

Learning Science in Immersive Virtual Reality

Jocelyn Parong and Richard E. Mayer
University of California, Santa Barbara

The goals of the study were (a) to compare the instructional effectiveness of immersive virtual reality (VR) versus a desktop slideshow as media for teaching scientific knowledge, and (b) to examine the efficacy of adding a generative learning strategy to a VR lesson. In Experiment 1, college students viewed a biology lesson about how the human body works either in immersive VR or via a self-directed PowerPoint slideshow on a desktop computer. Based on interest theory, it was predicted that students who learned in immersive VR would report more positive ratings of interest and motivation and would score higher on a posttest covering material in the lesson. In contrast, based on the cognitive theory of multimedia learning, it was predicted that students who learned with a well-designed slideshow would score higher on a posttest, although they might not report higher levels of interest and motivation. The results showed that students who viewed the slideshow performed significantly better on the posttest than the VR group, but reported lower motivation, interest, and engagement ratings. In Experiment 2, students either viewed a segmented VR lesson and produced a written summary after each segment or viewed the original, continuous VR lesson as in Experiment 1. Students who summarized the lesson after each segment performed significantly better on the posttest and the groups did not differ on reported interest, engagement, and motivation. These results support the cognitive theory of multimedia learning and demonstrate the value of generative learning strategies in immersive VR environments.

Educational Impact and Implications Statement

The findings from Experiment 1 showed that a PowerPoint slideshow may be more effective for teaching scientific information than an equivalent lesson in an immersive virtual reality environment, but may be less enjoyable and motivating for students. This suggests that the conversion of multimedia lessons into virtual reality may not yet be warranted. However, because increasing student interest and motivation may be a critical factor in classrooms, it is worthwhile to examine what supplemental learning strategies could be added to a medium that students enjoy. Experiment 2 showed that asking students to summarize segments of the immersive virtual reality lesson increased their learning outcomes without diminishing their motivation and interest. Taken together, the 2 experiments showed that students' interest can be primed with new and exciting technology while still being an effective medium to convey scientific information comparable to traditional PowerPoint slideshow lessons.

Keywords: virtual reality, multimedia learning, summarizing, science simulation, learning strategies

As new and innovative technologies emerge, so do ideas of incorporating them into classrooms to enhance student learning. One example gaining popularity among consumers is immersive virtual reality (VR) consoles with head-mounted displays and hand controllers, such as the Oculus Rift (Oculus VR LLC, Facebook Inc., Menlo Park, CA) and HTC Vive (HTC Corporation, Xindian, New Taipei City, Taiwan). Previous researchers have made claims about the usefulness of VR technologies (Blascovich & Bailenson,

2011). For example, Brelsford (1993) claimed that “VR allows the educational task to become much more intuitive; information is passed between the environment and the student with increased efficiency and selectivity” (p. 1287). As VR systems have become more affordable, using VR at the consumer level and in education has become more feasible. These VR technologies allow opportunities for educators to offer students easy and intuitive ways to interact with multimedia lessons. Students would be able to experience a fully immersive sensory experience in almost any space imaginable, which may encourage them to engage in deeper learning. However, there is still a gap between claims for the usefulness of VR in academic learning and scientific research testing these claims.

The goal of the present study is twofold. First, in a media comparison study in Experiment 1, we seek to compare the instructional effectiveness of immersive virtual reality versus a desktop slideshow for teaching scientific knowledge. This is a media comparison study because we wish to examine how well students

This article was published Online First January 25, 2018.

Jocelyn Parong and Richard E. Mayer, Department of Psychological and Brain Sciences, University of California, Santa Barbara.

This research was supported by Grant N000141612046 from the Office of Naval Research.

Correspondence concerning this article should be addressed to Jocelyn Parong, Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA 93106. E-mail: jocelyn.parong@psych.ucsb.edu

learn the same material when it is delivered via one medium versus another. In short, the first goal is to determine the efficacy of VR as a venue for science learning. Although media comparison studies have a history of methodological challenges (Clark, 2001), the present research question has important practical implications concerning whether it is worthwhile to convert conventional instruction to VR formats and theoretical implications concerning whether cognitive and motivational theories apply to learning in VR. Second, in a value-added study in Experiment 2, we seek to compare the instructional effectiveness of adding a prompt to write summaries at various points in a VR lesson. This is a value-added study because we start with a basic lesson and examine whether student learning is affected when we add one instructional feature to the lesson. In short, the second goal is to determine how to increase the effectiveness of VR as a venue for science learning.

Media Comparison: Using Virtual Reality for Learning Compared to a Slideshow

Because of the implications for education, it is important to examine the effectiveness of emerging VR technologies for teaching academic content. What are the underlying theories concerning the cognitive consequences of learning in immersive VR, and how would these new multimedia technologies fare in comparison to conventional methods of teaching, such as a PowerPoint slideshow? Theories based in multimedia learning and motivation provide a positive foundation for using VR in the classroom. Supporting these theories are newly emerging empirical studies of the efficacy of VR lessons on academic learning, as well as research on the use of related technologies in education, such as video games, simulations, and mixed reality games. However, because many educational games and simulations in immersive VR are also designed for entertainment, they may not necessarily adhere to all the principles of multimedia learning, and therefore may hinder learning compared with traditional methods.

It is customary to distinguish between two types of VR. Immersive virtual reality (immersive VR) typically includes a head mounted display (HMD) controlled by a computer where the learner can move through a 3-D virtual environment. Nonimmersive virtual reality includes a virtual world displayed on a computer screen where the learner can interact through an interface such as a mouse, touchscreen, touchpad, or handheld controls while seated in front of the screen (Lee & Wong, 2014). Although much more research has been done on educational simulations with desktop VR (e.g., Hilton & Honey, 2011; Merchant, Goetz, Cifuentes, Kenney-Kennicutt, & Davis, 2014), the growing availability of immersive VR technology for education creates a need to determine how immersive VR affects academic learning (Blasovich & Bailenson, 2011). The present study addresses this need by comparing learning about the human body in immersive VR (VR group) to learning with a self-paced PowerPoint slideshow containing the same words and images in static form on a desktop computer (slideshow group).

Multimedia Learning Theories

The case for using a well-designed slideshow for teaching of scientific material is grounded in the cognitive theory of multimedia learning (Mayer, 2009, 2014a) and cognitive load theory (Sweller,

Ayres, & Kalyuga, 2011) from which it is derived, which posit that adding material and features—such as the visual effects in immersive VR—can create extraneous processing in the learner—that is, cognitive processing that is not relevant to the instructional goal. Given that each learner has a limited amount cognitive processing capacity, if the learner allocates much of that capacity to extraneous processing, there may not be enough remaining capacity to engage in sufficient levels of cognitive processing aimed at making sense of the essential material.

Multimedia learning refers to learning from words and pictures (Mayer, 2009). The two media we compare in this study are a self-directed PowerPoint, where the learner is presented with written words and static pictures, and an interactive lesson in VR, where the learner is presented with spoken words and interactive animations. One approach to researching multimedia learning asks how a multimedia lesson can be adapted to reduce extraneous processing, which does not serve an instruction goal (Mayer & Fiorella, 2014), manage essential processing, which is necessary for holding and manipulating incoming information in working memory (Mayer & Pilegard, 2014), and foster generative processing, which is needed for a deeper understanding of the information (Mayer, 2014c). Empirical research has shown evidence for 12 principles for designing features within a multimedia lesson with educational content (Mayer, 2009), but the most relevant for the present study are coherence principle and the segmenting principle.

