MEDIA, SCAFFOLDS, CANVASES:

A QUANTITATIVE AND QUALITATIVE INVESTIGATION OF STUDENT CONTENT KNOWLEDGE OUTCOMES IN TECHNOLOGY-MEDIATED SEVENTH-GRADE HISTORY INSTRUCTION

A Dissertation

Presented to

The Faculty of the Curry School of Education

University of Virginia

Towards Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Thomas Chalmers Hammond, B.A.

May, 2007

ABSTRACT

In this quasi-experiment, two competing technologies for student creation of multimedia products were implemented in end-of-unit projects in a seventh-grade United States History class. The technologies (PrimaryAccess and PowerPoint) were each used by half the students on one unit and then reversed on a subsequent unit. The student outcomes, student behaviors, and instruction over four units of instruction are described and analyzed, with attention to those behaviors and outcomes emerging from the two units incorporating students' use of technology. No consistent differences were observed on the teacher-designed end-of-unit assessments, but consistent differences favoring the movie-making condition emerged on the semester exam. These differences were statistically and practically significant for the first intervention, and approached statistical significance for the second. A similar pattern of differences emerged in student and teacher actions during the related end-of-unit projects. Students who created movies used more teacher-selected resources, made fewer factual errors, and addressed more material covered on the semester exam. In the movie-making sections, the teacher provided both synchronous (in-class) and asynchronous (out-of-class, reflective) feedback and guided students through a process of iterative refinement as they developed their products. These differences in student and teacher behaviors by condition were strongest during the first intervention and weaker during the second. During the second intervention, student and teacher use of the movie-making application was reduced due to a disruption in internet access; the use of the application between the first and second interventions was therefore not equivalent. Due to low coefficients of reliability and irregularities in the implementation of the design, these results are exploratory and are not conclusive.

Leadership, Foundations, and Policy

Curry School of Education

University of Virginia

Charlottesville, Virginia

APPROVAL OF THE DISSERTATION

This dissertation, Media, Scaffolds, Canvases: A Quantitative and Qualitative Investigation of Student Content Knowledge Outcomes in Technology-Mediated seventh-Grade History Instruction, has been approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

____

Glen Bull, Advisor & Co-Chair

____

Stephanie van Hover, Co-Chair

____

Bruce Gansneder

____

Joe Garofalo

DEDICATION

This work is dedicated to my wife, Shannon Stokes, and my son, John Stokes Hammond. John was born within hours of the proposal, and he has been my constant companion during the research and writing of this dissertation. Of course, this dissertation would not be possible without the intelligence, energy, generosity, encouragement, and love of my wife.

ACKNOWLEDGMENTS

I would like to thank the members of my dissertation committee for accommodating the demands of an extended, classroom-based study, combined with a job search, combined with the distractions of a first-time parent, combined with the tendencies of an easily-distracted doctoral student. Glen Bull initiated this process three years ago when he called me (on his dime!) while I was in Saudi Arabia to tell me that I really wanted to enter the Curry doctoral program, not the Master's. At the time, I thought he was wrong. I've since learned that conversation was simply another instance of Dr. Bull seeing what is just over the horizon. His passion and dedication have proven infectious, and my time with him has been transformative. Like so many others, I'm grateful for what he's talked me into over the past three years, and I look forward to continued collaboration. Stephanie van Hover has invested her time and sanity into transforming me into a reputable thinker, researcher, writer, and teacher-educator. The entirety of my knowledge of social studies theory and education is due to her influence. Bruce Gansneder launched me in my journey towards becoming an education researcher, and his advice has proven true at every turn. Joe Garofalo has been my model for classroom-based research, as well as banjo-playing, guitar-collecting, and bonhomie. Collectively, their insight and good humor have been invaluable in allowing my research, analysis, and writing to mature and find the proper perspective to isolate and articulate the findings.

This line of research would not exist without the prodigious talents of Bill Ferster, the principal developer of PrimaryAccess. If there were an Emmy award for educational technology, he would win in a landslide. In addition to his technical prowess, he also contributed his critical eye to the findings and discussion.

Other faculty offered their time and expertise to the study. I would like to express my appreciation to Patrice Preston Grimes, for being an enthusiastic and insightful commentator on the measures. Meghan McGlinn-Manfra at Old Dominion University has been an invaluable collaborator on this and other projects. Andy Mink of the Virginia Center for Digital History provided his wisdom, and served as my original inspiration for what might be done with images, large-format displays, and powerful techniques for history education.

The participating teachers, administration, and students deserve a heartfelt thank you as well. The lead teacher of the team helped give birth to PrimaryAccess, and the team's continued cooperation and interest has made this research possible. They are outstanding, dedicated educators who deserve more recognition for their efforts than I can hope to give. I wish them and their students continued success.

Finally, I would like to express my gratitude for the extended family of the Center for Technology and Teacher Education. This supportive, highly-skilled group of staff, doctoral students, faculty, and alumni has made this project much more enjoyable.

TABLE OF CONTENTS

LIST OF TABLES

TABLE Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION TO THE STUDY

The Problem and Its Importance

> History, by apprising the people of the past, will enable them to judge of the future; it will avail them of the experience of other times and other nations; it will qualify them as judges of the actions and designs of men; it will enable them to know ambition under every disguise it May assume; and knowing it, to defeat its views (Jefferson, 1787, p. 148).

Jefferson published *Notes on the State of Virginia* in 1787 while serving as the American ambassador to France. The topic of the value and worth of historical knowledge was especially timely: in Philadelphia, the Constitutional Convention was re-writing–both literally and figuratively–the rules of American government, placing more power into the hands of the American electorate than before. The young nation was betting its future on its citizens' ability to judge "the actions and designs of men" (Ibid.).

In the following centuries, the American electorate has expanded, and Jefferson's theme of historical knowledge as a necessity for citizenship has been echoed and amplified (Barton & Levstik, 2004; Gutmann, 1987), with educational institutions playing a restorative role as the continual incubators of democracy (Barber, 1992). The Bradley Commission on History in the Schools (1988), for example, asserts that "[History] is vital for all citizens in a democracy…. History is the discipline that can best help them to understand and deal with change" (p. 5). Similarly, the National Center for

History in the Schools (NCHS) argues that "knowledge of history is the precondition of political intelligence" (NCHS, 1996, p. 41). The National Council for the Social Studies (NCSS) weaves this historical knowledge into its definition of an effective citizen, described as one who "has knowledge of the people, history, and traditions that have shaped our local communities, our nation, and the world" (NCSS, 2001, para. 7). History educators have embraced Jefferson's vision, and seek to prepare historically-literate citizens.

Unfortunately, student outcomes from history instruction are consistently poor (Paxton, 2003). On "the first large-scale test of factual knowledge in United States history," administered in 1915-1916, high school students scored an average of 33% (Wineburg, 2004, para. 5). In 1996 and 2001, the National Assessment of Educational Progress (NAEP) U.S. History examinations demonstrated that not much had changed: half of the 12th grade students scored below the Basic level (Lapp, Grigg, & Tay-Lim, 2002), unable "to identify the significance of many people, places, events, dates, ideas, and documents in U.S. history" (p. 12). Policy-makers look to these assessments not only to measure students' achievement but to spur improvement in education (Grant, 2006, p. 1). Efforts to improve these outcomes must draw upon careful observation of the ways in which students learn history.

Students learn about history both through the formal curriculum, experienced in K-12 and higher education classrooms, and through the informal curriculum, experienced through visits to museums and historical sites, by watching media with historical content, and in conversations with family and community members (Grant, 2003; National Commission on the Social Studies, 1989; Rosenzweig & Thelen, 1998; VanSledright, 2002; Wineburg, 2001). While the informal curriculum clearly affects students' knowledge of history, the K-12 history classroom is the only universal forum through which students can experience both the content of history and the skills of historical thinking (National Commission on the Social Studies, 1989; Ravitch & Finn, 1987).

The K-12 history teacher, therefore, serves as the last reliable agent of influence on students' knowledge of history (Seixas, 2000). Skillful history instruction can make a difference: for example, Smith and Niemi's (2001) analysis of questionnaire data collected during the 1994 NAEP examination found that

> The strongest effect of the history curriculum is tied to the nature of instruction. Methods that involve the increased use of complex writing tasks, in-depth reading, extensive use of student discussion, and the use of learning tools, are strongly related to higher student scores (pp. 33-34).

Smith and Niemi termed this style of teaching "active" instruction. Other conceptions of "best practice" in history instruction include Grant's (2003) call for "ambitious teaching and learning" of history, which takes place "(a) when teachers know well their subject matter…; (b) when teachers know their students well…; and (c) when teachers know how to create the necessary space for themselves and their students" (p. xi). Other researchers point to effective teachers' "wisdom of practice" (Yeager, 2005). These teachers

- Show a good grasp of content knowledge and pedagogical content knowledge…,
- Show enthusiasm for their content, model intellectual curiosity, and interact frequently with their students…,
- Promote critical thinking and/or problem solving…,

- Use different instructional approaches at different times, but whatever approach they take involves students in inquiry, meaningful issues, and classroom activities…,
- Bring in meaningful material beyond the textbook…, [and]
- Attend to their students' academic skills which engaging them in social studies content (p. 5).

