Marilyn C. Salzman
mcsalzm@advtech.uswest.com
Human Factors and Applied
Cognitive Psychology Department
George Mason University
Fairfax, VA 22030

Chris Dede
cdede@gmu.edu
Graduate School of Education
George Mason University

R. Bowen Loftin
bloftin@uh.edu
Virtual Environment Technology Lab
University of Houston
Houston, TX 77023

Jim Chen
jchen@cs.gmu.edu
Computer Science Department
George Mason University

A Model for Understanding How Virtual Reality Aids Complex Conceptual Learning

Abstract

Designers and evaluators of immersive virtual reality systems have many ideas concerning how virtual reality can facilitate learning. However, we have little information concerning which of virtual reality's features provide the most leverage for enhancing understanding or how to customize those affordances for different learning environments. In part, this reflects the truly complex nature of learning. Features of a learning environment do not act in isolation; other factors such as the concepts or skills to be learned, individual characteristics, the learning experience, and the interaction experience all play a role in shaping the learning process and its outcomes. Through Project ScienceSpace, we have been trying to identify, use, and evaluate immersive virtual reality's affordances as a means to facilitate the mastery of complex, abstract concepts. In doing so, we are beginning to understand the interplay between virtual reality's features and other important factors in shaping the learning process and learning outcomes for this type of material. In this paper, we present a general model that describes how we think these factors work together and discuss some of the lessons we are learning about virtual reality's affordances in the context of this model for complex conceptual learning.

1 Introduction

Understanding how to use immersive virtual reality (VR) to support the learning of abstract concepts presents a substantial challenge for the designers and evaluators of this emerging technology. In every aspect of our knowledge-based society, fluency in understanding complex information spaces is an increasingly crucial skill (Dede & Lewis, 1995; West, 1991). 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, find patterns in information for problem solving, and 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.

Presence, *Vol. 8, No. 3, June 1999,* 293–316

Whether in industry, research, or academia, people trying to understand complex information 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. They may need to translate among frames of reference, to imagine the dynamics of a model over time, and to reason qualitatively about physical processes (McDermott, 1991; White, 1993). They must be able to synthesize this information to build generic and runnable mental models (e.g., Larkin, 1983; Redish, 1993), and these mental models need to incorporate invisible factors and abstractions.

Unfortunately, real-life metaphors upon which to build these mental models may not exist, making it difficult for people to envision abstract phenomena (Frederiksen & White, 1992; Reif & Larkin, 1991). For example, learning electrostatics or quantum mechanics involves understanding phenomena that behave in ways that are very remote from direct experience. Additionally, people's real-life experiences are confounded with invisible factors that distort or contradict the principles they need to master. For example, the force of friction unobtrusively distorts objects' behaviors according to Newton's laws of motion. Faced with these mentally challenging tasks, people of all ages and occupations struggle with abstractions. Their lack of real-life referents for intangible phenomena, coupled with an inability to reify (''perceptualize'') abstract models, is an important aspect of this problem.

Immersive VR may support the type of learning environments people need. In fact, many researchers (e.g., Psotka, 1996; Winn, 1993) believe immersive VR has potential as a learning environment. If they are properly designed, three-dimensional, multisensory virtual ''worlds'' might be able to aid users in comprehending abstract information spaces by enabling them to rely on their biologically innate ability to make sense of physical space and perceptual phenomena. Researchers have identified the following VR features as promising:

- **Three-dimensional immersion:** Users develop the subjective impression that they are participating in a ''world'' that is comprehensive and realistic enough to induce the willing suspension of disbelief (Heeter, 1992; Witmer & Singer, in press). Additionally, some research suggests that users are intrigued by interactions with well-designed immersive ''worlds,'' inducing them to spend more time and concentration on a task (Bricken & Byrne, 1993). By engaging users in learning activities, immersion may make important concepts and relationships more salient and memorable, helping users to build more accurate mental models. Also, inside a head-mounted display, the user's attention is focused on the virtual environment without the distractions present in many other types of educational environments.
- **Frames of reference (FORs):** Spatial metaphors can enhance the meaningfulness of data and provide qualitative insights (Erickson, 1993). Psychological research on spatial learning, navigation, and visualization suggests that perspectives, or frames of reference, make salient different aspects of an environment and influence what people learn (e.g., Barfield, Rosenberg, & Furness, 1995; Ellis, Tharp, Grunwald, & Smith, 1991; Darken & Sibert, 1995; Presson, DeLange & Hazelrigg, 1989; Thorndike & Hayes-Roth, 1982). In virtual environments, 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. Enabling users to interact with spatial representations from various FORs may improve performance (McCormick, 1995; Wickens & Baker, 1995) and deepen learning by providing different and complementary insights.
- **Multisensory cues:** Via high-end VR interfaces, users can interpret visual, auditory, and haptic cues to gather information while using their proprioceptive system to navigate and control objects in the synthetic environment. This potentially deepens learning and recall (Nugent, 1982; Psotka, 1996).

Unfortunately, although researchers have many ideas concerning how VR might facilitate the understanding of complex concepts, the field has little information concerning which of virtual reality's features provide the most leverage for enhancing understanding or how to customize those affordances for different learning environments. In part, this reflects the complex nature of learning. Features of a learning environment do not act in isolation; other factors such as the concepts to be learned, individual characteristics, the learning experience, and the interaction experience all play a role in shaping the learning process and learning outcomes.

Through Project ScienceSpace, we have been trying to identify, use, and evaluate immersive virtual reality's features as tools to facilitate the learning of abstract concepts. In doing so, we have struggled to understand the interplay between VR's affordances and other important factors in shaping learning (both the learning process and learning outcomes). In this paper, we present a general model describing how we think these factors work together, explain how the model has guided our work, describe how our work has helped us to refine the model, and discuss some of the lessons we are learning about VR's affordances in the context of this model. Our goal is twofold. First, we want to provide designers and evaluators of VR learning environments a framework for thinking about the relationship between VR's features and learning. Second, we would like to share with you some of the insights we are gaining by applying various aspects of this model to our design and evaluation strategies.

2 The Model and How It Has Guided Our Research

Figure 1 shows a model describing how we believe VR's affordances work with other factors in shaping the learning process and learning outcomes. This is an initial model in which we tried to capture important factors and relationships. It lacks details about the specifics of key interrelationships (e.g., how a feature influences mastery or how a feature influences usability), because they depend on the specific feature being examined. Nevertheless, the model highlights important issues and is useful when generating research questions. The investigation of such questions has helped us refine and extend the model in general and with regard to specific features.

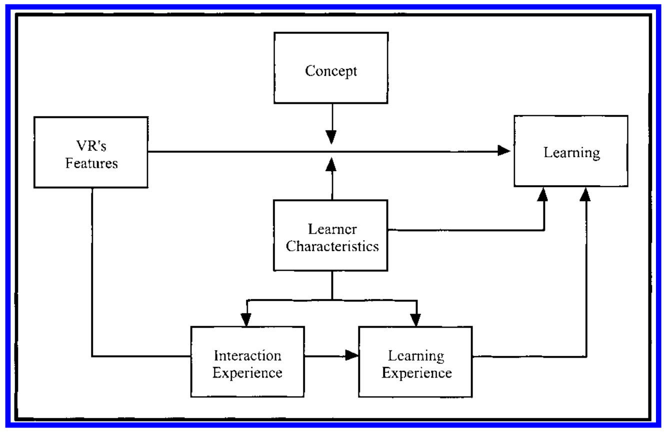


Figure 1. ***A general model describing how VR's features, the concept one is being asked to learn, learner characteristics, and the interaction and learning experiences work together to influence learning in VR learning environments.***

What does the model suggest? First, VR's features are likely to influence learning: both the learning process (or the kinds of information to which one attends) and learning outcomes (or the person's level of understanding after the lessons have been completed). Additionally, the concept one is trying to understand is likely to moderate how VR's capabilities influence the learning. In other words, the relative effectiveness of 3-D or multisensory representations may depend on the concept being learned. Individual characteristics of the learner, or learner characteristics (e.g., domain knowledge), should play a role in shaping the learning process and may also interact with VR's features in influencing learning. For example, the extent to which a feature supports learning may vary as a function of the learner's domain experience, spatial ability, or learning style. Finally, it is likely that VR's affordances, as well as individual characteristics, affect both the interaction experience (e.g., how easily the user can interact with the system) and the learning experience (e.g., motivation, perceived meaningfulness), which, in turn, influence learning.

As the model demonstrates, we believe the link between VR's affordances and learning occurs within a web of other relationships. This assumption has guided our research from the beginning. Early on, our emphasis focused on how VR's capabilities affected the interaction experience and learning. As our research progressed, we began to study the other factors identified in the model, to accommodate these factors in our designs, and to identify ways to capture these factors during the evaluation process. Research outcomes also helped us to identify which of VR's features have promise, which characteristics of the learner require careful attention, and which facets of the interaction and learning experiences play a substantial role in shaping learning outcomes. Thus, the model has evolved over time so that it specifies interrelationships in more detail and so that it is testable statistically. The model also has influenced both our design and evaluation activities.

2.1 Design

The model has helped guide our development processes. We use a combination of domain expertise, educational research, and research with students (e.g., questionnaires and focus groups) to identify the concepts we want to teach and to understand learner needs in those domains. Once the concepts are identified, we determine which of the concepts are most suitable for teaching through a VR environment, and which of VR's features are likely to facilitate the learning of those concepts. (With an understanding of learner needs comes a concomitant comprehension of the typical student and the range of learner characteristics in our target audience.) This information forms the foundation for our design. During the design process, our primary goal is to support the learning process given VR's capabilities and learners' needs; however, we also must attend to both the learning and interaction experiences in order to ensure that we are facilitating the learning process. We design to optimize the learning experience by sensorily immersing the student in an interactive learning process and relying on those capabilities of VR that we identified as promising for each of the concepts. We also design to optimize the interaction experience by minimizing interface problems.

2.2 Evaluation

The model has also influenced our evaluations. In our assessments of learning outcomes, our primary objective is to understand the link between VR's features and learning. In order to more fully understand this relationship, we also examine the other factors. First, we recognize that each learner is an individual; therefore, we collect background information that enables us to understand whether and how the learner's characteristics affect the learning process, as well as the learning and interaction experiences. Second, we design the learning assessments so that we can identify the strengths and weaknesses of VR's affordances. Third, we carefully monitor the learning process by asking students to make verbal predictions about an upcoming activity, observe the actual outcomes of the activity, and verbally compare their predictions to their observations. This cycle of predicting-observing-comparing helps us characterize the interaction between the learner and the VR learning environment. Finally, we try to measure relevant aspects of the learning and interaction experiences through questionnaires, student comments, and administrator observations.

