2015

Learning Biology through Innovative Curricula: A Comparison of Game- and Nongame-Based Approaches

Troy D. Sadler

William L. Romine

Deepika Menon

Richard E. Ferdig

Leonard Annetta

Published in *Science Education* 99:4 (2015), pp. 696–720; doi: 10.1002/sce.21171

Submitted May 9, 2014; revised December 22, 2014; accepted February 24, 2015; published online May 19, 2015.

Learning Biology through Innovative Curricula: A Comparison of Game- and Nongame-Based Approaches

Troy D. Sadler,[1] William L. Romine,[2] Deepika Menon,[1] Richard E. Ferdig,[3] and Leonard Annetta[4]

1. The ReSTEM Institute: Reimaging and Researching STEM Education, University of Missouri, Columbia, Missouri, USA
2. Department of Biological Sciences, Wright State University, Dayton, Ohio, USA
3. Research Center for Educational Technology, Kent State University, Kent, Ohio, USA
4. Graduate School of Education, George Mason University, Fairfax, Virginia, USA

Corresponding author – Troy D. Sadler, The ReSTEM Institute: Reimaging and Researching STEM Education, University of Missouri, Columbia, MO 65211, USA, email sadlert@missouri.edu

Abstract
This study explored student learning in the context of innovative biotechnology curricula and the effects of gaming as a central element of the learning experience. The quasi-experimentally designed study compared learning outcomes between two curricular approaches: one built around a computer-based game and the other built around a narrative case. The research questions addressed student learning of basic biological principles, development of interest in learning science, and how a game-based approach compared to a nongame-based approach in terms of supporting learning. The study employed a pre-post design with 1,888 high school students nested within the classes of 36 biology teachers. Results indicated that students participating in both approaches demonstrated statistically and practically significant gains on both proximal and distal assessments of biological content knowledge. Neither group demonstrated gains in science interest. The curriculum by time interaction was not statistically different, indicating that students in both groups showed similar results. Implications for game-based science learning and future research include building better awareness of technological and professional development challenges associated with implementing educational games, the

need for new strategies for understanding the impacts of games for learning, and the need for cost-benefit analyses in the planning of game-based educational approaches.

Introduction

Video games have become popular media for many segments of the population including children and adolescents (Entertainment Software Association, 2013; NPD Group, 2009). Technology use among school-aged individuals is higher now than ever before in history with more than 90% of students reporting access to and regular use of computers or mobile devices (Madden, Lenhart, Duggan, Cortesi, & Gasser, 2013). Over 95% of teens indicate that they have played video games and about half report daily game play (Lenhart et al., 2008). For well over a decade, educators have been considering ways in which computer-based games may be leveraged to support learning (Squire & Jenkins, 2003), and a growing body of evidence supports the conclusion that students can learn through game play and that games can be purposefully designed to foster desired learning outcomes (Honey & Hilton, 2011; Squire, 2011).

Research on educational games has documented the successful use of computer-based games for supporting a variety of learning outcomes important for science education. These outcomes include understandings of content (Clark et al., 2011), interest in science (Kuo, 2007), inquiry skills (Ketelhut, Nelson, Clarke, & Dede, 2010), creativity (Annetta, Cheng, & Holmes, 2010), scientific habits of mind (Steinkuehler & Duncan, 2008), and critical thinking (Squire, 2006). A recent report from the National Research Council (NRC) on games and education concludes that

> Simulation and games have potential to advance multiple science learning goals, including motivation to learn science, conceptual understanding, science process skills, understanding of the nature of science, scientific discourse and argumentation, and identification with science and science learning. (Honey & Hilton, 2011, p. 54)

The report goes on to call for additional research to better inform the nature and extent of learning through games in various educational settings such as classroom environments. So, while it appears that games *can* support learning, many questions remain unanswered, such as how games can be used in classrooms, how games may interact with various content domains, and how much learning takes place in games relative to other approaches.

This study addressed some of the unanswered issues regarding the use of games in science education. We explored the extent to which a biotechnology-based computer game supported high school students' development of interest in learning science and understanding of core concepts in biology. To better understand the relative impacts of the game, we compared learning outcomes between students who learned with the game and others who experienced an alternative curriculum similar in scope and sequence but without the game element.

Games and Education

What Is a Game?

Over the past few years, there has been extensive discussion about the use of games in education and the "gamification" of education (Kapp, 2012). While the use of computer-based games is a relatively new trend in education, educators have long used games facilitated through other media (Ferdig, 2009). Salen and Zimmerman (2004) offer a useful framework for considering what constitutes a game. Based on game-related work from diverse disciplines including history, anthropology, sociology, philosophy, and education, these authors define a game as "a system in which players engage in an artificial conflict, defined by rules, that result in a quantifiable outcome" (Salen & Zimmerman, 2004, Chap. 7). According to this framework, game systems define a boundary between an artificial game environment and the real world, and the artificial environment is partially defined by objects, attributes, rules, and a set of internal relationships. Games prescribe a conflict or contest that players negotiate through decision making and actions. As players make progress toward the resolution of the conflict or contest, they earn points or other quantifiable markers denoting success (or failure). Games can be delivered through multiple media requiring varying levels of technology (from no tech to high tech); in the current study, we explored a game mediated through a computer-based virtual environment.

A Framework for Games in Classrooms

An important question for educators has been whether or not students learn from game play. The simple answer to this question is a resounding yes; games have consistently been shown to support learning and, in some cases, development of expertise (Gee, 2007; Shaffer, 2006). The fact that students *can learn* through game play has been established, but questions have been raised about *what* students can learn through games and the extent to which learning facilitated by games corresponds to the kinds of learning outcomes valued and prioritized in schools (e.g., Barko & Sadler, 2013; Bosche, 2010; Gentile, Lynch, Linder, & Walsh, 2004). While some of the things students might learn through playing games are well beyond the canons of formal education, research has documented the potential for games to support student progress on many learning outcomes valued within school systems.

In the area of science education, researchers have explored numerous games and their potential to support a variety of science learning outcomes. Student experiences with several different video games have been linked to the learning of science content. For example, *River City*, a game which brings middle school students back in time to help a town struggling with infectious disease, has been shown to help students learn about the science of disease transmission (Nelson, 2007). In the game *Whyville*, middle school students navigate a virtual world and unknowingly spread a virtual virus. Here again, game play has been linked with student learning of disease transmission. Games such as *Mad City Mystery* and *Quest Atlantis* create opportunities for students to explore mysterious occurrences that are ultimately linked back to environmental problems. Research on these games suggests that they help students make significant gains in their understandings of human-environment interactions (Hickey, Ingram-Goble, & Jameson, 2009; Squire & Jan, 2007). These examples highlight learning related to the biological sciences, and these results are most pertinent to the current study because of its focus on biotechnology and biology; however, other research supports the use of games for learning physics (Clark et al., 2011), chemistry (Rastegapour & Marashi, 2012), geology (Mayer, 2011), and astronomy (Ruzhitskaya et al., 2013).

In addition to these findings on science content learning, studies have addressed the potential for games to support development of interest in science and in learning science. In one investigation of *Quest Atlantis*, upper-level elementary students expressed increased interest in learning science associated with game play (Barab, Sadler, Heiselt, Hickey, & Zuiker, 2007). In *Crystal Island*, a three-dimensional (3D) narrative-based game for middle school learners, players explore the nature and cause of an infectious disease. Successful game play was associated with gains in interest in science and science self-efficacy (Rowe, Shores, Mott, & Lester, 2010). *Go Go Bugs* is an online game for elementary students designed to teach basic ideas about entomology. Students learning through *Go Go Bugs* demonstrated significant gains in interest in learning science relative to a control group that learned about insects through a nongame curriculum (Kuo, 2007).

Given the sum of these findings, our research team is interested in moving beyond the question of whether games support learning to how games can be used productively in classrooms. Games designed for and implemented in classrooms necessarily become part of a broader curricular intervention; that is, rather than standing alone as an intervention strategy, games become part of a curriculum woven through the complex environment of a classroom community. In a recently published multicase study of teacher implementation of a game-based curriculum in science classrooms, Eastwood and Sadler (2013) theorized about the multiple interacting elements of game-based curricula. This work focused on classroom implementations of *Mission Biotech* (MBt), an educational game that challenges players to use tools and processes of biotechnology to stem the spread of an emerging epidemic (MBt is one of the interventions that serve as a focus in the current project). In describing a model for how MBt interactswith the broader curriculum, Eastwood and Sadler (2013) described MBt as the central focus or anchor for the curriculum but highlighted ways in which other elements connect to and shape use of and learning from the game. Some of these other pieces include written curriculum materials (including educative materials for teachers; Davis & Krajcik, 2005), assessments, classroom activities including small group work and laboratory exercises, and Web resources.