Superlative history instruction – whether described as active instruction, ambitious teaching and learning, or wise practice – stimulates students to engage historical content in multiple ways and achieves superior results.

In contrast, the instructional methods used in many history classrooms are uninspiring: lecture, assigned readings from the textbook, worksheets, and standardized assessments (Goodlad, 1984). Furthermore, history teachers are charged with covering an ever-expanding amount of content, whether through the passage of time or the additions of new historiography (Guzzetta, 1969, p. 399). Teachers who began their career covering "Plato to NATO" (Nash, Crabtree, & Dunn, 1997, p. 90) or "colonialism to Clinton" (Grant, 2003, p. 123) must now address terrorism, the wars in Afghanistan and Iraq, and onward. Little wonder, then, as David Lowenthal observes, "history may be harder to learn than is commonly thought" (Lowenthal, 2000, p. 63).

Purpose of the Study

What actions can history teachers take to improve student outcomes in historical content knowledge? Can elements of active instruction or ambitious teaching or wise practice be incorporated into classroom strategies, even by novice teachers? Researchers in technology and social studies argue that technological tools hold the promise of changing the nature of social studies instruction through access to digitized primary source images, geographic information systems (GIS), virtual field trips, and so forth (Cantu & Warren, 2003; Diem, 2000; Martorella, 1997; Milman & Heinecke, 1999, 2000; Whitworth & Berson, 2003). These tools can empower rich pedagogical strategies, such as historical inquiry (Friedman & Hicks, 2006), and may help support the rich teaching practices and positive student outcomes identified by Smith and Niemi (2001), Grant (2003), and others.

A technological tool built for the specific needs and purposes of history instruction may assist a wide cross-section of teachers in engaging students in elements of active instruction and ambitious teaching. The author has assisted in the design, development, and field-testing of such an application. This application, PrimaryAccess, capitalizes on several concurrent trends in technology: the proliferation of historical information on the internet (McMichael, Rozenzweig, & O'Malley, 1996), the increasing ubiquity of internet access (Rainie & Horrigan, 2005), and an explosion in digital media creation by young people (Lenhart & Madden, 2005). PrimaryAccess is a web-based video editor that allows teachers and students to create brief digital documentaries (Ferster, Hammond, & Bull, 2006). The impact of this technology upon teachers and students is not yet known. Does the production of web-based digital documentaries and/or other multimedia produce an observable effect upon students' content knowledge?

Rationale

While classroom teachers have in general been slow to adopt new technologies, or at least slower than their students (Levin & Arafeh, 2002), they have enthusiastically implemented one computer-based technology: PowerPoint slideware. This Microsoft product is well on its way to becoming a genericized trademark. PowerPoint is already ubiquitous in the business world (Tufte, 2006) and is approaching the same level of penetration in the education community. Practitioner conventions such as the National Educational Computing Conference feature dozens of workshops, presentations, and commercial products based on PowerPoint; a search in the H.W. Wilson Education Full Text database turns up hundreds of articles referring to the term. PowerPoint is one of the three most popular applications for grades 6-12 and is used in more than 50% of classrooms (Dynarski, Honey, & Levin, 2002, p. 15).

However, PowerPoint has its detractors (Tufte, 2006, pp. 157-185). The program can empower communication and spark discussion, but it can also encourage a focus on style over substance. A slick presentation can achieve the same effect as a television commercial, encouraging the audience to gloss over complex points or even accept completely inaccurate information. PowerPoint's designers were corporate programmers, not educators (Ibid.).

The expectation was that either PowerPoint (the current standard of use) or PrimaryAccess (the emerging, purpose-built tool) would prove to be more effective at promoting student learning. For example, in their survey of the emerging data on the science of learning, Bransford, Brown, & Cocking (2000) discussed the possibilities for technology to support learning. Specifically, they identified five functions that technology can fulfill (p. 207):

1. Integrate stimulating real-world problem into the curriculum;
2. Scaffold students' thinking as they undertake complex tasks;
3. Provide opportunities for teacher feedback to students and for student reflection and revision;
4. Build communities that bridge classrooms, content experts, and other stakeholders such as parents and administrators; and
5. Extend teachers' professional development.

Some technologies will afford more of these activities than others: more or fewer opportunities for feedback, reflection, and revision; more or less powerful options for scaffolding students' thinking, and so forth. One tool, either a purpose-built application designed for the content area (PrimaryAccess) or a generic application designed for business use (PowerPoint), may better serve the needs of K-12 history teachers as they implement the curriculum in their classroom instruction.

Research Questions

This study focused on instructional strategies integrating two competing technological tools and specifically examined students' historical content knowledge learning outcomes as dictated by the local curriculum guide and assessment practices. Because this study involved students' use of complex technological tools to address open-ended tasks, the research also encompassed the context of the students' actions, such as the technological infrastructure and patterns of student grouping. The specific research questions to address these topics are:

1. Do differences exist in student outcomes, as measured by pre- and posttest scores on the standards-driven assessment, between students who use PrimaryAccess vs. those who generate a PowerPoint slideshow?
2. If differences exist, does the effect on student outcomes vary by achievement-level grouping?
3. What environmental factors – such as the curricular context, teacher behaviors, student behaviors, or technological infrastructure – appear to inhibit or promote this effect?

Overview of the Methodology

One seventh-grade history teacher and his students participated in the study. This teacher had six sections of 10 to 25 students each, with each section meeting for 45 minutes daily. The sections were tracked by achievement-level grouping. The study spanned one semester and included four units of study.

The design followed an untreated control group design with multiple pretests and posttests (Shadish, Cook, & Campbell, 2002, p. 143). For one unit of study half of the sections used PrimaryAccess, and the other sections used PowerPoint. Those sections using PrimaryAccess composed short (less than two minutes in length) digital movies about topics within the unit of study. The teacher supplied the visuals for both conditions, generating a list of relevant images that students could choose from. The audio tracks for the digital documentaries generated in the PrimaryAccess groups were the students' own voices as they commented on the visual images used and explained their topic to the viewer. The sections using PowerPoint composed short (approximately ten slides) presentations that addressed the same topics and drew upon the same images as the PrimaryAccess sections. (See Figure 1.)

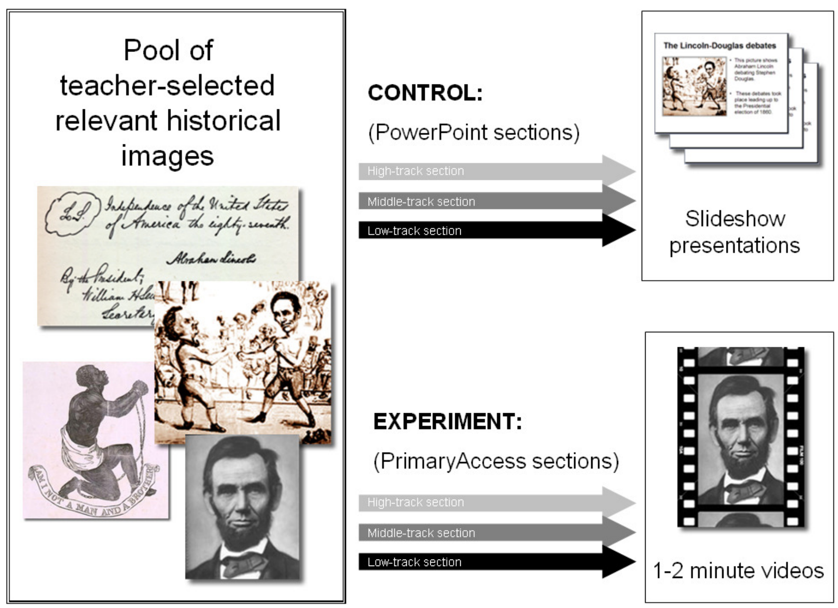


Figure 1: Control Versus Experimental Conditions. Three sections experienced the control condition, and three sections experienced the experimental condition. Each section accessed a teacher-selected pool of relevant historical images and produced a product. The control groups used PowerPoint to produce a slideshow; the experimental groups used PrimaryAccess to produce a brief digital movie.

Each section received the pretests and posttests assigned by the local school authorities. (See Table 1.) The first pretest was the relevant subscale from a beginning-of-year multiple-choice test covering the year's content. The second pretest was a series of open-ended questions covering the main topics of the unit. Each pretest was compared against the posttest, which contained similar or identical prompts.

Table 1

Units, Observations, Treatments for Each Unit of Instruction Within the Study

Achievement-level grouping	pretest A	pretest B	Instructional unit	posttest
High-track A	O	O	E (use PrimaryAccess)	O
Middle-track A	O	O	E (use PrimaryAccess)	O
Low-track A	O	O	E (use PrimaryAccess)	O
High-track B	O	O	C (use PowerPoint)	O
Middle-track B	O	O	C (use PowerPoint)	O
Low-track B	O	O	C (use PowerPoint)	O

Table 1: Untreated control group design with pre-test and post-test across six sections grouped by achievement level. The two pre-tests were: A.) a beginning-of-year multiple-choice test covering the year's content, and B.) a beginning-of-unit set of open-ended questions covering the unit's main topics.