The *coherence principle* states that people learn better when extraneous words (Mayer, Bove, Bryman, Mars, & Tapangco, 1996; Mayer & Jackson, 2005), sounds (Moreno & Mayer, 2000), and pictures (Harp & Mayer, 1997, 1998; Mayer, Heiser, & Lonn, 2001) are excluded rather than included. Extraneous material diverts attention from the important material, disrupts the process of organizing the material, and may prime the learner to integrate the material with prior knowledge in an inappropriate manner (Mayer, 2009). Extraneous features in an immersive VR lesson may include the constant animations that the learner is immersed in. However, a well-designed slideshow may not include extraneous visual stimuli. Therefore, an immersive VR lesson would lead to worse learning outcomes than a lesson that eliminates those extraneous features. This is foreshadowed by research showing that instruction with desktop VR in 2D was more effective than identical instruction with desktop VR in 3D, presumably because the perceptual realism in the 3D environment created more extraneous cognitive load that distracted the learners from the essential content (Moreno & Mayer, 2002, 2004; Richards & Taylor, 2015). These findings are consistent with research showing that cartoon-like characters are more effective than photo-realistic representations in educational games (Wouters & van Oostendorp, 2017) and static diagrams can be more effective than animations in multimedia lessons on how scientific systems work (Lowe & Ploetzner, 2017; Mayer, Hegarty, Mayer, & Campbell, 2005). However, it is worth noting that 3D VR may be more effective for some subjects, particularly those that teach spatial or motor skills (e.g., laparoscopic surgery; Torkington, Smith, Rees, & Darzi, 2001).

The *segmenting principle* states that people learn better when the lesson is presented in user-directed segments, rather than as a continuous unit, particularly when learning the steps in a process (Lee, Plass, & Homer, 2006; Mayer & Chandler, 2001; Mayer, Dow, & Mayer, 2003). The segmenting principle helps to manage

essential processing by allowing the learner to decide when to add more incoming information. When the lesson is not self-paced, the learner may still be processing previous information when new information is presented. Well designed-slideshows may be self-paced and each slide itself segments the lesson, whereas immersive VR animations are typically continuous and not user-paced, thereby violating this principle. This adds to the learner's working memory load and may inhibit some essential processing for the learner.

The foregoing analysis shows that VR environments are more susceptible to violating the coherence and segmenting principles. Additionally, there may be other factors within immersive VR lessons that may impact learning outcomes, so the theoretical groundwork of other unique features of VR must also be examined.

Motivational Theories

Student motivation plays a large role in deeper learning in the classroom; those who are more motivated are more likely to engage in the lesson or task, put in more effort for understanding the material, and be resilient when overcoming obstacles in understanding (Mayer, 2008; Wentzel & Miele, 2016). This motivation may cause the learner to stay focused even during continuous lessons and invest more energy in allocating cognitive resources to difficult parts of the lesson. The case for using immersive VR for teaching of scientific material is grounded in interest theory and self-efficacy theory (Dewey, 1913; Schiefele, 2009).

According to interest theory, students work harder when they value and are interested in the material, either intrinsically (individual interest) or as elicited by the situation (situational interest; Mayer, 2008; Schiefele, 2009; Wigfield, Tonks, & Klauda, 2016). A student with individual interest would be more motivated to learn about the subject. In a meta-analysis of the relationship between interest and academic achievement, there was a moderate correlation between self-ratings of how much a student liked a school subject and how well the student performed in school as measured by grades or test scores (Schiefele, Krapp, & Winteler, 1992). Individual motivation would not be uniquely primed by a multimedia lesson as a student's intrinsic interest may not be affected by the media through which he or she is learning. However, a novel, immersive technology in education, such as a VR lesson, may be exciting for students and could prime a learner's situational interest more than conventional lessons (Kintsch, 1980; Wade, 1992). By priming their situational interest in the VR lesson, students may be more likely to pay attention to the content of the lesson, actively interact with the lesson, and persist throughout the entire lesson, which would foster generative processing and deeper learning leading to better learning outcomes than conventional lessons in the classroom.

Self-efficacy theory states that students work harder when they see themselves as competent for the task (Schunk & DiBenedetto, 2016). Self-efficacy, as defined by Schunk (1991), is a person's judgments of his or her ability to perform a given action. Bandura (1977) claimed that self-efficacy affects the amount of effort and persistence that a person devotes to a task. Research has shown that a student's sense of self-efficacy is related to how active he or she is in the lesson (Zimmerman & Martinez-Pons, 1990) and how well he or she performs on academic measures (Chemers, Hu, &

Garcia, 2001; Pietsch, Walker, & Chapman, 2003; Schunk & Hanson, 1985). Schunk (1989) described the process of self-efficacy and achievement behaviors as a feedback loop. First, the student has his or her beliefs about his or her self-efficacy (e.g., I'm good at this). This self-efficacy then affects the student's task engagement (e.g., I will try hard). After the task, the student receives feedback (e.g., I did well on the task), and receives efficacy cues from the feedback (e.g., the instructor thinks I'm good at this). Finally, this aptitude feedback reshapes the student's self-efficacy. VR games that incorporate a feedback system for progress on academic content would be acting on this system. For example, certain interactions within a lesson could provide appropriate feedback that boosts the student's self-efficacy, which would in turn enhance a student's motivation for the lesson. Specifically, an action in which a learner must show progress in learning in VR could provide immediate and adaptive feedback, which may be an advantage over the feedback seen in traditional academic lessons as learners would have immediate updates in their self-efficacy.

According to Pintrich (2003), student motivation lies in a value-expectancy model, a combination of the two theories described, in which motivation depends on a learner's values, which gets the student started (e.g., liking the material), and expectancies, which keeps him or her going (e.g., feeling competent). To increase the student's motivation, first, the lesson may prime the student interest; then, the learner's interaction with the lesson may prime his or her self-efficacy to continue the lesson. In the case for VR, the stimulating, immersive experience may spark the learner's individual interest and the feedback from interacting with the lesson should keep him or her feeling competent to progress in the lesson. However, in line with Dewey's (1913) admonitions against situational interest more than 100 years ago, research on seductive details shows that adding interesting but irrelevant material to a multimedia lesson may hurt learning (Harp & Mayer, 1997, 1998; Mayer, Griffith, Naftaly, & Rothman, 2008).

Media Comparison Evidence

The foregoing analysis suggests that immersive VR could be explored as a venue for academic learning as a potential replacement for conventional media such as slideshow presentations. The first experiment is an example of a media comparison study, in which learning academic material with one medium is compared with learning the same material in another medium (Mayer, 2014b). There is a small but growing body of research comparing learning with games and simulations presented on desktop computers versus conventional media (Hilton & Honey, 2011; Mayer, 2014b). Our focus in the present study is on science learning. Science was chosen as the domain of interest because many topics, including chemistry and biology, are spatially oriented. Therefore, it may be important to examine whether an immersive environment is important for learning about spatially relevant information.

In the science domain, there is evidence that playing a computer game or simulation leads to better learning outcomes than a standard lesson, such as learning from a textbook, slideshow, or lecture. A recent meta-analysis of the effects of playing desktop games and simulations in education showed that games tend to improve learning more than conventional media in science ($d = .69$) based on 16 experimental comparisons (Mayer, 2014b). In one

study, Moreno, Mayer, Spire, and Lester (2001) presented high school and college students with a simulation computer game called *Design-a-Plant*, in which the players completed a botany lesson by choosing a plant's roots, stem, and leaves and comparing their survival rates in various environmental conditions. Compared with a control group that received the same information in the form of a noninteractive computer tutorial, those who played the game outscored those who received the tutorial on a transfer test after the lesson. Other experiments in the science domain have shown comparable results in positive learning outcomes of games in electronics (Anderson & Barnett, 2011), physics (Swaak, de Jong, & van Joolingen, 2004), biology (Adams, Mayer, MacNamara, Koenig, & Wainess, 2012), chemistry (Evans, Yaron, & Leinhardt, 2008), and animal life (Hwang, Wu, & Chen, 2012).

More recently, McLaren, Adams, Mayer, and Forlizzi (2017) reported that students who learned decimal arithmetic through playing a computer-based adventure game, *Decimal Point*, performed better on posttests than those who learned with a conventional computer-based tutoring system. Similarly, a recent review also reported small-to-medium effect sizes favoring learning with computer-games over learning with conventional media (Wouters & van Oostendorp, 2017).