Considering games for classroom science education as necessarily a part of a game-based curriculum is an improvement on simpler articulations of learning as an outcome of game play. However, the game-based curriculum model leaves out at least one significant element that shapes classroom implementation: teachers. The ways in which teachers interpret curricular materials, provide opportunities for students to engage with the game and other learning materials, structure classroom time, make use of assessment data, and so on necessarily influence the extent to which the intervention successfully promotes student learning (Fogleman, McNeill, & Krajcik, 2011; Suárez, Pias, Membiela, & Dapía, 1998; Vos, Taconis, Jochems, & Pilot, 2011). A number of factors will likely affect teachers' implementation decisions and practices including the norms and expectations of their schools and districts, their own experiences with the content of the intervention, their experiences with the intervention curriculum and pedagogy, and resources such as technology for supporting the intervention implementation. Professional development (PD) opportunities should also have an impact on teachers' implementation of game-based curricula, particularly when the PD is well aligned with the target intervention (Ferdig, 2010; Klopfer, Osterweil, & Salen, 2009). Figure 1 presents a graphic representation of a model that characterizes relationships among teachers, game-based curricula, and student learning and that informed our work in the current project.

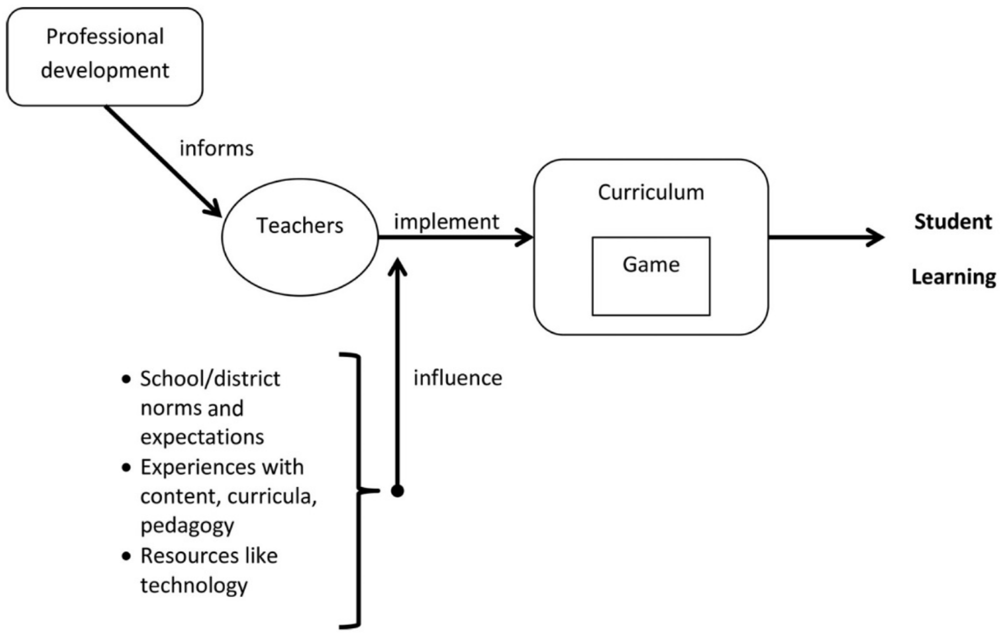

Figure 1. Graphic representation of a model for considering teacher implementation of a curriculum-embedded game for supporting student learning.

The theorized model of game implementation and outcomes helps to highlight gaps in the field's knowledge base of student learning relative to science education-oriented games. As a field, we know that students can learn science content through engagement with games (e.g., Clark et al., 2011; Sadler, Romine, Stuart, & Merle-Johnson, 2013).We also know that when games are implemented successfully as a part of science classes, they are connected to a broader game-based curriculum (Barab et al., 2007; Eastwood & Sadler, 2013), and that teacher enactment of these game-based curricula can be influenced significantly by their background experiences, PD opportunities, and school norms and expectations (Becker, 2007; Klopfer et al., 2009). We do not have enough evidence for how, what, and the conditions under which students can learn from game-based science curricula. Specifically, we are interested in investigations that capitalize on the nested nature of games within a broader curriculum and also appreciate the influences of teacher implementation decisions that have been informed through PD. For our research team, this is a nuanced but important point. Research studies related to the effects of games on student learning in science have been conducted in classroom environments, but there have not been enough research studies designed to allow for large-scale comparisons of student learning from games and nongame control curricula, delivered by teachers who have participated in PD and who make a wide range of implementation decisions.

Fidelity of Implementation

The conceptualization of games, teachers, and student learning offered in Figure 1 presents challenges for traditional notions of implementation fidelity. Fidelity of implementation provides a way of thinking about implementation of interventions motivated by the assumption that there is an ideal way of implementing a particular intervention. Measures of implementation fidelity tend to provide an estimate of how similar an actual implementation is to the ideal. High fidelity of implementation is prioritized in experimental analyses of interventions so that the interpreted patterns can be attributed to known aspects of the intervention as opposed to unanticipated or random dimensions. However, the implementation model described in Figure 1 suggests that the implementation of innovations within classrooms will be impacted by a variety of factors that cannot possibly be controlled or precisely factored into a measure of implementation fidelity. In the current project, we approached our work with teachers based on the assumption that teachers would (and should) make implementation decisions based on their local contexts and the students they served. These decisions necessarily change aspects of the curriculum and as a research team it is not possible for us to control for all of these changes. Our perspective on the position of teachers as a part of the design and implementation of classroom innovations and on the many factors which influence this work is consistent with views that have been described as a part of design-based implementation research (e.g., Penuel, Fishman, Cheng, & Sabelli, 2011).

Focus of the Research

Our research team was interested in understanding whether students can learn core ideas in science and develop interest in science through a game designed for and implemented in high school science classes.We were also interested in how much learning occurs relative to other comparable curricular interventions. Some of our team's preliminary work provided evidence that students could learn science through classroom implementations of MBt, a computer game designed around a biotechnology-themed virtual environment (Barko & Sadler, 2013; Sadler et al., 2013), but these studies were limited by relatively small sample sizes and lack of a control group. For the current work, we set up a quasi-experimental comparison between a game-based intervention and a nongame-based intervention (the control context) in a way that acknowledged and valued the complex interrelationships among games, their broader curricular context, teachers, and PD. To do this, we created two innovative biology curricula: the first was based around the MBt game; the second was a unit that focused on all the same content presented in MBt but featured a narrative case that served as the focal point. This control intervention was named Viral Quest (VQ). For this research, we studied student learning in response to two interventions (MBt as the experimental treatment and VQ as the control) that were implemented by teachers in their own school environments following similar PD experiences. As is evident in the research questions presented below, our research focuses on student learning and development of interest, but these processes occur within complex classroom environments and are shaped significantly by teachers who are informed by PD and influenced by a variety of factors. In this study, we do not present data on all of the teacher implementation decisions, impacts of PD, and impacts of local influencing factors; analyses of many of these issues are presented in an earlier publication (Eastwood & Sadler, 2013). However, the research design is based on an underlying assumption that these issues necessarily shape the ways in which innovative curricula are implemented. In other words, both the game-based and control curricula were implemented in complex classroom environments with teachers who made a variety of implementation decisions and dealt with external influences. We explored four research questions through the comparison of MBt and VQ:

1. Do innovative biotechnology curricula support student learning of basic biological principles?
2. Do innovative biotechnology curricula support development of interest in learning science and science careers among students?
3. How does a game-based biotechnology curriculum compare, in terms of its support for student learning of basic biological principles, to a nongame biotechnology curriculum?
4. To what extent does students' interest in learning science and science careers have an impact on their learning of biological principles in the context of biotechnology curricular innovations?

The Interventions

Biotechnology as a Context

The *Next Generation Science Standards* (NGSS Lead States, 2013) and the underlying framework (NRC, 2011) call for science learning opportunities that focus on disciplinary core ideas, scientific practices, and crosscutting themes. The integration of these dimensions of learning represents a new direction for science education and requires opportunities for learners to deeply engage in rich contexts that encourage thinking about big ideas in science and drawing connections to crosscutting themes while enacting meaningful scientific practices. Biotechnology offers a wide range of scenarios that provide the potential for rich contexts as the basis for next-generation science learning. Biotechnology was attractive for our research team because it is a vital aspect of the modern life sciences but tends to be underrepresented in schools. Biotechnology also connects to many of the core ideas of biology, including genetics and key molecular and cellular processes, and it deals with content and processes that can be difficult to feature in classrooms because of issues of scale, cost, and safety (Steele & Aubusson, 2004; Zeller, 1994). Both interventions (described below) were similarly aligned to NGSS content and principles.

Mission Biotech

TheMBt curriculum was built around a computer-based game featuring a biotechnology-themed virtual environment (http://virtualheroes.com/projects/mission-biotech). The MBt virtual world was created with the Unreal gaming engine and provided players with a first-person perspective on a 3D environment. Players could interact with several non-player characters who served as guides and could "walk" through several rooms of a virtual biotechnology facility including a large laboratory modeled after a working research and education biotech laboratory, offices, and a conference room (see Figures 2 and 3 for screen shots from the MBt environment). Upon entering the environment, players were welcomed as a new biotechnology researcher tasked with identifying the infectious agent causing a viral outbreak. Players progressed through four successive levels that increased in procedural and conceptual complexity. In the earliest level, players were introduced to the idea of using DNA sequence data to distinguish among different viruses and using virtual equipment and samples to isolate viral DNA. As the game progressed, students used the DNA they extracted in real-time polymerase chain reaction (PCR) analyses. They also were scaffolded in thinking about and ultimately setting up positive and negative controls for their reactions. In later levels of the game, RNA viruses were introduced as a possible culprit for the outbreak so players had to consider differences in RNA and DNA and the use of reverse transcription.