The quantitative analysis compared pretest and posttest data within achievement-level groupings to detect significant differences. The researcher evaluated the reliability and validity of the pretest and posttest instruments. This analysis addressed research questions 1 and 2, describing students' outcomes of historical content knowledge.

The qualitative data supplemented the quantitative data. This qualitative data was gathered through classroom observations, document analysis of student products, and interviews with the teacher. The analysis of the qualitative data explored research question 3, identifying the factors that affected observed differences in student outcomes.

The design was repeated on a subsequent instructional unit, with the two cohorts trading conditions: those who initially experienced the PowerPoint condition experienced the PrimaryAccess condition, and vice versa. This switching replication allowed the researcher to repeat the experiment and determine whether the effect observed during the first intervention was repeated during the second intervention, with conditions reversed for each group.

Significance of the Study

Dissertations on technology use in social studies have not focused on student outcomes, but rather teachers' use of technology (DeWitt, 2004; Mason, 1998; Friedman, 2004; Swan, 2004). When student outcomes are addressed, researchers have focused on students' historical thinking (Waring, 2003; Whitaker, 2003). However, assessments of students' historical content knowledge, while disputed as an indicator of students' true historical understandings (Grant, 2006), are the current benchmark of instructional success in American classrooms (VanSledright & Limon, 2006) and in the popular imagination (Paxton, 2003). By examining student outcomes in historical content knowledge from using a competing technology-integration strategies, this study will provide the comparison called for by Whitaker (2003, p. 99), using a design that allows for some inference of causation (Shadish, Cook, & Campbell, 2002) by the technology and teaching strategy used.

Furthermore, this study will provide a service to the local school authorities by exploring the reliability and validity of their assessments at the beginning and end of each unit of study.

Summary

Careful, discipline-specific uses of educational technology have the potential to improve student outcomes in history instruction. Most of the research on educational technology in history classrooms has explored teachers' use of technology and/or the development of students' higher-order thinking skills regarding history.

This study followed a quasi-experimental design. The control group used internet browsers and PowerPoint, common technologies in many technology-enriched history classrooms. The experimental group used a purpose-built tool, PrimaryAccess. The experiment was repeated on a subsequent unit of study, with conditions reversed for each group. Progress in all groups was assessed by beginning- and end-of-unit content assessments and beginning-of-semester and end-of-semester exams. Connections between students' performance on the tests and their experience of instruction and use of the tools was explored via classroom observations, interviews, and document analysis.

Following this chapter is a review of the literature regarding instructional technology, social studies education, and technology in social studies education.

CHAPTER 2

REVIEW OF THE LITERATURE

Introduction

> A substantive, thorough, sophisticated literature review is a precondition for doing substantive, thorough, sophisticated research (Boote & Beile, 2005, p. 3).

As the quotation above illustrates, to ensure that a proposed research project is responsive to previous studies, researchers must conduct a thorough review of the literature and provide a grounding in the frameworks that inform the research questions. The literature review must relate the findings of this previous research to the questions at hand and critically examine the methodology used to reach those conclusions (Krathwohl, 1998, Ch. 2).

This study explored three questions:

1. Do differences exist in student outcomes, as measured by pre- and posttest scores on the standards-driven assessment, between students who use PrimaryAccess vs. those who generate a PowerPoint slideshow?
2. If differences exist, does the effect on student outcomes vary by achievement-level grouping?

3. What environmental factors – such as the curricular context, teacher behaviors, student behaviors, or technological infrastructure – appear to inhibit or promote this effect?

The context for these questions–use of technology in classroom instruction to support a curriculum–is complex. This literature review examines several conceptual frameworks relevant to this scenario of students in a history classroom using technology to acquire historical content knowledge:

- The role of technology in American education, with attention to the intersection of learning theories and technology;
- The context of history education in America, exploring competing purposes and patterns in instruction and assessment;
- The impact of educational accountability movements and high-stakes assessments on classroom instruction and assessment;
- Previous efforts to integrate technology into history education, with attention to research examining student outcomes from the integration of technology into history education.

Technology and American Education

American education, and American society as a whole, has historically held a certain fascination with technology. Europeans invented the first programmable machines (i.e., the Jacquard loom in the nineteenth century), but an American, B.F. Skinner, conceived of a *teaching* machine. As new technologies are created–such as film, radio, television, and computers–enthusiasts rapidly apply (or misapply) them to educational contexts (Saettler, 1990). True to the American tradition of a market-oriented culture, promoters of these innovations present their wares as a solution to the problems of teaching and learning. According to one educational technology developer,

> Books will soon be obsolete in the schools. Scholars will soon be instructed through the eye. It is possible to teach every branch of human knowledge with the motion picture. Our school system will be completely changed in ten years (Saettler, 1990, p. 98).

The speaker was Thomas Edison, quoted in a New York newspaper in 1913. A more recent commentator, failing to learn from Edison's overstatement, asserted that

> We are at the onset of a major revolution in education, a revolution unparalleled since the invention of the printing press. The computer will be the instrument of this revolution. ... By the year 2000 the major way of learning at all levels, and in almost all subject areas will be through the interactive use of computers (Bork, 1980, p. 53).

These prophets were–and still are–wrong, or at least not entirely correct. American schools are not "completely changed." While futurists generally make the error of *under*stating how different the future will be from the present (Gilbert, 2006), those making predictions about the future of education tend to *over*state the difference.

American schools and schooling have demonstrated a remarkable resiliency since the advent of the common school in the 1800s. Well into the 20th century, American students and teachers have continued to use technologies that were familiar to the generations before Edison. Few or no other professions remain so untouched by the changes wrought over the past century (Papert, 1993). Goodlad (1984), in a survey of American public and private schools, characterized history education as a field that is particularly un-touched by technology. History students' class time is spent "listening, reading textbooks, completing workbooks and worksheets, and taking quizzes" (p. 213).

The use of these paper-based practices has persisted despite significant spending on educational technologies. Following the launch of the first satellite, Sputnik, by the Soviet Union in 1957, the American government and national opinion-leaders concluded that education was a strategic arena for competition in the Cold War: the communist bloc had launched a satellite into orbit before the United States because, among other reasons, the existing patterns in American education had failed to produce adequate numbers of scientists, mathematicians, and engineers. Accordingly, unprecedented sums were poured into developing educational media, with a focus on educational television (Saettler, 1990). Edison's individual enthusiasm was repeated on a societal scale, as the nation expressed its faith in the transformative power of technology and programmed instruction "to design effective and replicable instruction" (Heinich, 1995).

However, when subjected to the rigors of education research, the large-scale, technology-based systems for delivering instructional media such as educational television have failed to improve student outcomes. While a study in El Salvador did indicate improvements after introducing a nation-wide educational television system (Schramm, 1970), the improvements actually stemmed from the simultaneous introduction of a new curriculum, not from the use of educational technology (Clark, 1983). The conclusion, after decades of research, is that the use of technology during instruction does not alter students' learning outcomes as long as the instructional paradigm is based on a transmission model of delivering information from the instructor to the students (CTGV, 1996). In Clark's summary,

> The best current evidence is 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. Basically, the choice of vehicle might

> influence the cost or extent of distributing instruction, but only the content of the vehicle can influence achievement (Clark, 1983, p. 445).

Since his landmark meta-analysis, Clark's conclusion has been contested (Kozma, 1994), but also re-confirmed (Dynarski et al. 2007; Ely, Foley, Freeman, & Scheel, 1992).

In contrast to the transmission model, in which the student is viewed as an empty vessel to receive information and skills from the instructor, a competing philosophy in American education emphasizes student activity and the value of learning through experience. An American, John Dewey, was the pioneer and proponent of the method (1916), but a Russian psychologist, Lev Vygotsky, formulated the building blocks of the actual theory of learning, constructivism. Vygotsky attempted to explain the mental processes of knowledge-building that take place during learning (Hedegaard, 1990). The central insight of constructivism is that all learning is socially mediated; all new information is absorbed (or not) through integration with pre-existing schema shaped by cultural values and personal experiences (Doolittle & Hicks, 2003). Vygotsky also developed the concept of the zone of proximal development to describe tasks that the learner can execute with the proper assistance, or scaffolding (Hedegaard, 1990, p. 349). A Swiss psychologist, Jean Piaget, generated a similar and overlapping (but not identical) set of observations about student learning. Because Vygotsky's work remained little-known outside Russia until the 1960s, Paiget's description of stages of cognitive development provided the framework of constructivist theory in the West (Cole & Wertsch, 1996; Wertsch, 1998). Seymour Papert, a student of Piaget, has proposed students' use of technology as a natural forum for constructivist learning (Papert, 1981). By studying topics within technology, such as the LOGO programming language, students can discover academic concepts in other fields, such as geometry and algebra.