Although there is evidence for positive learning outcomes from games, simulations, and mixed reality multimedia lessons, there is scant empirical research on the specific use of immersive VR in education. A meta-analysis of the use of VR in K–12 classrooms reported that many studies rely on a descriptive analysis based on student surveys and interviews to determine if the virtual worlds helped them learn (Cooper, 2007; Hew & Cheung, 2010; Ligorio & van Veen, 2006). One study examined students in a computer graphics class who used an immersive 3D program to learn about function-based shape modeling. The researchers used exam scores throughout the course to determine the efficacy of VR and reported a 14% increase after one semester (Sourin, Sourina, & Prasolova-Førland, 2006). However, the researchers did not include a control group in their experiment, so the increase in exam scores cannot be attributed to solely the use of VR as it may be due to other factors within the classroom or students.

More recent empirical studies show promise in the use of VR in education, particularly in the science domain. In one study, Kozhevnikov, Gurlitt, and Kozhevnikov (2013) presented participants with a lesson on relative motion concepts in either an immersive virtual environment (IVE), using a head-mounted display, or a desktop virtual environment (DVE) presented on a computer. After the lesson, both groups showed a significant shift in their scientific understanding of relative motion, as well as a significant improvement on problem-solving tests. Additionally, the IVE group performed significantly better than the DVE group on solving two-dimensional relative motion problems, suggesting a further transfer of learning from the use of VR compared with the use of a computer. In another example, military participants learned about basic corrosion prevention and control either in an immersive VR environment using a combination of a head mounted display from WorldViz and movement tracking with the Microsoft Kinect or a desktop multimedia presentation. They found that both forms of instruction increased learning, but those in the VR environment had significantly higher gains in learning than those in desktop learning environments (Webster, 2016). Overall, this review suggests that more methodologically sound research is

needed on the effects of immersive VR in education, particularly in specific academic domains, such as science.

Criticisms of Media Comparison Studies

It is worth noting that media comparison studies have received criticisms. Since the early 1980s, there has been a media debate, sparked by Richard Clark and Robert Kozma, regarding the usefulness of studying the type of media used in learning tasks in an apples-to-oranges type of comparison. Clark (1983) claimed that “media are mere vehicles that deliver instruction but do not influence student achievement any more than the truck that delivers our groceries causes changes in our nutrition” (p. 445). Instead, he argued that many different media can accomplish the same learning goal, and therefore, because no single media attribute can contribute a unique cognitive effect on a learning task, other variables are more instrumental in learning gains (Clark, 1994).

However, Kozma (1994) argued that the question is more nuanced than whether media causes learning, but rather what is the relationship between media and learning, and how can we explain that relationship? Certain media may possess particular characteristics that may make them more or less suitable for learning tasks. For example, Salomon (1979) argued that media can be analyzed by the types of their cognitively relevant characteristics that affect the ways in which learners represent and process information. Some of these characteristics include the physical, mechanical, or electronic characteristics of the technology of the media, the symbol systems, including pictures, text, and spoken language, that are used in the media, and the processing capabilities of the media to act on the symbol systems. With this view, underlying theories about the way we process lessons, such as the cognitive theory of multimedia learning, can help inform the field about how particular features in multimedia affect learning, and therefore find value in media comparison studies. Additionally, to minimize the inherent differences between different media, we constructed the slide-show from the VR lesson to equate the lessons as much as possible by using the same words and images. Specifically, we transcribed the narration from the VR lesson and used the same words on the slides, and we took screenshots from the VR lesson to create still pictures on the slides.

Value-Added Research: Adding Generative Learning Strategies to VR

In addition to a media comparison study between immersive VR and a desktop slide show, we also examined the efficacy of adding a generative learning strategy to the existing VR lesson. The second experiment is an example of a value-added study, in which an additional feature is added to an existing lesson to promote learning and is compared with the original lesson (Mayer, 2014b).

Generative Learning Theory

According to *generative learning theory*, meaningful learning occurs when learners engage in appropriate cognitive processing during learning, including selecting (i.e., paying attention to relevant incoming information), organizing (i.e., mentally arranging the information into a coherent structure), and integrating (i.e., connecting the verbal and pictorial representations with each other

and with relevant prior knowledge activated from long-term memory). *Generative learning* is the process of taking incoming information and transforming it into usable information by engaging in appropriate selecting, organizing, and integrating (Fiorella & Mayer, 2015, 2016).

Along these same lines, Wittrock's (1989) generative model of learning stated that one important component of meaningful learning is generation, or the creation of connections within to-be-learned material as well as between to-be-learned material and existing knowledge. Wittrock's model of generative learning is closely tied to Mayer's (2009, 2014b) select-organize-integrate model, in which learners select relevant incoming information in sensory memory, organize the selected information in working memory, and integrate the organized information with prior knowledge in long term memory.

Value-Added Evidence

The extant research shows support for adding a feature, such as an additional worksheet, to a game-like lesson that prompts a generative learning strategy (namely, self-explanation) to help learners deeply understand the material (Fiorella & Mayer, 2015, 2016; Mayer, 2014b). For example, in one study by Pilegard and Mayer (2016), students played a computer game called *Cache 17*, which included instructional content about how wet cell batteries work. One group of students were asked to complete worksheets with questions about how wet cell batteries work during the game, while another group did not fill out worksheets. The worksheet group significantly outperformed the control group on a subsequent comprehension test. In another study, Fiorella and Mayer (2012) had participants play an educational game about electrical circuits, called *Circuit Game*. Students were asked to fill in worksheets about the features of each principle in an electrical circuit. Those who were able to correctly fill in the worksheets did significantly better than a control group on a transfer test.

Summarizing

The generative learning strategy of interest in this study is summarizing, or taking the important information from a lesson and putting it into one's own words (Fiorella & Mayer, 2016). According to the generative learning theory, the act of creating a well-written summary primes active cognitive processing during learning including selecting, organizing, and integrating. Summarizing promotes selecting by encouraging learners to choose the most important information for inclusion in the summary, organizing by encouraging learners to relate the pieces of information to each other within a coherent summary, and integrating by encouraging learners to use their own words to relate the material to relevant prior knowledge.

Research on the effects of asking students to summarize text they are reading has produced encouraging evidence of improvements on tests of learning outcome. Support for the idea that summarizing can be an effective generative learning strategy for learning from text comes from experiments where a group that generates summaries during or after a lesson performs better on subsequent tests than a group that does not summarize during or after the lesson (e.g., Annis, 1985; Coleman, Brown, & Rivkin, 1997; Peper & Mayer, 1986; Ross & Kirby, 1976). For example,

in a classic study by Doctorow, Wittrock, and Marks (1978), middle school students read a narrative either normally or were asked to generate a one-sentence summary of each paragraph. Those who wrote summaries performed better on a subsequent comprehension test than those who read the narrative normally, indicating that summarizing led to a deeper understanding of the material.

In a meta-analytic review of published experiments comparing learning by reading (i.e., control group) versus learning by reading along with summarizing (i.e., summarizing group), Fiorella and Mayer (2015) found that the summarizing group outperformed the control group on learning outcome tests in 26 out of 30 experimental comparisons, yielding a median effect size of $d = 0.50$, which is a medium sized effect. In contrast to previous work that focused on learning from printed text, the current study is the first to investigate the effectiveness of applying the generative learning strategy of summarizing to learning in an immersive VR environment.

Predictions

The VR lesson used for this experiment was an interactive biology simulation called *The Body VR: Journey Inside a Cell* (The Body VR, 2016). In this simulation, the player is taken through a tour of the blood stream and the inside of a cell. The simulation is narrated with descriptions and explanations of the parts and functions of the blood stream and cell and includes a full 360-degree view of animations within the blood stream and cell. The self-directed PowerPoint slideshow was adapted from this simulation to match the content.

In Experiment 1, based on interest and self-efficacy theory, if the motivational features of immersive VR encourage the learners to process the material more deeply, then we predict that VR group will have increased self-report measures of interest, engagement, motivation, and affect, and subsequently outperform the slideshow group on tests of learning outcomes. In contrast, if the extraneous features of immersive VR distract the learners from processing the material more deeply, then we predict that the slideshow group will outperform the VR group on tests of learning outcomes but not necessarily on self-report measures of interest, engagement, motivation, and affect.

