another class of potential problems with the use of immersive virtual worlds for education is the danger of introducing new or unanticipated misconceptions due to the limited nature of the "magic" possible via this medium. For example, learners will not feel their sense of personal physical weight alter, even when the gravity field in the artificial reality they have created is set to zero. The cognitive dissonance this mismatch creates, due to conflicting sensory signals, may create both physiological problems (e.g., simulator sickness) and possibly false intellectual generalizations. One part of our research is to examine the extent to which manipulating learners' visual, auditory, and tactile cues may induce subtle types of misconceptions about physical phenomena. The medium (virtual reality) should not detract from the message (learning scientific principles).

The Virtual Worlds of ScienceSpace

ScienceSpace is a collection of virtual worlds we have designed to explore the potential utility of physical immersion and multisensory perception to enhance science education [Dede, Salzman, & Loftin 1996]. ScienceSpace now consists of three worlds—NewtonWorld, MaxwellWorld, and PaulingWorld—in various states of maturity. All three worlds are built using a polygonal geometry. Colored, shaded polygons and textures are used to produce detailed objects. These objects are linked together and given behaviors through the use of NASA-developed software (VR-Tool) that defines the virtual worlds and connects them to underlying physical simulations. Interactivity is achieved through the linkage of external devices (e.g., a head-mounted display) using this same software. Finally, graphics rendering, collision detection, and lighting models are provided by other NASA-developed software.

Our hardware architecture includes a Silicon Graphics Onyx Reality Engine2 4-processor graphics workstation, Polhemus magnetic tracking systems (with a 3Ball or stylus), and a Virtual Research VR4 head-mounted display (HMD). One Polhemus tracker is in the 3Ball or stylus held by the participant in one hand; a second is mounted on a fixture and held in the other hand; and a third is mounted on the HMD. The hand holding the 3Ball or stylus is represented in the virtual world as a hand with the index finger extended (aligned with the user's hand). Attached to the second tracker is a menu system. Sound is produced by a Silicon Graphics Indy workstation and delivered via HMD headphones and external speakers. Vibrations are delivered to a subject's torso using a "vest" with embedded subwoofers. This interface enables us to immerse students in 3-D virtual worlds using the visual, auditory, and haptic senses. Students use a virtual hand (controlled by the 3Ball), menus, and direct manipulation to perform tasks in these immersive virtual environments.

NewtonWorld

NewtonWorld provides an environment for investigating the kinematics and dynamics of one-dimensional motion. In NewtonWorld, students spend time in and around an activity area, which is an open "corridor" created by colonnades on each side and a wall at each end (see Figure 1 on the next page). Students interact with NewtonWorld using a "virtual hand" and a menu system, which they access by selecting the small 3-ball icon in the upper left corner of the HMD. Students can launch and catch balls of various masses and can "beam" (teleport) from the ball to cameras strategically placed around the corridor. The balls move in one dimension along the corridor, rebounding when they collide with each other or the walls. Equal spacing of the columns and lines on the floor of the corridor aid learners in judging distance and speed. Signs on the walls indicate the presence or absence of gravity and friction.

Multisensory cues help students experience phenomena and direct their attention to important factors such as mass, velocity, and energy. For example, potential energy is made salient through tactile and visual cues, and velocity is represented by auditory and visual cues. Currently, the presence of potential energy before launch is represented by a tightly coiled spring, as well as via vibrations in the vest. As the ball is launched (Figure 2) and potential energy becomes kinetic energy, the spring uncoils and the energy vibrations cease. The balls now begin to cast shadows whose areas are directly proportional to the amount of kinetic energy associated with each ball. On impact, when kinetic energy is instantly changed to potential energy and then back to kinetic energy again, the shadows disappear and the vest briefly vibrates. To aid students in judging the velocities of the balls relative to one another, the columns light and chime as the balls pass.

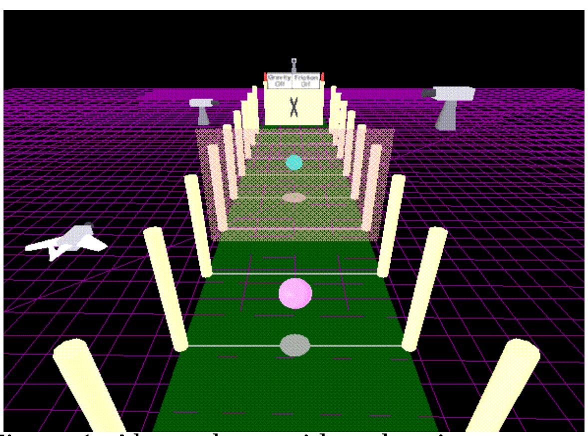

Figure 1. Above the corridor, showing cameras, balls with shadows, and the far wall

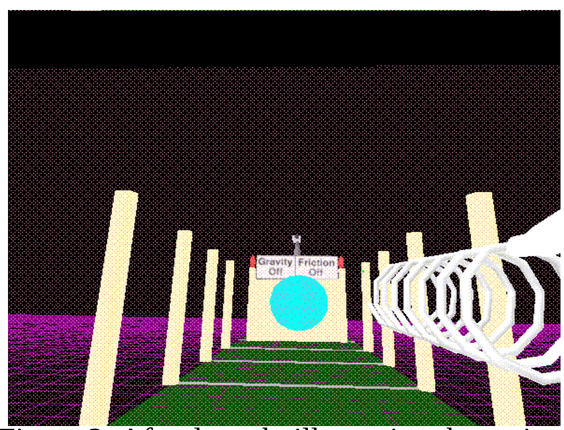

Figure 2. After launch, illustrating the spring-based launching mechanism

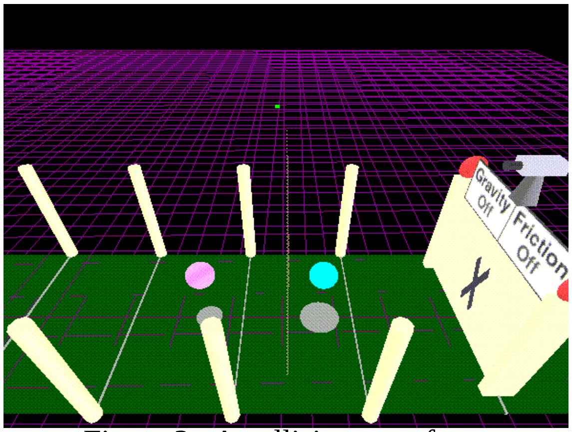

Figure 3. A collision seen from the center-of-mass reference frame

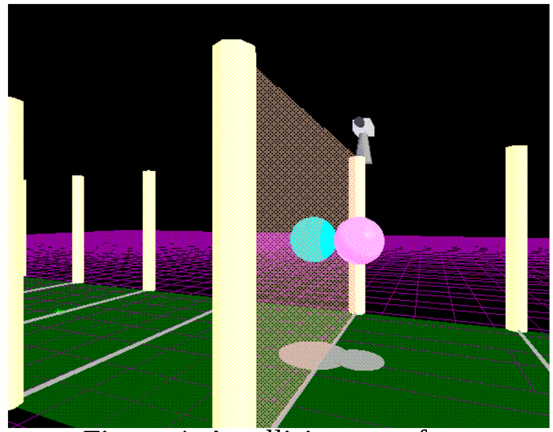
Figure 4. A collision seen from just outside a colonnade

Additionally, we provide multiple representations of phenomena by allowing students to assume the sensory perspectives of various objects in the world. For example, students can become one of the balls in the corridor, a camera attached to the center-of-mass of the bouncing balls (Figure 3), a movable camera hovering above the corridor, etc. Figure 4 shows a collision seen from just outside one colonnade. These features aid learners in understanding the scientific models underlying Newton's three laws, potential and kinetic energy, and conservation of momentum and energy.

NewtonWorld was the first virtual environment we built, so its current interface does not incorporate sophisticated features we developed in designing MaxwellWorld and PaulingWorld. Accordingly, we are redesigning NewtonWorld to take advantage of these new capabilities. On the next page are two sketches illustrating our redesign, at present under construction. New features include a "scoreboard" (Figure 5) to aid learners in relating qualitative and quantitative representations, an improved interface based on a "roadway" metaphor (Figure 6), three levels of interaction that support progressively more complex types of learning activities, and the inclusion of perfectly elastic and perfectly inelastic collisions. Dede, Salzman, Loftin, & Ash [in preparation] provides additional detail on our design strategies and early research results for NewtonWorld.

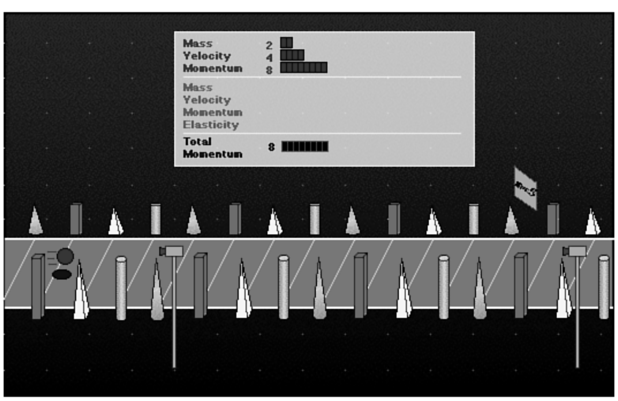

Figure 5. Level 1 of redesigned NewtonWorld showing "scoreboard" and "roadway"

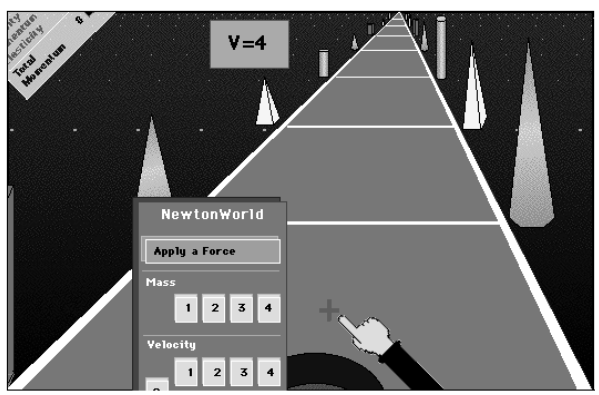

Figure 6. Within the "roadway" view

PaulingWorld

PaulingWorld enables the study of molecular structures via a variety of representations, including quantum-level phenomena.