3 Design of ScienceSpace's Learning Environments

ScienceSpace consists of three learning environments—NewtonWorld (NW), MaxwellWorld (MW), and PaulingWorld (PW). Central to the design of ScienceSpace's learning environments are the three features afforded by VR technology: immersive 3-D representations, multiple FORs, and multisensory cues. We believe each of these can offer unique advantages in the learning of abstract information. How these features are used in each of the worlds is driven by a detailed assessment of the concepts we are trying to convey, as well as the kinds of support learners need in order to master those concepts. In this section, we begin with a general discussion of ScienceSpace's design. Then, we discuss each of the learning environments in more detail, describing how the concepts to be taught and learner needs have driven our design choices regarding the usage of VR's capabilities.

ScienceSpace's physical interface is typical of current high-end immersive VR. Its hardware architecture includes a Silicon Graphics Onyx Reality Engine2 four-processor graphics workstation with a Multichannel Option Board, a Silicon Graphics Indy workstation with an Iris audio processor, a Crystal River Engineering Acoustetron II, a Polhemus 3Space Fastrak System (with a 3Ball, three trackers, and a magnetic tracking device), a Virtual Research VR4 head-mounted display (HMD), and an Aura Interactor Virtual Reality Game Wear Vest. The Onyx is used for processing and rendering the stereoscopic images that are sent to the HMD. The Indy is used for playing auditory feedback cues (e.g., menu selection, object placement, etc.). The Acoustetron is used to generate 3-D auditory informational cues (e.g., the direction and strength of the force). Both information and feedback sounds are delivered via HMD headphones or external speakers. Auditory cues can also be translated into vibrations and delivered to the user's torso using the vest. The remaining equipment (the Fastrak system and HMD) enables the user to interact with the learning environment. On his or her head, the user wears the HMD. In one hand, the user holds the 3Ball, which is represented in the virtual environment as a virtual hand. In the other hand, he or she holds a device for displacing the interface menu, which is represented in the virtual environment as a hand holding a menu system. Via trackers attached to these devices, the Polhemus Fastrak system monitors the location of the HMD, the 3Ball, and the menu-display device.

The software interface relies on 3-D models and qualitative representations controlled through NASA-developed physical simulation software. The models are built using polygonal geometry. Colors, 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 environments and connects them to underlying physical simulations. Interactivity is achieved through the linkage of external devices (e.g., an HMD) using this same software. Finally, graphics rendering, collision detection, and lighting models also are provided by NASA-developed software.

These hardware and software interfaces enable us to immerse students in 3-D learning environments using the visual, auditory, and haptic senses. Students control where they are looking by turning their heads. They use the virtual hand (controlled by the 3Ball), menus, and direct manipulation to perform tasks in these immersive virtual environments. By attaching the menu to one of the user's hands, we allow him or her to remove the menu from his or her field of view, while keeping it immediately accessible.[1]

Users 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. Thus, menu selection in ScienceSpace's learning environments is similar to menu selection on two-dimensional interfaces in which users manipulate the options with a cursor controlled by a mouse. Figure 2 shows a user manipulating the 3Ball and tracker to control a virtual hand and menu system in MaxwellWorld; Figure 3 shows what she sees. Direct manipulation in ScienceSpace's learning environments enables students to interact directly with objects in the space. For example, MW enables users to place source charges in a 3-D space, move them around, and delete them; NW enables users to beam among cameras located in frames of reference and to throw and catch balls; PW enables users to directly grab and rotate molecular structures. Figure 4 shows an example of direct manipulation in MW. In all worlds, 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 to move in that direction.

3.1 NewtonWorld's Design

NewtonWorld (NW) addresses many of the typical misconceptions students carry with them as they enter—and leave—physics courses (Halloun & Hestenes, 1985a). Clement (1982) refers to such misconceptions as "conceptual primitives," and these reflect erroneous generalizations from personal experience about the nature of mass, acceleration, and momentum, as well as Newton's laws and the laws of conservation. Conceptual

1. In the original version of NewtonWorld, the menu was attached to the HMD; however, as we discuss later in the article, this interfered with interaction and learning. We adopted the latter menu interface to overcome the shortcomings of our early designs.

Figure 2. *A student immersed in ScienceSpace's MaxwellWorld.*

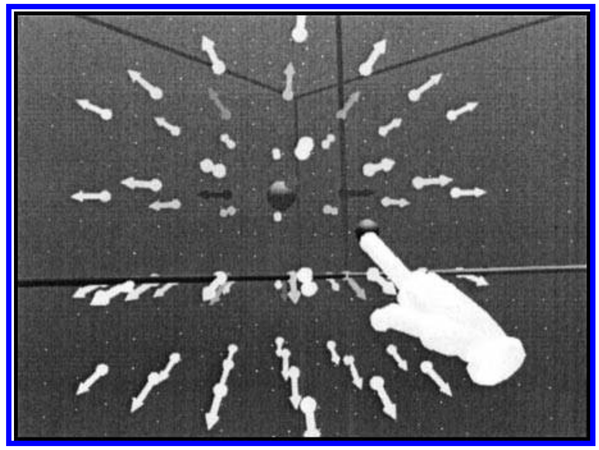


Figure 4. *Direct manipulation of a source charge in MaxwellWorld.*

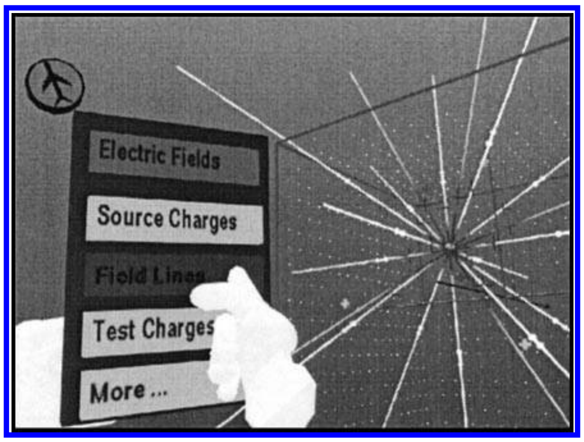


Figure 3. *The menu in ScienceSpace's MaxwellWorld.*

primitives form mental constructs, the understanding of which is a basic prerequisite for many higher-order concepts. Among common misconceptions about motion documented by Halloun & Hestenes (1985b) are the ''motion implies force'' notion, the ''impetus'' theory (an object's past motion influences the forces presently acting on it), and ''position-speed confusion'' (i.e., ahead = faster). Not only are these misconceptions strongly held by students entering physics courses, but they are very difficult to change with conventional approaches to instruction. Reinforced by their own real-world experiences, learners persist in believing that motion requires force (rather than that a change in motion requires force), that constant force produces constant velocity (rather than producing constant acceleration), and that objects have intrinsic impetus (rather than moving based on instantaneous forces). Thus, making these factors and their relationships salient is critical to the teaching of Newton's laws and the laws of conservation.

In NW, we rely on sensorial immersion to enhance the saliency of important factors and relationships and to provide experiential referents against which learners can compare their intuitions. Learners can be ''inside'' moving objects; this 3-D, egocentric frame of reference centers attention on velocity as a variable. Multisensory cues are used to further heighten the saliency of crucial factors such as force, energy, and velocity. Students begin their guided inquiry in a virtual reality in which gravity and frictional forces are set to zero, allowing observation of Newton's three laws operating without other superimposed phenomena clouding their perceived effects. Studying the collision of objects also enables the introduction of other scientific principles, such as conservation of momentum and of energy and reversible conversions between kinetic and potential energy.

In NW, students are immersed in a 3-D activity area that contains an open ''corridor'' created by colonnades on each side and a wall at each end. (See Figure 5.) Users 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 di-

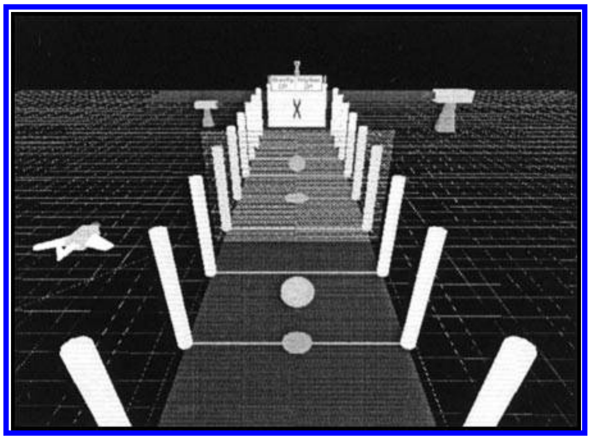


Figure 5. ***Above the corridor, showing cameras, balls with shadows, and the far wall.***

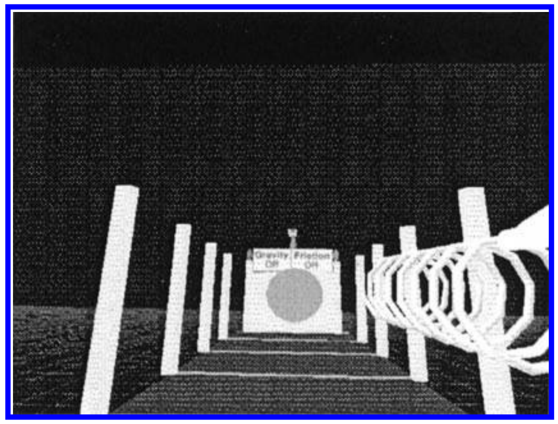


Figure 6. ***After launch, illustrating the spring-based launching mechanism.***

mension 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.

We use multisensory cues to direct user attention to important variables such as mass, velocity, and energy. For example, we use tactile and visual cues to represent potential energy, and auditory and visual cues to represent velocity. 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 and potential energy becomes kinetic energy, the spring uncoils and the energy vibrations cease. (See Figure 6.) 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 users in judging the velocities of the balls relative to one another, the columns light and chime as the balls pass.

Additionally, we provide multiple frames of reference 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 reference frame of the bouncing balls, a movable camera hovering above the corridor, etc. These features aid learners in understanding the scientific models underlying Newton's three laws, potential and kinetic energy, and conservation of momentum and energy.

3.2 MaxwellWorld's Design

MaxwellWorld (MW) is designed to help learners understand the difficult concepts underlying electrostatics: electric field (force) and electric potential (energy). Electrostatics concepts are three-dimensional, abstract, and have few observable real-life metaphors. Our early work with students and with our domain expert, Dr. Edward Redish of the University of Maryland, indicated that learners have trouble visualizing these 3-D phenomena. They also confuse the concepts of forces and energy, demonstrating that they do not understand the true meaning of the representations that are traditionally used (2-D field lines, 2-D equipotential lines, etc.) to convey information about these concepts. In addition, learners have trouble understanding how the electric field would propel an imaginary charged particle (a test charge) through the field. This may be because they lack the ability to visualize the distribution of forces throughout a vector field and to relate how that distribution of force translates into the motion of the test charge. This is an example of an instance in which learners lack a real-life referent to which they can anchor scientific phenomenon. They also lack an interactive, sensorial environment in which to test their ideas.