The second aim (examined in Experiment 2) is to determine whether adding summarizing prompts to the VR lesson increases learning outcomes. Based on generative learning theory as well as previous research using text-based learning environments, we predict that students who view the VR lesson and summarize each segment will perform better on a posttest than those who view the original VR lesson. Although generative learning theory makes no predictions about motivation and interest, we expect that groups will not differ on measures of interest, engagement, motivation, and affect because they are exposed to the same immersive lesson.

There are no published experiments that examine the effects of applying a generative study aid to an immersive virtual reality lesson, although this has been shown to improve learning with online science games (Fiorella & Mayer, 2012; Pilegard & Mayer, 2016). We note the value of determining whether study aids can be successfully added to a VR lesson without diminishing the motivational benefits of new media, which is a new research question involving new media. We also note that replication of the sum-

marizing principle, which was based on studies involving reading text (Fiorella & Mayer, 2015), in a new venue involving immersive VR is a worthwhile educational goal according to the focus on replication in Shavelson and Towne's (2002) NRC report, *Scientific Research in Education*, especially because no previous studies on study strategies involved learning in virtual reality. In short, this study is in line with calls for replication of educational intervention results (Hattie, 2009; Phye, Robinson, & Levin, 2005), particularly involving new media (Mayer, 2014b).

Method

Experiment 1

Participants and design. The participants were 55 college students recruited from a subject pool at a university in the western United States (38 women, ages 18–30, $M = 19.18$, $SD = 1.90$). Participants fulfilled a class requirement by participating. Twenty-seven participants were randomly assigned to the immersive VR condition (VR group) and 28 were assigned to the slideshow condition (slideshow group).

Materials. The instructional materials consisted of lessons on how cells in the human bloodstream work rendered in immersive VR simulation (VR lesson) or as a PowerPoint slideshow (slideshow lesson). The VR lesson was an interactive biology simulation, called *The Body VR: Journey Inside a Cell* (The Body VR, 2016), which contained narration and immersive animations of the circulatory system and parts of cells. The simulation “shrunk” the learner, so the learner could travel through an artery on a moving platform while a narrator explained the purpose of the cells within it. The learner then traveled into a cell and the parts and functions of the cell were described. The entire tour included a full 360-degree view of animations within the blood stream and cell. The learner was also occasionally presented with a close-up view of a part of the blood stream or cell, and he or she could physically touch, move, and rotate these objects (e.g., a close-up of a red blood cell or a mitochondrion). The VR lesson lasted approximately 12 min. Screenshots from the VR simulation are shown in Figure 1.

The slideshow lesson was adapted from *The Body VR: Journey Inside a Cell*; the same spoken words from the simulation were transcribed and printed in a slideshow format. Corresponding screenshots for each of the body parts shown in the simulation were also included in the slides to make the lessons as similar as possible (as shown in Figure 2). The PowerPoint slideshow was self-paced and participants completed it in about 8 min on average, ($M = 7$ min, 44 s; $SD = 3$ min, 50 s).

The paper-based materials consisted of a prequestionnaire, postquestionnaire, and posttest. The prequestionnaire solicited basic demographic information such as the participant's age, gender, and year in school. It also asked participants to indicate the science classes they completed in high school and college, to rate their knowledge of the human body on a 5-point scale from “very low” to “very high”, and to check any of the following statements that applied to them: “I have participated in science programs or research fairs,” “Science is my favorite subject in school,” “I earned mostly As and Bs in my science classes in high school and college,” “I enjoy watching science documentaries about biology or anatomy,” “I would like to have a career in a science-related



Figure 1. Screenshots of *The Body VR* (2016). See the online article for the color version of this figure.

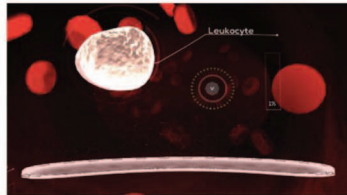
field,” “I can name most of the components of the circulatory system from memory,” “I have attended a course on cardiopulmonary resuscitation (CPR) training,” “I sometimes find myself on the internet looking up science related topics in my free time,” “I can draw the structure of the eukaryotic cell from memory,” and “I took advanced (AP, IB, Honors) science classes in high school.” The prequestionnaire was used to control for preexisting differences between the two groups. We used a self-reported background knowledge questionnaire rather than a pretest about the material in the lesson because we did not want to create a testing effect in which the pretest primes learners to construct answers before the lesson as a form of instruction (Fiorella & Mayer, 2015).

The postquestionnaire asked students to make self-ratings on a 7-point scale from 1 (*strongly disagree*) to 7 (*strongly agree*) of 3 statements of effort and understanding (e.g., “I used a lot of mental effort in the lesson”), a statement on motivation (e.g., “I felt motivated to understand the material”), 4 statements on interest for the subject (e.g., “I am interested in learning more about this subject”), engagement with the lesson (e.g., “I felt that the lesson was engaging), and 6 statements about affect during the lesson (e.g. “I felt happy during the lesson”). The postquestionnaire also had an open-ended question asking if the participant had any additional comments about the lesson.

Finally, the posttest consisted of 20 questions based on the lesson, including 16 factual questions in multiple-choice format and 4 conceptual questions in short-answer format to examine the

White blood cells

- White blood cells, or leukocytes, take up less than 1% of the blood's total volume.
- Their main function is to protect our body from infection.



Cell membrane

- On the outside of the macrophage is a typical cell membrane structure.
- There are thousands of receptors proteins on the surface of the cell.
- Some of these proteins are tasked with transferring information and others with transferring cargo.

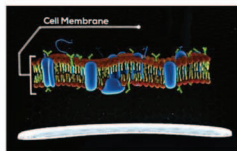


Figure 2. Screenshots of the PowerPoint adapted from *The Body VR* (2016). See the online article for the color version of this figure.

learning outcome of students (Cronbach's $\alpha = .78$). A large portion of the lesson consisted of factual statements, which led the posttest to include more factual questions than comprehension questions. An example of a factual question is, "In what process are ribosomes involved? (A) ATP production. (B) Protein synthesis. (C) Protein transportation. (D) Both B and C." An example of a conceptual question is, "What could cause kinesin motor protein to move slower?"

Apparatus. The VR lesson was presented to participants using Steam Software on a Dell Alienware computer and an HTC Vive, a virtual reality system that included a head-mounted display and two wireless hand controllers. The controllers allowed the user to interact with the virtual environment using intuitive gestures, and users received haptic feedback (i.e., vibrations) for certain interactions. The console also included wall-mounted sensors in the room to allow the software to map the space in which the user could move. The self-paced slideshow lesson was presented on a Dell computer with a 20-inch color screen.

Procedure. Participants signed up for the experiment via a subject recruitment website as part of a requirement for lower division psychology courses. Participants were randomly assigned to the VR group or slideshow group. Participants in the VR group were tested individually in a lab with a large space of approximately 12×12 feet to allow for walking around in the immersive

VR lesson. Participants in the slideshow condition were tested individually in cubicles containing a computer workstation in a lab setting. First, the experimenter described the study and the participant signed the informed consent form. Second, the participant completed the prequestionnaire. Third, the biology lesson was presented. In the VR condition, the participant began by putting on the head-mounted display and standing in the middle of room. The participant could move around the room throughout the 12-min lesson. In the slideshow condition, each participant was seated in an individual cubicle and went through a self-paced PowerPoint slideshow lesson that was presented on a desktop computer. Fourth, participants completed a postquestionnaire about their experiences with the lesson and a posttest on the material they viewed during the lesson, with no time limit. Finally, participants were thanked and dismissed. We followed guidelines for ethical treatment of human subjects and obtained IRB approval for the study.

Experiment 2

Participants and design. The participants were 57 students recruited from the same subject pool as in Experiment 1 (36 women, ages 18–22, $M = 19.56$, $SD = 1.15$). Twenty-nine participants were randomly assigned to the immersive VR condition (VR group) and 28 were assigned to the immersive VR plus summarizing condition (VR+ group). The participants in Experiment 2 were different from those in Experiment 1.