Papert gave a clear statement of learning and the road forward for education: “children learn by doing and by thinking about what they do. And so the fundamental ingredients of educational innovation must be better things to do and better ways to think about oneself doing these things” (1980, p. 161).

While constructivism has proven useful as a theory of learning, its application to teaching is more complex (Marlowe & Page, 2005). One difficulty is the problem of transfer: learning that has taken place in one context does not always translate into knowledge that is useful in other contexts (Bransford, Brown, & Cocking, 2000). A study of elementary students by the Cognition and Technology Group at Vanderbilt found “no evidence that learning LOGO can aid in transfer to situations that differ substantially from the LOGO environment” (Littlefield et al., 1988, p. 111). Students in the experimental group did perform better than those in the control group on tasks that were very similar to programming (e.g., given a map, write a set of directions to move from one point to another) but showed no difference on tasks that were removed from this context (e.g., an organization of dots test). Careful teaching strategies can promote transfer. When instruction provides students with not just information but structures for these concepts, learners can more readily apply their knowledge to new settings (Bransford, Brown, & Cocking, 2000).

Since the time of Edison’s comment, American society has displayed unflagging optimism about the power of technology to improve education (Shrock, 1995). The funding poured into educational technology during the post-Sputnik era demonstrates that, when provided with a powerful stimulus, the government is willing to invest in technologies for schools. The body of research on technology in education suggests that the focus ought to be placed not just on the presence or absence of technology in the classroom or the use or non-use of technology during instruction, but on the instructional paradigm in which the technology is used. Traditional, transmission-oriented instruction does not appear to be improved by the use of technology (CTGV, 1996). Constructivist or constructionist instruction can be improved by the inclusion of educational technology, but the effect depends upon skillful teaching-for-transfer (Bransford, Brown & Cocking, 2000) and content-specific technological tools (Hickey, Moore, & Pellegrino, 2001). A third possibility is a return to the Skinnerian ideal of a teaching machine, a simple stimulus-response device (Kingsley, 2005). Three roads lie open for the further application of technology to American education. Which path should the educational technology community follow?

The question arises because American society is yet again in the midst of placing a very large bet on technology for education (Paige, 2005). Current national spending on educational technology is well over $6 billion annually (Angulo, 2001), and includes billions from the federally-mandated E-Rate program to wire all public schools to the internet (Goolsbee & Guryan, 2006). The effect of these expenditures is visible in the rising rates of school access to the internet and dropping student-to-computer ratios (Wells & Lewis, 2006). However, a longstanding criticism has been that educational spending on technology goes to naught: computers can be put in classrooms, and classrooms can be wired to the internet, but this capacity goes unused (Cuban, 2001) or is used for academically un-productive purposes (Oppenheimer, 2003). According to Edward Tufte (2006), some technology tools for communication actually impede clear expression, and clear thinking.

While the use of generic technologies for content-specific purposes has not succeeded, other strategies show more promise. One component to a new, more productive strategy is a shift in the understanding of technology itself. Rather than view technology as a product to be delivered, such as programmed instruction, some instructional technologists have come to view technology as a process, a tool that supports users as they accomplish their goals (Heinrich, 1995). A second component of a more productive approach to education technology is the use of purpose-built, content-specific technologies. In mathematics education, for example, students using the Jasper Woodbury application achieved significant gains on tests of mathematical problem-solving compared to students receiving traditional instruction (Hickey, Moore, & Pellegrino, 2001). A third component is the use of new, more powerful technological tools. The current push to wire schools with internet access and place computers in the classroom comes at the same time as a dramatic increase in what internet users are able to do (O'Reilly, 2005). The internet had been conceived of as a read-write medium, in which users could create and consume information with equal facility (Berners-Lee, 1999). However, the first widespread use of the internet was as a read-only medium: web browsers, much like the viewers of Edison's educational films, could only receive information. With the advent of Web 2.0, however, users can both read and write the web, through tool such as wikis and blogs, and can express themselves in an increasing variety of media, such as podcasts or video blogs (O'Reilly, 2005). Papert's 1980 call for "better things to do and better ways to think about oneself doing these things" (p. 161), at least in a mechanical sense, is being addressed.

Social studies teacher-educators recognize that constructivism is a useful theoretical foundation to guide the meaningful integration of technology into social studies education (Doolittle & Hicks, 2003). Through the lens of constructivist pedagogy, teachers can move beyond using technology to replicate existing patterns of instruction or viewing technology as an efficiency tool that allows them to cover more content more quickly (Ibid.). Constructivism provides teachers with a framework for exploring how technology can create more powerful educational experiences. Specifically, technology can empower inquiry-driven learning, engagement of students' prior knowledge, feedback regarding students' learning, interaction within and beyond the classroom community, access to information and authentic historical artifacts, and the expression of students' "autonomous, creative, and intellectual thinking" (p. 92).

However, the previously-noted challenges regarding the integration of technology into instruction still await. Students may learn more about the technology itself than the intended content (Ely, Foley, Freeman, & Scheel, 1992), or they may show no gains in learning at all (Clark, 1983). After all, an activity in which students create a digital documentary about a historical event can be not only an opportunity to engage in historical thinking but also to be "seduced by the bells and whistles of the technology and [lose] sight of the primary goal of the assignment" (Hofer & Owings-Swan, 2005, p. 107).

In 1997, two veterans of social studies education and technology surveyed possible intersections between these two domains. Peter Martorella, excited by the tremendous potential of technology to empower social studies teaching and learning, seized upon the metaphor of technology as "a sleeping giant in the social studies curriculum" (p. 511). This image of a sleeping giant has been cited by many, but as Doolittle and Hicks note, "the sleeping giant has been having quite a long nap" (2003, p. 74). In the same year, Charles White sounded a cautionary note about the application of technology to social studies:

> The American social studies scholar Shirley Engle reminded us...of a comment by Albert Einstein—that, for the first time in history, there is a surplus of means over ends. Technology is a means to an end, and is of limited usefulness in helping us decide on important ends. The field of social studies education is best served by attending first to the ends—to what is worth knowing, what is worth contemplating, what is worth doing. Then we can determine what tools will help us extend our reach to these ends (White, 1997, pp. 9-10).

Before we can evaluate the use of an educational technology, we must consider the instructional purpose to which it is being applied. The next section, therefore, explores the purpose and practice of history education in American schools.

History Education in American Schools

History education in the United States is a highly political topic. In 2004, for example, the U.S. Department of Education destroyed 300,000 copies of a 57-page booklet entitled "Helping Your Child Learn History." The department made minor modifications to three sentences and published 300,000 replacement copies. The instigator was not the Secretary of Education but a politically-connected private citizen, the wife of the sitting Vice President (Nash, 2004). This incident is only the latest in a series of attacks and counter-attacks over history, the history curriculum, and materials for history instruction, stretching back to the early 19th century (Nash, Crabtree, & Dunn, 1997).

History education is the "most contentious field of the curriculum" (NCHS, 1996, p. ix) because it is the cultural high ground. As sociologist James Loewen asserts, "More than any other topic, [history] is about *us*. Whether one deems our present society wondrous or awful or both, history reveals how we arrived at this point" (1995, p. 2). Every event, seemingly, carries some connection to the history curriculum. In a speech to the Dallas Institute of Humanities and Culture less than a month after the attacks of September 11, 2001, Lynne Cheney asserted that

> At a time of national crisis, I think it is particularly apparent that we need to encourage the study of our past. Our children and grandchildren—indeed, all of us—need to know the ideas and ideals on which our nation has been built. We need to understand how fortunate we are to live in freedom. We need to understand that living in liberty is such a precious thing that generations of men and women have been willing to sacrifice everything for it. We need to know, in a war, exactly what is at stake (Cheney, 2001, para. 9).

A historically-literate citizenry is part of the common assumptions underpinning American government (Bradley Commission, 1988; NCHS, 1996; NCSS, 2001; Ravitch & Finn, 1987; Virginia Department of Education, 2001). But what does it mean to be historically literate? What knowledge and what skills are integral to this definition?

History educators have no shortage of instructional frameworks from which to choose. The federal government has adopted one set of content standards (NAEP U.S. History Consensus Project, 2001) after hotly rejecting another (Nash, Crabtree, & Dunn, 1997). At the state level, 49 states and the District of Columbia have adopted some definitions of the social studies knowledge and skills students are to have. In addition to these governmental authorities, organizations such as NCSS and NCHS have created and documented their own visions of national standards (Hill, 2006; NCSS, 1994; NCHS, 1996). Even private citizens, such as English Literature professor E.D. Hirsch, have devised their own expectations of what students should know (Hirsch, 1995).