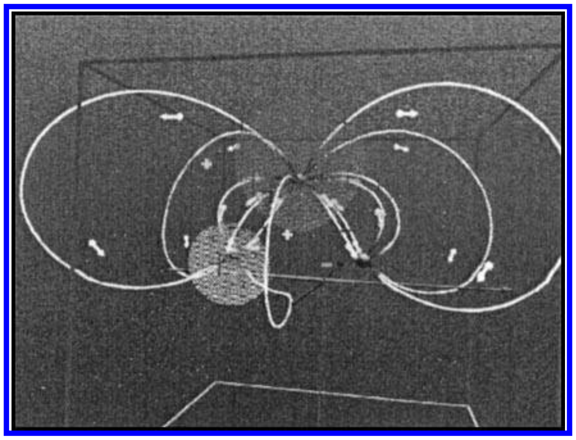


Figure 7. *Electric field lines, test charge traces, and equipotential surfaces describing the electric field and electric potential for an arrangement of three source charges.*

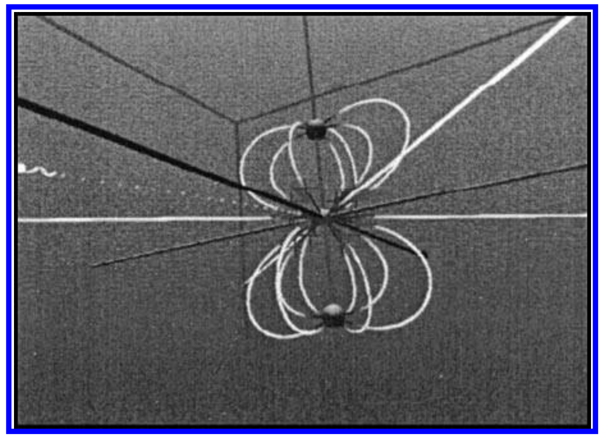


Figure 8. *A test charge being propelled by the electric field.*

For these reasons, we created a 3-D immersive learning environment in which users can explore electrostatic forces and fields, learn about the concept of electric potential, explore how test charges would move through the space, and ''discover'' the nature of electric flux. We created 3-D qualitative representations to emphasize the relationship between force and energy, as well as enhancing visual representations with multisensory cues to increase the saliency of force and energy as crucial variables. We allowed users to change between frames of reference by navigating around, scaling the world, or even becoming a tiny charged particle within the world. When making design decisions, we worked to minimize usability problems and simulator sickness problems that would degrade the interaction experience. We also wanted students to feel immersed in and motivated by their learning experiences.

In MaxwellWorld, learners work within a ''fieldspace,'' where they can create and explore electric fields. The fieldspace occupies a cube approximately one meter on a side, with Cartesian axes for convenient reference. The small size of the world produces large parallax when viewed from nearby, making its 3-D nature quite apparent. Students can place both positive and negative charges of various relative magnitudes into the fieldspace. Once a charge configuration is established, learners can instantiate, observe, and interactively explore the space using 3-D representations of the force on a positive test charge, potential (energy) of a positive test charge, electric field lines, equipotential surfaces, and lines of electric flux through surfaces. Figure 7 illustrates some of these representations. 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 traces 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. (See Figure 8.) Alternatively, users can ''become'' a test charge and explore the field or be propelled through the field by the electric forces.

An electric field line can also be attached to the virtual hand. Learners can then move their hands 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. 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 potential—or the energy—a test charge would have at any point 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. By default, the sur-

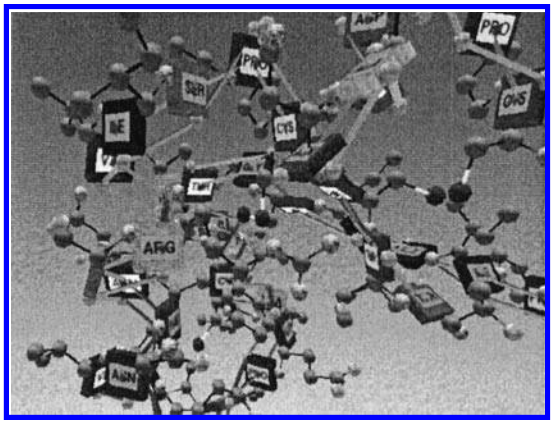


Figure 9. *Ball-and-stick with some amino acids.*

Figure 10. *Space-filling model.*

faces 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 learners 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 the representation of many aspects of the complex scientific models that underlie vector field phenomena.

3.3 PaulingWorld's Design

PaulingWorld (PW) is still under development and has not yet undergone formative evaluation. Currently, 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 9 and 10 for examples of these models.) We are working on extending PW to address concepts underlying quantum-mechanical bonding—another phenomenon that has no observable everyday referents, is difficult to represent, and is hard for students to understand via common-sense intuitions. These concepts include probability density and wave functions, ionic versus covalent bonding, and determinants of bond angles and bond length. As with the other ScienceSpace learning environments, we are exploring how we can use VR's affordances to support learner needs. 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 (www.qsad.bu.edu).

4 Evaluations of ScienceSpace's Learning Environments

4.1 NewtonWorld Evaluations

Our earliest evaluations were conducted on NewtonWorld (NW). Because our design and evaluation process is iterative, NW evolved slightly from evaluation to evaluation. Therefore, when we talk about the implications of each evaluation, we briefly describe how it impacted NW's design as well as how it expanded our knowledge regarding the model of how VR can enhance learning.

4.1.1 NW's Interaction Experience. In our first evaluation, we focused on the interaction experience; however, outcomes also shed light on issues relating to the learning experience, learning process, and outcomes, as well as a potential trade-off between designing for interaction versus designing for learning. This early version of NW contained no sound or tactile cues and no visual cues representing energy or velocity. Additionally, this version provided only two points of reference: the ball and a movable camera. Also, the menubar for accessing menu items was displayed at all times in the upper-right field of view in the HMD, as opposed to our later placement of the menu on the user's hand.

Nine high-school students (five female and four male) participated in this study; two of these students served as pilot subjects. Participants had a range of science, computer, and video experience to ensure that our sample was representative. Using each of four variations of the user interface (menu based, gesture based, voice based, and multimodal), participants performed a series of "typical" and "critical" activities, thinking aloud as they performed them. Participants conducted activities such as becoming a ball, using the menus, selecting masses of the balls they were to launch (throw), launching balls, catching balls, and changing camera views. To characterize the interaction experience, we recorded task strategies, task completion, error frequency, and student comments as participants attempted each of the tasks. Following the sessions, we also asked participants to rate the ease of use of various aspects of the interaction, to rank interaction alternatives, and to list what they liked and disliked about the system. Below is a summary of the lessons we learned from this evaluation:

- **VR's features and learning (process and outcomes).** Student comments suggested that the ability to observe phenomena from multiple viewpoints was crucial to understanding and that even more visual, auditory, or tactile cues could be used to help them focus on important information.
- **Learning experience.** Student comments suggested that the ability to observe phenomena from multiple viewpoints (FORs) was motivating.
- **Interaction experience.** Participants were comfortable with the bouncing-ball metaphor, liked the virtual hand, and intuitively understood that it enabled them to interact with objects in the world. Seven of nine students (77%) ranked the multimodal interface (mean ranking = 1.42) above the others, and eight students (89%) used one or more of the interaction alternatives (voice, gestures, and menus) available to them while using the multimodal interface. Of the interaction alternatives, voice was the most preferred, most used, and most error-free method of interaction. Menus and gestures were less preferred and more error prone. However, students appreciated the value of the menus for supporting a more flexible interaction than gestures. When given a multimodal interface, students used voice most frequently (for 56.7% of the tasks), then menus (for 33% of the tasks), and finally gestures (for 9.9% of the tasks).

  When asked what would help smooth the interaction, students indicated that additional visual, auditory, or tactile cues would have helped. Finally, all students experienced some symptoms of simulator sickness (oculomotor discomfort in particular) after wearing the HMD for approximately 1.25 hours (even with one to two breaks during that period). The simulator sickness score (and standard deviation) on Kennedy and Fowlkes' (1992) Simulator Sickness Questionnaire (SSQ) was 19.40 (11.15).
- **Interaction experience versus learning.** Several students interpreted the size of the ball as a cue for mass. If we focus on the interaction from a purely usability perspective, this might suggest size as an indicator of mass. However, from an educational perspective, this is problematic, because it reinforces the misconception that larger objects are more massive. Using color to distinguish the balls—and labels or color intensity as cues for mass—meets both usability and learning criteria.

This study yielded certain insights into our model. First, it indicated that flexible FORs increased the motivation of the learning experience. By doing so, FORs were likely to facilitate learning in a way we had not anticipated (the FORs' effects being mediated through motivation). (During the design, we had focused on how FORs might affect what students observe, but not how they felt about their experiences.) Second, it suggested that the usability of the interface and simulator sickness were important aspects of the interaction experience. Finally, it highlighted the fact that designing for learning and designing for interaction are not necessarily congruous tasks. Although we would expect interaction to facilitate learning in many cases, the last finding listed above demonstrates that maximizing certain aspects of the interaction can, at times, hinder learning. To summarize, this study suggested that we needed to pay careful attention to the relationship among VR's features, the learning experience, and learning outcomes, and that we needed to look for potential interactions between different dimensions (e.g., usability and simulator sickness) of the interaction experience.

The study's outcomes also helped us to enhance NewtonWorld. We maintained the ball metaphor and the general nature of the activities. We expanded the possible viewpoints from two to five and implemented the more flexible "beaming" method for moving among views. We also implemented sound cues to supplement visual cues and eliminated gesture-based commands.