Materials, apparatus, and procedure. The same materials and apparatus from Experiment 1 were used for Experiment 2. The procedure for the VR condition was identical to Experiment 1. In the VR+ condition, participants followed a similar procedure as the VR condition, except that they viewed the lesson in 6 segments. After each segment, the participant removed the headset and was asked to write a summary of the segment they just viewed on a sheet of paper. Participants were told to write a summary as completely as they could and were given unlimited time to answer the prompt, which stated to summarize key features of the segment (e.g., "Summarize the functions of red blood cells, white blood cells, platelets, and monocytes."). The postlesson questionnaire also included two additional items about being distracted and overwhelmed during the lesson.

Results

Scoring

The prequestionnaire was scored with a point given for each science class participants completed in high school and college and a point for each science-related activity they checked off. For the self-rating of knowledge of the human body, zero points were given for a rating of very low and somewhat low, a point was given for average, two points for somewhat high, and three points for very high. The posttest was scored out of 20 points, with a point given for each correct multiple-choice and short-answer question; half-points were given for partially correct answers on short-answer question. The short answer questions were scored based on a rubric that indicated the words and phrases required for 1 point or 1/2 point.

Do the Groups Differ on Basic Characteristics?

The first step was to determine whether there were any preexisting differences between the groups. In Experiment 1, the VR and slideshow groups did not differ significantly in mean age, $t(53) = 0.13, p = .899$, mean year in school, $t(53) = 1.27, p = .210$, and mean science-knowledge score, $t(53) = 1.33, p = .190$. The two groups also did not differ in the proportion of men and women, $\chi^2(1, N = 55) = 1.87, p = .171$. We conclude that groups did not differ on basic characteristics.

In Experiment 2, the VR and VR+ groups did not differ significantly in mean age, $t(55) = 1.22, p = .227$, and mean science-knowledge score, $t(55) = 0.18, p = .857$. The two groups also did not differ in their proportion of men and women, $\chi^2(1, N = 57) = 0.30, p = .862$. The VR+ group ($M = 2.21, SD = 1.13$) was marginally further along in school based on year in school than the VR group ($M = 1.76, SD = 0.83$), $t(55) = 1.74, p = .088$; therefore, year in school was included as a covariate in subsequent analyses in Experiment 2.

Does Learning in Immersive VR Lead to Better Learning Outcomes Than Viewing a Corresponding Slideshow?

A primary media-comparison question in Experiment 1 concerns whether students learn better from the VR lesson or slideshow lesson. Table 1 shows the mean score on the posttest (as well as the mean score on the factual items and conceptual items separately) for the two groups. An analysis of covariance (ANCOVA) was conducted on the posttest scores of the VR and slideshow groups with background science-knowledge scores as covariates. As shown in the top line of Table 1, the slideshow group ($M = 13.54, SD = 3.55$) scored significantly better than the VR group ($M = 10.17, SD = 3.80$) on the posttest overall, $F(1, 52) = 9.54, p = .003, d = 0.92$. As shown in the next two lines of Table 1, additional ANCOVAs revealed that the slideshow group performed significantly better than the VR group on the factual questions (slideshow group: $M = 11.00, SD = 2.74$; VR group: $M = 7.74, SD = 3.07$), $F(1, 52) = 14.94, p < .001, d = 1.12$; but not on the conceptual questions (slideshow group: $M = 2.53, SD = 1.17$; VR group: $M = 2.43, SD = 1.19$), $F(1, 52) = 0.03, p = .870, d = 0.08$. The same pattern of significant differences was found when we conducted t tests (i.e., without covariates). It is worth noting that the slideshow group spent less time on the lesson, but performed better on the posttest. This further shows

that the slideshow was more efficient at conveying scientific information than the VR lesson. We conclude that students learned less from a biology lesson when it was presented in immersive VR than when it was presented as a slideshow on a desktop computer.

Do Students Give More Positive Ratings to Learning in Immersive VR Than Learning From a Corresponding Slideshow?

A secondary media-comparison question in Experiment 1 concerns whether students give more positive self-report ratings to the VR lesson or the slideshow lesson. Table 2 shows the mean rating (and standard deviation) on each of 15 items for the VR and slideshow groups. As summarized in Table 2, t tests (corrected for multiple comparisons, $p < .001$) showed that the VR group gave significantly better ratings on items involving enjoyment, engagement, and motivation. Additionally, concerning affect, the VR group was significantly happier, more excited, and less bored during the lesson than the slideshow group. We conclude that the students have more positive feelings about learning in immersive VR than learning from a slideshow.

Do the Groups Differ in Their Comments About the Lesson?

As a secondary analysis, we also examined participants' general comments about their experiences with the lesson. The participants in the VR group (4 out of 28) were significantly more likely to mention being distracted or having a hard time focusing on the dialogue than the participants in the slideshow group (0 out of 27), $\chi^2(1, N = 55) = 4.16, p = .041$. For example, one VR participant wrote, "I was somewhat distracted by the excitement of experiencing a new technology. Because of that, I wasn't able to completely focus on the lesson."

Does Summarizing a Segmented VR Lesson Lead to Better Learning Outcomes than Viewing a Continuous VR Lesson Without Summarizing?

A primary value-added issue addressed in Experiment 2 concerns whether students learn more when they are asked to write written summaries after each segment of the VR lesson. Table 3 shows the mean score on the posttest (as well as the mean score on the factual items and conceptual items separately) for the VR and VR+ groups in Experiment 2. An ANCOVA was run on the posttest scores between the VR and VR+ groups with background science-knowledge scores and year in school as covariates. As shown in the top row of Table 3, the VR+ group ($M = 13.84, SD = 3.20$) scored significantly better than the VR group ($M = 10.31, SD = 3.10$) on the posttest overall, $F(1, 53) = 17.81, p < .001, d = 1.12$. As shown in the next two lines of Table 3, additional ANCOVAs revealed that the VR+ group performed significantly better than the VR group on the factual questions (VR+: $M = 10.89, SD = 2.59$; VR: $M = 8.07, SD = 2.33$; $F(1, 53) = 17.41, p < .001, d = 1.14$) and the conceptual questions (VR+: $M = 2.95, SD = 1.09$; VR: $M = 2.24, SD = 1.07$; $F(1, 53) = 4.58, p = .037, d = 0.66$). The same pattern of significant differences was found when we conducted t tests, without covariates. We conclude that students learned more from a biology

Table 1
Mean Scores and Standard Deviations on Posttest for Virtual Reality (VR) and Slideshow Groups

Test score	VR group ($N = 27$)		Slideshow group ($N = 28$)		d
	M	SD	M	SD	
Total test score (out of 20)	10.17	3.80	13.54**	3.55	0.92
Factual questions (out of 16)	7.74	3.07	11.00***	2.74	1.12
Conceptual questions (out of 4)	2.43	1.19	2.53	1.17	0.08

** $p < .01$. *** $p < .001$.

Table 2
Mean Ratings of Interest, Motivation, Engagement, and Affective States During the Lesson
Between by the Virtual Reality (VR) and Slideshow Groups

Post-Questionnaire Item	VR group (<i>N</i> = 27)	Slideshow group (<i>N</i> = 28)
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
"I used a lot of mental effort in the lesson"	4.00 (1.21)	4.00 (1.16)
"I felt that the subject matter was difficult"	3.30 (1.14)	2.89 (1.23)
"I have a good understanding of the material"	4.85 (1.03)	5.00 (1.09)
"I enjoyed learning this way"	5.89 (1.40)	3.61 (1.37)***
"I would like to learn this way in the future"	5.81 (1.52)	3.32 (1.42)***
"I am interested in learning more about this subject"	5.70 (1.17)	4.96 (1.29)
"I felt that the lesson was engaging"	6.11 (1.05)	3.32 (1.34)***
"I found the lesson to be useful to me"	5.56 (1.40)	4.61 (1.40)
"I felt motivated to understand the material"	5.93 (1.07)	4.11 (1.45)***
"I felt happy during the lesson"	5.67 (1.49)	3.43 (1.07)***
"I felt excited during the lesson"	5.81 (1.21)	3.07 (1.22)***
"I felt bored during the lesson"	1.81 (1.00)	4.25 (1.51)***
"I felt confused during the lesson"	2.15 (1.10)	2.39 (1.20)
"I felt sad during the lesson"	1.07 (0.27)	1.71 (0.94)
"I felt scared during the lesson"	1.48 (1.09)	1.25 (.044)

Note. A 7-point rating scale from 1 (*strongly disagree*) to 7 (*strongly agree*) was used.