Figure 2. Screen shot from MBt showing the virtual laboratory featured in the game.

Figure 3. Screen shot from MBt showing a laboratory workstation within the virtual environment. Items on the lab bench include various solutions necessary for DNA extraction, a micropipettor, and a centrifuge. A nonplayer character, who serves as a lab assistant for players, can be seen in the background.

Use of the game by students was supported by numerous classroom activities including laboratory exercises, brief lectures, formative assessments, small group activities, and whole-class discussions. For example, prior to student exposure to the PCR component of the game, several teachers presented a short lecture on DNA structure and replication. After students completed virtual PCR within the game, the classes participated in an activity in which students "replicated" paper models of DNA sequences in a PCR simulation. Additional details regarding the MBt curriculum are available elsewhere (Eastwood & Sadler, 2013; Sadler et al., 2013). Given the focus of this study on the effects of a game (i.e., MBt), it is worth noting that the PCR activity was similar to a game in that it created an artificial environment defined by objects, rules, and a set of internal relationships. However, the PCR activity was a simulation rather than a game because it did not involve a conflict, contest, or quantifiable markers of success or failure. A more important distinction to note is that MBt provided an anchoring experience for the entire unit, whereas the PCR simulation was a brief activity that transpired over one or two class periods.

Viral Quest

The VQ unit was designed as a control curriculum and as such covered the same basic biological content and addressed the same science learning standards as MBt. Table 1 presents a list of the state science standards with which both curricula were aligned. In both cases, VQ and MBt, the curricula were designed as stand-alone units of instruction for teachers to implement within high school biology classes. Whereas the MBt unit used the game as a central organizing feature, VQ used a narrative case that featured a virology researcher tracking the emergence of a new disease with similar symptom sets as human immunodeficiency virus (HIV) and human papilloma virus (HPV). As the unit unfolded, students followed the scientist and the procedures and tools she used to solve the viral mystery. In the process, students learned about and engaged in activities related to DNA extraction, PCR, and reverse transcription. Most of the nongame elements of the MBt curriculum were incorporated in the VQ unit. For example, when the narrative case progressed to the point at which the research scientist in the featured case employed PCR, the students completed the PCR simulation referenced earlier. Table 2 presents a comparison of the MBt and VQ units as implemented over a standard 2-week period.

Table 1. Florida State Science Standards Aligned with *Mission Biotech*

Thematic Set	Standard
DNA structure and replication	SC.912.L.16.3: Describe the basic process of DNA replication and how it relates to the transmission and conservation of the genetic information.
	SC.912.L.16.4: Explain how mutations in the DNA sequence may or may not result in phenotypic change. Explain how mutations in gametes may result in phenotypic changes in offspring.
Transcription, translation, and protein structure	SC.912.L.16.5: Explain the basic processes of transcription and translation, and how they result in the expression of genes.
	SC.912.L.18.4: Describe the structures of proteins and amino acids. Explain the functions of proteins in living organisms. Identify some reactions that amino acids undergo. Relate the structure and function of enzymes.
Genetic technologies	SC.912.L.16.11: Discuss the technologies associated with forensic medicine and DNA identification, including restriction fragment length polymorphism (RFLP) analysis.
	SC.912.L.16.12: Describe how basic DNA technology (restriction digestion by endonucleases, gel electrophoresis, polymerase chain reaction, ligation, and transformation) is used to construct recombinant DNA molecules (DNA cloning).
Pathogens and immune responses	SC.912.L.14.52: Explain the basic functions of the human immune system, including specific and nonspecific immune response, vaccines, and antibiotics.
	SC.912.L.16.7: Describe how viruses and bacteria transfer genetic material between cells and the role of this process in biotechnology.

Table 2. Standardized Instructional Sequences for the MBt and VQ Curricula

Day[a]	Mission Biotech *Primary Focus*: Instructional activity	Viral Quest *Primary Focus*: Instructional activity
1	*Biotechnology tools, processes, and safety:* Student identification of equipment; minilecture (biotech tools, processes, and safety)	*Introduce VQ:* Students explore the case
2	*Introduce MBt:* Game play—begin Level 1 (develop avatars; learn game controls and mechanics)	*Viruses*: Video on viruses and viral diseases; minilecture; Students rewrite video script
3	*DNA extraction:* Lab activity (extract DNA from strawberries); minilecture (DNA location and function)	*Viruses and media literacy:* Jigsaw activity in which learners explore information about HIV and HPV
4	*DNA extraction:* Game play—complete Level 1 (DNA extraction and introduce PCR)	*DNA as evidence:* Small groups consider use of DNA as forensic evidence; class discussion; students hypothesize about use of DNA to identify viruses
5	*DNA structure and PCR:* Brief video and minilecture; small group questions; whole-class discussion	*DNA extraction:* Lab activity (extract DNA from strawberries); minilecture (DNA location and function)
6	*PCR:* Game play—begin Level 2 (extract DNA and conduct PCR)	*PCR process:* Small group activity—simulation of PCR process
7	*PCR process:* Small group activitysimulation of PCR process; class discussion	*PCR process:* Finish PCR simulation; class discussion
8	*PCR analysis:* Game play—complete Level 2 (conduct and analyze real-time PCR results)	*PCR analysis:* Minilecture; students work independently to interpret real-time PCR data
9	*PCR analysis:* Minilecture; students work independently to interpret real-time PCR data	*Reverse transcription:* Simulation activity for reverse transcription
10	*Reverse transcription:* Game play—Level 3 (reverse transcription, conduct and analyze PCR results)	*Case wrap-up and Careers:* Students discuss use of biotechnology tools relative to the case and explore careers featured

[a]Each "day" represents approximately 1 hour of instructional time.

The Dissemination Model

Both interventions were disseminated to life science teachers through a biotechnology-focused summer PD academy. Teachers who had been identified by administrators or other teachers as teacher leaders and who were interested in incorporating biotechnology into their classes participated in the program. The summer academy, hosted at a research university, provided opportunities for teachers to conduct biotechnology experiments, learn about new developments in biotech and biomedical research, and develop strategies for integrating biotechnology-themed instruction into their classes. The summer academies took place over a 2-week period during which teachers stayed on the university campus.

In the academic year following the summer academy, the program facilitated teacher interactions through an online community, program staff visited classrooms, and teachers came together for at least one face-to-face meeting in which they reported on ways in which they integrated biotechnology in their classes. MBt was a featured element of one summer academy; VQ was a featured element of the next summer academy (with a different set of teachers). Teachers were introduced to the game (MBt) or narrative case (VQ), had opportunities to participate in the game or engage in some of the case-based activities, explored supporting materials and assessments, and considered ways in which they could integrate the unit in their own classes. Teachers then chose whether or not they would implement the intervention. Teachers were encouraged to implement a relatively standardized sequence of instruction for each intervention (corresponding to the sequences presented in Table 2). Teachers who implemented the sequences and who were able and willing to partner with our team for data collection and completion of consent processes (with schools, students, and parents) participated in the study. Partner teachers received an honorarium for supporting our research efforts (collecting student data, consent forms, completing surveys, and daily logs) in addition to stipends for completion of the summer academy.

As suggested in the introductory sections, our team assumed that teachers would make modifications to the curriculum to meet the needs of their local contexts. For example, participant schools adopted a wide variety of scheduling models, and this impacted ways in which teachers implemented the suggested instructional sequences. Among the MBt teachers, several had already discussed safety issues with their students, so they chose to skip the first suggested lesson. We considered these as relatively minor modifications to the curricula. In other cases, MBt teachers wanted to implement the unit, but did not want to use class time for students to play the game and made game play a requirement for homework. We did not include data in the current analyses from teachers who made these kinds of major modifications, which our team deemed as too far removed from the intent of MBt. Similar exclusions were made with teachers implementing VQ. The decisions regarding what constituted negligible versus exclusionary modifications were based on analyses of the daily logs that teachers kept during their implementation of the interventions and a teacher survey completed after the unit.

Methods

Sample

A total of 1,888 students from the biology classrooms of 36 different teachers participated in this study. Of these students, 1,058 from the classes of 20 teachers participated in the MBt intervention, and 830 students from the classes of 16 teachers participated in the VQ intervention. The 20 MBt teachers participated in the summer PD programming in 1 year; the 16 VQ teachers participated in the next year. The summer academies followed the same structure, provided access to the same resources and activities, and used the same means of recruiting teachers. The biggest observable difference in the PD academies was that Year 1 featured MBt and Year 2 featured VQ. In Year 1, the summer PD academy was held twice, and a total of 46 teachers participated. In Year 2, only one summer academy was offered, and 29 teachers participated. Rates of participation in our research study are higher among teachers from the Year 2 academy (55%) than teachers from the Year 1 academies (43%). However, we have evidence that this difference is due more to technical constraints than underlying differences in the teacher samples.

There were at least six teachers from the academies featuring MBt who expressed serious interest in using the intervention but whose teaching context prohibited implementation.

For example, we worked with several teachers who did not have access to the hardware (i.e., computers or laptops) or network connectivity necessary for implementation. Most teachers from both groups (MBt and VQ) were the sole implementers of the curricula in their schools, but there were some cases in both groups where two or three teachers from a single school participated. None of these teachers formally collaborated in the implementation of the MBt or VQ curricula, but we cannot rule out the possibility of informal communications that may have affected implementation. In most schools involved in the study, only one teacher implemented the curricula; therefore, it was not possible to explore possible school-level variation.