4.1.2 NW's Potential. In our second evaluation, we surveyed 107 physics educators and researchers who used NewtonWorld at a meeting of the American Association of Physics Teachers. The goal of this evaluation was to gather information from experts in the field regarding the interaction experience, learning experience, learning process and expected learning outcomes, as well as to identify any potential trade-offs among these factors in NewtonWorld. Participants observed a ten-minute demonstration of NewtonWorld via a computer monitor, then received a personal demonstration while immersed in the virtual learning environment. After the demonstration, they completed a survey that focused on their interactive experiences, recommendations for improving the system, and perceptions of how effective this 3-D learning environment would be for demonstrating Newtonian physics and conservation laws. Below is a summary of survey outcomes:

- **VR's features and learning (process and outcomes).** On average, participants felt that NewtonWorld would be an effective tool for demonstrating Newtonian physics and dynamics, giving it an average rating of 1.5 on a scale from -3 (ineffective) to $+3$ (effective). The majority of participants were enthusiastic about the 3-D and immersive nature of this learning environment and appreciated the ability to observe phenomena from a variety of viewpoints (FORs). However, several participants expressed concerns regarding the limitations of the prototype and encouraged expanding the activities, environmental controls, and multisensory cues provided.
- **Interaction experience.** On average, survey participants found the basic activities easy to perform, giving it an average rating of 1.6 on a scale from -3 (difficult) to $+3$ (easy). Like participants in the first evaluation, many survey participants experienced difficulty using the menus and focusing on the HMD optics.
- **Learning experience versus interaction experience.** Several participants felt a broader field of view would have improved their learning experiences; however, some reported slight symptoms of simulator sickness (oculomotor discomfort in particular). Thus, identifying an appropriate solution to this problem was difficult because increasing the field of view could result in interaction problems due to eyestrain and nausea.

This survey yielded additional insights into our model. First, it suggested that teachers and disciplinary experts felt that three of VR's features had promise for learning the concepts addressed in NewtonWorld: 3-D immersion, FORs, and multisensory cues. The teachers' feedback also indicated that we needed to do more with these features in our learning environments. This survey also made it clear that usability and simulator sickness required careful attention in our designs and that we'd sometimes face trade-offs between the quality of the learning experience and interaction experience.

Outcomes of the study also helped us continue iterating on NewtonWorld's design. Participant feedback indicated that, while we had improved upon the version of NewtonWorld tested in the first evaluation, we needed to more fully utilize the multisensory feature of VR. Therefore, we expanded the interface to include a haptic vest and more-extensive visual and sound cues. We also refined the menus to make selecting their items easier. Finally, because the menus were not used during the observation portion of activities, we changed the menubar to a small 3-Ball menu icon, increasing the visual field of view and improving users' abilities to experience motion and immersion and to see important visual cues.

4.1.3 NW's Features and Learning. Our next evaluation focused on how multisensory cues and frames of reference affected the interaction experience, learning process, and learning outcomes in NewtonWorld. Thirty high-school students (eighteen males and twelve females) participated in this evaluation. Students ranged from sixteen to eighteen years old and had completed at least one year of high-school physics. Pre-tests indicated that the majority of students confused velocity and acceleration and had trouble inferring object behaviors after collisions.

Students were assigned to one of three groups: (1) visual cues only, (2) visual and auditory cues, or (3) visual, auditory, and haptic cues. Groups were differentiated by controlling the visual, tactile, and auditory cues that students received while performing learning tasks. Students in each group participated in an individual learning session that lasted 2.5 to 3 hours. During the session, students completed a lesson in NewtonWorld that consisted of a series of learning tasks that focused on relationships among force, mass, velocity, momentum, acceleration, and energy during and between collisions. The lessons required between 1 and 1.25 hours to complete. For each learning task, students first predicted what the relationships or behaviors would be, then they experienced the behaviors, and finally the students assessed their predictions based on what they had experienced.

To help us assess the value of multisensory cues for learning, we relied on the students' comments during the learning process (predictions, observations, and comparisons), administrator observations during the lessons, usability questionnaires, interview feedback, and pre- and post-test knowledge assessments. To assess the value of the FORs, we examined the learning process, usability questionnaires, and interview feedback. Below is a summary of the outcomes of these evaluations:

- **VR's features and learning (process and outcomes).** Overall, students did not demonstrate significant learning from pre- to post-lesson knowledge assessments, and no significant differences were found among groups, suggesting that single-session usage of NW was not enough to transform users' mental models. (Post-lesson scores were as follows (out of 100): visual = 61.4; visual and auditory = 58.0; visual, tactile, and auditory = 65.45.) Although no differences were found in overall performance, there were some nonsignificant trends on different portions of the tests. Students who received haptic cues in addition to sound and visual cues performed slightly better than students in other groups on questions relating to velocity and acceleration. Additionally, lesson administrators observed that students receiving haptic and sound cues were more attentive to these factors than students without these cues. However, those same students performed slightly worse on predicting the behavior of the system. One possible explanation is that haptic cues may have caused students to attend more to factors at play just before, during, and after collisions—and less to the motions of the balls.
- **Learner characteristics and learning.** Performance on the pre-lesson test was highly predictive of post-lesson performance. Scores on the pre-lesson test explained 37% ($R^2 = 0.37$, $F = 15.35$, $p < 0.05$) of the variance in post-lesson scores. This indicates the importance of measuring relevant learner characteristics when trying to understand learning outcomes.
- **Learning experience.** Most students found the activities interesting and enjoyed their learning experiences. Additionally, many users stated that they felt NW provided a good way to explore physics concepts. When asked to list the features they liked most, almost all participants cited NW's FORs and multisensory cues. They liked the ability to beam to various cameras and to navigate in the movable cam-

era. They felt that the multisensory informational cues used to represent velocity, energy, and collisions added to the learning experience. Students suggested that we could improve the learning experience by expanding the features and representations used in NewtonWorld, as well as by adding more variety to the nature of the learning activities.

- **VR's features and the learning experience.** Participants who received multisensory cues (auditory and haptic) appeared to be more engaged in activities during the sessions than students receiving only visual cues. These participants also rated the egocentric reference frame as more meaningful. (Ratings were as follows on a scale from -3 to $+3$: visual $= 0.33$; visual and auditory $= 0.60$; visual, tactile, and auditory $= 1.00$.)
- **Interaction experience.** Overall, students felt the environment was fairly easy to use. The mean rating (and standard deviation) for NW's overall usability was 1.67 (1.42) on a scale of -3 (very difficult) to $+3$ (very easy). Additional ratings and performance indicated the major sources of difficulty were the menus and navigation. As in earlier tests, several students experienced symptoms of simulator sickness, including oculomotor discomfort, nausea, and disorientation. The mean score (and standard deviation) on SSQ was 17.69 (12.69). At times, these usability and simulator sickness problems appeared to distract users from the learning activities and to contribute to fatigue.
- **Interaction experience and learning.** Although neither usability nor simulator sickness significantly predicted post-lesson performance beyond pre-lesson test scores, usability had a positive relationship and simulator sickness had a negative relationship with post-lesson scores. Thus, negative aspects of the interaction experience may have interfered slightly with learning.
- **VR's features and the interaction experience.** Students receiving sound or sound plus haptic cues rated NewtonWorld as easier to use than participants who received visual cues only. (Ratings were as follows on a scale from -3 to $+3$: visual $= 0.56$; visual and auditory $= 2.10$; visual, tactile, and auditory $= 1.82$.)
- **Learner characteristics and the interaction experience.** We observed substantial individual variability in students' abilities to interact in NW, and their subjective ratings of its usability were highly variable, particularly for ratings concerning the menus and navigation. Ratings for the usability of the menus ranged from -2 to $+3$; ratings for navigation ranged from -3 to $+3$. Additionally, students were highly variable in their susceptibility to symptoms of simulator sickness. Total simulator sickness scores on the SSQ ranged from 1 to 55.40.

This study yielded additional insights into our model. Although we were unable to detect significant differences in overall learning outcomes, outcomes of this study did help us understand the role of VR's features (multisensory cues and FORs) in shaping learning. An examination of participant feedback on their interaction and learning experiences, as well as examining the learning process, yielded substantial evidence to suggest that the multisensory cues and FORs enhanced the learning experience and affected learning. In this study, the multisensory cues directed learners' attention toward important relationships, but, in doing so, they may also have prevented them from noticing others. This experiment also suggested that using features in combination could have potentially powerful effects on the learning experience, the learning process, and learning outcomes. For example, we found that the egocentric FOR (being the ball) was more meaningful to students receiving haptic and auditory cues than it was for students receiving only visual cues. This emphasizes the importance of considering and testing how these VR's affordances work in combination and how they affect attention. Finally, this study highlighted the crucial factor that the characteristics of each learner plays in assessing the strengths and weaknesses of our environments.

In addition to helping us elaborate upon our model of how VR enhances learning, this study guided us in refining NW's interface and learning activities. We moved the menu from its fixed location in the HMD's field of view to the user's second virtual hand, allowing users to freely adjust menu position and to judge menu location based on the physical position of their own hands. We also reconceptualized NewtonWorld to focus the emphasis of educational activities on the relationships among mass, velocity, and momentum for colliding objects. Additionally, we carefully considered ways to use the multisensory cues to increase the saliency of these factors and to appropriately direct the students' attention. Finally, we determined to target slightly younger students with our new design. Analysis of our data, as well as the educational literature, suggested that younger users might gain more from virtual experiences in sensorily immersive Newtonian environments than high-school students. Via VR experiences, early interventions that undercut Aristotelian mental models just at the time when young learners are developing these misconceptions might become a foundation for a less difficult, accelerated transition to a Newtonian paradigm. See Dede, Salzman, Loftin, & Ash (in press) for additional details on NewtonWorld's redesign.

4.2 MaxwellWorld

We have completed two evaluations of MaxwellWorld (MW) and are currently conducting a third. Because of our iterative design and evaluation approach, MW has evolved slightly from one study to the next. For each evaluation, we discuss what the outcomes say about our model of how VR influences learning, and we discuss how we used the evaluative outcomes to shape our design.

4.2.1 MW's Features and Learning. Recall that MW is a learning environment for teaching electrostatics concepts. In our first evaluation, we wanted to demonstrate that students could learn about electric fields, electric potential, and Gauss's law in this environment and to identify the strengths and weaknesses of its features and interface. Therefore, we focused on learning (process and outcomes) and the interaction and learning experiences.

Fourteen high-school and four college students completed from one to three lessons in MaxwellWorld. Thirteen of the fourteen high-school students had recently completed their senior year, and one student had recently completed his junior year. All students had completed one course in high-school physics. Pre-test interviews suggested that, despite having recently completed an introduction to electric fields, students understood very little about electric force, Gauss's law, and electric potential.

Each student session lasted approximately two hours. Students were scheduled on consecutive days for the first two sessions (which covered electric force and Gauss's law), while the third session (covering electric potential) was conducted approximately two weeks later. We examined students' comments during the learning process (predictions, observations, and comparisons), administrator observations during the learning process, usability questionnaires, interview feedback, and pre- and post-lesson knowledge assessments. For more detail concerning this study, see Dede, Salzman, Loftin, & Sprague (in press).

Below is a brief overview of some of the findings:

- **Learning (process and outcomes).** Pre- and post-lesson knowledge assessments showed that students developed a significantly more in-depth understanding of the distribution of forces in an electric field, as well as representations such as test charge traces and field lines, while using MW. Students cited immersive 3-D representations, interactivity, the ability to navigate to multiple perspectives (FORs), and the use of color as characteristics of MW as important to the learning process. Additionally, observations made during the learning process indicated that manipulating the field in 3-D 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 MW gave clear descriptions during the post-lesson interviews and demonstrations.
- **Learning experience.** Overall, students were very enthusiastic about their experiences. They rated MW as stimulating, giving an average rating of 2.1 on a scale from -3 (very dull) to $+3$ (very stimulating). Additionally, most students indicated that they thought MW was a more effective way to learn

about electric fields than either textbooks or lectures.