*** $p < .001$.

lesson when it was presented in a segmented immersive VR lesson with summarization prompts than when it was presented as continuous VR lesson. In short, students had better learning outcomes from a VR simulation when they were asked to engage in a summarizing learning strategy. We note that adding the summarization prompt allowed the VR+ group to achieve a level of test performance similar to the slideshow group in Experiment 1.

Does Adding Summary Prompts to a VR Lesson Diminish Positive Ratings Toward the Lesson?

A secondary value-added issue concerns whether adding summarization prompts to a VR simulation lesson affects how students rate the lesson. Table 4 shows the mean rating (and standard deviation) on each of 17 items for the VR and VR+ groups. As summarized in Table 4, *t* tests (corrected for multiple comparisons, $p < .001$) showed that the VR group and VR+ group did not give significantly different ratings on interest, motivation, engagement, or affective states after the lesson. The conclusion applies when we conducted *t* tests without correcting for multiple comparisons. We conclude that adding summary prompts to a VR lesson does not diminish positive ratings toward the lesson.

Table 3
Mean Scores and Standard Deviations on Posttest for
Summarizing Group and Control Group

Test score	VR+ group (<i>N</i> = 28)		VR group (<i>N</i> = 29)		<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Total test score (out of 20)	13.83	3.21	10.31***	3.10	1.12
Factual questions (out of 16)	10.89	2.59	8.07***	2.33	1.14
Conceptual questions (out of 4)	2.95	1.09	2.24*	1.07	0.66

* $p < .05$. *** $p < .001$.

Discussion

Empirical Contributions

Experiment 1 was a media comparison study showing that students who viewed an immersive VR lesson reported significantly higher ratings of motivation, interest, engagement, and affect than students who viewed a slideshow lesson covering the same material, but scored significantly worse on a posttest, particularly on the factual questions.

Experiment 2 was a value-added study showing that adding a generative learning strategy, summarizing, to the existing VR lesson significantly improved learning outcomes compared with the original VR lesson, but did not significantly change ratings of motivation, interest, engagement, or affect. This is consistent with previous literature showing the effect of summarizing during a lesson (Fiorella & Mayer, 2015, 2016).

Theoretical Contributions

The results from Experiment 1 indicate that the instructional features used in the immersive VR lesson may have not been as effective as those in the slideshow lesson. This could have been due to three possibilities: the coherence principle, the segmentation principle, or higher learner control. First, the coherence principle, which calls for eliminating extraneous material and features, seemed to play the biggest role in hindering the students' learning in immersive VR. The learner is immersed in constant animations in a 360-degree view that are not essential for some of the narration, thereby violating the coherence principle. For example, in the part of the lesson in which the learner was in the blood stream, various blood cells were constantly moving past the learner and the learner could look in any direction to see these movements. These animations could have added to the learner's cognitive load as he or she had to also pay attention to the narration; they may have

Table 4
*Mean Ratings of Interest, Motivation, Engagement, and Affective States During the Lesson
 Between by the Virtual Reality (VR) and VR+ Groups*

Post-Questionnaire Item	VR group (N = 29)	VR+ group (N = 28)
	M (SD)	M (SD)
"I used a lot of mental effort in the lesson"	4.38 (1.12)	4.61 (1.64)
"I felt that the subject matter was difficult"	3.21 (1.15)	3.43 (1.53)
"I have a good understanding of the material"	4.76 (1.12)	4.79 (1.29)
"I enjoyed learning this way"	6.07 (1.10)	5.64 (1.57)
"I would like to learn this way in the future"	5.59 (1.45)	5.46 (1.80)
"I am interested in learning more about this subject"	5.31 (1.26)	5.64 (1.47)
"I felt that the lesson was engaging"	6.07 (0.96)	5.79 (0.96)
"I found the lesson to be useful to me"	5.24 (1.22)	5.04 (1.32)
"I felt motivated to understand the material"	5.66 (1.17)	5.61 (1.26)
"I felt happy during the lesson"	5.59 (1.15)	5.54 (1.07)
"I felt excited during the lesson"	6.03 (1.02)	5.86 (.97)
"I felt bored during the lesson"	1.62 (1.08)	2.29 (1.30)
"I felt distracted during the lesson"	3.66 (1.95)	4.64 (1.77)
"I felt confused during the lesson"	2.55 (1.43)	2.79 (1.50)
"I felt overwhelmed during the lesson"	3.24 (1.79)	3.57 (1.57)
"I felt sad during the lesson"	1.34 (0.77)	1.36 (0.56)
"I felt scared during the lesson"	1.93 (1.49)	1.46 (0.92)

Note. A 7-point rating scale from 1 (*strongly disagree*) to 7 (*strongly agree*) was used.

diverted attention from the important material, disrupting the process of organizing the material, and may have primed the learner to integrate the material with prior knowledge in an inappropriate manner.

Second, the VR simulation did not adhere closely to the segmenting principle because it was a continuous lesson that was not under learner control, whereas the segments were short and under learner control in the self-directed slideshow. As the learner listened to and watched a lesson, he or she may still have been processing previous information when new information was presented, adding to his or her essential processing load. Additionally, because the PowerPoint was user-paced, it added a higher element of learner control, mitigating the learner's essential processing during the lesson.

In short, the slideshow lesson may have incorporated research-based principles of instructional design more effectively than did the VR lesson. The off-the-shelf educational simulation used in this study was meant to be an entertaining interactive experience, so it was likely not necessarily designed with the principles of multimedia learning in mind. Additionally, this particular lesson contained mostly factual knowledge, which animations may not have enhanced, but rather added extraneous visual input. Supporting this conclusion were a few comments written by participants in the VR condition, which mentioned that they were distracted during the lesson, whereas no participants in the slideshow condition mentioned they were distracted. It appears that immersive VR creates situational interest, but as has long been noted, adding interesting features to a lesson may not be enough to enhance learning (Dewey, 1913). In short, immersive VR may create so much extraneous cognitive processing that the learner does not have sufficient cognitive resources left to learn the essential material in the lesson.

The results from Experiment 2 provide evidence for generative learning theory in that promoting meaningful learning through asking learners to engage in the learning strategy of

summarizing caused enhanced learning gains from a lesson. Creating summaries during breaks in the VR lesson prompted the learners to select, organize, and integrate the information from the lesson into their existing knowledge structures. This work shows that generative learning strategies that have been shown to be effective in non-VR environments can also apply to learning in VR environments. This study is the first to show that conventional study strategies such as summarizing can be applied to learning in VR without harming student interest, which extends research on study strategies and broadens the domain of impact for generative learning theory.

Practical Contributions

On the practical side, these results suggest that there may not be strong research evidence to invest in the costly job of converting conventional science instruction or even simulations in desktop virtual reality into simulations in immersive virtual reality. When the goal is to help students learn basic scientific knowledge, we cannot recommend wholesale replacing of conventional media with immersive VR. However, in situations where VR is the instructional medium, the effectiveness of VR can be increased by prompting students to use generative learning strategies, such as summaries, without diminishing the learner's motivation, interest, engagement, and affect while using the new technology. Because VR is a new popular technology that students find more motivating than conventional media, using it in tandem with these strategies may be useful in sparking interest among learners while still maintaining learning outcomes comparable to conventional media.