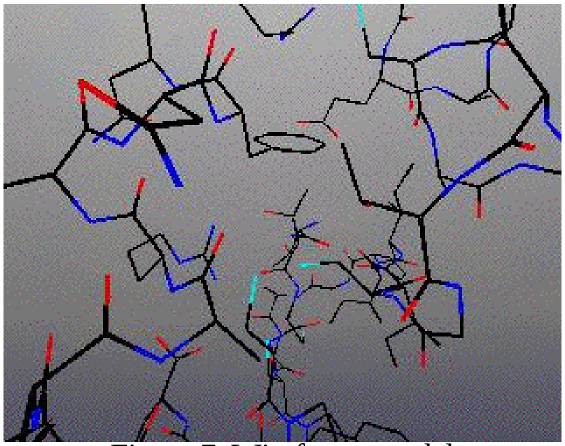
Figure 7. Wireframe model


Figure 8. Backbone model

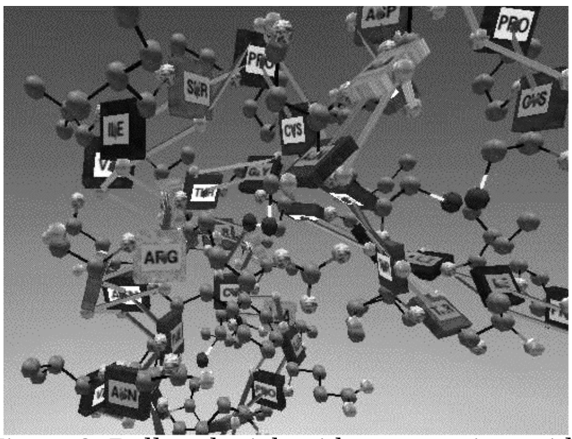
Figure 9. Ball-and-stick with some amino acids

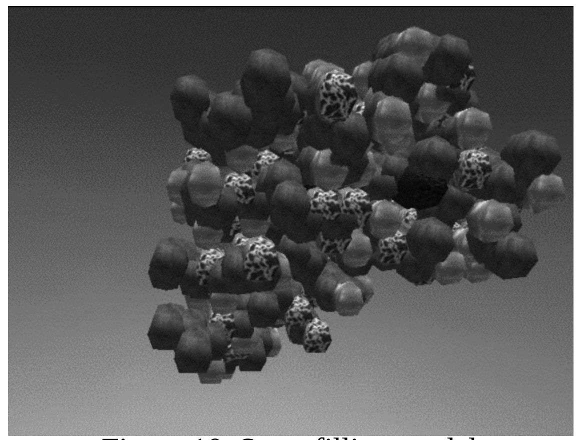
Figure 10. Spacefilling model

PaulingWorld is in its early stages of development. Learners can view, navigate through, superimpose, and manipulate five different molecular representations: wireframe, backbone, ball-and-stick, amino acid, and space-filling models. See Figures 7 through 10 for examples of these models. To design the immersive multisensory representations and underlying scientific models we will use for quantum-mechanical bonding phenomena, we are working with a NSF-funded project, “Quantum Science Across the Disciplines,” led by Peter Garik at Boston University (http://qsad.bu.edu/).

MaxwellWorld

Although we will discuss examples from all three of our virtual worlds, this chapter centers on our design and evaluations of MaxwellWorld (described below) as an illustration of how models based on multisensory immersion can aid in learning complex scientific concepts. To date, we have collected more research data on learning in MaxwellWorld than in our other virtual environments, and MaxwellWorld also illustrates some particularly interesting applications of scientific modeling to education.

MaxwellWorld allows students to explore electrostatic forces and fields, learn about the concept of electric potential, and "discover" the nature of electric flux. The fieldspace in this virtual world occupies a cube approximately one meter on a side, with Cartesian axes displayed for convenient reference. The small size of the world produces large parallax when viewed from nearby, making its three-dimensional nature quite apparent.

Students use a virtual hand, menu, direct manipulation, and navigation to interact with this world (see Figure 11 on the next page). The virtual hand is attached to the 3Ball, which is held in one hand. The menu is attached to the tracker held by the other hand. Attaching the menu to user's other hand allows students to remove the menu from their field of view, while keeping it immediately accessible. Students select menu items by holding up the menu with one hand, pointing to the menu option with the virtual hand, and depressing the 3Ball button (see Figure 12). Thus, menu selection in MaxwellWorld is similar to menu selection on two-dimensional interfaces in which users manipulate the menu with a cursor controlled by a mouse. MaxwellWorld also utilizes direct manipulation. For example, once users have selected objects from the menu, they can place them in the world, move them around, and delete them. Finally, users can change their location by selecting the navigation mode via the menu, pointing the virtual hand in the desired direction, and depressing the 3Ball button.

Our design of which vector field phenomena and representations to incorporate into MaxwellWorld was based on the advice of our domain expert, Dr. Edward Redish from the University of Maryland. Using the virtual hand, students can place both positive and negative charges of various relative magnitudes into the world. Once a charge configuration is established, learners can instantiate, observe, and interactively control model-based scientific representations of the force on a positive test charge, electric field lines, potentials, surfaces of equipotential, and lines of electric flux through surfaces. For example, a small, positive test charge can be attached to the tip of the virtual hand. A force meter associated with the charge then depicts both the magnitude and direction of the force of the test charge (and, hence, the electric field) at any point in the workspace. A series of test charges can be "dropped" and used to visualize the nature of the electric field throughout a region. In our most recent version of MaxwellWorld, learners can also release a test charge and watch its dynamics as it moves through the fieldspace (Figure 13), then "become" the test charge and travel with it as it moves through the electric field.

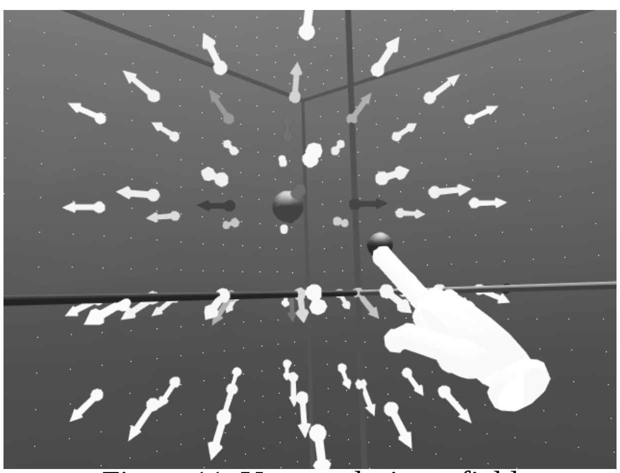

Figure 11. User exploring a field with test charges and field lines.

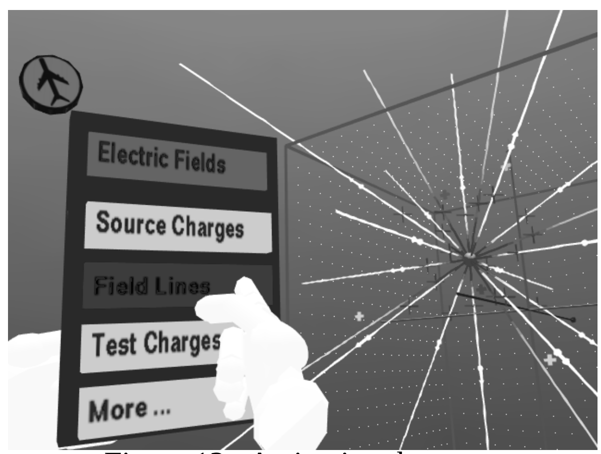

Figure 12. Activating the menu via the virtual hand.

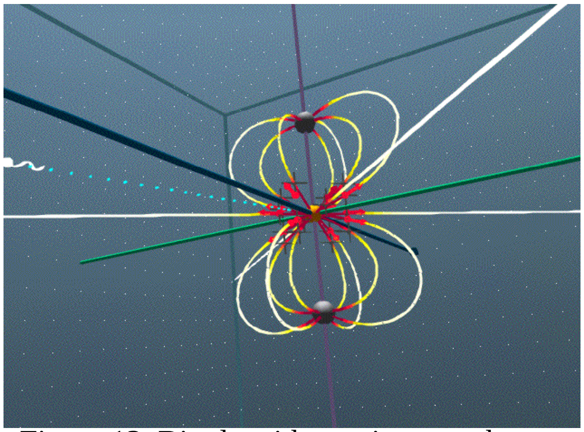

Figure 13. Bipole with moving test charge.

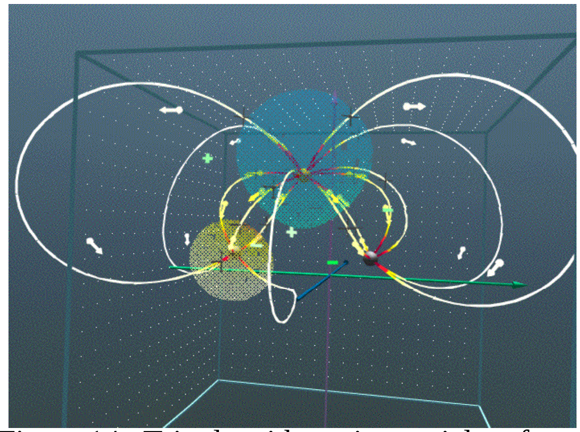

Figure 14. Tripole with equipotential surface.