While acknowledging the debate over curriculum, researchers in history education have chosen to focus not on *what* history to teach and learn, but *how* history is taught and learned (Cantu & Warren, 2003; Donovan & Bransford, 2005; Levstik & Barton, 1997). Barton and Levstik (2004), for example, identify four purposes that practitioners pursue through history education.:

1. Identification: History-for-identification turns on the premise that "History tells us who we are." Teachers present content to students that encourages them to identify with a particular community, most commonly the national community—as defined by the presenters of the curriculum.
2. Moral response: History-as-a-moral-response seeks to engage students in forming judgments about the past. Students are encouraged to celebrate the good, condemn the bad, and remember those who came before.
3. Analysis: History-as-analysis requires students to deconstruct historical information, breaking down sources or master narratives, and to engage in the construction of their own historical accounts, as they form their own interpretations of evidence and identify connections and patterns. This practice comes closest to emulating the practice of professional historians.
4. Exhibition: History-as-exhibition is the centerpiece of most classroom assessment, holding students accountable for knowing terms and being able to answer questions. Exhibition is also practiced outside the classroom curriculum, through both formal institutions such as museums and less formal

arrangements such as historical re-enactments or private collections of historical artifacts.

These four stances allow history education researchers to tease apart the more political elements from the more academic elements. Identification and moral responses are lightening rods for political conflict; Lynne Cheney's call to "know what is at stake" (2001) is a call for students to identify with America and "to celebrate the good things in history and condemn the bad" (Barton & Levstik, 2004, p. 91). Analysis and exhibition, while bearing strong political implications, are the primary objects of interest for historians and history educators (Ashby, 2005; Barton, 1997a, 1997b; Barton & Levstik, 1996; Booth, 1993; Brush & Saye, 2002; Donovan & Bransford, 2005; Lee & Molebash, 2004; Shemilt, 1987; VanSledright, 2004). When the authors of the *National Standards for History* framed the standards in terms of "historical thinking skills" and "historical understandings" (NCHS, 1996, p. 42), they signaled that the focus of the history classroom should be analysis and exhibition, not identification and moral response. Analysis requires the exercise, and exhibition demands the display, of both historical thinking skills and historical understandings.

Intended Student Outcomes from History Education

History education is a profoundly political arena (Lowen, 1995; Nash, Crabtree, & Dunn, 1997; Symcox, 2002; Zimmerman, 2002); similarly, history as a knowledge domain spurs its own debates (VanSledright & Limon, 2006, p. 546). Every framework for history instruction–whether promulgated by state or district authorities, the National

Assessment Governing Board, or a professional association of historians or history educators–contains its own definition of the intended student outcomes. These outcomes are a combination of *historical thinking skills* and *historical understandings*.

History educators and history education researchers have struggled to establish a clear definition of historical thinking. R.G. Collingwood, a philosopher of history, described the discipline as being, "like…natural science,…a special form of thought" (in Cantu & Warren, 2003, p. 20). Like natural science, historical thinking requires analysis, especially "when working with evidence to construct historical explanations or accounts" (Barton & Levstik, 2004, p. 69). British researcher Martin Booth (1987) asserts that historical thinking goes beyond the scientific process and instead is "a form of speculation, directed imagination or vicarious living" (p. 27). Wineburg prefers a different term, a historical "way of knowing" that allows the historical thinker to "navigate the tension between the familiar and the strange" (2001, p. 17).

One strategy for defining historical thinking is to identify its sub-components. The NCSS and NCHS, for example, do not attempt an omnibus definition of historical thinking but instead distill the concept into several distinct skills (see Table 2). British

Table 2

Definitions of Historical Thinking by NCSS and NCHS

National Council on the Social Studies' Teacher Standards Volume I: Disciplinary Standards (1997), "History"	National Center for History in the Schools' *National Standards for History* (1996), pp. 6-7
"Historical thinking skills enable learners to • evaluate evidence, • develop comparative and causal analyses, • interpret the historical record, and • construct sound historical arguments and perspectives	[There are] five types of historical thinking: • Chronological thinking, • Historical comprehension, • Historical analysis and interpretation, • Historical research capabilities, [and] • Historical issues-analysis and decision-making.

education researchers have gone a step further, not only identifying a set of historical thinking skills (e.g., time, causation, change and continuity, motivation, evidence) but identifying a sequence of stages of progression in each skill (Donovan & Bransford, 2005; Sansom, 1987).

Efforts to develop a coherent curriculum around progression through these skills have been hindered by the fact that development is not, contrary to Piaget's implication, age-dependent. Instead, progression is idiosyncratic and inconsistent from one student to the next. For example, individual students can skip levels in the sequence (Donovan & Bransford, 2005). One solution is to measure groups, not individuals. When measuring a series of age-defined groups, researchers observe a trend of progression in historical thinking and understanding by age (Barton & Levstik, 1996; VanSledright & Limon, 2006, p. 549), with some exceptions (Bell & McCollum, 1917; Brophy & Alleman, 2000; Guzzetta, 1969). However, British researchers have identified a persistent finding of "the 7-year gap": "the conceptual understandings of *some* 8-year-old students [can] be more advanced than those of *many* 14-year-olds" (Donovan & Bransford, 2005, p. 82). An intensive study of 4th- and 6^{th}-grade students found that "on the whole, the 6^{th}-grade students operate a higher conceptual level than those in the 4^{th} grade, but the conceptual understanding of *some* 4^{th}-grade students is more advanced than that of *some* 6^{th}-grade students" (p. 166). Even when history education researchers conducted an experiment in instruction designed to encourage progression in historical thinking, "a significant minority [of the control group] achieved results as good as those of the experimental group" (Sansom, 1987, p. 118). History educators, therefore, continually operate in mixed-ability grouping on historical thinking, regardless of tracking, and "must accommodate a ‘7-year-gap’ between the ideas of the lowest- and highest-attaining students” (Donovan & Bransford, p. 82).

In contrast to the difficulty and uncertainty in identifying historical thinking skills, historical content knowledge is more easily defined. Historical content knowledge, also referred to as *first-order substantive knowledge*, is the specific facts that students are to know: “descriptions and explanations about what occurred, who was involved, when things happened, in what larger historical context, and what it means when taken together” (VanSledright & Limon, 2006, p. 546). Each curricular framework contains its own specification of these particulars, whether “the reasons for and results of the Spanish-American War” (Virginia Department of Education, 2001, p. 24) or the more detailed requirement of “causes and consequences of the Filipino insurrection” (NCHS, 1996, p. 109). In addition to answering who-what-when-where-how-why questions, historical content knowledge also encompasses substantive concepts such as *war*, *empire*, *imperialism*, and *insurrection*. Students’ understanding of these substantive concepts requires “knowing a rule (what makes something a migration, for example) and being able to identify instances of that rule” (Donovan & Bransford, 2005, p. 61).

While history education seeks to strengthen both students’ historical thinking and their historical understandings, most textbooks and most assessments “exclusively sample students’ recall or recognition” of specific items of historical content knowledge (VanSledright & Limon, 2006, p. 546). Historically, classroom instruction has favored content coverage, and teacher-designed tests have stressed factual recall (Cuban, 1991). A study of students’ historical content knowledge outcomes must therefore consider the context of its assessment in the United States.

Assessing Students' Historical Content Knowledge

Historical content knowledge has traditionally been assessed using some form of standardized test (Kurfman, 1991). The tradition of assessing Americans' historical content knowledge dates back to 1915-1916, when J. Carleton Bell and David F. McCollum tested more than 1,500 Texas students in elementary school, high school, and college. Their test, "the first large-scale test of factual knowledge in United States history" (Wineburg, 2004, para. 5), consisted of 40 open-ended prompts asking students to identify significant events associated with dates (e.g., 1861), events associated with men (Abraham Lincoln), men associated with events (the President of the United States during the Civil War), and historic terms (the Nullification Ordinance of South Carolina, the Emancipation Proclamation). The high school and college students were additionally asked to

- Make a list of all the political parties that have arisen in the United States since the Revolution, and state one political principle advocated by each….
- Indicate the great divisions or epochs of United States history, [and]
- On an outline map of the United States (supplied)[,] draw the land boundaries of the United States at the close of the Revolution and indicate the different acquisitions of territory since that date (Bell & McCollum, 1917, p. 259)

The researchers selected these topics and tasks based upon "their own opinions, supplemented by the advice of several high school and college teachers of history" (Ibid.). They acknowledged that this selection lacked "any objective basis" but hoped that the questions "would range from easy to difficult, …[and] would touch upon as many different phases of history teaching as possible…" (Ibid.). The researchers felt that these questions "could be answered briefly,…easily scored as right or wrong, and…could be attempted by all pupils in the thirty-minute period ordinarily given to history in the elementary schools, or the forty-minute period of the high schools" (Ibid.) The results ranged from the predictable (the most-recognized item was the date 1492) to the culturally-interesting (many students identified Jefferson Davis, not Abraham Lincoln, as the United States president during the Civil War), to the inexplicable (freshmen and juniors at the University of Texas decisively out-scored their senior school-mates) (pp. 268-271).