Methodological Contributions

Clark (2001) and Saettler (1990, 2004) have warned researchers about the perils of media comparison research in which learning

with one medium is compared against learning the same content with another medium. An ever-present threat is that the two treatments differ not only in instructional medium but also in terms of instruction method and even instructional content. The present study offered some experimental control by insuring that both lessons contained the same words (albeit in spoken vs. printed form) and same graphics (albeit in dynamic vs. static form). Thus, we tried to mitigate some of the more gross problems in media comparison studies.

Limitations and Future Directions

A limitation of this study is that it involved one, short off-the-shelf instructional lesson delivered in a lab setting with an immediate test. Although this study gives us a preliminary view of the effectiveness of immersive VR for teaching scientific material and contributes to the small research base, future work should involve other types of material, more authentic learning contexts, and delayed tests. A second limitation concerns the generalizability of this study to other content domains. Further research is needed to examine whether these results are domain, or even lesson, specific and whether they are generalizable to other technologies. In addition, future studies are needed to examine whether applying the principles of multimedia learning design would also help learners in immersive VR. For example, would better incorporation of the coherence principle or the segmenting principle in an interactive lesson, like *The Body VR: Journey Inside a Cell*, be effective at increasing learning outcomes to the same levels as learning from a PowerPoint slideshow? Additionally, it would be useful to examine in a value-added type study whether something else could be added to a VR lesson to make it more effective. For example, adding other generative learning activities, such as drawing (Fiorella & Mayer, 2015), to an existing VR lesson may help enhance deeper learning, or adding an interactive activity with feedback during the lesson may help enhance the learner's self-efficacy along with his or her motivation for the lesson. It may be useful in future research to investigate the benefits of using immersive VR as a pre-lesson activity to spark situational interest that may support subsequent learning from a conventional lesson.

The goal of experiment 2 was to determine whether adding a generative learning strategy (in this case, summarizing) to an immersive reality lesson could improve academic learning without hurting motivation and interest. This is a worthwhile issue because the exciting features of a new instructional medium like immersive VR may cause learners to be less reflective, so interventions aimed at promoting deeper processing may be useful. Our goal was not to examine the relative power of summarizing with two different media (slideshows and VR), although this is certainly a question worthy of future research.

References

- Adams, D. M., Mayer, R. E., MacNamara, A., Koenig, A., & Wainess, R. (2012). Narrative games for learning: Testing the discovery and narrative hypotheses. *Journal of Educational Psychology, 104*, 235–249. <http://dx.doi.org/10.1037/a0025595>
- Anderson, J., & Barnett, M. (2011). Using video games to support pre-service elementary Teacher learning of basic physics principles. *Journal of Science Education and Technology, 20*, 347–362. <http://dx.doi.org/10.1007/s10956-010-9257-0>
- Annis, L. F. (1985). Student-generated paragraph summaries and the information-processing theory of prose learning. *Journal of Experimental Education, 54*, 4–10. <http://dx.doi.org/10.1080/00220973.1985.10806390>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191–215. <http://dx.doi.org/10.1037/0033-295X.84.2.191>
- Blascovich, J., & Bailenson, J. (2011). *Infinite reality*. New York, NY: HarperCollins.
- Brelsford, J. W. (1993, October). *Physics education in a virtual environment*. Paper presented at proceedings of the human factors and ergonomics society 37th meeting, Boston, MA.
- Chemers, M. M., Hu, L., & Garcia, B. F. (2001). Academic self-efficacy and first-year college student performance and adjustment. *Journal of Educational Psychology, 93*, 55–64. <http://dx.doi.org/10.1037/0022-0663.93.1.55>
- Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*, 445–459. <http://dx.doi.org/10.3102/00346543053004445>
- Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development, 42*, 21–29. <http://dx.doi.org/10.1007/BF02299088>
- Clark, R. E. (2001). *Learning from media*. Greenwich, CT: Information Age Publishing.
- Coleman, E. B., Brown, A. L., & Rivkin, I. D. (1997). The effect of instructional explanations on learning from scientific texts. *Journal of the Learning Sciences, 6*, 347–365.
- Cooper, T. (2007). Nutrition game. In D. Livingstone & J. Kemp (Eds), *Proceedings of the second life education workshop* (pp. 47–50). Chicago, IL: Second Life Education.
- Dewey, J. (1913). *Interest and effort in education*. Boston, MA: Houghton Mifflin. <http://dx.doi.org/10.1037/14633-000>
- Doctorow, M., Wittrock, M. C., & Marks, C. (1978). Generative processes in reading comprehension. *Journal of Educational Psychology, 70*, 109–118. <http://dx.doi.org/10.1037/0022-0663.70.2.109>
- Evans, K. L., Yaron, D., & Leinhardt, G. (2008). Learning stoichiometry: A comparison of text and multimedia formats. *Chemistry Education Research and Practice, 9*, 208–218. <http://dx.doi.org/10.1039/B812409B>
- Fiorella, L., & Mayer, R. E. (2012). Paper-based aids for learning with a computer-based game. *Journal of Educational Psychology, 104*, 1074–1082. <http://dx.doi.org/10.1037/a0028088>
- Fiorella, L., & Mayer, R. E. (2015). *Learning as a generative activity: Eight learning strategies that promote understanding*. New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781107707085>
- Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review, 28*, 717–741. <http://dx.doi.org/10.1007/s10648-015-9348-9>
- Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology, 89*, 92–102. <http://dx.doi.org/10.1037/0022-0663.89.1.92>
- Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90*, 414–434. <http://dx.doi.org/10.1037/0022-0663.90.3.414>
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.
- Hew, K. F., & Cheung, W. S. (2010). Use of three-dimensional (3-D) immersive virtual worlds in K–12 and higher education settings: A review of the research. *British Journal of Educational Technology, 41*, 33–55. <http://dx.doi.org/10.1111/j.1467-8535.2008.00900.x>