Working with 36 different teachers in different schools situated in multiple districts across the state created a number of logistical challenges. One issue related to gaining research approvals from the university institutional review board (IRB) as well as the many districts and schools.We were not able to gain permissions to collect student-level race and ethnicity data in many of the districts we partnered with. We were able to survey teachers regarding class-level demographics.White students were the majority group in all classes studied, and students from ethnic or racial minority groups were a part of all classes. Minority student participation in each class ranged from just under 50% to approximately 10% depending on the school. Students from African American, Hispanic, and multiracial backgrounds made up the largest proportion of minority student groups, although some teachers also reported low percentages (less than 5%) of Asian and Native American students. Given the limitations of our data collection, it is not possible to directly compare demographic variables between the MBt and VQ groups; however, data provided by the teachers through postintervention surveys suggested no systematic differences by intervention in student demographic characteristics.

Assessments

The research questions call for assessment of student understanding of basic biological principles, interest in science, and interest in careers in science. For the assessment of content learning, a multilevel assessment framework was employed (Hickey & Pellegrino, 2005; Klosterman & Sadler, 2010; Ruiz-Primo, Shavelson, Hamilton, & Klein, 2002). Proximal and distal assessments were administered to students before and after the intervention. The proximal assessment contained questions directly tied to the content covered in the MBt and VQ interventions. This assessment contained 23 items that measured with a pretest reliability of 0.686 and a posttest reliability of 0.827 on a scale between 0 and 23. The distal assessment contained 18 items that were aligned with the eight content standards on which the interventions were based (see Table 1). The items for the distal assessment were derived from publically released portions of several state (California, Florida, New York, and Texas) and national (International Baccalaureate and National Assessment of Educational Progress) examinations. The 18 items measured the students in our sample with a reliability of 0.839 on the pretest and 0.868 on the posttest using a scale between 0 and 18. Details on the construction and validity evidence for these assessments can be found in Sadler et al. (2013).

In addition to exploring potential impact of the biotechnology curricula on student learning of biological content, we were interested in possible effects on student interest in science and careers in science. We used the Student Interest in Technology and Science (SITS) instrument (Romine, Sadler, Presley, & Klosterman, 2014) to measure student interest. Five subscales comprise the SITS instrument: interest in learning science (L1), interest in using technology to learn science (L2), interest in science careers (C1), interest in technology careers (C2), and attitudes toward biotechnology (Bio). The SITS instrument contains a total of 25 items (five items for each subscale) using a 5-point Likert scale yielding total scores for interest ranging between 0 and 75. Pretest reliability was 0.909, and posttest reliability was 0.929 for the current sample of students.

In previously published work, our research team describes development, piloting, and reliability and validity analyses for the instrument. Expert panels were used to establish content validity of the SITS. Construct validity and instrument precision were quantified through classical test and item response theory frameworks (Romine et al., 2014). Follow-up investigations have demonstrated the utility of the SITS instrument for documenting change in student interest in response to educational interventions (Romine & Sadler, 2014).

Data Analysis

The effects of the MBt and VQ interventions, classroom, and interest on students' proximal and distal knowledge were evaluated through linear mixed modeling with STATA 11 using a quasi-experimental pre-post design. In our linear mixed models, the intercept and time were treated as random effects, meaning that these parameters were free to vary about the mean for each participant and teacher. Taking student- and classroom-level cluster effects into account allows for hypothesis tests that are freed from biases associated with student- and classroom-level differences, such as those associated with ethnogeographic variables, that would otherwise be impossible to control for in the experimental design.

A second advantage of a mixed model format is increased flexibility with missing data. In a repeated measures analysis of variance (ANOVA) context, data from a student not measured at both time points get excluded from the analysis. However, treatment of time as a random effect in a linear mixed model format allows information from students with missing data to be incorporated into the model. We welcomed the opportunity to integrate data, which would otherwise be excluded since doing so reduces biases inherent in data exclusion. Despite this advantage, the inclusion of participants with missing data carries the inherent assumption that their nonpresence is random, meaning that the characteristics of participants with missing data are similar to other participants in the study. Assuming participants are missing-at-random when this is not the case can also yield biased results (Hedeker & Gibbons, 1997). In the case of this study, the rate of missingness was under 10%. Since this is comparable to what one would expect from a high school classroom, we chose to utilize the simplifying missing-at-random assumption.

In defining our linear mixed models for the dependent variables of proximal and distal knowledge, we nested within-student, between-student, and between-classroom variance and treated time and the intercept as random effects, meaning their effects were free to vary across students and classrooms. In this nested approach, individual students (n = 1,888) were clustered according to teacher (n = 36). This allowed us to partition variance associated with individual students as well as with groups of students working in the classes of particular teachers. In the fixed effects part of the model, time (pre = 0, post = 1), the intervention (VQ = 1, MBt = 0), and interest (continuous between 0 and 75) were added to the model as main effects. We chose to add the interaction between the intervention and time as a test of difference of gains between interventions, and the interaction between intervention and interest as a test ofwhether the relationship between interest and knowledge was different for treatment and control groups. Finally, an interest-by-time interaction term was added to test the effect of interest on student knowledge gains.

The z-test (α =.05) was used to test the null hypothesis of zero slope. Random effects, the intercept and time, were modeled using a diagonal or independent structure. In this formulation, each random effect is assumed to have a unique variance, and zero between-effect covariance. In addition, the effect of time was also used to model the residual variance using an independent structure. This allowed the model to account for the heteroskedasticity that often occurs in pre-post or longitudinal designs. Wald's Z, which is asymptotically equivalent to the likelihood ratio test (Engle, 1984), was used as a test for significance of each variance component. Wald's Z and the likelihood ratio test will be approximately equivalent for the sample size in this study, and Wald's Z is computationally more efficient since it requires that only one model be estimated.

Results

Highly significant t tests indicate that scores on both the proximal unit test and the distal standards-based test improved with time for both the MBt and VQ groups (Table 3 and Figures 4 and 5). As would be expected from a multilevel assessment framework, proximal gains were more pronounced than distal gains. Using Cohen's (1988) rough guidelines, distal gains were small but detectable, whereas proximal gains were moderate to large.

Pre- and postintervention scores for each of the five interest sub-constructs from the SITS instrument (interest in learning science, interest in learning science with technology, interest in science careers, interest in technology careers, and attitudes toward biotechnology) are presented in Table 4. Means and standard errors are very similar for students participating in both interventions at both the pre- and postintervention administrations. In early stages of our analysis, we explored possible differences in interest patterns between the interventions and time points using multivariate analysis of variance (MANOVA) procedures, paired t tests, and indicators of practical significance (Cohen's d and partial η^2). The analysis suggested that a few individual comparisons of interest were statistically significantly different (e.g., the time-by-curriculum interaction for student interest in science careers). However, given the impact of the large sample sizes on tests of statistical significance, it is crucial to attend to practical significance, and there is no evidence of practically significant differences among any of the interest comparisons made (e.g., partial η^2 for the time by curriculum interaction referenced above is 0.003 where a value of 0.01 is generally considered to mark the lower boundary for detecting small effect sizes). This indicates that as a whole, the MBt and VQ interventions had negligible effects on interest.

Table 3. Pre- and Posttest Comparisons of Content Tests for the Two Interventions (MBt and VQ)

Group	Test	Pretest Mean (*SE*)	Posttest Mean (*SE*)	*T**	*d*
MBt	Proximal ($n = 938$)	9.99 (0.13)	14.36 (0.16)	31.7	0.71
	Distal ($n = 935$)	11.29 (0.15)	12.30 (0.15)	9.6	0.16
VQ	Proximal ($n = 759$)	9.16 (0.13)	14.58 (0.18)	28.4	0.89
	Distal ($n = 740$)	11.02 (0.17)	12.74 (0.17)	11.0	0.26

*All values $\alpha \ll .001$.

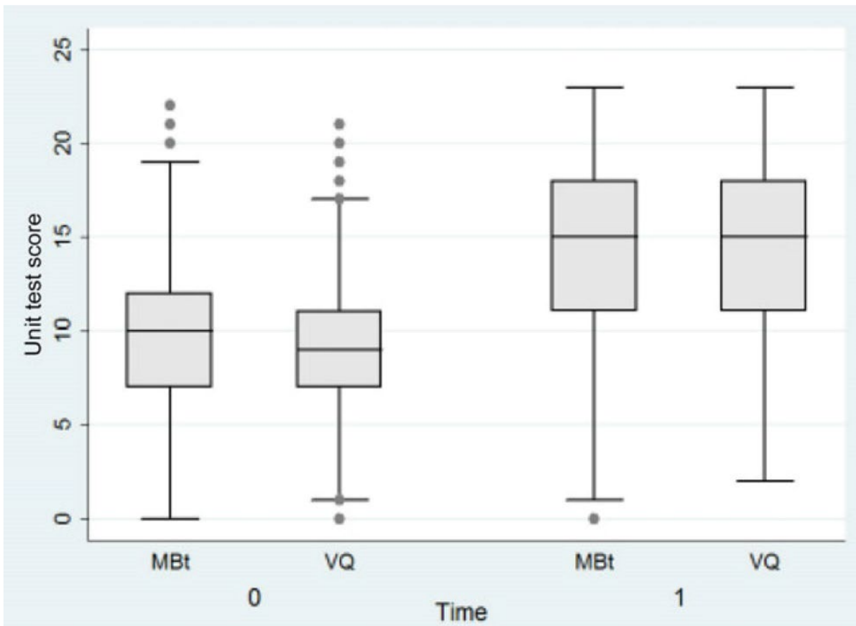

Figure 4. Differential effects of the MBt and VQ interventions on proximal knowledge gains.