The researchers' commentary on the results displays an appreciation for the complexity of historical information: "Jefferson Davis was frequently confused with Thomas Jefferson.… The Articles of Confederation were ascribed to the Confederacy" (p. 271). Detailed understandings, such as the Nullification Ordinance of South Carolina and "the fact that the Emancipation Proclamation was a war measure and applied only to the states then in rebellion[, were] quite generally overlooked" (Ibid.). The researchers also highlighted the difficulty of evaluating student responses on the more complex prompts regarding the political parties, historical eras, and the identification of map regions. Despite their intention to elicit clear, right-wrong answers, they had to improvise the scoring protocols on these questions.

Bell and McCollum's study highlights what have proven to be three enduring trends in the assessment of historical content knowledge:

1. The test itself is questionable, and
2. The results are unstable and possibly uninterpretable.

3. Regardless, the results of these assessments are presented as conclusive, rather than tentative. They ignore the shortcomings of the test and focus on the bottom line, i.e., the posited finding on student knowledge.

First, in assessing historical content knowledge, the testing process is itself a limiting factor (Grant, 2003, p. 92). Assessors' desire for easy-to-administer and easy-to-score items runs at odds with the complex understandings involved in even seemingly simple historical identifications, such as the Emancipation Proclamation's limits of authority. Students may understand a concept but lack the vocabulary called for by the question (Brophy & Alleman, 2000). Furthermore, students' responses are more complex than simple right/wrong scorings can capture (Wineburg, 2004). A student who confuses Jefferson Davis for Thomas Jefferson may be only mixing up names; a student who confuses Jefferson Davis and Abraham Lincoln may have misunderstood the question, thinking that it asked about the leader of the Confederacy, not the federal government. Conversely, a student may be able to offer a correct answer despite a lack of understanding. A 1995 study providing an in-depth analysis of students in three elementary and middle school classrooms concluded "that students with the same scores on the achievement test were unlikely to know, or have learned, the same content" (Nuthall & Alton-Lee, 1995, p. 192). Finally, the content of the test is often arbitrary and may not be aligned with local curricula or reflect what students truly know and understand.

A second theme illustrated by the 1917 report is that students' performances on historical content knowledge assessments exhibit tremendous variability. While between-group differences may appear, individuals' scores span a tremendous range. Bell and

McCollum noted that "Each class and group studied shows a wide variation in individual scores…. Some seventh grade pupils make as good a score as the best university senior" (p. 274). Further exploration of the data turns up counter-intuitive results. For example, the authors examined gender differences, anticipating that girls would out-perform boys, since "girls usually get higher school marks in history than boys, and [since] in co-educational colleges where many elective courses are offered the girls seem to prefer history and literature" (p. 272). They were surprised to observe that boys outperformed girls, both as a whole and within each school, in almost every category—with one exception. On the historical periods question, the girls outperformed the boys both as a whole and within four out of five schools. Furthermore, the girls were markedly better in identifying the event associated with 1492. Bell and McCollum merely report these findings without attempting to explore their meaning. If nothing else, these inexplicable patterns may suggest that the test is flawed. The test items, with the exception of the historical periods question, may display a systematic gender bias favoring boys. Alternatively, the scoring of this test item may have been influenced by gender bias on the part of the scorer. The explanation of the scoring protocol noted that "the divisions of history are so much a matter of individual taste and opinion that the only course left for the examiner was to give a good mark in case a *reasonable* [italics added] outline of United States history was given" (p. 272). The students' names were written on the back of the tests, and all tests were scored by the same person; a gender-biased scorer could therefore routinely assign higher grades to girls than to boys for the same answer, and thus be the cause of at least some of the puzzling gender differences in performance.

Considering these threats to the test's validity, readers would be advised to take Bell and McCollum's findings lightly.

Finally, the 1917 study, like many subsequent efforts, maintains that the result of the historical content knowledge assessment is valid and conclusive, with the conclusion being that American students are historically ignorant (Paxton, 2003; Wineburg, 2001). The average scores for each cohort (elementary, high school, and university) were 16%, 33%, and 49%, respectively. In the spirit of the deficit model of instruction, Bell and McCollum comment that "This does not show a very thorough mastery of *basic historical facts* [italics added]" (p. 274). Focusing on the high school results, the researchers note that while several of the schools within their study "have the reputation of being very well administered and of having an exceptionally high grade of teachers,…Surely a grade of 33 in 100 on the *simplest and most obvious facts of American history* [italics added] is not a record in which any high school can take great pride" (pp. 268-269). Furthermore, the authors do not limit their interpretation to the tested historical content knowledge but extrapolate to more general skills: "ability is [*sic*] responding to tests of historical facts is a fair index of general historical ability" (p. 258). Gustave Flaubert coined a term that captures this behavior: *la rage de vouloir conclure*—the rage to conclude, to ignore the "implications, complexities, and uncertainties of primary evidence" (Tufte, 2006, p. 154).

Subsequent large-scale assessments of historical content knowledge have borne out the same pattern of weak testing methodology, erratic test performance, and yet thunderous conclusions. In 1986, Diane Ravitch and Chester Finn oversaw the first NAEP in U.S. History. They assessed a nation-wide sample of 7,812 students–from all regions of the country and from both public and private schools–over 141 multiple-choice questions in history. The authors describe the items as

> cognitive or ‘knowledge’ questions…. The questions were not difficult. Students did not have to analyze or interpret a passage, perform a calculation, intuit a relationship, construct an analogy, or puzzle out a multistage problem. The assessment gauged students’ knowledge of *basic information* [italics added] in history (Ravitch & Finn, 1987, p. 43).

Like Bell and McCollum, Ravitch and Finn draw a connection between content knowledge and deeper understandings: “any fact worth knowing illumines at least one important concept: knowing what Magna Carta is helps to explain the evolution of the concept of limitations on the power of the sovereign and the origins of political democracy” (p. 17). A committee of 7 (including Ravitch and Finn) oversaw the composition of the items, in consultation with 59 “professionals in the field of history and social studies” (p. 28). The shift to a multiple-choice question format expedited the test administration and eliminated the ambiguities Bell and McCollum faced in scoring their open-ended items. However, “guessing ensued and this tended to inflate the results” (p. 46). Even accepting any inflation, the results were, in Ravitch and Finn’s view, “extremely weak” (Ibid.): the nation-wide average score was 54.5%. If one were to take Bell and McCollum’s findings at face value, this is a dramatic improvement over high school students’ performance in 1917. However, Ravitch and Finn chose the conventional American classroom scale (90-100% = A, 80-89% = B, and so forth) as their basis for judgment and characterized this 54.5% as “failing” (p. 45). A careful examination of the scores reveals patterns that suggest a flawed test. As on Bell and McCollum’s exam, boys outscored girls by an average of 4.5%. “For example, 31 percent of boys know that Lincoln was president between 1860 and 1880, but only 18 percent of the girls do; 80 percent of the boys know that Germany and Japan were our principal enemies during World War II, but only 60 percent of the girls do" (p. 130). The researchers present no explanation for the gender gap, and they do not feel that its existence undermines the credibility of the test or allow it to color their interpretation of the results.

Subsequent NAEP assessments of US History in 1988, 1994, and 2001 introduced further modifications in the testing format but delivered similar results. The 1988 test, conforming to standard NAEP practices in other subjects, sampled 4^{th}-, 8^{th}-, and 12^{th}-grade students. The 1994 assessment added open-ended items, scored by rubric. Both the 1994 and 2001 assessments employed the same content framework. This framework identified targeted Proficient and Basic achievement levels. More importantly, the standard framework allowed for comparisons across the two sets of results. In both years, the researchers highlighted students' poor performance. The two lead findings from the 1994 assessment are that fewer than 20% of the sampled students in the targeted grades scored at the desired Proficient level and that more than half of the sampled 12^{th}-grade students failed to even reach the Basic level (Beatty, Reese, Persky, & Carr, 1996, p. x). The report on the 2001 assessment leads off with increases in scores at the 4^{th}- and 8^{th}-grade levels, but notes that fewer than 20% of students at any grade reached the Proficient level, and exactly the same percentage (57%) of 12^{th}-graders scored below the Basic level in both years (Lapp, Grigg, & Tay-Lim, 2002, p. x). In contrast to the 1917 and 1987 reportings of superior performance by boys, there was no statistically significant difference between the boys' vs. girls' performances on the 1996 and 2001 tests (p. 25). Further explorations of sub-group comparisons present odd contrasts: boys, but not girls, displayed statistically significant improvements between the 1996 and 2001 tests at the 8th-grade level (p. 24); 4th-graders in the Northeast and 8th-graders in the Southeast made statistically significant improvements, while at all other grade levels and in all other regions scores stayed the same or displayed non-significant changes. Again, the researchers make no effort to explain these differences or discuss their possible implications for the validity of the test or the interpretability of the results.