In a like manner, an electric field line can be attached to the virtual hand. A student can then move his or her hand to any point in the workspace and see the line of force extending through that point. MaxwellWorld can also display many electric field lines to give students a view of the field produced by a charge configuration (Figure 11). In another mode of operation, the tip of the virtual hand becomes an electric "potential" meter that, through a simple color map and a "=" or "-" sign on the finger tip, allows students to explore the distribution of potential in the fieldspace. Via the production and manipulation of equipotential surfaces, learners can watch how the shapes of these surfaces alter in various portions of the fieldspace (Figure 14). By default, the surfaces are colored to indicate the magnitude of the potential across the surface; however, the student can also choose to view the electric forces as they vary across the surface. This activity helps students to contrast the concepts of electric force and potential.

Via the production of a "Gaussian" surface, the flux of the electric field through that surface can be visually measured. Gaussian surfaces can be placed anywhere in the workspace by using the virtual hand to anchor the sphere; the radius (small, medium, large) is selected from the menu. This representation enables students to explore flux through a variety of surfaces when placed at various points in the field. All these capabilities combine to enable representing many aspects of the complex scientific models underlying vector field phenomena.

Conducting Research on ScienceSpace

We have developed elaborate, customized assessment methodologies for evaluating the usability and learnability of our ScienceSpace worlds [Salzman, Dede, & Loftin, 1995]. Although infrequent, potential side effects such as “simulator sickness” mandate the inclusion of special questions and protections to ensure users’ comfort. Moreover, because each person evolves a unique psychomotor approach to interacting with the three-dimensional nature of physical space, individuals appear to have much more varied responses to 3-D, multimodal interfaces than to the standard 2-D graphical user interface with menus, windows, and mouse. Evaluating the multisensory dimensions of an immersive virtual world also adds an additional level of complexity to the assessment process.

Thus, portions of our protocols center on calibrating and customizing the virtual world’s interface to that particular learner. Throughout the sessions, we carefully monitor the learning process and record student comments and insights. We also videotape the hours of time we spend with each subject so that we can study these records for additional insights. Finally, our protocols are designed so as to help us capture various aspects of the learner’s experience, in addition to assessing educational outcomes. By focusing on the students’ experience as well as their learning, we gain insights that guide the refinement of the user interface and help us understand how to leverage VR’s features for modeling science..

Below is a summary of the four issues our protocols are designed to assess.

- *The VR experience*. The VR experience can be characterized along several dimensions. We have focused on usability, simulator sickness, meaningfulness of our models and representations, and motivation. Our most recent addition has been the inclusion of questions to assess how immersed students feel in the modeling environment.
- *Learning.* Our focus in learning is to determine whether and how students progress through learning tasks within the virtual environment, to assess their mastery of concepts at both the descriptive and the causal levels (discussed later), and to assess whether their learning can be generalized to other domain-specific problems.
- *The VR experience vs. learning.* Our focus in this contrast is to understand the relationship between the VR experience and learning and to identify when the two goals may conflict.
- *Educational utility.* To demonstrate that the system is a better (or worse) teaching tool than other pedagogical strategies, comparing the quality and efficiency of learning among different alternatives of varying cost, instructional design, and pedagogical strategy.

This careful evaluation strategy is generating data from which we are gaining insights into how multisensory immersion can enhance learning, as well as how virtual reality's usability can be enhanced. Many of the strategies underlying these assessment methodologies and instruments are also generalizable to a wide range of synthetic environments beyond virtual reality.

MaxwellWorld: Formative Evaluation

In Summer, 1995, we assessed our initial version of MaxwellWorld as a tool for 1) remediating misconceptions about electric fields, and 2) teaching concepts with which students are unfamiliar. During the sessions, we administered one to three lessons centering on the construction and exploration of electric fields (electric force, superposition, test charges, field lines); electric potential (potential and kinetic energy, potential difference, work, potential vs. force); and the concept of flux through open and closed surfaces, leading up to Gauss's Law.

Our observations during these sessions, students' predictions and comments, usability questionnaires, interview feedback, and pre- and post-test knowledge assessments helped us to determine whether this early version of MaxwellWorld aided students in remediating any of their pre-existing misconceptions and in learning underlying scientific concepts with which they were unfamiliar. Additionally, these experiences were valuable in developing modifications to MaxwellWorld to enhance learning outcomes.

The findings below are based on 14 high school students and 4 college students who participated in these evaluations. Thirteen of the 14 high school students had recently completed their senior year; 1 student had recently completed his junior year. All students had completed 1 course in high school physics. Each session lasted for approximately 2 hours. Students were scheduled on consecutive days for the first two sessions, while the third session was conducted approximately 2 weeks later.

All of the students enjoyed learning about electric fields in MaxwellWorld. When asked about their general reactions to MaxwellWorld, a majority of the students commented that they felt it was a more effective way to learn about electric fields than either textbooks or lectures. Students cited the 3-D representations, the interactivity, the ability to navigate to multiple perspectives, and the use of color as characteristics of MaxwellWorld that were important to their learning experiences.

Pre- and post-lesson evaluations showed that lessons in MW helped students deepen their understanding of the distribution of forces in an electric field, as well as their comprehension of the scientific models interrelating representations such as test charge traces and field lines. Manipulating models of the vector fields in three dimensions appeared to play an important role in their learning. For example, several students who were unable to describe the distribution of forces in any electric field prior to using MaxwellWorld gave clear descriptions during the post-test interviews and demonstrations. Also, manipulating field lines and traces in three dimensions helped students visualize the distribution of force. As an illustration, one student expected field lines to radiate from a single charge along a flat plane and was surprised to see that they radiated in three dimensions. Another student expected to see field lines cross, but found this could not occur.

Although this initial version of MaxwellWorld helped students qualitatively understand 3-D superposition, students had difficulty applying superposition when solving post-test problems. Learners appeared to understand the concept of superposition during the lessons and particularly enjoyed the demonstrations of superposition (moving the source charges dynamically changes the traces and field lines), often alluding to this during the post-testing. However, many of them exhibited difficulties in applying superposition to post-test demonstrations and sketches, indicating the need to refine our modeling and instruction.

This early version of MaxwellWorld extended traditional 2-D representations to include 1) the third dimension; 2) the ability to manipulate representations as a means of understanding the dynamics of electrostatic models; and 3) two color schemes to measure and distinguish the magnitude of the force on and the potential experienced by test charges, field lines, and equipotential surfaces. These representational capabilities helped students to deepen their understanding of physics concepts and models. The post-test outcomes showed that students were able to learn about flux through open and closed surfaces using MaxwellWorld. All students performed well during post-testing, demonstrating an understanding of important and difficult-to-master concepts such as Gauss's law, field vs. flux, and directional flux.

Although only four of the students used MaxwellWorld to learn about electric potential, all of them demonstrated that they could visualize the distribution of potential for basic charge arrangements, interpret the meaning of a distribution of potential, identify and interpret equipotential surfaces, relate potential difference and work, and describe some of the differences between electric force and electric potential. All were particularly surprised to see 1) 3-D representations of equipotential surfaces, particularly in the case of a bipole (two charges of the same size and magnitude), and 2) the varying nature of forces over an equipotential surface.

We observed significant individual differences in the students' abilities to work in the 3-D environment and with 3-D controls, as well as their susceptibility to symptoms of simulator sickness (eye strain, headaches, dizziness, and nausea). While some students learned to use the menus, manipulate objects, and navigate very rapidly, others required guidance throughout the sessions. Most students experienced nothing more than slight eyestrain; however, two students experienced moderate dizziness and slight nausea during the first session, and consequently did not return for the second session. No student complained of any symptoms during the first 30-45 minutes of the lesson, reinforcing our strategy of using multiple, short learning experiences.

These “lessons learned” from an early formative evaluation of MaxwellWorld are consistent with evaluative data collected on our other ScienceSpace worlds [Salzman, Dede, & Loftin, 1995]:

- Enabling students to experience phenomena from multiple perspectives appears to help learners understand complex scientific concepts and models.
- Multisensory cues appear to engage learners, direct their attention to important behaviors and relationships, help them understand new sensory perspectives, prevent errors through feedback cues, and enhance ease of use.
- Simulator sickness and system usability pose potential threats to the learning process.
- Talk-aloud protocols employing a cycle of prediction-observation-comparison [White, 1993] are highly effective for administering lessons and for monitoring usability and learning in VR modeling environments.

Our early evaluations of MaxwellWorld indicate that using this type of scientific modeling helped students to deepen their understanding of and to remediate misconceptions concerning electric fields and potential. However, these studies did not establish the extent to which students' learning was due to (a) the method of instruction (the lessons), (b) scientific models and representations that could equally well have been utilized in a conventional 2-D modeling environment, or (c) the unique features of multisensory immersion in virtual reality.

MaxwellWorld: Multisensory Immersion versus Conventional 2-D Representations

In January, 1996, we initiated an extended study designed to accomplish two goals: (1) compare learning and usability outcomes from MaxwellWorld to those from a highly regarded and widely used two-dimensional microworld, EM Field™, which covers similar material, and (2) assess the usability and learnability of an enhanced version of MaxwellWorld with additional modeling and representational capabilities suggested by results from the initial formative evaluation above.

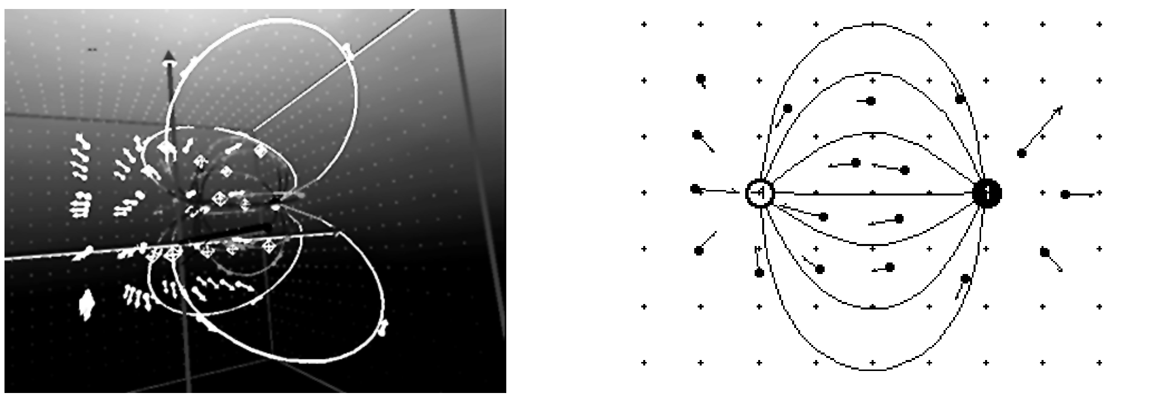

Figure 15. A dipole with field lines and test charge traces in MW and EMF.