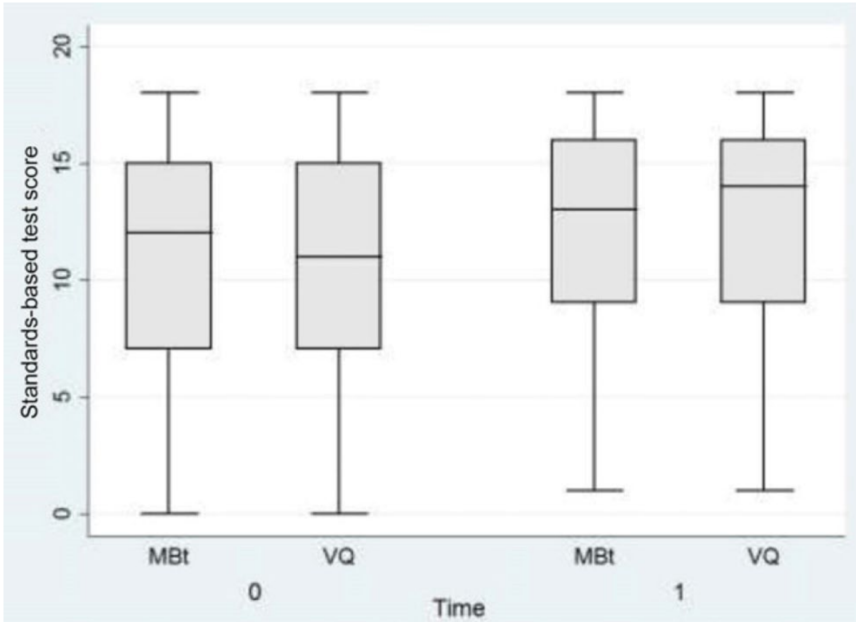

Figure 5. Differential effects of the MBt and VQ interventions on distal knowledge gains.

Table 4. Pre- and Posttest Comparisons of Interest for the Two Interventions (MBt and VQ)

Interest Construct	MBt		VQ	
	Pretest Mean (*SE*)	Posttest Mean (*SE*)	Pretest Mean (*SE*)	Posttest Mean (*SE*)
L1: Interest in learning science	8.98 (0.093)	8.81 (0.099)	9.27 (0.11)	9.16 (0.12)
L2: Interest in learning science with technology	10.14 (0.084)	9.67 (0.097)	9.96 (0.097)	9.60 (0.11)
C1: Interest in science careers	6.23 (0.12)	6.29 (0.12)	6.39 (0.13)	6.73 (0.14)
C2: Interest in technology careers	8.29 (0.098)	8.18 (0.11)	8.44 (0.11)	8.42 (0.12)
Bio: Attitudes toward biotechnology	11.04 (0.082)	11.09 (0.091)	11.08 (0.095)	11.21 (0.11)
Total score	44.42 (0.33)	43.82 (0.38)	44.88 (0.38)	44.85 (0.44)

Tests of fixed effects for the standards-aligned test (ST) and the unit test (UT) are presented in Table 5. In the multilevel model, the effect of time was significant for both the UT and the ST at the $\alpha = .01$ level for the student sample as a whole. This suggests that students across both interventions demonstrated gains in knowledge of biological principles as assessed through both proximal and distal measures. UT gains ($b = 5.35$, $SE = 0.76$, $r^2 = .026$) for the entire group were approximately two times greater than ST gains ($b = 2.88$, $SE = 0.57$, $r^2 = .013$), indicating the more pliable nature of knowledge demonstrated through proximal assessments as compared to distal assessments. The curriculum-by-time interaction term offers an analysis of the extent to which students may have performed differentially depending on the intervention (MBt or VQ) in which they participated. These comparisons were not statistically significant for either the UT or ST. This result suggests that gains made by students on both the proximal and distal assessments were not significantly different between the MBt and VQ groups.

Table 5. Fixed Effects Parameters

Variable	b	SE_B	z	95% CI[b]	r^2
			Distal assessment		
Time	2.88	0.57	5.09**	1.77–3.99	.013
Interest	0.08	0.01	8.14**	0.059–0.097	.033
Curr[a]	–0.25	0.88	–0.29	–1.98 to 1.48	.000
Curr × time	–0.53	0.51	–1.04	–1.53 to 0.47	.001
Curr × interest	–0.01	0.01	–0.74	–0.034 to 0.015	.001
Interest × time	–0.03	0.01	–2.93**	–0.042 to –0.008	.005
Intercept	7.21	0.71	10.17**	5.82–8.60	.052
			Proximal assessment		
Time	5.35	0.76	7.04**	3.86–6.85	.026
Interest	0.07	0.01	7.15**	0.048–0.085	.025
Curr[a]	0.15	0.77	0.19	–1.36 to 1.66	.000
Curr × time	–0.97	0.73	–1.34	–2.40 to 0.45	.001
Curr × interest	–0.02	0.01	–1.79	–0.045 to 0.002	.002
Interest × time	–0.004	0.01	–0.41	–0.23 to 0.015	.000
Intercept	6.70	0.59	11.38**	5.55–7.85	.064

*α = .05, **α = .01.
[a]"Curr" denotes the compared curricula, MBt and VQ.
[b]"CI" stands for confidence interval.

The effect of interest was positive and highly significant for both tests (ST: b = 0.08, SE = 0.01, r^2 = .033; UT: b = 0.07, SE = 0.01, r^2 = 0.025), indicating a positive effect of interest on test scores, and this relationshipwas similar for both proximal and distal assessments. The curriculum-by-interest interaction terms were insignificant, indicating that the relationship between knowledge and interest was invariant between interventions for both knowledge assessments. The interaction effect between interest and time was negative and statistically significant for the ST (b = –0.03, SE = 0.01, r^2 = 0.005). This finding suggests that the effect of the interventions on the distal assessment was significantly greater for students with lower interest levels. However, the low r^2 value indicates that while the relationship exists, its importance is superseded by other elements in the model. While the interest by time interaction parameter is also negative for the UT, it is both statistically and practically insignificant. The intercept terms indicate the pretest score that would be predicted for a student with no interest in science. The intercepts are statistically significantly above zero, indicating that students came into the interventions with prior knowledge related to the basic biological principles underlying biotechnology.

The mixed modeling approach employed in this study allowed us to partition variance associated with individual students as well as groups of students working in the classes of particular teachers. Variance components for these variables, which are treated as random effects, are displayed in Table 6. Allowance for randomly varying intercepts at the student and teacher levels in addition to the assumption of heteroskedasticity were important considerations as evidenced by highly significant Wald Z tests for all variance components. At the between- and within-classroom levels, variance of the intercept provides a measure of the variability of pretest scores. Distal pretest scores exhibit more variability (var_{cls} = 8.17, var^{sdt} = 6.01) than proximal pretest scores (var_{cls} = 4.01, var_{sdt} = 3.43). However, variability in gains in proximal knowledge across the intervention (var_{cls} = 7.46) was much greater than the variability in gains for distal knowledge (var_{cls} = 2.72). These estimates reflect some important differences regarding how proximal and distal knowledge relate to innovative curricula. It makes sense that greater variability would exist between distal knowledge (of content connected to state standards) than proximal knowledge (of more narrow content addressed in a specific activity) since none of the students had significant prior experience with the specific biological content associated with MBt and VQ. However, the greater variability in gains in proximal knowledge demonstrates that it is easier to change in the context of an intervention. Variability in gains in both proximal and distal knowledge likely reflects between-classroom differences in the ways in which the curricula were implemented in individual classrooms. This result supports the conceptualization of classroom-based implementation environments that reflect multiple, interacting elements (see Figure 1) as opposed to a more simplified model that only considers a single element of an intervention.

Table 6. Random Effects Parameters

Grouping Variable	Variance	Estimate	SE_{est}	Wald z	95% CI[a]
	Distal assessment				
Teacher	Time	2.72	0.76	3.58*	1.58–4.70
	Intercept	8.17	2.04	4.01*	5.40–13.33
Student	Intercept	6.01	0.33	18.21*	5.40–6.70
Residual	Error (pre)	5.31	0.32	16.59*	4.72–5.97
	Error (post)	6.47	0.35	18.49*	5.82–7.19
	Proximal assessment				
Teacher	Time	7.46	1.95	3.83*	4.47–12.46
	Intercept	4.01	1.05	3.82*	2.40–6.68
Student	Intercept	3.43	0.30	11.43*	2.88–4.07
Residual	Error (pre)	6.46	0.36	17.94*	5.80–7.19
	Error (post)	9.46	0.43	22.00*	8.65–10.34

*α = .01.
[a]"CI" stands for confidence interval.