Over the past 100 years, historical content knowledge has been assessed using standardized tests employing multiple choice items and/or open-ended prompts graded with a rubric (Beatty, Reese, Persky, & Carr, 1996; Bell & McCollum, 1917; Kurfman, 1991; Lapp, Grigg, & Tay-Lim, 2002; Ravitch & Finn, 1987). The content assessed has been determined by an authority, often a committee of history educators (Bell & McCollum, 1917; Ravitch & Finn, 1987). The statistical analysis used on test results has been rudimentary, rarely extending beyond reporting means for groups and sub-groups. Examination of sub-group results often reveals puzzling differences in performance that seem to be unexplainable (Beatty, Reese, Persky, & Carr, 1996; Bell & McCollum, 1917; Lapp, Grigg, & Tay-Lim, 2002; Ravitch & Finn, 1987). These seemingly arbitrary sub-group performances suggest that the tests have limited validity or reliability (Nuthall & Alton-Lee, 1995; Wineburg, 2004). Interpretation of test results ignores these possibilities, however, and takes the scores as valid and reliable measures of students' historical content knowledge (Kurfman, 1991; Paxton, 2003; Wineburg, 2004).

High-Stakes Testing and History Education

Despite the limitations inherent in assessments of historical content knowledge (Horn, 2006), 23 states make decisions about school accreditation and student promotion based on end-of-year examinations in social studies. Because of the consequences attendant upon student performance on these tests, they are known as *high-stakes tests* (Kurfman, 1991). When in place as state-level education policy, these high-stakes tests serve to heighten the anxieties about curriculum, student and teacher expectations, and assessment (Grant, 2006).

Education policy researchers such as David Berliner criticize high-stakes testing, asserting that they trigger “a social science version of Heisenberg's Uncertainty Principle…. The more important that any quantitative social indicator becomes in social decision-making, the more likely it will be to distort and corrupt the social process it is intended to monitor” (Amrein & Berliner, 2002, para. 15). As states institute high-stakes tests, schools make every effort to secure higher scores by teaching to the test (Haney, 2000). For example, teachers may substitute an assessment’s *sampling* of curricular content for the actual curriculum framework itself (Kurfman, 1991). In many instances, school districts may go a step further and, as in West Virginia, use the actual assessment items as practice (Cannell, 2006). Such practices diminish, if not destroy, the validity of the test, and reverse the conventional educational paradigm of curriculum, instruction and assessment:

> Instead of evaluation serving and refining the other elements in curriculum development, it becomes the driving force, in effect the determiner of curriculum and instruction. In illogical circularity, the test provides both direction to the

> curriculum and the measure of success for curriculum and instruction (Kurfman, 1991, p. 317).

In a high-stakes testing environment, schools use assessments not to improve their instruction but to increase their statistics (Ibid.).

Despite these and other criticisms, the passage of the federal No Child Left Behind (NCLB) act in January, 2002 enshrined high-stakes testing in American K-12 education. According to the United States Department of Education (n.d.), this law requires all states

> to implement statewide accountability systems covering all public schools and students. These systems must be based on challenging State standards in reading and mathematics, annual testing for all students in grades 3-8, and annual statewide progress objectives ensuring that all groups of students reach proficiency within 12 years (para. 4).

The interpretation of the law is up to each state government. Each state composes its own standards, and therefore each state in essence writes its own test. In the process of approving standards, states have often chosen to underscore their emphasis on content knowledge and faith in testing, as when the Florida legislature mandated that "history shall be viewed as factual, not as constructed, shall be viewed as knowable, teachable, and testable" ” (K-20 Education Code, 1003.42.f).

NCLB specifies only reading and mathematics as required content areas for evaluation, but 23 states include social studies in their state-wide student assessments. Among these 23 states, some (e.g., Kentucky) apply the accountability only to the school and not to the student; student test scores are used to evaluate the school’s yearly progress toward meeting the state standards, but a poor performance does not impact the student. Ten states apply the accountability to the individual student as well, using test scores to make student promotion decisions or, as in the Commonwealth of Virginia, requiring a passing grade on a state history test to graduate (Grant & Horn, 2006). In addition to this graduation sanction, the Virginia Department of Education requires "student pass rates of 70 percent or above in all four content areas [English, Mathematics, Science, History]" for schools to maintain their accreditation (Virginia Department of Education, School accreditation ratings description, para 1).

The combination of state standards and high-stakes tests has altered history teachers' decisions about what content to teach, seeking to "make content choices that will most advantage their test-taking students" (Grant, 2006, p. 307). Less impact is observed on teacher's practices in instruction: teachers continue to use a wide variety of instructional strategies, such as debates and group work, to prepare students for the state-wide assessments. Little or no impact is observed on history teachers' classroom assessment practices, because "Most teachers already use test questions that mirror those on state history exams—multiple choice, short answer, and essays" (Grant, 2006, p. 308). As one Virginia teacher explained, "[when I write a test] I think multiple-choice—we just need to practice with that" (van Hover, 2006, p. 213).

An extensive literature has developed to analyze history standards (Nash, Crabtree, & Dunn, 1997; Symcox, 2002) and testing (Horn, 2006; Wineburg, 2004), and an emerging literature base is exploring state-level testing and its impact on classroom teaching and assessment (Friedman, 2006; Grant, 2003, 2006; Grant & Horn, 2006; Salinas, 2006; van Hover, 2006; van Hover & Heinecke, 2005; Yeager & van Hover, 2006). An as-yet unexplored area is a contextualized description of the student outcomes from the teaching taking place within the high-stakes, content-knowledge testing paradigm.

Technology in History Education

Developments in instructional technology have opened new opportunities for student assessment (Cantu & Warren, 2003). In fact, Virginia students take their end-of-year high-stakes assessments on custom-configured computers. However, history educators are far more interested in whether technology can be used to improve student outcomes (Martorella, 1997; Diem, 2000).

The discipline of history is notoriously slow to change (Milman & Heinecke, 2000, p. 538), possibly due to its focus on textual sources and products (Staley, 2003). History teacher-educators have not been eager adopters of technology (Martorella, 1997, p. 512; Dawson, Bull, & Swain, 2000, p. 594). For example, while the NCSS' College and University Faculty Association (CUFA) has established guidelines for integrating technology into pre-service teacher education (Mason et al., 2000), a survey of CUFA members reveals that "Regular use of technology is infrequent among most social studies faculty members" (Bolick, Berson, Coutts, & Heinecke, 2003, p. 304). Tellingly, the control groups for contemporary experiments and quasi-experiments in uses of technology in social studies still use "textbook and lecture" as the benchmark teaching practice to be measured against (Kingsley, 2005).

A body of research in *digital history* is emerging (Clarke & Lee, 2004; Friedman, 2006; Lee, 2002; Milman & Heinecke, 1999). Its advocates focus on historical thinking skills far more than historical content knowledge (Clark & Lee, 2004; Friedman, 2004, 2006; Hicks, Doolittle, & Lee, 2004; Swan, 2004). Some professional historians and educators, such as Ed Ayers, feel that “history may be better suited to digital technology than any other humanistic discipline” (1999, para. 4) and are already pioneering a new form of electronic scholarship. Repositories such as the Virginia Center for Digital History (VCDH) amass, organize, and digitize archival material. VCDH then makes this material available online for use by professional historians, history educators, and the general public. Thomas and Ayers have written historical accounts that “translate the fundamental components of professional scholarship—evidence, engagement with prior scholarship, and a scholarly argument—into forms that take advantage of the possibilities of electronic media” (Thomas & Ayers, 2003, para. 1). In these digital publications, historical arguments are linked to the digitized primary sources they rest upon, allowing the reader full access to the evidentiary record and not just the finished interpretation.

Digital history may have only a limited impact on K-12 history instruction. Research on the diffusion of innovations indicates that a key variable in determining the rate of adoption is the innovation’s compatibility with existing values (Rogers, 2003), and innovations in social studies education are no exception (Dawson, Bull, & Swain, 2000). The aspirations of those who would change history education through the innovative use of technology are frustrated by the fact that computers are used to support existing classroom practices, not to re-invent them, such as replacing paper-based media with online media (Hicks, Doolittle, & Lee, 2003). Technology can play a causal role in shifting the classroom paradigm to a more student-oriented, critical-thinking model, but to date such instances require exceptional teachers, exceptional levels of resources, or both (Lowther, Ross, & Morrison, 2003; Milman & Heinecke, 2000). Digital history will not become a widespread phenomenon, therefore, until either a paradigm shift takes place in American history classrooms or technology provides some improvement in students' historical thinking and/or understanding. Currently, however, the discussion of technology in history education has highlighted the potential of technology more it has provided actual implementation strategies or assessments of impacts on teachers and students (Clarke & Lee, 2004; Friedman & Hicks, 2006). Social studies education publications catalog the ever-expanding list of internet resources that teachers can draw upon but do not investigate the effectiveness of these resources or describe their best use (Whitworth & Berson, 2003).