The first phase of this study compared MaxwellWorld (MW) and EM Field (EMF) on the extent to which representational aspects of these simulations influenced learning outcomes. EM Field runs on standard desktop computers and presents learners with 2-D representations of electric fields and electric potential, using quantitative values to indicate strength [Trowbridge & Sherwood, 1994]. To make the two learning environments comparable, we removed some of MaxwellWorld's more powerful features and designed lessons to utilize only those features of MaxwellWorld for which EM Field had a counterpart; this limited version of MaxwellWorld we designated MW_L. Under these conditions constraining the functionality of the VR environment, the primary differences between the simulations were representational dimensionality (EMF's 2-D vs. MW_L's 3-D) and type (EMF's quantitative vs. MW_L's qualitative). (See Figure 15 below.)

In the second phase of the study, we utilized MaxwellWorld's full range of capabilities (including multisensory input) to ascertain the value these features added to the learning experience. Through the pre-test for phase two, we also examined the extent to which students, after a period of five months, retained mental models learned in either environment. Through this two-stage approach, we hoped to separate the relative contributions of 3-D representation vs. multisensory stimulation as instrumental to the learning potential of virtual reality.

Initial Hypotheses

Our initial hypotheses for this two-phase study were:

Learning: Learning can occur along three dimensions. First, there is *conceptual understanding*, students' ability to define key concepts and describe interrelationships among significant representations in the scientific model. Second, there is *2-D understanding*, students' abilities to create and interpret 2-D representations of the phenomena (the ability to illustrate concepts on paper). Third is *3-D understanding*, students' abilities to create and interpret full 3-D representations of the phenomena (reflecting the ability to visualize the true three-dimensional nature of the concept). We expected students to learn along all three dimensions while completing lessons in either MW_L or EMF.

Learning in EMF vs. MW_L *:* Our previous experience with VR learning environments indicated 3-D simulations are likely to facilitate the construction of more complete and accurate mental models of intrinsically three-dimensional phenomena. Therefore, we hypothesized that students who used MW_L would perform better on conceptual questions than those who used EMF. Additionally, students in earlier studies of MW had demonstrated the ability to represent phenomena using 2-D sketches after working in the 3-D simulation. However, because students learning in MW_L would need to translate 3-D information into 2-D information for the tests, we did not expect their performance to exceed the performance of students working in a 2-D environment, who would not need to perform this translation. Therefore, we hypothesized that students who used MW_L would not perform significantly worse on questions requiring two-dimensional understanding than those who use EMF. Finally, we felt working in 3-D would result in a better three-dimensional understanding of the phenomena than working in 2-D. Therefore, we hypothesized that students who used MW_L would demonstrate significantly better three-dimensional understanding than those who used EMF.

Learning in the full version of MW: In phase two of the study, with the constraints on MaxwellWorld's performance removed, we hypothesized that students would identify the multisensory representations as valuable for their learning.

Retention: We hypothesized that, over a five month period, students would experience significantly greater retention of learning in MW_L over learning in EMF.

We also identified several additional factors that would likely influence the learning experience and outcomes:

Simulator sickness: Our work with virtual realities suggested that many students would experience mild symptoms of simulator sickness, particularly eye strain, at some point during their use of the system. We hypothesized that MW_L and MW students would experience significantly more simulator sickness symptoms than EMF students. Additionally, we expected simulator sickness to interfere with learning.

Nature of the learning experience: Inherent to any human-computer interaction are the subjective experiences of usability, motivation, and ability to understand the representations. We expected students to find MW_L and MW more motivating and meaningful than EMF. We hypothesized that such greater motivation might result in increased learning (through such factors as increased student attention and concentration). We also expected students to find using MW_L and MW more difficult to use than EMF. However, because careful design of the interface and lessons had greatly reduced usability problems in MW, we hypothesized that these interface challenges characteristic of VR would not interfere with learning.

Stage One of the Comparison Study

Fourteen high school students (12 males and 2 females) completed the first stage of this study. All students had 1-1/2 years of high school physics and were recruited from a physics class in a local high school. (The gender disparity in the sample population was caused by the relative paucity of women who take high school level physics.) Students' performance in their science and math classes varied; their grades ranged from A through C (As and Bs were the norm). Although students were advanced in their knowledge of physics relative to the typical high school population, the pre-test they received at the start of this study indicated that most remembered little about electric fields and electric potential—this confirms the limits of conventional approaches to teaching this type of scientific material. Students participated in two, two-hour learning experiences, completing lessons in either EMF or MW_L. The lessons focused on electric fields and electric potential, mirroring concepts covered in our initial formative evaluation of MW.

Procedure and materials

Evaluations were conducted in our virtual reality lab, where we videotaped and logged student-administrator interactions. Male and female students were assigned randomly to one of two groups: EMF, and MW_L. Both groups were equivalent in terms of their science background, and both groups of students participated in two sessions. Prior to the sessions, students were asked to complete Recruitment questionnaires. During Session 1, students completed a Background questionnaire and Pre-lesson test, and then they completed Lesson 1. During Session 2, students completed Lesson 2, which was followed by a Post-lesson test, Experience questionnaire, and Interview. Immediately following each lesson, students also completed a Simulator sickness questionnaire. Lessons required approximately 1 hour 15 minutes to complete, and sessions lasted for approximately 2 hours. Students were given a break approximately half-way through each lesson.

Background questionnaire & Recruitment questionnaires. These questionnaires were used to characterize our sample for this study. They gathered information about the students' demographic information, educational backgrounds, attitudes towards learning science, and motion sickness history.

Lessons. Lessons were constructed so that the informational content and learning activities were the same for both groups. Lesson 1 focused on the construction and exploration of electric fields, while Lesson 2 focused on electric potential and its relationship to the electric field. These lessons were administered verbally by the test administrator. Learning activities in the lessons relied on a cycle of "predict-observe-compare." This served two purposes: to help us gauge the students' understanding and progress, and to prime them for the upcoming activity. Each successive learning activity built on the previous activities, increasing both in level of complexity and in the information integration necessary.

Pre- and Post-lesson tests. We used two versions of a Pre/Post-lesson test to assess learning. Half of the students in each group were randomly assigned to receive version A for the pre-test and version B for the post-test, and vice versa. These tests examined three dimensions of understanding for each concept: conceptual understanding (ability to define concepts); two-dimensional understanding (abilities to create and interpret 2-D representations); and three-dimensional understanding (ability to create and interpret 3-D representations). The first two sections were administered on paper. The third section was administered verbally, and students used physical 3-D manipulatives to demonstrate their understanding.

Experience questionnaire. This questionnaire was used to assess the nature of the learning experience. It consisted of a series of 7 point anchored rating scales (e.g., Using the menu system was... very difficult -3 to +3 very easy) relating to usability, motivation, and ability to understand the representations.

Simulator sickness. The Simulator sickness questionnaire (SSQ) [Kennedy, Lane, Berbaum, & Lilienthal, 1993] consisted of a series 4 point ratings of symptoms associated with simulator sickness. It can be analyzed to yield oculomotor, disorientation, nausea, and overall simulator sickness scores.

Interviews and qualitative data. To help us understand the nature of the statistical outcomes, as well as to diagnose strengths and weaknesses of EMF and MW and of our lessons, we collected the following qualitative data: students' predictions and observations throughout the lessons, their comments, likes and dislikes, suggestions for improvement, and reflections on the learning process.

Analyses and Results

Stage one of our study yielded the following outcomes:

Learning. As anticipated, using either MW_L or EMF, students learned as a result of completing the lessons. Students were better able than before to (a) define concepts, (b) describe concepts in 2-D, and (c) demonstrate concepts in 3-D. Table 1 summarizes these learning outcomes.

Learning in MW_L vs. EMF. MW_L students were better able to define concepts than EMF students (see Table 2). Also, MW_L students were not any worse than the EMF students at sketching concepts in 2-D. A close examination of the sketches shows that, while MW_L students performed better on the force sketches, they performed worse on the sketches relating to potential, resulting in total sketch scores that were similar for the two groups. Finally, MW_L students were better able than EMF students to demonstrate concepts and their underlying scientific models in 3-D. Despite the inherent three-dimensionality of the demonstration exercises (as well as our use of the terms "surface" and "plane" in the lessons), EMF students typically restricted answers to a single plane; drew lines when describing equipotential surfaces; and used terms such as "circle," "oval," and "line." In fact, only one of the seven students in the EMF group described the phenomena in a three-dimensional manner. In contrast, MW_L students described the space using 3-D gestures and phrases such as "sphere" and "plane" referring to equipotential surfaces.

Learning	**EMF**		**MW_L**	
	Pre	Post	Pre	Post
Concepts	.12 (.11)	.58 (.08)	.14 (.18)	.70 (.15)
	F(1,6) = 135.70, p < .05*		F(1,6) = 48.37, p < .05*	
2-D sketches	.32 (.24)	.78 (.12)	.38 (.28)	.82 (.09)
	F(1,6) = 39.75, p < .05*		F(1,6) = 26.61, p < .05*	
3-D demos	.34 (.16)	.69 (.15)	.22 (.19)	.85 (.07)
	F(1,6) = 70.15, p < .05*		F(1,6) = 69.72, p < .05*	

Table 1. Mean pre- and post-lesson scores (and standard deviations) and outcomes for F-tests; significant outcomes are marked with an *.