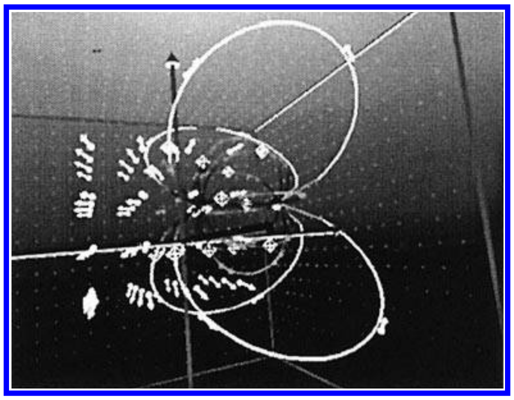
- **Interaction experience.** Overall, students felt the environment was fairly easy to use and gave it an average rating (with standard deviation) of 1.73 (1.12) on a scale of −3 (very difficult) to +3 (very easy). Additional ratings, student comments, and observations during the sessions indicated that usability problems typically occurred when navigating, using menus, and deleting source charges. A few students were susceptible to symptoms of simulator sickness (oculomotor discomfort, disorientation, and nausea). The mean total sickness scores (and standard deviations) on the SSQ were 11.00 (8.25) for the first session and 4.05 (7.80) for the second session. 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.
- **Learner characteristics and the interaction experience.** As with NewtonWorld, we observed substantial individual variability in the students' abilities to work in the 3-D environment and with 3-D controls (usability). Ratings for the menus and navigation (the two types of interaction with which students had the most trouble) reinforce this. Ratings for the usability of the menus ranged from −2 to +3; ratings for navigation ranged from −3 to +3. We also noticed variability in the students' susceptibility to symptoms of simulator sickness.

This study reinforced some of the lessons we had learned from NewtonWorld regarding the model. First, it implicated VR's 3-D immersive representations and FORs as capabilities that supported learning. Second, it indicated that students could be strongly motivated by immersive learning environments. Third, it demonstrated that the interaction experience can interfere with the learning process. Finally, this experiment reinforced the finding that interaction is highly variable and dependent to some extent on individual learner characteristics. For example, highly variable interaction experiences might be explained by individual differences in experience with computers or in susceptibility to motion sickness. To summarize, this study implicated a relationship between VR's features, motivation, and learning, and a relationship between individual characteristics and the interaction experience.

We used student feedback from this study to refine the interface and to implement additional features that would address some of the misconceptions the students demonstrated in these sessions. For example, we discovered that several students thought the field line representation indicated the path along which a test charge would travel if it were free to move through the electric field. (This is not the case.) Therefore, we added the ability to free test charges in the electric field and observe their motion.

4.2.2 MW's Features Versus Features of a Non-Immersive 2-D Learning Environment. Although this first evaluation yielded insights into the learning in MaxwellWorld, it did not allow us to establish whether learning was due to MaxwellWorld or to the lessons. Therefore, our next evaluation was designed to compare the learning, learning experience, and interaction experience in MaxwellWorld to those of a highly regarded and widely used two-dimensional learning environment, EM Field™, while tightly controlling the learning activities and the instructional content of the lessons.

The first stage of this study compared MaxwellWorld (MW) and EM Field (EMF) on the extent to which representational aspects of these environments influenced learning, learning experience, and interaction experience 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 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. Thus, the primary difference between the two learning environments was representational dimensionality (EMF's non-immersive 2-D versus MW_L's immersive 3-D). (See Figure 11.)



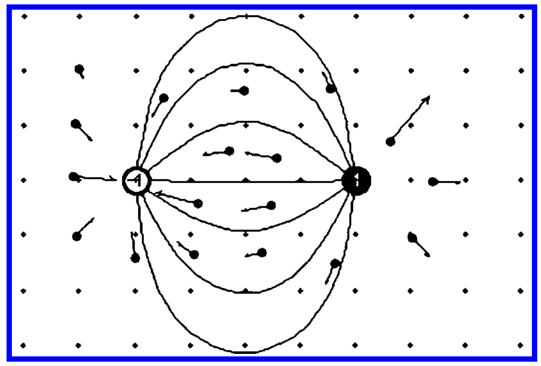


Figure 11. *A dipole with field lines and test charge traces in MW and EMF.*

In the second stage of the study, we utilized MaxwellWorld's full range of capabilities (including multisensory input) to ascertain the value these affordances 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 versus multisensory stimulation as instrumental to the learning potential of virtual reality.

Fourteen high-school students completed lessons in MW_L or EMF. Lessons leveraged the visual representations used in EMF and MW_L. During the second stage, we examined the "value added" by unique VR features (e.g., multisensory cues) supported by MaxwellWorld. Seven EMF and MW_L students returned for the second stage approximately five months after participating in the first stage. All students received an additional lesson in the full version of MW that utilized multisensory cues as well as visual representations. During both stages, we examined pre- and post-lesson understanding for each of the groups. We also assessed first-stage retention for those students that returned for stage two. Additionally, we gathered information about individual differences such as domain experience, computer and gaming experience, and motion sickness history. Finally, we examined whether learning experience (motivation and meaningfulness of the representations) and interaction experience (simulator sickness and usability) differed between the groups and whether they predicted learning outcomes.

Below is a summary of the first-stage outcomes.

- **VR's features and learning.** Both groups demonstrated significantly better conceptual 2-D and 3-D understanding post-lesson than pre-lesson. (All *F*-tests were significant at $p < 0.05$.) Therefore, lessons in EMF and MW_L were meaningful. However, MW_L students were better able to define concepts than EMF students. Although not statistically significant, differences also occurred in the students' ability to describe electric fields in 3-D on the five-month retention test. (See Table 1.)
  MW_L students were not any worse than the EMF students at sketching concepts in 2-D. 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. An explanation for this outcome may be that representations of force (lines and arrows) are more easily translated from 3-D to 2-D than representations of potential (surfaces).
  MW_L students were better able to demonstrate concepts in 3-D than EMF students. For example, despite the inherent three-dimensionality of the lessons and demonstration exercises, all but one EMF student restricted answers to a single plane, drew lines when describing equipotential surfaces, and used terms such as "oval" and "line." In contrast,

Table I. ***Adjusted Post-Lesson Retention Means and ANCOVA Outcomes for Stage 1 (covariate = pre-lesson scores)***

	Post-Lesson		Retention	
Learning	EMF	MW_L	EMF	MW_L
Concepts	0.58	0.70	0.69	0.66
	$F(1,11) = 3.17$*		$F(1,5) = 0.27$	
2-D sketches	0.80	0.82	0.42	0.43
	$F(1,11) = 0.24$		$F(1,5) = 0.00$	
3-D demos	0.67	0.87	0.31	0.57
	$F(1,11) = 9.99$*		$F(1,5) = 2.40$)	

"*" indicates F is significant at $p < 0.05$.

MW_L students described phenomena using 3-D gestures and phrases such as "sphere" and "surface." Although not statistically significant, differences also occurred in the students' ability to describe electric fields in 3-D on the retention test. (See Table 1.)

- **Learning experience.** Student ratings indicated that they felt significantly more motivated by MW_L than EMF ($F(1,12) = 7.66$). On a scale from -3 to $+3$, mean ratings (and standard deviations) were 2.03 (0.29) and 1.11 (0.82), respectively. Mean ratings of the meaningfulness of the representations did not vary significantly between groups. On a scale from -3 to $+3$, mean ratings (and standard deviations) were 2.36 (0.38) and 1.93 (1.13), respectively.
- **Learning experience versus learning outcomes.** Because motivation differed between the groups, one might hypothesize that higher motivation explains the superior learning outcomes in MW_L. However, neither motivation nor meaningfulness of the representations significantly predicted learning outcomes. Additionally, the group students were in (EMF versus MW_L) significantly predicted post-lesson scores beyond motivation ($R^2_{change} = 0.28$, $p < 0.05$), suggesting that an additional factor beyond motivation accounted for the differences in each group's learning.
- **Interaction experience.** Students experienced significantly greater simulator sickness symptoms in MW_L than EMF ($F(1,12) = 6.94$, $p < 0.05$) and had significantly more trouble using MW_L than EMF ($F(1,12) = 4.77$, $p < 0.05$). Mean simulator sickness scores (and standard deviations) were 7.11 (5.10) for MW_L and 1.70 (1.89) for EMF. On a scale from -3 to $+3$, mean usability ratings (and standard deviations) were 1.77 (0.57) for MW_L and 2.41 (0.52) for EMF.
- **Interaction experience versus learning outcomes.** Neither simulator sickness nor usability significantly predicted learning outcomes.
- **Learning experience versus interaction experience.** In this study, the learning experience and interaction experience were negatively related and each provided different information about the relationship between VR's features and learning. Although students had more trouble using MW than EMF, they rated their MW learning experiences more positively than they did their EMF learning experiences. For us, this has increased our interest in characterizing various facets of these experiences.
- **Learner characteristics versus interaction experience.** As in earlier experiments, there was a substantial range in participants' facility with the VR interface. To better understand this issue, we analyzed the relationship between the time that learners spent using computers and electronic games and their ratings of perceived usability. We found that time spent using computers explained marginally significant variance in the usability ratings beyond which group participants were in ($R^2_{change} = 0.18$, $F = 3.71$, $p = 0.08$). The more time the participant spent using computers, the higher usability was rated ($\beta = 0.43$).

Data for the second stage yielded the following outcomes:

- **VR's features and learning.** Students demonstrated significantly better understanding of concepts, 2-D sketches, and 3-D demos post-lesson than pre-lesson. (All F-tests were significant at $p < 0.05$.) Students learned from the visual and multisensory representations used in the lesson.
- **Learning versus learner characteristics.** Ratings concerning multisensory representations (haptic and

sound), learning outcomes, and student comments all suggested that some students (those who experienced difficulty with the concepts we were discussing) found multisensory cues more valuable than others.

- **Learning and interaction experiences.** Mean motivation, simulator sickness, and usability ratings (1.80, 6.99, and 1.78, respectively) were similar to the ratings for MW_L in stage one, indicating that our assessments of the learning and interaction experiences are fairly reliable.