Discussion

Learning Biology in the Context of Biotechnology

This study provides evidence for the effectiveness of innovative curricula situated in biotechnology as tools for supporting student learning of biological content knowledge. In this study, we created two curricular interventions that leverage biotechnology as a rich context for engaging learners in the negotiation of basic principles of biology. The NGSS emphasize the need for students to develop understandings of core ideas in science, that is, ideas that can "have broad importance across multiple sciences . . . or be a key organizing principle" and can "provide a key tool for understanding or investigating more complex ideas and solving problems," and "relate to the interests and life experiences of students" (NRC, 2011, p. 31). Genetics and heredity certainly qualify as core ideas in the discipline of biology. As biology educators consider the challenge of moving from covering the breadth of biological knowledge to engaging students in deeper explorations of a more limited set of core ideas, like genetics, they should simultaneously consider that a range of rich and engaging contexts may be employed to support learning of these core ideas. Biotechnology has revolutionized study of the life sciences and provides many opportunities to draw connections between the significance of science for societal and personal purposes (Borgerding, Sadler, & Koroly, 2013; Kidman, 2010). The results of this study offer empirical support justifying the use of curricula designed to feature biotechnology as a context for learning about core ideas in biology. In the case of this study, curricula that made use of either a computer-based game or a narrative case that featured biotechnology proved to be effective in supporting student learning consistent with NGSS principles.

Interest in Science and Science Careers

Supporting the development of interest in science and science-related careers is an important goal for science education, particularly in light of recent emphases on the expansion of the science, technology, engineering and mathematics (STEM) workforce (Rukavina, Zuvic-Butorac, Ledic, Milotic, & Jurdana-Sepic, 2012). In this study, we found no evidence to support the notion that unit-based curricular interventions related to biotechnology can support the development of interest in science among learners. Of course, we cannot conclude that biotechnology curricula cannot support development of interest in science, science careers, and attitudes toward biotechnology; only that the interventions we implemented showed no evidence of change with the instrumentation we employed. Personal interest in science, the primary construct assessed by the SITS instrument, is an enduring characteristic and tends to be resistant to change (Alexander & Jetton, 1996). However, given the positive reports of games for learning and student affect (Kuo, 2007; Rowe et al., 2010; Squire, 2011), it was reasonable to investigate development of interest as a possible outcome. Although we found no evidence of interest changing as a result of engagement with either of the curricular interventions, interest in science and careers in science was positively related to student performance on the tests of biological content both before and after the interventions. This finding is consistent with other work linking interest and understanding (Nieswandt, 2007; Silvia, 2006). In developing the SITS, we included a separate subsection that focused on attitudes toward biotechnology. Given the enduring nature of personal interest, we hypothesized that students would likely show greater changes on the attitudinal subscale. However, our study produced no evidence that attitudes toward biotechnology responded any differently in association with the interventions than the interest constructs.

One of the interesting findings of the work was that lower interest was associated with slightly greater gains in learning, at least with respect to the distal assessment. The evidence here is limited, but it is consistent with findings from the broader MBt project. In a previous study of MBt implementation, the game was implemented in classes that ranged in terms of the academic levels of the students. Students from lower-level classes tended to show greater gains in learning than their peers from higher-level classes (Sadler et al., 2013). This result was supported by feedback from teachers (through their daily logs and informal communications), many of whom expressed surprise that lower-level students benefited more from the game than their more advanced students, despite initial concerns from many of these same teachers that MBt and VQ may be too conceptually difficult for their students. We hypothesize that students who demonstrated relatively low interest in learning science may have been more motivated to engage in science learning when it was facilitated through the MBt game or VQ narrative. Both interventions could have offered more compelling learning opportunities for students less enthused about learning science in more traditional formats, and this may account for increased performance on the content exams. We offer this discussion as a possible explanation for the observed patterns with full recognition that the pattern requires more thorough investigation before it can be considered a tentative conclusion.

Games to Support Science Learning

A primary goal for this study was to extend the field's knowledge base regarding the use of games to support science learning. As a field, we know that students can learn through engagement with computer-based games (for reviews, see Honey & Hilton, 2011; Squire, 2011) and several studies provide evidence of science-specific learning (e.g., Annetta, Minogue, Holmes, & Cheng, 2009; Neulight, Kafai, Kao, Foley, & Galas, 2007; Steinkuehler & Duncan, 2008). But we know less about how effective games are for supporting learning relative to other educational approaches, particularly when the approaches are situated in broader systems of teacher PD, teacher implementation decisions and actions, and the complexity of science classrooms. Our preliminary work with MBt provided evidence that this particular game could serve as the anchor for a game-based curriculum (Barko & Sadler, 2013a) and that teachers could successfully implement the game in high school science classrooms (Eastwood & Sadler, 2013). We also generated evidence suggesting that students could learn significant biology associated with MBt game play (Sadler et al., 2013). We were not, however, able to say much about the effectiveness of learning through a game like MBt relative to other approaches, particularly when the treatment and control conditions incorporated implementation within schooling systems that included PD, teacher implementation decisions, and classroom environments, and this became a primary focus of the current work.

In considering a reasonable comparison group, we initially considered using a "business-as-usual" group. After documenting a few classes using teacher-developed units to teach material they saw as analogous to the MBt content focus, we concluded that the variation in teacher-led approaches would be too great to draw meaningful conclusions from the analyses. Early analyses of student-level data from teacher-created units showed no gains on the content instruments employed in this study. Therefore, we created and implemented VQ as a control curriculum. In doing this, we set up a comparison between a game-based curriculum and a nongame-based curriculum, both of which used biotechnology as a context. In the case of MBt, the game was the central element of a broader curriculum; for VQ, a narrative case served as the central element of a broader curriculum that shared most other dimensions of the former curriculum. These shared elements included similar student materials, class exercises, teacher resources, laboratory activities, and formative assessment prompts. Both interventions were featured in the same teacher PD program (in successive years) and teachers served similar roles in terms of making decisions regarding implementation of the curricula, dealing with the challenges of implementing a new unit, and modifying the sequences suggested in the curricula to fit their local contexts and needs. The incorporation of processes for teacher PD and teacher implementation decisions in the comparison of a game-based curriculum with a nongame control curriculum distinguishes this study from other investigations that have explored games for learning in systems with reduced complexity.

Figure 1 presents the implementation model that guided our work. In setting up the comparison between the game- and nongame-based curricula, we sought to recognize the complex systems which include PD programming, teacher implementation decisions, and classroom contexts. For both interventions, teachers were informed by PD and influenced by a variety of factors (including school norms, experience with the content, and school resources). Results indicate that both curricula lead to significant learning gains in both the proximal and distal assessments of content knowledge. The curriculum-by-time interaction effect was not statistically different, indicating that students in both groups showed similar results. Therefore, we have no evidence that the game afforded learning beyond that which was accomplished with the nongame comparison curriculum.

It is important to point out that the MBt and VQ curricula shared many similarities (by design) but that the central components of both interventions were quite different and these components occupied a significant portion of their respective interventions. MBt students spent approximately half of their class time within the unit engaged in game play. We did work with teachers who decided not to incorporate the MBt game to the same extent, but these teachers were not included in the analysis. VQ offered students an opportunity to read about a case, dealing with contemporary issues (i.e., sexually transmitted diseases), embedded in the context of using biotechnology to address a social issue; however, VQ was not a game according to the criteria for operationalizing games as presented by Salen and Zimmerman (2004). Like some games including MBt, VQ featured a compelling narrative, but unlike games, it did not create boundaries between an artificial environment and the real world; it did not prescribe a conflict or contest; students did not make decisions or take actions to affect the progress of the experience; and students did not track success through quantifiable markers. Therefore, despite some similarities in the biotechnology context and similar learning experiences like common lab exercises and class activities, the central elements of the MBt and VQ curricula were quite distinct. Despite these differences, the results were similar in terms of student learning gains.

Limitations

While we argue that this research yields new insights regarding the use of computer-based games in science education, we also recognize limitations in the study design. Despite the large sample size and diversity of classrooms sampled, the quasi-experimental approach impacts the generalizability of findings. Without a randomized control trial or utilization of carefully matched intervention groups, we cannot guarantee that the patterns observed are due only to the difference in the interventions (game-based curriculum vs. the nongame, control curriculum). Our control group serves as a reasonable comparison given the similarities in teachers who participated in both groups, the PD experiences, and general class-level characteristics. However, we were not able to randomly assign students or teachers to intervention groups, and IRB issues prevented us from collecting student-level demographic data that could have strengthened our claims regarding the equivalency of the two groups. In the interest of research ethics, the best we could do was control for student and classroom-level cluster effects in the statistical models.

We acknowledge these limitations, but we also suggest that studying the impacts of innovative curricula within the complex contexts of modern schooling systems forces some of these limitations. For us, the value of exploring teaching and learning in complex school environments offsets limitations that these contexts often impose in terms of research design. We suggest that careful analyses of these learning environments, including the use of sophisticated quantitative techniques, which control for student- and classroom-level cluster effects, can offer important insights into teaching and learning processes even if the generalizability of findings is limited (Shaffer & Serlin, 2004).