Simulator sickness scores. As we anticipated, MaxwellWorld's immersive VR environment induced more symptoms associated with simulator sickness than EM Field's monitor-based 2-D environment. (See Table 3.) In MW_L, overall simulator sickness scores tend to be slightly higher on day 1 than day 2; we suspect this may be due to an adaptation to the VR environment. Consistent with our earlier research findings, there also appear to be large individual differences in the way students react to the VR environment. We found simulator sickness scores did *not* significantly predict learning outcomes; while a minor nuisance, simulator sickness did not interfere with mastering the material.

Learning Concepts	EMF	MW_L
	.58	.70
	$F(1,11) = 3.17$, $p < .05$*	
2-D sketches	.80	.82
	$F(1,11) = .24$, $p > .05$	
3-D demos	.67	.87
	$F(1,11) = 9.99$, $p < .05$*	

Table 2: Adjusted mean Post-lesson test scores for each group; and outcomes for the ANCOVA (group by Pre-lesson test score covariate); significant outcomes are marked with an *.

Simulator Sickness	EMF		MW_L	
	Day 1	Day 2	Day 1	Day2
Total SSQ score	1.0 (1.0)	2.39 (3.21)	9.03 (7.57)	5.33 (6.42)
	$F_{group}(1,11) = 6.01$, $p < .05$* $F_{day}(1,11) = .46$ $p > .05$ $F_{group*day}(1,11) = 2.25$ $p > .05$			

Table 3. Mean total SSQ scores (and standard deviations) for each group by day; and outcomes for the ANOVA (group by day); significant outcomes are marked with an *.

The nature of the learning experience. Table 4 summarizes how students rated various aspects of their learning experiences: usability, motivation, and understanding. Students rated MW_L as more motivating than EMF. Ideally, we would like to see the ratings for motivation even higher. However, we suspect that the intensity and the controlled nature of the lessons may prevent students from feeling extremely motivated during the learning experience. We found that motivation scores did not significantly predict learning outcomes, and it was not motivation alone that accounted for the differences in each groups' learning.

Students found using the various features of MW_L significantly more difficult than using EMF. Further, ratings for the ability to understand MW_L 's representations were slightly, though not significantly, higher than the ratings for EMF. Notice also that the variability in ratings was greater for EMF than for MW_L , suggesting that there were more individual differences in ability to extract information from the EMF representations than from the MW_L representations.

Nature of experience	EMF	MW_L
Usability	2.41 (.52)	1.77 (.57)
	F(1,12) = 4.77, p < .05*	
Motivation	1.11 (.82)	2.03 (.29)
	F(1,12) = 7.66, p < .05*	
Understand	1.93 (1.13)	2.36 (.38)
	F(1,12) = .90, p > .05	

Table 4. Mean usability, motivation, and understanding ratings (and standard deviations) for each group, and outcomes for the ANOVA; significant outcomes are marked with an *.

Student comments. Students' comments provide further insight into the nature of the learning experience. Overall, students described MW_L as easy to use, interesting, and informative. They especially liked the three-dimensional representations, the ability to see phenomena from multiple perspectives, and the interactivity of the system. MW_L students found using the 3-Ball and virtual hand somewhat challenging and indicated that the responsiveness of MW_L was problematic at times. Students described EMF as very easy to use, but somewhat boring. They found the simplicity of its graphics both a strength and a weakness. Additionally, MW_L students indicated that they found it easier to remain attentive during sessions than EMF students.

Stage Two of the Comparison Study

Procedure and materials

During stage two, we examined the "value added" by the full power of MaxwellWorld's multisensory representations. Seven EMF and MW_L, students returned for stage two, conducted approximately 5 months after stage one. All students experienced the full power of MaxwellWorld, receiving an additional lesson (built upon the concepts taught in earlier lessons) that relied on multisensory cues to supplement the visual representations. The auditory and haptic representations used provided simultaneous, redundant information to that expressed through the visual sensory channel.

We also assessed stage one retention at the beginning of stage two. (The retention test was an abbreviated version of the post-lesson test used in stage one.)

Analyses and Results

As seen as in Table 5, no statistically significant differences were observed in retention outcomes. However, this stage of the study had very low power (with only seven participants). Our limited data are suggestive that, with a larger number of subjects, retention of 3-D understanding might be significantly higher for MaxwellWorld participants than for EM Field participants.

Data for stage two did yield insights into the value of multisensory representations. Students learned from visual and multisensory representations used in the lesson and demonstrated significantly better understanding of concepts, 2-D sketches, and 3-D demos post-lesson than pre-lesson. Ratings concerning multisensory representations (haptic and sound), post-lesson understanding, and student comments all suggest that learners who experienced difficulty with the scientific concepts found that multisensory representations helped them understand more than did purely visual representations.

Learning Concepts	EMF	MW_L
	.69	.66
	$F(1,5) = .27$, $p > .05$	
2-D sketches	.42	.43
	$F(1,5) = .0.00$, $p > .05$	
3-D demos	.31	.57
	$F(1,5) = 2.40$, $p > .05$	

Table 5. Adjusted retention means and ANCOVA outcomes for stage one (covariate = pre-lesson scores); significant outcomes are marked with an *.

Summary of this Comparison Study

Both stages lend support to the thesis that immersive 3-D multisensory representations can help students develop more accurate and causal mental models than 2-D representations. Learning outcomes for stage one show that MW_L learners—more than EMF learners—were able to understand the space as a whole, recognize symmetries in the field, and relate individual visual representations (test charge traces, field lines, and equipotential surfaces) to the electric field and electric potential. MW_L students appeared to visualize the phenomena in 3-D, while EMF students did not.

Subjective ratings for stage one yielded converging evidence that the virtual worlds' representational differences were responsible for differences in learning. In stage 1, students rated the representations used in MW_L as easier to understand than representations used in EMF. Second, differences in learning could not be attributed soley to motivation (which was higher in MW_L than EMF). Additionally, MW_L students learned more even though they experienced more usability and simulator sickness problems. Finally, during interviews students cited MW_L's immersive 3-D representations as one of its key strengths. In stage two, the enhancement of visual representations with multisensory cues appeared to enhance learning, especially for students who had trouble grasping the concepts.

Outcomes this study support the following findings related to modeling scientific concepts:

- Virtual modeling experiences such as those provided by EMF and MW_L should be integrated with initial instruction to avoid forming misconceptions difficult to remediate later. Although students in both the EMF and MW_L groups demonstrated a better overall understanding of the topics on the post-test than on the pre-test, some students with a moderate knowledge of electrostatics at pre-test benefited less than students demonstrating little or no knowledge at pre-test. In addition, some of the more advanced students who had misconceptions appeared to have a difficult time overcoming them despite experiences in the virtual worlds.

- Immersive 3-D multisensory representations such as those used in MW_L may facilitate the students' development of more complete and runnable mental models than the 2-D representations of EMF. Learning outcomes, subjective ratings, and comments from both stages 1 and 2 all provide evidence that supports this finding.

- Although no new types of misconceptions were introduced by conducting the learning experiences in an immersive environment, students have a number of misconceptions about electrostatic phenomena and some of them are difficult to remediate. Working with the students yielded insights into the nature of their preexisting misconceptions. For example, learners had a strong tendency to think of charges in an electric field independently and had trouble describing the nature of superpositional fields and potential for sets of charges. Experiences in both MW_L and EMF clearly helped students to think about this issue, but they still had some difficulty understanding regions between sets of charges. In addition, field line representations are notoriously difficult to comprehend. Even after use of EMF or MW_L, several students continued to have misconceptions about the meaning of a field lines, although most learners gained a greater understanding of this representational formalism. Upon concluding the lessons in either system, how the electric field influences charged objects and the interrelationship between potential and force were also not completely understood by some students. Thus, modeling environments and activities must be carefully designed to target these kinds of misconceptions.

Although the subject population is small, the results of this study suggest that the three-dimensional nature of VR aids with learning and that the virtual reality experience can be more meaningful and motivating for students than comparable 2-D microworlds. Given that many capabilities of MaxwellWorld were suppressed in this study, these finding are a promising indication of the potential of immersive scientific models to enhance educational outcomes.

Next Steps in Our Virtual Reality Research

Over the next two years, we plan to extend our current research on the ScienceSpace worlds along several dimensions. For example, we will conduct two studies in MaxwellWorld (on immersive frames-of-reference and on multiple sensory channels) to examine the contribution of “perceptualization” to scientific model-based learning. Using the revised version of NewtonWorld, we also intend to examine how, by facilitating innovative types of student collaborations, virtual reality may enhance the nature of social constructivist learning. These three planned studies are described in more detail below.

Further, to examine challenges in curriculum integration and in classroom implementation, we will move our VR worlds out of laboratory environments into pre-college settings. In public school classrooms (fifth grade for NewtonWorld, twelfth grade for MaxwellWorld), we plan to integrate VR experiences into science instruction. Through this, we can assess whether our laboratory results on learning and usability hold up in the more complex environment of schools, as well as determine how students and teachers adapt VR environments to their needs and interests.

Understanding Frames-of-Reference as a Means of “Perceptualization”

We believe that transforming current scientific visualization tools into "perceptualization" experiences may augment their power for learning. We have documented that adding multisensory perceptual information aided students struggling to understand the complex scientific models underlying NewtonWorld and MaxwellWorld. Providing experiences that leverage human pattern recognition capabilities in three-dimensional space (e.g., shifting among various frames-of-reference and points-of-view) also extends the perceptual nature of a visualization. These enhanced visualization (or perceptualization) techniques facilitate student experiences that increase the saliency and memorability of abstract scientific models, potentially enhancing the learning process.

By using frames-of-reference (FORs) in virtual reality, we can provide learners with experiences that they would otherwise have to imagine. For example, we can enable students to become part of a phenomenon and experience it directly. Alternatively, we can let learners step back from the phenomenon to allow a global view of what is happening. One frame-of-reference may make salient information that learners might not notice in another frame-of-reference. Further, multiple frames-of-reference might help students to fill in gaps in their knowledge and to become more flexible in their thinking.