Both stages shed light on the model. First, they tell us something about VR's features. The first-stage learning outcomes suggest that VR's 3-D immersive representation can help students develop more accurate and causal mental models than 2-D non-immersive representations. Subjective ratings characterizing the learning and interaction experiences for stage one yielded converging evidence that representational differences were responsible for differences in learning. Motivation, though higher in MW_L than in EMF, was not a predictor of learning. Although surely motivation is important, an additional factor beyond the motivational differences between the two virtual environments caused the learning differences. Additionally, despite MW_L's usability and simulator sickness problems, students learned more using this virtual environment than they did using EMF.

In the second stage, the enhancement of visual representations with multisensory cues appeared to facilitate learning, especially for students who had trouble grasping the concepts. Second, outcomes from the two stages tell us something about learner characteristics, as well as the interaction and learning experiences. These factors are critical to understanding the relationship between VR's features and learning. Take motivation as an example; by examining motivation versus group as a predictor of learning, we were able to rule out motivation as an explanation of learning differences. Alternatively, we could have found that motivational differences were the best predictor of learning. Had this been the case, we would have had to then try to determine if VR's representational features were responsible for the motivational differences, or if the outcome was due to some other aspect of the VR learning environment.

Concerning MW's design, the outcomes of this study were highly encouraging. We had developed a learning environment that was more powerful than a very successful commercial tool. However, we also recognized that we could do more to take advantage of VR's features. Accordingly, we have increased the flexibility of MW's frames of reference. MW now allows students to build and explore electric fields as a tiny test charge inside the electric field, as well as to build and explore the fields from outside the fieldspace (the box in which electric fields are built). Students can also scale down the fieldspace, attach it to their hand, and rotate it around for inspection. In addition to enhancing MW's FORs, we have incorporated more-sophisticated multisensory cues, relying on 3-D sound and haptic cues. MW's new capabilities will enable us to support more-detailed investigation of the relationships between FORs, multisensory cues, and learning.

4.2.3 MW's FORs, Learning, and Other Factors in the Model.

Our current study is designed to investigate the relationship between FORs, learning (both process and outcomes), the learning experience, the interaction experience, and individual differences. We are examining three FORs: the *egocentric* FOR in which students build and explore electric fields as a tiny test charge within that MW's fieldspace, the *exocentric* FOR in which students build and explore electric fields from outside the fieldspace, and the *bicentric* FOR in which students switch between the egocentric and exocentric FORs. Figure 12 illustrates MaxwellWorld's egocentric and exocentric FORs. It is our thesis that MW's egocentric FOR will make salient information that learners might not notice in the exocentric frame of reference, and vice-versa. Further, using the two FORs in combination might help students to fill in gaps in their knowledge and to become more flexible in their thinking. We also believe that FORs might affect both the interaction and learning experiences, which in turn influence learning. Finally, earlier studies have lead us to believe that individual differences in learner characteristics will be very important to understanding the true relationship between FORs and learning. For example, char-

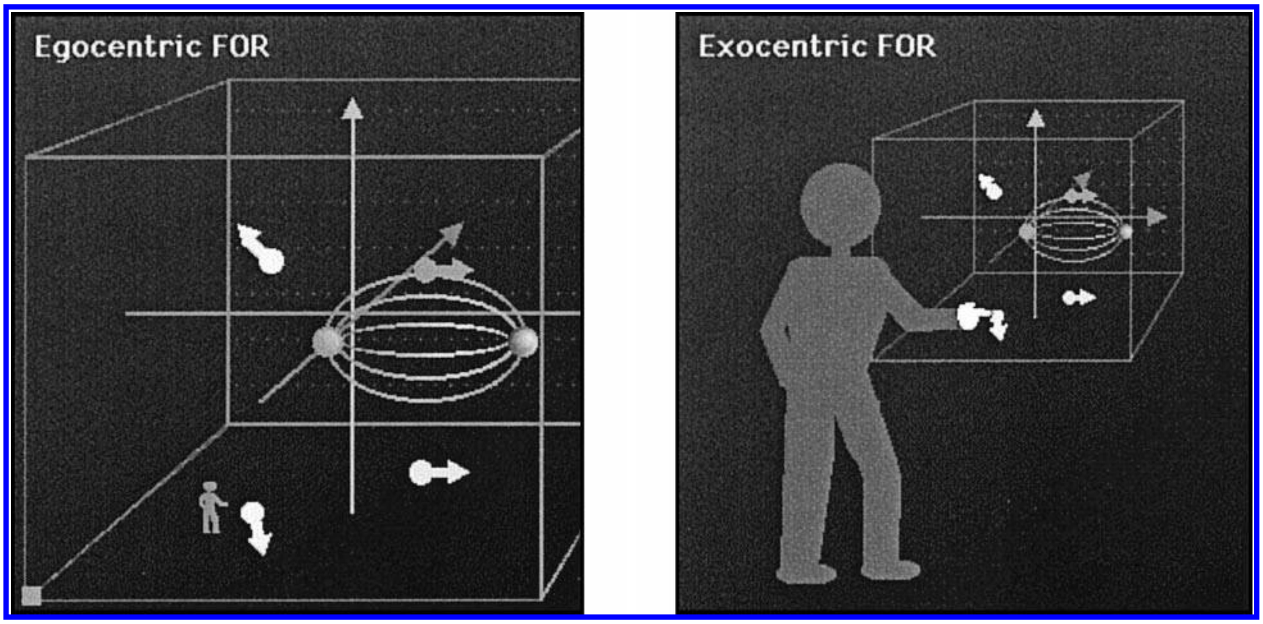


Figure 12. *Exocentric versus Egocentric Frames of Reference in MaxwellWorld.*

acteristics such as spatial abilities, gender, and domain experience might moderate how effective a FOR is.

We expect approximately 48 high-school juniors and seniors to participate in the study. Students are being assigned randomly to a FOR group such that groups are proportionally balanced as to gender. As part of the recruiting process, students complete several questionnaires concerning learner characteristics (e.g., gender, spatial abilities, domain experience, motion sickness history, and immersive tendencies). Then, they complete a lesson in MW that covers two concepts: the distribution of force in electric fields, and the motion of test charges through electric fields. These two concepts were selected because they differ in the extent to which global and local knowledge are important, the kinds of knowledge believed to be afforded by different FORs (Thorndike & Hayes-Roth, 1982). Comprehending distribution depends more heavily on global judgments than local judgments, while understanding motion requires more-local judgments than global judgments.

Throughout the lessons, we are monitoring the learning process via recording student's predictions, observations, and comparisons. We are also monitoring the interaction experience by capturing task times, error rates, physical interface problems, etc. Immediately after the students finish the lesson, they complete questionnaires concerning the interaction experience (usability and simulator sickness) and learning experience (motivation and presence). A few days after the lesson is completed, students return for post-lesson testing. We are examining several aspects of their learning as a means of assessing the strengths and weaknesses of the different FORs in supporting learning.

As the above description demonstrates, this study provides the most complete coverage of the model we presented earlier. Therefore, this experiment should provide considerable information to help us understand the interplay of VR's features, the learning experience, the interaction experience, individual characteristics, and learning at a finer-grained level of detail. In doing so, it should bring us one step closer to understanding how we can use VR's features to support the learning of complex information.

5 Lessons Concerning the Model and VR's Features

Through iterative design and evaluation, we have gained considerable insight into the complex relation-

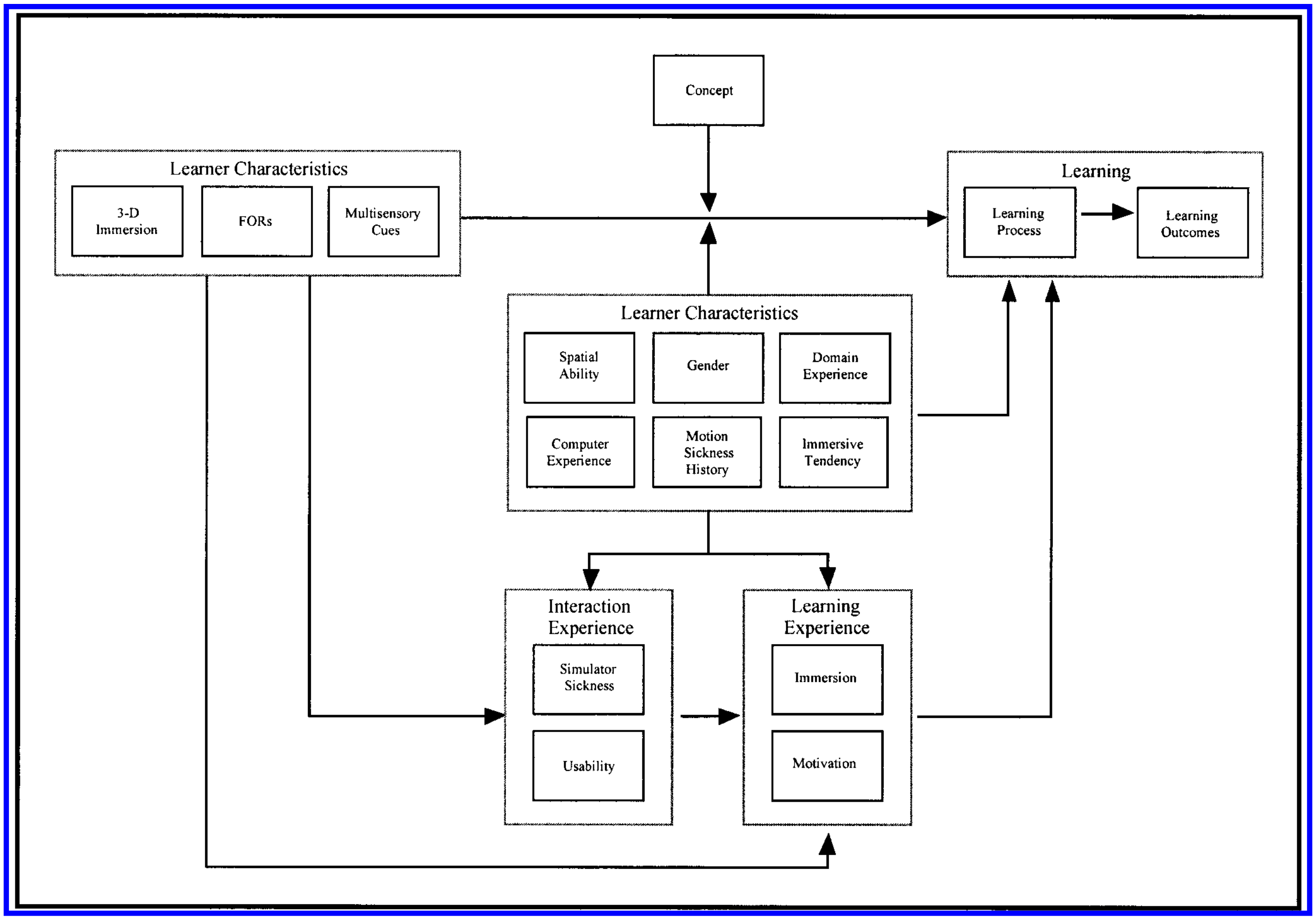


Figure 13. *A hypothetical model describing how VR's features, the concept one is being asked to learn, learner characteristics, and the interaction and learning experiences work together to influence the learning process and learning outcomes in VR learning environments.*

ship between VR's features and learning. The evaluations in particular have helped us discover ways to elaborate upon the factors in the model. Figure 13 shows a more detailed view of the model. Notice that each of the original factors can be broken down to reflect our increased understanding of these determinants. The model serves to highlight the types of relationships that are important to examine rather than to specify whether the relationships are positive or negative and strong or weak, because the relationships will differ depending on the specific nature of the virtual environment, the concepts being taught, and the learners. For example, we found that dimensions of the interaction and learning experiences were sometimes, but not always, positively correlated. We also found that, although the overall relationship between usability and learning was generally positive, there were instances in which it could have been negative.