- Hilton, M. A., & Honey, M. L. (2011). *Learning science through computer games and simulations*. Washington, DC: The National Academy Press.
- Hwang, G. C., Wu, P. H., & Chen, C. C. (2012). An online game approach for improving students' learning performance in Web-based problem-solving activities. *Computers & Education*, 59, 1246–1256. <http://dx.doi.org/10.1016/j.compedu.2012.05.009>
- Kintsch, W. (1980). Learning from text, levels of comprehension, or: Why anyone would read a story anyway. *Poetics*, 9, 87–98. [http://dx.doi.org/10.1016/0304-422X\(80\)90013-3](http://dx.doi.org/10.1016/0304-422X(80)90013-3)
- Kozhevnikov, M., Gurliitt, J., & Kozhevnikov, M. (2013). Learning relative motion concepts in immersive and non-immersive virtual environments. *Journal of Science Education and Technology*, 22, 952–962. <http://dx.doi.org/10.1007/s10956-013-9441-0>
- Kozma, R. B. (1994). The influence of media on learning: The debate continues. *School Library Media Research*, 22. Retrieved from http://www.ala.org/aasl/sites/ala.org.aasl/files/content/aaslpubsandjournals/slr/edchoice/SLMQ_InfluenceofMediaonLearning_InfoPower.pdf
- Lee, E. A., & Wong, K. W. (2014). Learning with desktop virtual reality: Low spatial ability learners are more positively affected. *Computers & Education*, 79, 49–58. <http://dx.doi.org/10.1016/j.compedu.2014.07.010>
- Lee, H., Plass, J. L., & Homer, B. D. (2006). Optimizing cognitive load for learning from computer-based science simulations. *Journal of Educational Psychology*, 98, 902–913. <http://dx.doi.org/10.1037/0022-0663.98.4.902>
- Ligorio, M. B., & van Veen, K. (2006). Constructing a successful cross-national virtual learning environment in primary and secondary education. *AACE Journal*, 14, 103–128.
- Lowe, R., & Ploetzner, R. (Eds.). (2017). *Learning from dynamic visualization*. New York, NY: Springer. <http://dx.doi.org/10.1007/978-3-319-56204-9>
- Mayer, R. E. (2008). Teaching by priming students' motivation to learn. In R. E. Mayer (Ed.), *Learning and instruction* (pp. 458–522). Upper Saddle River, NJ: Pearson Education.
- Mayer, R. E. (2009). *Multimedia learning*. New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511811678>
- Mayer, R. E. (2014a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139547369.005>
- Mayer, R. E. (2014b). *Computer games for learning: An evidence-based approach*. Cambridge, MA: MIT Press.
- Mayer, R. E. (2014c). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 345–368). New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139547369.017>
- Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology*, 88, 64–73. <http://dx.doi.org/10.1037/0022-0663.88.1.64>
- Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology*, 93, 390–397. <http://dx.doi.org/10.1037/0022-0663.93.2.390>
- Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology*, 95, 806–812. <http://dx.doi.org/10.1037/0022-0663.95.4.806>
- Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 279–315). New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139547369.015>
- Mayer, R. E., Griffith, E., Naftaly, I., & Rothman, D. (2008). Increased interestingness of extraneous details leads to decreased learning. *Journal of Experimental Psychology: Applied*, 14, 329–339. <http://dx.doi.org/10.1037/a0013835>
- Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied*, 11, 256–265. <http://dx.doi.org/10.1037/1076-898X.11.4.256>
- Mayer, R. E., Heiser, H., & Lonn, S. (2001). Cognitive constraints on multi-media learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, 93, 187–198. <http://dx.doi.org/10.1037/0022-0663.93.1.187>
- Mayer, R. E., & Jackson, C. I. (2005). The case for coherence in science explanations: Quantitative details hurt qualitative understanding. *Journal of Educational Psychology*, 100, 380–386. <http://dx.doi.org/10.1037/0022-0663.100.2.380>
- Mayer, R. E., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 316–344). New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139547369.016>
- McLaren, B. M., Adams, D. M., Mayer, R. E., & Forlizzi, J. (2017). A computer-based game that promotes mathematics learning more than a conventional approach. *International Journal of Game-Based Learning*, 7, 36–56. <http://dx.doi.org/10.4018/IJGBL.2017010103>
- Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K–12 and higher education. *Computers & Education*, 70, 29–40. <http://dx.doi.org/10.1016/j.compedu.2013.07.033>
- Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia messages. *Journal of Educational Psychology*, 92, 117–125. <http://dx.doi.org/10.1037/0022-0663.92.1.117>
- Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality environments: Role of methods and media. *Journal of Educational Psychology*, 92, 724–733. <http://dx.doi.org/10.1037/0022-0663.92.4.724>
- Moreno, R., & Mayer, R. E. (2004). Personalized messages that promote science learning in virtual environments. *Journal of Educational Psychology*, 96, 165–173. <http://dx.doi.org/10.1037/0022-0663.96.1.165>
- Moreno, R., Mayer, R. E., Spire, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, 19, 177–213. http://dx.doi.org/10.1207/S1532690XC11902_02
- Peper, R. J., & Mayer, R. E. (1986). Generative effects of note-taking during science lectures. *Journal of Educational Psychology*, 78, 34–38. <http://dx.doi.org/10.1037/0022-0663.78.1.34>
- Phye, G. D., Robinson, D. H., & Levin, J. (Eds.). (2005). *Empirical methods for evaluating educational interventions*. San Diego, CA: Elsevier Academic Press.
- Pietsch, J., Walker, R., & Chapman, E. (2003). The relationship among self-concept, self-efficacy, and performance in mathematics during secondary school. *Journal of Educational Psychology*, 95, 589–603. <http://dx.doi.org/10.1037/0022-0663.95.3.589>
- Pilegard, C., & Mayer, R. E. (2016). Improving academic learning from computer-based narrative games. *Contemporary Educational Psychology*, 44, 12–20. <http://dx.doi.org/10.1016/j.cedpsych.2015.12.002>
- Pintrich, P. R. (2003). Motivation and classroom learning. In W. M. Reynold & G. E. Miller (Eds.), *Handbook of psychology: Vol. 7. Educational psychology* (pp. 103–122). New York, NY: Wiley.

- Richards, D., & Taylor, M. (2015). A comparison of learning gains when using a 2D simulation tool versus a 3D virtual world: An experiment to find the right representation involving the marginal value theorem. *Computers & Education*, 86, 157–171. <http://dx.doi.org/10.1016/j.compedu.2015.03.009>
- Ross, S. M., & Kirby, F. J. (1976). Oral summary as a review strategy for enhancing recall of textual material. *Journal of Educational Psychology*, 68, 689–695. <http://dx.doi.org/10.1037/0022-0663.68.6.689>
- Saettler, P. (1990/2004). *The evolution of American educational technology*. Greenwich, CT: Information Age Publishing.
- Salomon, G. (1979). *Interaction of Media, Cognition, and Learning*. San Francisco: Jossey-Bass.
- Schiefele, U. (2009). Situational and individual interest. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation in school* (pp. 197–223). New York, NY: Taylor & Francis.
- Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 183–212). Hillsdale, NJ: Erlbaum.
- Schunk, D. H. (1989). Self-efficacy and achievement behaviors. *Educational Psychology Review*, 1, 173–208. <http://dx.doi.org/10.1007/BF01320134>
- Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, 26, 207–231. <http://dx.doi.org/10.1080/00461520.1991.9653133>
- Schunk, D. H., & DiBenedetto, M. K. (2016). Self-efficacy theory. In K. R. Wentzel & D. B. Miele (Eds.), *Handbook of motivation at school* (2nd ed., pp. 34–54). New York, NY: Routledge.
- Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influences on children's self-efficacy and achievement. *Journal of Educational Psychology*, 77, 313–322. <http://dx.doi.org/10.1037/0022-0663.77.3.313>
- Shavelson, R. J., & Towne, L. (Eds.). (2002). *Scientific research in education*. Washington, DC: National Academy Press.
- Sourin, A., Sourina, O., & Prasolova-Førland, E. (2006). Cyber-learning in cyberworlds. *Journal of Cases on Information Technology*, 8, 55–70. <http://dx.doi.org/10.4018/jcit.2006100105>
- Swaak, J., de Jong, T., & van Joolingen, W. R. (2004). The effects of discovery learning and expository instruction on the acquisition of definitional and intuitive knowledge. *Journal of Computer Assisted Learning*, 20, 225–234. <http://dx.doi.org/10.1111/j.1365-2729.2004.00092.x>
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory*. New York, NY: Springer. <http://dx.doi.org/10.1007/978-1-4419-8126-4>
- The Body VR. (2016). *The Body VR: Journey inside a cell* [Video Game]. New York, NY: The Body VR LLC.
- Torkington, J., Smith, S. G., Rees, B. I., & Darzi, A. (2001). Skill transfer from virtual reality to a real laparoscopic task. *Surgical Endoscopy*, 15, 1076–1079. <http://dx.doi.org/10.1007/s004640000233>
- Wade, S. E. (1992). How interest affects learning from text. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 255–278). Hillsdale, NJ: Erlbaum.
- Webster, R. (2016). Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments*, 24, 1319–1333. <http://dx.doi.org/10.1080/10494820.2014.994533>
- Wentzel, K., & Miele, D. B. (Eds.). (2016). *Handbook of motivation at school* (2nd ed.). New York, NY: Routledge.
- Wigfield, A., Tonks, S., & Klauda, S. L. (2016). Expectancy-value theory. In K. R. Wentzel & D. B. Miele (Eds.), *Handbook of motivation in school* (2nd ed; pp. 55–74). New York, NY: Routledge.
- Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist*, 24, 345–376. http://dx.doi.org/10.1207/s15326985ep2404_2
- Wouters, P., & van Oostendorp, H. (2017). Overview of instructional techniques to facilitate learning and motivation of serious games. In P. Wouters & H. van Oostendorp (Eds.), *Instructional techniques to facilitate learning and motivation of serious games* (pp. 1–16). New York, NY: Springer. http://dx.doi.org/10.1007/978-3-319-39298-1_1
- Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82, 51–59. <http://dx.doi.org/10.1037/0022-0663.82.1.51>

Received June 5, 2017

Revision received August 10, 2017

Accepted September 13, 2017 ■