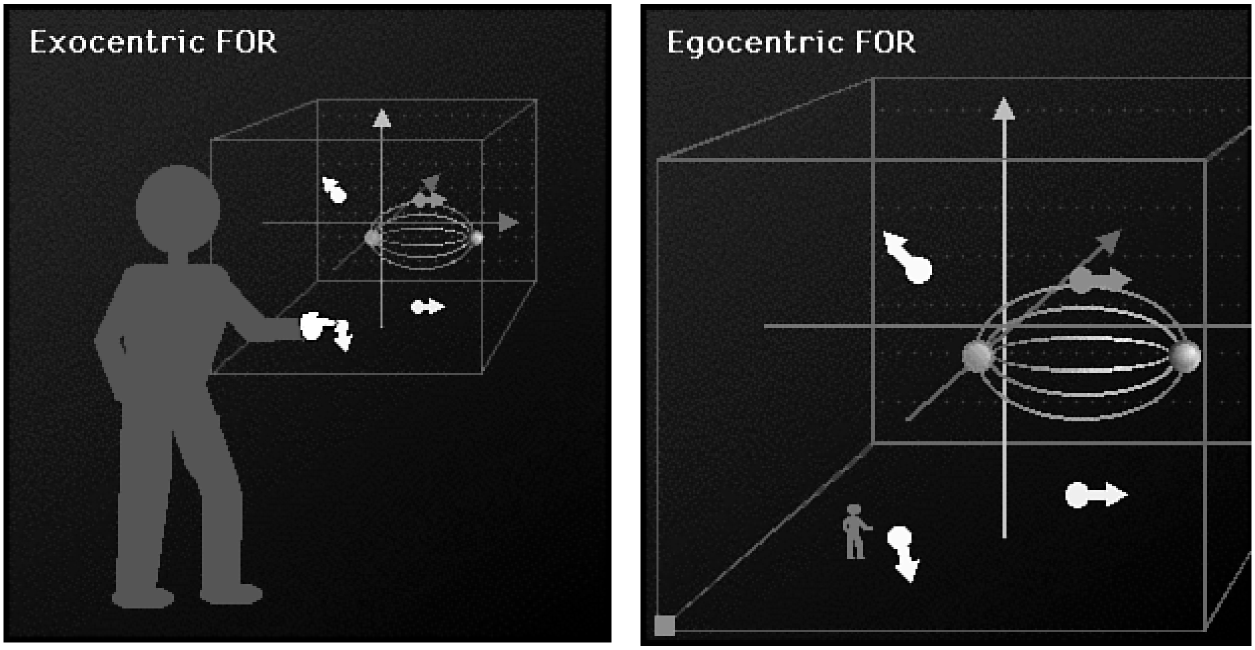

Figure 16. Exocentric versus Egocentric Frames of Reference in MaxwellWorld

Although there are numerous FORs, many can be classified as one of two types: exocentric or egocentric [McCormick, 1995; Wickens & Baker, 1995]. (See Figure 16.) In our MaxwellWorld study on FORs and perceptualization, the two concepts learners will be asked to master are 1) the distribution of force in electric fields and 2) the motion of test charges through electric fields. Comprehending distribution depends more heavily on global judgments than local judgments, while understanding motion requires more local judgments than global judgments. We will examine how the egocentric FOR, the exocentric FOR, and the bi-centric FOR (utilization of both frames of reference) shape mastery of these two types of material.

Mastery of scientific models can be assessed on two levels: descriptive and causal. If an individual can describe what he or she is examining and identify patterns in the data, his/her mastery can be described as “descriptive.” If that person can further interpret the meaning of the patterns and manipulate the information for problem solving, his/her mastery is “causal.” The latter reflects a deeper understanding of the information and is what we seek to accomplish in teaching learners about scientific models. We will examine both descriptive and causal performance on FOR learning tasks as a means of providing insights into the strengths and weaknesses of FORs.

Multisensory Cues and Perceptualization

Through a study of visual, auditory, and haptic (touch/pressure) sensory cues, we will extend our explorations on how multisensory immersion influences learning. Various sensory modalities can provide similar, mutually confirming input or can increase the amount of information conveyed to the learner through each sensory channel conveying different data. Little is known about what level of redundancy in sensory input is optimal for learning and about how much information learners can process without sensory overload. Moreover, each sense uniquely shapes the data it presents (e.g., perceived volume and directionality of sound is nonlinear, varies with the pitch of the input, and is idiosyncratic to each person). This poses complex considerations in deciding which sensory channel to use in presenting information to learners. Virtual reality provides a good research environment for exploring these design issues, as well as for exploring how multisensory immersion shapes collaborative learning.

Immersive Collaborative Learning as a Means of Enhancing Social Constructivism

As a near-term research initiative in our ScienceSpace worlds, we will investigate the effectiveness of collaborative learning situations in which three students in the same location rotate roles among (1) interacting with the world via the headmounted display, (2) serving as external guide, and (3) participating as a reflective observer. We also plan to experiment with collaborative learning among distributed learners inhabiting a shared virtual context. The student would act and collaborate not as himself or herself, but behind the mask of an "avatar": a surrogate persona in the virtual world. Loftin [1997] has already demonstrated the capability of two users simultaneously manipulating a shared immersive environment using communications bandwidth as low as a standard ISDN telephone line. By adapting military-developed distributed simulation technology, we may eventually scale up to many users in a shared, interactive virtual world.

Collaboration among learners' avatars in shared synthetic environments may support a wide range of pedagogical strategies (e.g., peer teaching, Vygotskian tutoring, apprenticeship). In addition, adding a social dimension aids in making technology-based educational applications more intriguing to those students most motivated to learn when intellectual content is contextualized in a social setting. However, in virtual environments, interpersonal dynamics provide leverage for learning activities in a manner rather different than typical face-to-face collaborative encounters [Dede, 1995]. We believe that our ScienceSpace worlds offer an intriguing context for extending work on social constructivism in virtual environments. Physical immersion and multisensory stimulation may intensify many of the psychological phenomena above, and "psychosocial saliency" may be an interesting counterpart to perceptual saliency in enhancing learning. Important questions to be answered include the relative value of providing learners with graphically-generated bodies and the degree to which the "fidelity" of this graphical representation affects learning and interaction (here fidelity is not simply visual fidelity, but also the matching of real body motions to the animation of the graphical body).

Lessons Learned to Date on Learnability and Usability in Virtual Reality

What can be generalized about model-based science learning from our research to date with immersive multisensory virtual environments? Based on lessons learned from all our ScienceSpace worlds, we are developing design heuristics, assessment methodologies, and insights, some of which are applicable to a range of educational modeling environments beyond virtual reality.

Learning and Knowledge Representation

Our goal is to develop an overarching theory of how learning difficult, abstract material can be strongly enhanced by scientific models instantiated via multisensory immersion and perceptualization. Illustrative themes applicable across all the virtual worlds we have created are:

- Multisensory cues can engage learners, direct their attention to important behaviors and relationships, prevent interaction errors through feedback cues, and enhance perceived ease of use.
- The introduction of new representations and perspectives can help students gain insights for remediating misconceptions formed through traditional instruction (e.g., many representations used by physicists are misleading for learners), as well as aiding learners in developing correct mental models. Our research indicates that qualitative representations

(e.g., shadows showing kinetic energy in NewtonWorld) can increase saliency for crucial features of both phenomena and traditional representations.

- Three-dimensional representations seem to aid learners in understanding phenomena that pervade physical space. Being immersed in a 3-D environment is also motivating for learners.
- Learner motivation is high in virtual reality environments, even when novelty effects wear off. The inclusion of interactivity; constructivist pedagogy; and challenge, curiosity, fantasy, and beauty [Malone & Lepper, 1984] all seem to augment students' interest and involvement.
- Initial experiences in working with students and teachers suggest collaborative learning may be achievable by having two or more students working together and taking turns "guiding the interaction," "recording observations," and "experiencing activities" in the virtual reality. Extending this to collaboration among multiple learners co-located in a shared synthetic environment may further augment learning outcomes, as may features (such as a "Hall of Fame") that provide social recognition for learner achievements.
- In addition to pre- and post-test assessments of learning, continuous evaluation of progress through lessons is critical to diagnosing the strengths and weaknesses of the virtual worlds. We have found talk-aloud protocols employing a cycle of prediction-observation-comparison are highly effective for monitoring usability and learning.

Based on these early results, we feel strongly encouraged on the potential utility of VR for facilitating certain types of scientific model-based learning more effectively than do any other pedagogical modalities.

Challenges in Current Virtual Reality Interfaces

We have identified the following usability issues characteristic of virtual reality interfaces:

- Limitations of the physical design and optics in today's head-mounted displays may cause discomfort for users. Since the visual display is an integral part of interaction and communication of information in these learning environments, these limitations are a current hindrance to usability and learning. Delays in VR system response time can also be a factor with complex environments. Both of these problems are steadily improving as hardware technology advances.
- Immersion does present some challenges for lesson administration (for example, students in the head-mounted display are not able to access written instructions or to complete written questions). We have found that verbal interaction works well.
- Students exhibit noticeable individual differences in their interaction styles, abilities to interact with the 3-D environment, and susceptibility to simulator sickness.
- To help learners utilize educational virtual worlds, calibrating the display and virtual controls for each individual is important. Additionally, monitoring and systematically measuring simulator sickness is vital, as malaise signals interface problems and also can explain why a student is having trouble with certain learning activities.
- Spreading lessons over multiple VR sessions appears to be more effective than covering many topics in a single session. For example, while students began to challenge their misconceptions during a single 3-hour NewtonWorld session, many had trouble synthesizing their learning during post-testing. We believe that factors such as fatigue and cognitive overhead in mastering the interface influenced these outcomes. In contrast, our MaxwellWorld evaluations were completed over multiple sessions, tackling fewer topics during each session and dedicating less time per session to pre- or post-testing. Reviews

and post-tests demonstrated that students were better able to retain and integrate information over multiple lessons.

In our judgment, none of these interface challenges precludes developing compelling learning experiences in virtual reality.