As in the original model, the links of greatest interest are those that relate VR's features, the learning process, and learning outcomes. However, the concepts to be covered in any VR learning environment are also critical in determining the effectiveness of VR's affordances. Another important factor is learner characteristics. As our research has shown, an assessment of learner needs can provide some guidance, but it may not be sufficient. Individual learner characteristics may serve to moderate the relationship between VR's features and learning, and they may influence the learning and interaction experiences.

Finally, the quality of learning and interaction experiences also are important. When individuals interact with a VR learning environment, each has a unique experience in that environment. For example, one may find the environment easy to use and highly motivating, while another may find the experience to be negative in some way. This unique experience can be influenced by the characteristics of the individual as well as the features of environment. Therefore, measuring different aspects of the learning and interaction experiences may help explain learning outcomes beyond what VR's affordances can explain, and help understand the strengths and weaknesses of VR's capabilities in shaping the learning process and learning outcomes.

As the above discussion suggests, we have a long way to go to understand how the pieces of the model work together. However, lessons learned throughout our iterative design and evaluation process have helped us better understand the factors within the model.

- **VR's features.** Our work has focused on three capabilities: 3-D immersion, FORs, and multisensory cues. Our experience has convinced us that it is necessary to study various aspects of the model in order to truly understand the relationship between VR's affordances and learning. Evaluation outcomes have reinforced the notion that VR's features affect not only learning, but the quality of the interaction and learning experiences. Specifically, the outcomes have shown that 3-D immersive representations can be motivating and can support learning beyond 2-D non-immersive representations, that flexible FORs can be motivating and facilitate the learning process, and that multisensory cues can be used to direct students' attention and enhance the quality of the learning and interaction experiences. Finally, they have shown that VR features have potentially combinatorial effects. For example, different multisensory cues may be appropriate for different FORs.
- **Concepts.** Our design and evaluation experience suggests that different VR features are appropriate for different concepts. Generally, we analyze the concepts to be mastered as a guide for our decisions concerning the appropriate usage of VR's affordances. However, we also have examined this issue in some of our studies. We have found that VR features sometimes support the learning of one concept, but hinder the learning of another.
- **Individual characteristics.** Over time, we have identified several learner characteristics likely to be important: gender, domain experience, spatial ability, computer experience, motion sickness history, and immersive tendencies. Although our research to date has not focused on the role of each of these factors, we have developed some theories by combining our observations with information gathered from the literature. The first three factors are important because they might influence a person's aptitude for mastering abstract science concepts (Dillon, 1985; Dillon & Schmeck, 1983; Halpern, 1992). Of these, spatial ability may be particularly important in influencing how effectively students use the information provided by VR features (Egan & Gomez, 1985; Norman, 1995). Spatial ability is the cognitive ability that enables individuals to perceive patterns and to manipulate and rotate that information relative to one's own position in space (McGee, 1979). Numerous researchers have implicated spatial ability as one of the strongest predictors of performance in math and science (Halpern, 1992). Others have demonstrated that spatial ability affects mental manipulation, a process important to scientific reasoning (Hegarty & Sims, 1994). Spatial ability, along with computer experience and motion sickness history, might also affect the interaction experience. In our own work, we have found that computer experience can be predictive of how students rate the usability of the learning environment. Other researchers have found a strong positive correlation between spatial ability and computer-based performance (e.g., Egan, 1988; Gomez, Egan, & Bowers, 1986; Vincent, Hayes, & Williges, 1987) and between motion sickness history and simulator sickness (Kennedy, Lane, Berbaum, & Lilienthal, 1993). Finally, the last factor, immersive tendencies, may be useful for explaining how immersed a student will become in the VR learning environment (Singer & Bailey, 1994).
- **Learning experience.** Evaluations to date suggest that the learning experience is affected by VR's features. Although we have not yet studied how factors

such as motivation and presence affect learning, this is a promising topic for future research. Therefore, we have expanded the extent to which we focus on the quality of the learning experience and have worked to improve our measures of motivation, as well as adding a measure of presence to our evaluations. For example, in the current study of frames of reference, we are using Singer and Witmer's (1996) Presence Questionnaire.

- **Interaction experience.** Our evaluations have demonstrated that usability and simulator sickness can be affected by VR's features. However, our studies also highlight that usability can sometimes support learning and, at other times, interfere with it. As shown in our model, we are continuing to focus on usability and simulator sickness as indicators of the interaction experience. Additionally, we are exploring ways to measure performance usability as well as subjective usability. For example, in our current study, we will use task times, error types, and error rates as indicators of actual performance.
- **Interactions and trade-offs.** As some of the above discussion suggests, the interplay between factors sometimes can force us to make design trade-offs. One such trade-off was highlighted in some of our evaluations: the trade-off between designing for interaction and for learning.

6 Conclusions

Given that the ability to understand abstract and complex information is increasingly important in research, industry, and education, learning environments that support these skills are in growing demand. In order for designers of VR learning environments to meet this need, capabilities such as 3-D immersion, frames of reference, and multisensory cues are potentially crucial, but we need to know when and how to use them for supporting different learning tasks and various learner needs. A critical step towards achieving an informed design of VR learning environments is the investigation of the interplay among VR's affordances and other factors, such as the concepts to be learned, learner characteristics, the learning experience, and the interaction experience. By understanding how these factors work together to shape learning, we will be better able to target learning and visualization problems with the appropriate affordances and to maximize the benefits of this emerging technology.

As demonstrated through our work on Project ScienceSpace, a model that incorporates VR's features with other important factors can be a useful guide for both design and evaluation activities. We have used such a model to guide our development processes. Each factor in the model has impacted our design choices: which concepts we have attempted to address with our environments, which features we have utilized, and how we have designed the environments' interfaces. We have also used the model to guide the data collected and the analyses conducted in our evaluations. Although our primary research goal has been to assess the link between VR's characteristics and learning, we recognized that this could be achieved without careful attention to individual learner characteristics, interaction experiences, learning experiences, and the learning process. Had we not been concerned with these factors, we might have missed several potentially important findings. In NewtonWorld, for example, we might have missed the subtle but important links between multisensory cues or FORs and learning. In MaxwellWorld, we would have been less able to identify the benefits of 3-D immersion for learning.

In this paper, we have demonstrated how a model of VR's affordances and their impact on learning has guided our research activities and synthesized what we have learned about the factors in the model. The lessons we have shared in this paper should be useful for informing the design and evaluation of immersive VR learning environments. However, we recognize that the lessons we have learned to date provide only initial insights into a very complex web of relationships. Substantial additional research is necessary to elaborate and expand upon the model. For example, we need to study the relationship between VR's features and concepts more carefully so that we can appropriately match capabilities to content. We also need to further investigate the role of individual characteristics in determining the effectiveness of VR learning environments. Potentially, VR's affordances can be used in various ways to support the needs of different kinds of learners. Additionally, we need to understand interface issues surrounding the implementation of VR's features so that we can minimize negative interaction outcomes such as simulator sickness. Research done by Bob Kennedy relating simulator sickness to characteristics of VR environments (e.g., Kennedy, Drexler, & Berbaum, 1994) is a good example of work toward this goal.

In our work, we plan to continue to explore the relationships among factors of the model and to incorporate additional factors as necessary. We urge other designers and researchers of immersive VR learning environments to do the same: to study VR's potential in the context of these factors and to think about additional factors that also may be important. The success or failure of VR learning environments in practice critically depends upon the web of relations among VR's features, the concepts to be learned, learner characteristics, the learning experience, the interaction experience, and more. Informed design will help ensure that such environments succeed.

References

Barfield, W., Rosenberg, C., & Furness, T. A. (1995). Situation awareness as a function of frame of reference, computer graphics eyepoint elevation, and geometric field of view. *International Journal of Aviation Psychology, 5*(3), 233–256.

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

Clement, J. (1982). Students' Preconceptions in Introductory Mechanics. *American Journal of Physics, 50,* 66–71.

Darken, R. P., & Sibert, J. L. (1995). Navigating large virtual spaces. *International journal of human-computer interaction, 8* (1), 49–71.

Dede, C., Salzman, M., Loftin, B., & Ash, K. (in press). NewtonWorld: Utilizing multisensory immersion to aid in learning science. In M. Jacobson & R. Kozma (Eds.), *Learning the sciences of the 21st century: Research, design, and implementation of advanced technological learning environments.* Hillsdale, NJ: Lawrence Erlbaum.

——, ——, ——, & Sprague, D. (in press). In N. Roberts, W. Feurzeig, & B. Hunter (Eds.), *Computer Modeling and Simulation in Science Education*. New York: Springer-Verlag.

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

Dillon, R. (Ed.). (1985). *Individual Differences in Cognition*. New York: Academic Press.

——, & Schmeck, R. (Eds.). (1983). *Individual Differences in Cognition*. New York: Academic Press.

Egan, D. E. (1988). Individual differences in human-computer interaction. In Helander, M. (Ed.), *Handbook of Human-Computer Interaction* (pp. 543–568). New York: Elsevier Science Publishing Company.

——, & Gomez, L. M. (1985). Assaying, isolating, and accommodating individual differences in learning a complex skill. In R. F. Dillon (Ed.), *Individual Differences in Cognition, vol. 2* (pp. 173–217). New York: Academic Press.

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.

Frederiksen, J., & White, B. (1992). Mental models and understanding: A problem for science education. In E. Scanlon & T. O'Shea (Eds.), *New Directions in Educational Technology* (pp. 211–226). New York: Springer Verlag.

Gomez, L. M., Egan, D. E., & Bowers, C. (1986). Learning to use a text editor: Some learner characteristics that predict success. *Human-Computer Interaction, 2,* 1–23.

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.

——, & ——. (1985b). The initial knowledge state of college students. *American Journal of Physics, 53,* 1043–1055.