Implications for Game-Based Science Learning and Future Research

Our results provided evidence that educational games can be successfully developed and implemented as a part of science classroom curricular interventions. Like previous research, our findings also support the notion that significant science learning can occur in such situations. However, the study provided no evidence that student learning through engagement with the game exceeded learning in a nongame curriculum with a similar content focus. Said differently, highly prepared and qualified educators can deliver pedagogically sound and engaging content in meaningful ways for students without necessarily having to implement a technology-based game curriculum.

Three implications immediately emerge. First, *game-based learning should continue to be pursued for research and curricular activities; however, those engaging in such endeavors should be aware of the technological and PD challenges of implementing such projects*. The technology infrastructure of schools became an important aspect of the deployment of MBt. Issues such as network connectivity, the availability of hardware, and security protocols had significant impacts on how students and teachers interacted with MBt (Eastwood & Sadler, 2013). Our team worked to minimize the effects of these issues; for example, we worked with school technology coordinators to overcome network security issues, provided technical support to individual teachers, and made a class set of laptops available to teachers who lacked access to computers in their schools. Despite these efforts, the inevitable challenges in gaining access to and using technology in schools surely had a greater impact on teachers and students in the MBt intervention as compared to their peers in the VQ intervention, which relied significantly less on technology.

There are also PD challenges. The first time teachers implement a new curriculum, we can expect a certain degree of challenge (Fogleman et al., 2011; Pintó, 2005; Schneider, Krajcik, & Blumenfeld, 2005; Valanides & Angeli, 2008). In fact, the issue of teacher familiarity and experience with an intervention is an explicit component of the implementation model showcased in Figure 1. In this study, MBt and VQ teachers were implementing the curricula for the first time, so both sets of teachers had to negotiate the challenges of adopting new materials and activities in their classes. This is one of the reasons that drove our decision to create VQ as a comparison curriculum; teachers from both groups were required to deal with initial implementation challenges. While this strategy helps to control for initial implementation effects, we cannot say, at this time, anything about how these effects may or may not wane with successive implementations. It may be the case that both curricula will become easier for teachers to implement in a second year, but they may also show differential patterns over time. These are trends that our group will have to monitor after multi-year implementations.

Second, as researchers and educators continue to explore game-based learning, *we need better tools and strategies to understand the impact of such implementations longitudinally*. In terms of learning basic science content, the interventions performed in similar fashion. There may be dimensions of learning not accounted for in this study that are differentially supported by games, but our research design, which included proximal and distal assessments of knowledge and several interest constructs, did not uncover these dimensions. On the one hand, it is good news that a game-based innovative curriculum as well as a reform-oriented nongame-based intervention both succeeded. Conversely, we need to better understand if those students in the game-based section are more or less likely to engage with science content, classes, or careers because of this interaction. Research has provided evidence that technology can support teachers who do not have the pedagogical or content expertise to provide reform-oriented instruction (Ferdig, Roehler, & Pearson, 2002). However, in cases where the pedagogy and the curriculum is sound, we need further evidence that warrants and justifies the technological and financial costs of game-based implementations.

One area related to technology interventions like MBt and issues of assessment is the potential of learning analytics (Suthers & Chu, 2012). As students engage with technology-mediated environments, huge data sets based on student actions and choices can be generated and analyzed. These data have the potential to yield invaluable insights regarding student progress through systems and how performance trajectories signal mastery of core concepts and development of competencies as well as ideas or decision points that present particular challenges. Ideally, these data will make it possible to more accurately track student successes and also to customize feedback to learners at optimal time points for the support of learning (Goggins, Mascaro, & Valetto, 2013). We have examined trace data developed through student engagement within MBt with some interesting results related to latent cognitive attributes and science learning (Lamb, Annetta, Vallet, & Sadler, 2014). One of the takeaway messages from this work is that in order for these kinds of data to be most useful for both research and teaching purposes, learning analytics must be a part of overall system design. When we developed MBt, we carefully considered issues such as learning objectives, standards-alignment, game mechanics, and embedded assessments. We did not deliberately prepare for a learning analytics system with target data points and probable algorithms for making sense of these data. This process of designing technology-mediated educational interventions like games with more systematic attention to the kinds of data that could be collected, analysis strategies, and how these data may be used to support learning will continue to be an important opportunity for the field.

Third, *researchers and educators interested in game-based learning should understand the financial costs of such endeavors and plan accordingly.* This project was funded by a relatively large grant from the National Science Foundation ($1.5 million). Well over a third of the grant went to the design and development of MBt, whereas development of VQ required only a fraction of these resources. Science educators interested in game-based learning should continue to find external commercial, foundation, and government-based organizations to partner with to share the financial burden. However, educators and researchers can also explore the use of commercial-off-the-shelf games for science learning (Van Eck, 2006) as well as less financially intensive minigames and apps accessed through mobile devices (Sánchez & Olivares, 2011). Finally, game researchers have made convincing empirical arguments for the role of student creation of games rather than just consumption—another cost-saving possibility given the availability of open source game and simulation development tools (Kafai & Peppler, 2011).

Conclusions

A number of researchers and practitioners, including members of our team, have lauded the potential of games for learning (Honey & Hilton, 2011; Squire, 2011). Given the place of games in modern society (Madden et al., 2013), they will continue to be a part of today's educational landscape. Our findings provide empirical support for the potential of games to support the kinds of learning targeted by the science education community. However, this study also highlights the importance of considering costs and benefits associated with innovative curricula. Professional game design studios spend millions of dollars on the development of new games (Nadolski et al., 2008); it seems unlikely that educators will ever be able to compete with these levels of funding in producing new educational games. The results of this study suggest that educators may be able to design other kinds of learning experiences for a fraction of the costs associated with game development and achieve at least some of the same learning outcomes. Science educators may want to consider innovating for learning with existing technologies (including games and other genres), as opposed to resource-intensive creation of new technologies, as a means of maximizing the use of limited resources.

Acknowledgments – Many individuals contributed to the design, development, and implementation of the MBt/VQ project, including Mary Jo Koroly, Richard Snyder, Michelle Klosterman, Tim Barko, Tamara Mandel, Julie Bokor, Julie Brown, and Felipe Echeverri. We partnered with Virtual Heroes to design and develop the MBt game. We appreciate the many schools, teachers, and students who participated in the project. This material is based upon work supported by the National Science Foundation under Grant No. 0833521. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Alexander, P. A., & Jetton, T. L. (1996). The role of importance and interest in the processing of text. Educational Psychology Review, 8(1), 89–121.

Annetta, L. A., Cheng, M. T., & Holmes, S. (2010). Assessing twenty-first century skills through a teacher created video game for high school biology teachers. Research in Science and Technological Education, 28(2), 101–114.

Annetta, L. A., Minogue, J. A., Holmes, S., & Cheng, M. T. (2009) Investigating the impact of video games on high school students' engagement and learning about genetics. Computers & Education, 54(1), 74–85.

Barab, S. A., Sadler, T. D., Heiselt, C., Hickey, D. T., & Zuiker, S. (2007). Relating narrative, inquiry, and inscriptions: Supporting consequential play. Journal of Science Education and Technology, 16, 59–82.

Barko, T., & Sadler, T. D. (2013). Learning outcomes associated with classroom implementation of a biotechnology-themed video game. The American Biology Teacher, 75(1), 29–33.

Becker, K. (2007). Digital game-based learning once removed: teaching teachers. British Journal of Educational Technology, 38(3), 478–488.

Borgerding, L. A., Sadler, T. D., & Koroly, M. J. (2013). Teachers' concerns about biotechnology education. Journal of Science Education and Technology, 22(2), 133–147.

Bosche, W. (2010). Violent video games prime both aggressive and positive cognitions. Journal of Media Psychology: Theories, Methods, and Applications, 22(4), 139–146.

Clark, D. B., Nelson, B. C., Chang, H. Y., Martinez-Garza, M., Slack, K., & D'Angelo, C. M. (2011). Exploring Newtonian mechanics in a conceptually integrated digital game: Comparison of learning and affective outcomes for students in Taiwan and the United States. Computers & Education, 57(3), 2178–2195.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. Educational Researcher, 34(3), 3–14.

Eastwood, J. L., & Sadler, T. D. (2013). Teachers' implementation of a game-based biotechnology curriculum. Computers & Education, 66, 11–24.

Engle, R. F. (1984). Wald, likelihood ratio, and Lagrange multiplier tests in econometrics. Handbook of Econometrics, 2, 775–826.

Entertainment Software Association (2013). Essential facts about the computer and video game industry. Retrieved from http://www.theesa.com/facts/pdfs/esa_ef_2013.pdf. Accessed August 12, 2014.

Ferdig, R. E. (Ed.) (2009). Handbook of research on effective electronic gaming in education. Hershey, PA: Information Science Reference.

Ferdig, R. E. (2010). Continuous quality improvement through professional development for online K–12 instructors. Lansing: Michigan Virtual University. Retrieved from http://www.mivu.org/Portals/0/RPT_PD_Ferdig_Final.pdf. Accessed August 12, 2014.

Ferdig, R. E., Roehler, L., & Pearson, P. D. (2002). Scaffolding preservice teacher learning through web-based discussion forums: An examination of online conversations in the Reading Classroom Explorer. Journal of Computing in Teacher Education, 18(3), 87–94.