Implications of Our Work for Using Models in Science Education

The results of our research can inform larger debates within the science education community on best practice in using models and simulations to aid students in learning complex scientific concepts. Issues of active discussion among researchers studying the utility of models for learning science are listed below [Feurzeig, 1997]. After each topic, our beliefs are presented about what ScienceSpace research contributes to the issue's resolution.

The tension between computer-based modeling activities versus real-world observation and laboratory experimentation.

The debate: In interacting with a model, learners are manipulating a representation of reality, one that can simplify complex scientific concepts and their interrelationships. However, unless carefully designed, models can oversimplify reality in a manner that later makes deeper understanding of phenomena harder to attain. Still, models that go beyond simulation to allow learners to change underlying variables and relationships—to illustrate how the idealized phenomenon functions by altering it in ways not possible in reality—can enable a kind of meta-understanding not possible via real world experimentation. Yet real world phenomena are more "real" to learners: more believable, more fully sensory. On the other hand, some complex scientific concepts (e.g., relativity, quantum mechanics) involve intangible phenomena unobservable in the everyday macroscopic settings to which learners have access. For these types of content, models are the only means by which students form non-abstract impressions of these phenomena. Given these relative strengths and limits, what should be the pedagogical balance between interacting with models and experiencing reality itself?

Our contribution: Models based on multisensory immersion give learners experiences closer to the perceptual aspects of reality than any other simulation medium. Our research suggests that virtual reality is a potentially powerful means of bridging the gap between models and real world experimentation through combining strengths of each: the sensorial, immersive involvement of real world experiences , and the emphasis on crucial variables for understanding that models can provide (in our work, through perceptual saliency). In our research so far, we have not found that carefully designed "almost real" models induce new types of learner misconceptions. However, we do believe transitional learning experiences that gradually remove the affordances of models to reveal the full complexity and confusion of reality are important for generalizability and transferability of learning. The best pedagogical strategy may involve beginning with real world experiments to show the complexity and counter-intuitive nature of phenomena, then using models to simplify the situation and to enhance comprehension via interactive representations, and finally combining and extending the models to show how the complexity of real world behavior emerges from a multiplicity of simultaneous underlying causes.

The tension between modeling in science research versus modeling in science education.

The debate: This issue concerns the differences between modeling by experts and modeling by novices, in particular between the modeling tools used by scientists and those used by precollege students. Some researchers claim that, under the guidance of professionals, typical students (especially at the secondary school level) can learn scientific concepts by using the same models and supercomputing facilities used by research scientists. Others insist that all but the brightest high-school students need specially designed modeling tools and applications to introduce them to model-based inquiry.

Our contribution: In our design of representations for virtual reality, we have noted that part of the difficulty in mastering complex scientific concepts is the misleading representational formalisms and terminology that have emerged historically in science and now are entrenched as standard professional notation. Students come to us with misconceptions that appear to be linked to these traditional representations. We find that, despite our best efforts to compensate for the shortcomings of these formalisms, students sometimes remain confused about how to relate conventional representations to reality and how to use standard scientific terminology to convey their ideas. Two examples from electrostatics illustrate this point.

First, from their prior physics instruction, many of our learners in MaxwellWorld have initial misconceptions linked to the “field line” representation. For experts, field lines are a quick way of ascertaining the direction of a vector field along a series of points. However, novices understandably develop several intuitive misconceptions through analogical reasoning: field lines illustrate the path an untethored test charge would take through the field, the force does not vary from point to point along the field line, field lines can cross, etc. Additionally, learners often have difficulties relating field lines to another common representation of force: test charge traces. In MaxwellWorld, we attempt to overcome the shortcomings of the traditional field line representation by adding several enhancements. First, field lines are colored according to the strength of the force along them, helping students visualize how the force varies from point to point. Second, our “enhanced” field lines can be continuously manipulated in 3-D. By grabbing a point on a field line and moving it, students can see how characteristics of the field line (both the shape and the strength of the field along it) change from point to point, and they can verify that field lines will never cross. Finally, by releasing a test charge on a field line, learners can see that the test charge moves along the field line only when the line does not curve.

Second, another example of a problematic representation is the “equipotential surface,” which indicates a set of points across which a test charge's electric potential (or energy) would remain constant. In 2-D, this surface appears to be a line, creating difficulties for students in distinguishing equipotential surfaces from field lines. Further, the standard formalism for equipotential surfaces does not convey information about the magnitude of the surface’s potential. In addition, this representation does not aid students in relating the concepts of potential and force on the surface (this is also a problem with field lines). Consequently, students have trouble remembering which representation tells them about electric field (or force) and which tells them about electric potential (or energy). For example, we have observed a number of students describing field lines when asked to describe equipotential surfaces and vice versa. At a deeper level, students have trouble distinguishing the concept of electric field (or force) from electric potential (or energy). For example, when students are asked whether the force on a test charge would vary or remain constant as they move it along an equipotential surface in a complex field, they most commonly predict that it will be constant. We have enhanced the equipotential surfaces displayed in MaxwellWorld to attempt to compensate for these shortcomings of the standard formalism.

In general, these traditional scientific representations share one thing in common: they fail to make salient to the novice information that may be obvious to the expert. The missing data often is crucial in providing the foundation for understanding how these models represent reality. Our approach has been to enhance traditional representations, adding new information and affording an investigation of the interrelationships among them. However, we have sometimes found ourselves limited in the extent to which we can build on these conventional formalisms; and even our enhanced versions are subject to some of the same misinterpretations.

We believe that researchers in the modeling community need to investigate the strengths and limits of both enhanced and entirely unique representations that are less subject to misinterpretation than those scientific formalisms that have emerged historically, before the availability of visualization tools. As we have found in our work, new notational systems may enable students to learn the underlying scientific concepts more readily. Unfortunately, learning with models based on new representations does not intrinsically convey the standard formalisms used by scientists to represent concepts. Therefore, until the scientific community is willing to alter historic formalisms to alternative, equally accurate representations that enable easier comprehension by novices, our research suggests that many students will need specially designed modeling applications that focus on making salient otherwise "cognitively opaque" notational systems.

The tension between computer visualization of a model's output behavior versus computer visualization of a model's structure and component processes.

The debate: Computational modeling programs often employ visual representations of the model's behavior—animated displays of the outputs generated in the course of running the model. (Indeed, for many researchers, computational models are used primarily for obtaining visualizations of model behavior, and modeling is thought of as almost synonymous with visualization.) Typically, scientists conducting computational modeling research with sophisticated visualization facilities (e.g., at supercomputer centers) are content with programs that visualizing a model's output behavior ("data visualization"), but not its internal structure and component processes. Researchers disagree about whether this “output only” approach to visualization should be followed in science education.

Our contribution: As discussed earlier, mastery of scientific models can be assessed on two levels: descriptive and causal. Descriptive mastery indicates that an individual remembers representations and their behavioral interrelationships; causal mastery shows a deeper understanding about what these descriptive dynamics imply about the nature of reality. In our evaluations of multisensory immersion’s educational utility, we are careful to define causal mastery as the true goal and are not overly impressed when students exhibit descriptive mastery (even though the ability to describe a phenomenon’s dynamics is a richer type of learning than presentational instruction typically achieves). Based on our experiences with educational modeling, both inculcating causal understanding in students and measuring their attainment of this capability would be far more difficult with “output only” models than the structure-and-processes instructional design we use, which allows real-time manipulation of causal factors to observe secondary effects.

The tension between learning to use models versus learning to design and build models.

The debate: Beyond students learning from preconstructed models, researchers differ as to how much (and how) pupils can learn to design and build their own models. Some argue that, if students don't learn how to create models in classroom settings, how can one expect them to develop fluency at model building in workplace contexts? Further, constructivist learning theorists argue that students can comprehend much about model-based inquiry from engagement in the process of building models and simulations—indeed, that the process of designing and building models is an essential part of learning to use models as investigative tools and of understanding models’ strengths and limits as a means for representing reality.

Our contribution: For the very difficult scientific concepts on which our research is based, material that warrants the “sledgehammer” power of multisensory immersion to enhance learning, we find designing appropriate representations, interactive interfaces, and educational experiences is very challenging. That naive students could rapidly construct meaningful models of these complex phenomena is unlikely, however well designed the authoring tools they utilize. Our studies suggest that—at least for this type of counterintuitive, abstruse material—the use of preconfigured models for guided inquiry is much more efficient and probably equally effective compared to learners creating models from scratch.

Beyond these issues of current debate, we wish to raise a weakness of most current approaches to model-based science learning: the lab-like nature of the learner experiences. Controlled manipulations of a phenomenon, as in a scientific laboratory setting, are vital for understanding its nature—yet are unmotivating to learners unless they already are interested in science. Beginning with more playful and gamelike exploration is important for motivating most students, and ending with these types of activities probably also aids the transferability and generalizability of learning. At this point, our ScienceSpace worlds are as subject to this criticism as most other science-based educational models, yet we believe a major strength of multisensory immersion will be its capacity to support playful exploration in fantastical settings. As we evolve our worlds, we plan to incorporate activities that support game-like competitions; enable explorations of curiously configured, beautiful environments (for example, Mandelbrot spaces); and contextualize scientific phenomena within an "edutainment" context (e.g., MaxwellWorld-like field spaces within the worp engines in a StarTrek virtual environment). We believe this is an important and challenging next frontier for model-based instructional design: making these science learning environments more motivating and intriguing without weakening their educational value.

Conclusions drawn from an incipient set of studies on virtual reality as a modeling medium certainly do not provide definitive, generalizable answers about model-based instructional design. However, our experiences and research results provide a different perspective on the strengths and limits of model-based learning, and further studies to explore the potential power of multisensory immersion certainly seem indicated.

Conclusion

The virtual reality interface has the potential to complement existing approaches to model-based instruction about science. An overarching theme in our ScienceSpace research is to develop a theory of how multisensory immersion aids learning. In our virtual worlds, we can simultaneously provide learners with 3-D representations; multiple perspectives and frames-of-reference; a multimodal interface; simultaneous visual, auditory, and haptic feedback; and types of interaction unavailable in the real world (e.g., seeing through objects, flying like Superman, teleporting). With careful design, these capabilities all can synthesize to create a profound sense of motivation and concentration conducive to mastering complex, abstract material.