Halpern, D. (1992). *Sex Differences in Cognitive Abilities.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Heeter, C. (1992). Being there: The subjective experience of presence. *Presence: Teleoperators and Virtual Environments, 1* (1), 262–271.

Hegarty, M., & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory and Cognition, 22* (4), 411–430.

Kennedy, R. S., Drexler, J. M., & Berbaum, K. S. (1994). Methodological and measurement issues for identification of engineering features contributing to virtual reality sickness. *Proceedings of the IMAGE VII Conference* (pp. 244–254). Tempe, AZ: Image Society, Inc.

——, & Fowlkes, J. E. (1992). Simulator sickness is polygenic and polysymptomatic: Implications for research. *The International Journal of Aviation Psychology, 2*(1), 23–38.

——, 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.

Larkin, J. (1983). The role of problem representation in physics. In D. Gentner & A. Stevens (Eds.), *Mental Models* (pp. 75–98). Hillsdale, NJ: Lawrence Erlbaum Associates.

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.

McGee, M. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin, 86*(5), 889–918.

Norman, K. (1995). *Interface Apparency and Manipulatability: Cognitive Gateways through the Spatial Visualization Barrier in Computer-Based Technologies.* Available at http://www.lap.umd.edu/LAPFolder/NSFIA/proposal.html.

Nugent, G. (1982). Pictures, audio, and print: Symbolic representation and effect on learning. *Educational Technology Research and Design, 30* (3), 163–174.

Presson, C. C., DeLange, N., & Hazelrigg, M. D. (1989). Orientation specificity in spatial memory: What makes a path different from a map of the path? *Journal of Experimental Psychology: Learning, Memory and Cognition, 15,* 887–897.

——, & Hazelrigg, M. D. (1984). Building spatial representations through primary and secondary learning. *Journal of Experimental Psychology: Learning, memory, and cognition, 10,* 716–722.

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.

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.

Singer, M., & Bailey, J. (1994). Development of 'presence' measures for virtual environments. Poster presented at *The 38th Annual Meeting of the Human Factors and Ergonomics Society.*

Singer, M. J., & Witmer, B. G. (1996). *Presence Measures for Virtual Environments: Background & Development.* (ARI draft research note and questionnaires). Alexandria, VA: Simulator Systems Research Unit, U.S. Army Research Institute for the Behavioral and Social Sciences.

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

Thorndike, P. W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology, 14,* 560–589.

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

Vincent, K. J., Hayes, B. C., & Williges, R. C. (1987). Assaying and isolating individual differences in searching a hierarchical file system. *Human Factors, 29,* 349–359.

West, T. G. (1991). *In the Mind's Eye: Visual Thinkers, Gifted People with Learning Difficulties, Computer Images, and the Ironies of Creativity.* Buffalo, NY: Prometheus Books.

Wickens, C. D., & Baker, P. (1995). Cognitive issues in virtual reality. In W. Barfield & T. Furness (Eds.), *Virtual Environments and Advanced Interface Design* (pp. 515–541). New York: Oxford University Press.

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

Winn, W. (1993). A conceptual basis for educational applications of virtual reality. (Tech. Rep. No. R-93-9). Washington: University of Washington, HITL. Available at http://www.hitl.washington.edu/publications.

Witmer, B. B., & Singer, M. J. (in press). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments.*

This article has been cited by:

1. Young Hoan Cho, Su Yon Yim, Sunhee Paik. 2015. Physical and social presence in 3D virtual role-play for pre-service teachers. *The Internet and Higher Education* **25**, 70-77. [CrossRef]
2. Simone Borsci, Glyn Lawson, Simon Broome. 2015. Empirical evidence, evaluation criteria and challenges for the effectiveness of virtual and mixed reality tools for training operators of car service maintenance. *Computers in Industry* **67**, 17-26. [CrossRef]
3. Rustin Webster. 2015. Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments* 1-15. [CrossRef]
4. Yi-Hsuan Lee, Chan Hsiao, Chin-Husan Ho. 2014. The effects of various multimedia instructional materials on students' learning responses and outcomes: A comparative experimental study. *Computers in Human Behavior* **40**, 119-132. [CrossRef]
5. Nicholas PolysInformation Visualization in Virtual Environments: Trade-Offs and Guidelines 1265-1294. [CrossRef]
6. Chris Fowler. 2014. Virtual reality and learning: Where is the pedagogy?. *British Journal of Educational Technology* n/a-n/a. [CrossRef]
7. Jeffrey Jacobson. 2013. Digital Dome versus Desktop Display in an Educational Game. *International Journal of Gaming and Computer-Mediated Simulations* **3**:10.4018/jgcms.20110101, 13-32. [CrossRef]
8. Kikuo Asai, Norio Takase. 2013. Learning Molecular Structures in a Tangible Augmented Reality Environment. *International Journal of Virtual and Personal Learning Environments* **2**:10.4018/jvple.20110101, 1-18. [CrossRef]
9. Min Jou, Jingying Wang. 2013. Investigation of effects of virtual reality environments on learning performance of technical skills. *Computers in Human Behavior* **29**, 433-438. [CrossRef]
10. Zahira Merchant, Ernest T. Goetz, Wendy Keeney-Kennicutt, Oi-man Kwok, Lauren Cifuentes, Trina J. Davis. 2012. The learner characteristics, features of desktop 3D virtual reality environments, and college chemistry instruction: A structural equation modeling analysis. *Computers & Education* **59**, 551-568. [CrossRef]
11. Dawei Jia, Asim Bhatti, Saeid Nahavandi. 2012. The impact of self-efficacy and perceived system efficacy on effectiveness of virtual training systems. *Behaviour & Information Technology* 1-20. [CrossRef]
12. Francisco Rebelo, Emília Duarte, Paulo Noriega, Marcelo SoaresVirtual Reality in Consumer Product Design 381-402. [CrossRef]
13. Tassos A. Mikropoulos, Antonis Natsis. 2011. Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education* **56**, 769-780. [CrossRef]
14. Nicholas F. Polys, Doug A. Bowman, Chris North. 2011. The role of Depth and Gestalt cues in information-rich virtual environments. *International Journal of Human-Computer Studies* **69**, 30-51. [CrossRef]
15. Eric D. Ragan, Ajith Sowndararajan, Regis Kopper, Doug A. Bowman. 2010. The Effects of Higher Levels of Immersion on Procedure Memorization Performance and Implications for Educational Virtual Environments. *Presence: Teleoperators and Virtual Environments* **19**:6, 527-543. [Abstract] [PDF] [PDF Plus]
16. Elinda Ai-Lim Lee, Kok Wai Wong, Chun Che Fung. 2010. How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education* **55**, 1424-1442. [CrossRef]
17. Daniel A. Guttentag. 2010. Virtual reality: Applications and implications for tourism. *Tourism Management* **31**, 637-651. [CrossRef]
18. Barney Dalgarno, Mark J. W. Lee. 2010. What are the learning affordances of 3-D virtual environments?. *British Journal of Educational Technology* **41**:10.1111/bjet.2010.41.issue-1, 10-32. [CrossRef]
19. Sarah Cornelius, Phil Marston. 2009. Towards an understanding of the virtual context in mobile learning. *Research in Learning Technology* **17**, 161-172. [CrossRef]
20. Lakshmi Goel, Sonja Prokopec. 2009. If you build it will they come?—An empirical investigation of consumer perceptions and strategy in virtual worlds. *Electronic Commerce Research* **9**, 115-134. [CrossRef]
21. K. C. Yu, K. Williams, D. Neafus, L. Gaston, G. Downing. 2009. Gaia Journeys: a museum-based immersive performance exploration of the Earth. *International Journal of Digital Earth* **2**, 44-58. [CrossRef]
22. Jeremy Bailenson, Nick Yee, Jim Blascovich, Andrew Beall, Nicole Lundblad, Michael Jin. 2008. The Use of Immersive Virtual Reality in the Learning Sciences: Digital Transformations of Teachers, Students, and Social Context. *Journal of the Learning Sciences* **17**, 102-141. [CrossRef]
23. Sue V. G. Cobb. 2007. Virtual Environments Supporting Learning and Communication in Special Needs Education. *Topics in Language Disorders* **27**, 211-225. [CrossRef]

24. E. Richard, A. Tijou, P. Richard, J.-L. Ferrier. 2006. Multi-modal virtual environments for education with haptic and olfactory feedback. *Virtual Reality* **10**, 207-225. [CrossRef]
25. Tassos A. Mikropoulos. 2006. Presence: a unique characteristic in educational virtual environments. *Virtual Reality* **10**, 197-206. [CrossRef]
26. Tassos Mikropoulos, Apostolos Katsikis, Eugenia Nikolou, Panayiotis Tsakalis. 2003. Virtual environments in biology teaching. *Journal of Biological Education* **37**, 176-181. [CrossRef]
27. Fabrizia Mantovani, Gianluca Castelnuovo, Andrea Gaggioli, Giuseppe Riva. 2003. Virtual Reality Training for Health-Care Professionals. *CyberPsychology <html_ent glyph="@amp;" ascii="&"/> Behavior* **6**:10.1089/CPB.2003.6.issue-4, 389-395. [CrossRef]
28. M SCAIFE. 2001. Informing the design of a virtual environment to support learning in children. *International Journal of Human-Computer Studies* **55**, 115-143. [CrossRef]
29. H NEALE. 2001. Theme-based content analysis: a flexible method for virtual environment evaluation. *International Journal of Human-Computer Studies* **55**, 167-189. [CrossRef]
30. Eva L. Waterworth, John A. Waterworth. 2001. Focus, Locus, and Sensus: The Three Dimensions of Virtual Experience. *CyberPsychology <html_ent glyph="@amp;" ascii="&"/> Behavior* **4**:10.1089/CPB.2001.4.issue-2, 203-213. [CrossRef]
31. Kevin L. Dean, Xylar S. Asay-Davis,, Evan M. Finn, Tim Foley, Jeremy A. Friesner, Yo Imai, Bret J. Naylor, Sarah R. Wustner, Scott S. Fisher, Kent R. Wilson. 2000. Virtual Explorer: Interactive Virtual Environment for Education. *Presence: Teleoperators and Virtual Environments* **9**:6, 505-523. [Abstract] [PDF] [PDF Plus]
32. Kikuo Asai, Norio TakaseLearning Molecular Structures in a Tangible Augmented Reality Environment 1-18. [CrossRef]
33. Jeffrey JacobsonDigital Dome versus Desktop Display in an Educational Game 13-34. [CrossRef]
34. Mohd Fairuz Shiratuddin, Alen Hajnal3D Collaborative Virtual Environment to Support Collaborative Design 185-210. [CrossRef]
35. José Neto, Paulo MendesGame4Manager 108-134. [CrossRef]