Fogleman, J., McNeill, K. L., & Krajcik, J. (2011). Examining the effect of teachers' adaptations of a middle school science inquiry-oriented curriculum unit on student learning. Journal of Research in Science Teaching, 48(2), 149–169.

Gee, J. P. (2007).What video games have to teach is about learning and literacy (rev. and updated ed.). New York: Macmillan.

Gentile, D. A., Lynch, P. J., Linder, J. R., & Walsh, D. A. (2004). The effects of violent video game habits on adolescent hostility, aggressive behaviors, and school performance. Journal of Adolescence, 27(1), 5–22.

Goggins, S. P., Mascaro, C., & Valetto, G. (2013). Group informatics: A methodological approach and ontology for sociotechnical group research. Journal of the American Society for Information Science and Technology, 64(3), 516–539.

Hedeker, D., & Gibbons, R. D. (1997). Application of random-effects pattern-mixture models for missing data in longitudinal studies. Psychological Methods, 2(1), 64–78.

Hickey, D. T., Ingram-Goble, A. A., & Jameson, E. M. (2009). Designing assessments and assessing designs in virtual educational environments. Journal of Science Education and Technology, 18(2), 187–208.

Hickey, D. T., & Pellegrino, J. W. (2005). Theory, level, and function: Three dimensions for understanding the connections between transfer and student assessment. In J. P. Mestre (Ed.), Transfer of learning from a modern multidisciplinary perspective (pp. 251–293). Greenwich, CT: Information Age.

Honey, M. A., & Hilton, M. (Eds.). (2011). Learning science through computer games and simulations. Washington, DC: National Academies Press.

Kafai, Y. B., & Peppler, K. A. (2011). Youth, technology, and DIY: Developing participatory competencies in creative media production. Review of Research in Education, 35(1), 89–119.

Kapp, K. M. (2012). The gamification of learning and instruction: Game-based methods and strategies for training and education. San Francisco: Wiley.

Ketelhut, D. J., Nelson, B. C., Clarke, J., & Dede, C. (2010). A multi-user virtual environment for building and assessing higher order inquiry skills in science. British Journal of Educational Technology, 41(1), 56–68.

Kidman, G. (2010). What is an "interesting curriculum" for biotechnology education? Students and teachers opposing views. Research in Science Education, 40(3), 353–373.

Klopfer, E., Osterweil, S., & Salen, K. (2009). Moving learning games forward. Cambridge, MA: The Education Arcade.

Klosterman, M., & Sadler, T. D. (2010). Multi-level assessment of content knowledge gains in the context of socioscientific issues based instruction. International Journal of Science Education, 32(8), 1017–1043.

Kuo, M. J. (2007). How does an online game based learning environment promote students' intrinsic motivation for learning natural science and how does it affect their learning outcomes? Paper presented at the First IEEE InternationalWorkshop on Digital Game and Intelligent Toy Enhanced Learning, 2007 (DIGITEL'07), Jhongli, Taiwan.

Lamb, R., Annetta, L., Vallet, D., & Sadler, T. D. (2014). Cognitive diagnostics using serious educational video games. Computers & Education, 70, 92–104.

Lenhart, A., Kahne, J., Middaugh, E., Macgill, A. R., Evans, C., & Vitak, J. (2008). Teens, video games and civics: Teens' gaming experiences are diverse and include significant social interaction and civic engagement. Pew Internet & American Life Project. Washington, DC: Pew Research Center.

Retrieved from http://www.pewinternet.org/Reports/2008/Teens-Video-Games-and-Civics.aspx. Accessed August 12, 2014.

Madden, M., Lenhart, A., Duggan, M., Cortesi, S., & Gasser, U. (2013). Teens and technology 2013. Pew Internet & American Life Project. Washington, DC: Pew Research Center. Retrieved from http://www.pewinternet.org/Reports/2013/Teens-and-Tech.aspx. Accessed August 12, 2014.

Mayer, R. E. (2011). Multimedia learning and games. In S. Tobias & J. D. Fletcher (Eds.), Computer games and instruction (pp. 281–305). Charlotte, NC: Information Age.

Nadolski, R. J., Hummel, H. G., VanDen Brink, H. J., Hoefakker, R. E., Slootmaker, A., et al. (2008). EMERGO: A methodology and toolkit for developing serious games in higher education. Simulation & Gaming, 39(3), 338–352.

National Research Council. (2011). A framework for K–12 science education: Practices, crosscutting concepts, and core ideas. Washington, DC: National Academies Press.

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

NGSS Lead States. (2013). Next Generation Science Standards: For states, by states. Washington, DC: National Academies Press.

Nieswandt, M. (2007). Student affect and conceptual understanding in learning chemistry. Journal of Research in Science Teaching, 44(7), 908–937.

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

NPD Group. (2009). More Americans play video games than go out to the movies. Retrieved from http://www.npd.com/press/releases/press_090520.html. Accessed August 12, 2014.

Penuel, W. R., Fishman, B. J., Cheng, B. H., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. Educational Researcher, 40, 331–337.

Pintó, R. (2005). Introducing curriculum innovations in science: Identifying teachers' transformations and the design of related teacher education. Science Education, 89(1), 1–12.

Rastegapour, H., & Marashi, P. (2012). The effect of card games and computer games on learning of chemistry concepts. Procedia—Social and Behavioral Sciences, 31, 597–601.

Romine, W., & Sadler, T. D. (2014). Measuring changes in interest in science and technology at the college level in response to two instructional interventions. Research in Science Education. DOI: 10.1007/s11165-014-9452-8

Romine, W., Sadler, T. D., Presley, M., & Klosterman, M. L. (2014). Student Interest in Technology and Science (SITS) Survey: Development, validation and use of a new instrument. International Journal of Science and Mathematics Education, 12, 261–283.

Rowe, J. P., Shores, L. R., Mott, B. W., & Lester, J. C. (2010). Individual differences in gameplay and learning: A narrative-centered learning perspective (pp. 171–178). Paper presented at the Proceedings of the Fifth International Conference on the Foundations of Digital Games, Monterey, CA.

Rukavina, S., Zuvic-Butorac, M., Ledic, J., Milotic, B., & Jurdana-Sepic, R. (2012). Developing positive attitude towards science and mathematics through motivational classroom experiences. Science Education International, 23(1), 6–19.

Ruiz-Primo, M. A., Shavelson, R. J., Hamilton, L., & Klein, S. (2002). On the evaluation of systemic science education reform: Searching for instructional sensitivity. Journal of Research in Science Teaching, 39(5), 369–393.

Ruzhitskaya, L., Speck, A., Ding, N., Baldridge, S., Witzig, S., & Laffey, J. (2013). Going virtual . . . or not: Development and testing of a 3-D virtual astronomy environment. Communicating Science: A National Conference on Science Education and Public Outreach, 473, 255.

Sadler, T. D., Romine, W., Stuart, P. E., & Merle-Johnson, D. (2013). Game-based curricula in biology classes: Differential effects among varying academic levels. Journal of Research in Science Teaching, 50(4), 479–499.

Salen, K., & Zimmerman, E. (2004). Rules of play: Game design fundamentals. Cambridge, MA: MIT Press.

Sánchez, J., & Olivares, R. (2011). Problem solving and collaboration using mobile serious games. Computers & Education, 57(3), 1943–1952.

Schneider, R. M., Krajcik, J., & Blumenfeld, P. (2005). Enacting reform-based science materials: The range of teacher enactments in reform classrooms. Journal of Research in Science Teaching, 42(3), 283–312.

Shaffer, D. W. (2006). How computer games help children learn. New York: Macmillan.

Shaffer, D. W., & Serlin, R. C. (2004). What good are statistices that don't generalize? Educational Researcher, 33, 14–25.

Silvia, P. (2006). Exploring the psychology of interest. Oxford, UK: Oxford University Press.

Squire, K. (2006). From content to context: Videogames as designed experience. Educational Researcher, 35(8), 19–29.

Squire, K. (2011). Video games and learning: Teaching and participatory culture in the digital age. New York: Teachers College Press.

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

Squire, K., & Jenkins, H. (2003). Harnessing the power of games in education. InSight, 3(1), 5–33.

Steele, F., & Aubusson, P. (2004). The challenge in teaching biotechnology. Research in Science Education, 34(4), 365–387.

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

Suárez, M., Pias, R., Membiela, P., & Dapía, D. (1998). Classroom environment in the implementation of an innovative curriculum project in science education. Journal of Research in Science Teaching, 35(6), 655–671.

Suthers, D., & Chu, K.-H. (2012). Multi-mediated community structure in a socio-technical network. In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (pp. 43–53). New York: ACM.

Valanides, N., & Angeli, C. (2008). Professional development for computer-enhanced learning: A case study with science teachers. Research in Science & Technological Education, 26(1), 3–12.

Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. EDUCAUSE Review, 41(2), 16.

Vos, M. A. J., Taconis, R., Jochems, W. M., & Pilot, A. (2011). Classroom Implementation of Context-based Chemistry Education by Teachers: The relation between experiences of teachers and the design of materials. International Journal of Science Education, 33(10), 1407–1432.

Zeller, M. F. (1994). Biotechnology in the high school biology curriculum: The future is here! The American Biology Teacher, 56(8), 460–464.