By themselves becoming part of phenomena, learners gain direct experiential intuitions about how the natural world operates. Instructional design can make those aspects of virtual environments that are useful in understanding scientific principles salient to learners' senses; and multisensory cues can heighten this saliency. Our experimental results indicate that transducing data and abstract concepts into mutually reinforcing multisensory representations is a valuable means of enhancing understanding of scientific models. Providing experiences that leverage human pattern recognition capabilities in three-dimensional space, such as shifting among various frames-of-reference (points of view), also extends the perceptual nature of a visualization. In addition, the social construction of knowledge among students immersed in a shared virtual environment may enable innovative, powerful types of collaborative learning.

Overall, we believe that these various aspects of multisensory immersion, when applied to scientific models, can provide learners with experiential metaphors and analogies that aid in understanding complex phenomena remote from their everyday experience and can help in displacing intuitive misconceptions with alternative, more accurate mental models. Studying this new type of learning experience to chart its strengths and its limits is an important frontier for cognitive science research, scientific modeling, and constructivist pedagogy.

Beyond its implications for model-based learning of science, we believe that our research illuminates larger issues related to students understanding complex information spaces. In every aspect of our knowledge-based society, fluency in utilizing complicated representations of information is an increasingly crucial skill. Comprehending models that include sophisticated interrelationships, such as non-linearities and feedback loops, is important not only for scientists, but also for workers and citizens. Such complex behaviors are typical of many crucial phenomena in modern civilization, and our well-being vitally depends on understanding the strengths and limits of the decision making models we create of those situations. Inculcating in students model-assessment skills such as sensitivity analysis is not simply a way of meeting discipline-based science standards as educational outcomes; these are survival skills necessary for our time, just as irrigation and planting skills were for agricultural economies. The next generation of educational standards will likely focus beyond knowledge of various isolated disciplines to integrated skills central to 21st century work and citizenship. Model-based learning has much to contribute in understanding how to conceptualize and achieve these next generation educational standards.

Further information, including Quicktime™ and Quicktime VR™ files for "viewing" the worlds we have developed, can be obtained from our website: http://www.virtual.gmu.edu.

References

Arthur, E.J., Hancock, P.A., & Chrysler, S.T. (1994). Spatial orientation in virtual worlds. In Proceedings of the 37th annual meeting of the human factors and ergonomics society (pp. 328-332). Santa Monica, CA: Human Factors Society.

Bricken, M. & Byrne, C. M. (1993). Summer students in virtual reality. In Wexelblat, A. (Ed.), Virtual Reality: Applications and Exploration (pp. 199-218). New York: Academic Press, Inc.

Chi, M.T.H., Feltovich, P.J., & Glaser, R. (1991). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, **5**, 121-152.

Clark, A.C. (1973). Technology and the Limits of Knowledge. In D. Boorstin, Technology and the Frontiers of Knowledge. New York: Doubleday.

Dede, C., Salzman, M., Loftin, B., & Ash, K. (in preparation). Using virtual reality technology to convey abstract scientific concepts. In M. Jacobson & R. Kozma (Eds.), Learning the sciences of the 21st century: Reseach, design, and implementation of advanced technological learning environments. Hillsdale, NJ: Lawrence Erlbaum.

Dede, C., Salzman, M., & Loftin, B. (1996). ScienceSpace: Virtual realities for learning complex and abstract scientific concepts. Proceedings of IEEE virtual reality annual international symposium 1996 (pp. 246-253). New York: IEEE Press.

Dede, C. (1995). The evolution of constructivist learning environments: Immersion in distributed, virtual worlds. *Educational Technology* **35** (5), 46-52.

Dede, C. and Lewis, M. (1995). Assessment of Emerging Educational Technologies That Might Assist and Enhance School-to-Work Transitions. Washington, DC: National Technical Information Service.

Ellis, S.R., Tharp, G.K., Grunwald, A.J., & Smith, S. (1991). Exocentric judgments in real environments and stereoscopic displays. In Proceedings of the 35th annual meeting of the human factors society (pp. 1442-1446). Santa Monica, CA: Human Factors Society.

Erickson, T. (1993). Artificial realities as data visualization environments. In Wexelblat, A. (Ed.), Virtual Reality: Applications and Explorations (pp. 1-22). New York: Academic Press Professional.

Feurzeig, W. (1997). Personal communication to the authors. Cambridge, MA: BBN Labs.

Fosnot, C. (1992). Constructing constructivism. In T.M. Duffy & D.H. Jonassen (eds.), Constructivism and the technology of instruction: A conversation (pp. 167-176). Hillsdale, NJ: Lawrence Erlbaum.

Gordin, D.N. & Pea, R.D. (1995). Prospects for scientific visualization as an educational technology. *The Journal of the Learning Sciences*, **4** (3), 249-279.

Halloun, I.A., & Hestenes, D. (1985a). Common sense concepts about motion. *American Journal of Physics*, **53**, 1056-1065.

Halloun, I.A., & Hestenes, D. (1985b). The initial knowledge state of college students. *American Journal of Physics* **53**, 1043-1055.

Heeter, C. (1992). Being there: The subjective experience of presence. *Presence: Teleoperators and virtual environments* **1** (1), 262-271.

Kalawsky, R.S. (1993). The Science of Virtual Reality and Virtual Environments. New York: Addison-Wesley Publishing Company, Inc.

Kennedy, R.S., Lane, N.E., Berbaum, K.S., & Lilienthal, M.G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology* **3** (3), 203-220.

Kohn, M. (1994). Is this the end of abstract thought? *New Scientist*, September 17, 37-39.

Krueger, M. (1991). Artificial Reality II. New York: Addison-Wesley Publishing Company, Inc.

Loftin, R.B. (1997). Hands Across the Atlantic. *IEEE Computer Graphics & Applications* **17** (2), 78-79.

Malone, T.W., & Lepper, M.R. (1984). Making learning fun: A taxonomy of intrinsic motivations for learning. In Snow, R.E., & Farr, M.J. (eds.) Aptitude, learning and instruction. Hillsdale, N.J.: Erlbaum.

McCormick, E.P. (1995). Virtual reality features of frames of reference and display dimensionality with stereopsis: Their effects on scientific visualization. Unpublished master's thesis, University of Illinois at Urbana-Champaign, Urbana, Illinois.

McDermott, L.C. (1991). Millikan lecture 1990: What we teach and what is learned - closing the gap. *American Journal of Physics*, **59**, 301-315.

Papert, S. (1988). The conservation of Piaget: The computer as grist for the constructivist mill. In G. Foreman & P.B. Pufall (eds.), Constructivism in the computer age. Hillsdale, NJ: Lawrence Erlbaum. 3-13.

Perkins, D. (1991). Technology meets constructivism: Do they make a marriage? *Educational Technology* **31**, 5 (May), 18-23.

Piantanida, T., Boman, D.K., & Gille, J. (1993) Human perceptual issues and virtual reality. *Virtual Reality Systems* **1** (1), 43-52.

Psotka, J. (1996). Immersive training systems: Virtual reality and education and training. *Instructional Science* **23** (5-6), 405-423.

Redish, E. (1993). The implications of cognitive studies for teaching physics. *American Journal of Physics*, **62** (9), 796-803.

Regian, J.W., Shebilske, W., & Monk, J. (1992). A preliminary empirical evaluation of virtual reality as a training tool for visual-spatial tasks. *Journal of Communication* **42**, 136-149.

Reif, F., & Larkin, J. (1991) Cognition in scientific and everyday domains: Comparison and learning implications. *Journal of Research in Science Teaching* **28**, 743-760.

Rieber, L.P. (1994). Visualization as an aid to problem-solving: Examples from history. In *Proceedings of Selected Research and Development Presentations at the 1994 National Convention of the Association for Educational Communications and Technology* (pp. 1018-1023). Washington, DC: AECT.

Salzman, M.C., Dede, C., & Loftin, R.B. (1995). Learner-centered design of sensorily immersive microworlds using a virtual reality interface. In J. Greer (Ed.), Proceedings of the 7th International Conference on Artificial Intelligence and Education (pp. 554-564). Charlottesville, Virginia: Association for the Advancement of Computers in Education.

Smith, R.B. (1987). Experiences with the alternate reality kit: an example of the tension between literalism and magic (pp. 324-333). In Proceedings of CHI+GI 1987. New York: Association for Computing Machinery.

Stuart, R., & Thomas, J.C. (1991). The implications of education in cyberspace. *Multimedia Review* **2**, 17-27.

Studt, T. (1995). Visualization revolution adding new scientific viewpoints. *R & D Computers & Software*, October, 14-16.

Trowbridge, D., & Sherwood, B. (1994). EMField. Raleigh, NC: Physics Academic Software.

Wenzel, E.M. (1992). Localization in virtual acoustic displays. *Presence* **1**(1), 80-107.

White, B. (1993). Thinkertools: Causal models, conceptual change, and science education. *Cognition and Instruction* **10**, 1-100.

Wickens C.D., & Baker, P. (1995). Cognitive issues in virtual reality. In W. Barfield & T. Furness (Eds.), Virtual environments and advanced interface design. New York: Oxford Press.

Wickens, C. (1992). Virtual Reality and education. *IEEE spectrum*, 842-47.

Witmer, B.B., & Singer, M.J. (1994). Measuring Presence in Virtual Environments (ARI Tech Report No. 1014). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Acknowledgments

This work is supported by NSF's Applications of Advanced Technology Program, Grants RED-93-53320 and 95-55682, and by NASA through a grant (NAG 9-713) and through access to equipment and computer software. The authors gratefully acknowledge the aid of Katy Ash, Brenda Bannan, Craig Calhoun, Jim Chen, Erik Geisler, Jeff Hoblit, Pat Hyde, Deirdre McGlynn, Edward (Joe) Redish, Saba Rofchaei, Chen Shui, and Susan Trickett.