Research on technology in social studies education has typically focused on teachers and/or teacher-educators (Bolick, Berson, Coutts, & Heinecke, 2003; DeWitt, 2004; Friedman, 2004, 2006; Hicks, Doolittle, & Lee, 2004; Keiper, Harwood, & Larson, 2000; Mason, 1998; Mason & Berson, 2000; Merryfield, 2000; Saye, 1998; Saye & Brush, 2006; Swan, 2004). One consistent finding from the research is that teachers' use of technology in the classroom is inhibited by many contextual factors: limits on instructional time, disincentives from the state testing regimen, lack of hardware and software, lack of technological knowledge, lack of technical support, lack of appropriate and/or credible information online, difficulty locating information, and concerns over plagiarism (Cantu & Warren, 2003; Friedman, 2004, 2006; Hanson & Carlson, 2005; Swan, 2004). As a result, many teachers have positive views of technology's potential impact in the classroom but do not use it themselves (Hicks, Doolittle, & Lee, 2004, p. 219).

The research base on student use of technology during social studies instruction is small and often focuses on non-academic constructs such as motivation (Heafner, 2004), comfort level (Crowe, 2004), or classroom dynamics (Milman & Heinecke, 2000; Waring, 2003). Educational technology commentators, however, have called for more research on students' academic outcomes from history instruction (Friedman & Hicks, 2006; Roblyer, 2005; Roblyer & Knezek, 2003). Three such studies – Kingsley (2005), Brush and Saye (2002), and Lee and Molebash (2004) – provide the strongest examples of experimental research on student outcomes from technology-mediated history instruction.

Kingsley (2005) assessed the impact of an online multimedia package on learning outcomes for seventh-grade students. This quasi-experimental study observed four teachers, each of whom administered the treatment to one class and the held another class in the control condition. The control treatment was traditional instruction: teacher lectures, reference to the textbook, and occasional use of worksheets focusing on key topics. The experimental treatment consisted of traditional instruction supplemented with the use of the online multimedia package, Ignite!Learning. This application is described as a "small-scale independent learning system" (p. 42), covering topics in early American history (1492-1877) through 15 sequential units. Each unit contains "songs, animation, short video clips, text, matching problems, stories, maps, illustrations, documents, timelines, and interactive games to teach students" (Ibid.). Students in the experimental group used this program individually in their school's computer lab approximately one day per week, working at their own pace. According to the researcher's observation, "In each fifty-minute class period where the Ignite! software was used, students were usually able to finish one full lesson [viewing all media in the lesson] and its accompanying Topic Review [six multiple-choice items]" (p. 57). The researcher assessed student outcomes, defined as “student knowledge of early American history” (p. 67), with a 50-question, multiple-choice instrument based on the students’ history curriculum. The researcher administered the instrument as a pretest in September and again as a posttest in March. Students in both the control and experimental conditions scored high on the pretest (respectively: mean = 33.6, SD = 5.3; mean = 30.95, SD = 6.12) and achieved limited gains on the posttest (mean = 36.66, SD = 5.58; mean = 37.04, SD = 5.51). In other words, over seven months, students in the control group increased their scores by an average of 6.1%, while students in the experimental group increased their scores by an average of 12.2%. While the pattern of scores raises questions about the nature of the instrument relative to students’ previous knowledge and subsequent instruction, the difference in score increases is statistically significant at the 0.01% level (pp. 74-75).

Kingsley’s research is a classic media-comparison study, exploring the possibilities provided by using computer-based media rather than the traditional paper-based media. The use of control groups allows for inference of causation, and the pattern of results–in which the experimental group starts below and ends above the control group–is highly interpretable (Shadish, Cook, & Campbell, 2002). According to Clark’s 1983 meta-analysis, “media do not influence learning under any condition” (p. 445). Contrary to Clark’s dictum, Kingsley convincingly demonstrates a media effect. However, the results are not practically significant: both the experimental and control groups moved from scoring a D (60-69%) to a C (70-79%) on an exam *after seven months of instruction*. True, the experimental group improved by twice as much as the control group, but this result must be considered a very heavily qualified success. One response to Clark asserts that while media alone do not influence learning, “media *and* *methods* [italics added] influence learning, and they frequently do it by influencing each other" (Kozma, 1994). Unfortunately, Kingsley's qualitative observation data does not provide sufficient description of how the teachers and students in the experimental group used the software to describe the teaching methods used. An unexplored possibility is whether more skillful scaffolding, intermixed with the media, would allow for greater improvements in students' historical content knowledge.

Over the course of nine years, Brush and Saye have conducted a series of uncontrolled experiments with a tool of their own devising called Decision Point! (Brush & Saye, 2001, 2002, 2005; Saye & Brush, 1999, 2002, 2004a, 2004b, 2006). The software supports problem-based learning, presenting high school students with a task and providing them with an "interactive hypermedia database" containing multimedia resources about a historical topic, specifically "the African-American Civil Rights Movement of the 1950s and 60s" (Brush & Saye, 2002, p. 3). Students are to "work collaboratively to develop a solution for the unit problem: What strategies should be pursued in 1968 [following the assassination of Dr. Martin Luther King, Jr.] to continue the struggle for a more just, equal society?" (p. 4). Students present their analysis through "a persuasive multimedia presentation that explains possible actions, evaluates the likely consequences of each alternative, and defends their solution as the best course of action" (Ibid.) and through an end-of-unit essay.

The focus of Brush and Saye's research has been the development and refinement of *hard* and *soft* scaffolds. Soft scaffolds are dynamic, on-the-fly actions by the teacher or a peer to assist a student's learning.

> For example, if students fail to discern differences in the messages of two civil rights figures, a teacher might help them think more deeply about the texts by

> asking questions such as: "What does Lewis mean when he says _ ? Why do you think he uses the word __? Do you find similar words in King's speech? Do you notice any difference in his tone and King's?" Once students discover that differences exist, the teacher might refer them to other documents that could help them understand the origins of those differences (Saye & Brush, 2002, p. 82).

Soft scaffolds in the Decision Point! project included teacher meetings with each cooperative group to answer questions and offer formative feedback and debriefing sessions following the presentations and before the final essay. In contrast to the dynamic nature of soft scaffolds, hard scaffolds are static, pre-planned interventions, such as hyperlinks between primary source documents and secondary descriptions of the events to which they are related, or a template for constructing the presentation. By integrating media and scaffolds, Brush and Saye capitalized on Kozma's observation that media *and* methods (i.e., the scaffolds) can influence learning.

In addition to civics education, Brush and Saye are interested in students' thinking skills, specifically students' ability to reason critically about ill-structured social problems. However, their evaluation of the success of students' work in the Decision Point! projects included an examination of the historical content knowledge displayed: in their presentations, "groups had to present accurate information with no major omissions or misstatements" (p. 88). As Brush and Saye increased the hard and soft scaffolding available to students during the task, student performance on this criterion increased. Since their research to date has been in a single-classroom, uncontrolled format, causation cannot be inferred.

Lee and Molebash (2004) conducted a small-scale experiment to explore competing media-and-scaffold strategies. They placed the participants, 30 graduate social studies education students, into one of three conditions: using an internet search engine

(Google), using a large (275 documents) online archive, and using a pre-selected subset of online documents. As students consulted their assigned media, they used a sourcing heuristic to analyze the documents. The researchers gave the students the prompt, "How was the Cuban Missile Crisis resolved?" A pretest in the form of a brief essay explaining students' previous knowledge about the resolution of the crisis indicated that only three students knew the answer as part of their prior knowledge. The participants were then given 45 minutes to improve their answers, using the media and scaffolds provided. Following this activity, the essay prompt was repeated. Nearly every participant in the search engine and pre-selected documents groups was able to answer the question on this second attempt; in the online archive group, only one participant provided the answer following the working session. Finally, Lee and Molebash repeated the prompt two months later. At this point, a difference emerged between the search engine group and the pre-selected documents group, with the latter recalling far more contextual detail regarding the resolution of the Cuban Missile Crisis.

Lee and Molebash's experiment helps establish the causal link that Brush and Saye's work (2002) has suggested: skillful choices of media and integration of scaffolds can improve students' historical content knowledge. Extrapolating from this line of research, the creation of a digital video editor built specifically for social studies (Ferster, Hammond, & Bull, 2006) provides an additional opportunity to explore the integration of media and scaffolds in a student-authorship environment. This integrated tool, PrimaryAccess, may provide further statistically and practically significant gains in students' historical content knowledge. Asking students to become creators, rather than consumers, of historical accounts may provide yet another avenue for them to acquire new knowledge (Greene, 1994, p. 166). An experiment or quasi-experiment conducted with students using an integrated technological tool will allow history educators and educational technologists to gain further insight on how teachers can apply these emerging technological capabilities to improve student outcomes on historical content knowledge.

Summary

History education is a contested field in American society, conflating political and academic motives. Setting aside the political motives, the academic motives of the history curriculum include building students' historical thinking skills and their historical understandings. To date, most of the serious research in history education has focused on historical thinking skills. Historical content knowledge, however, is the primary target of K-12 instruction and assessment. Technology holds the potential to improve the teaching and learning of history. To date, research on technology education has focused on teachers, not students, and has emphasized historical thinking skills, not content knowledge. However, the small body of research on student use of technology to acquire historical content knowledge suggests that a strategy incorporating a skillful combination of media and appropriate pedagogical methods may produce a statistically and practically significant improvement in students' historical content knowledge outcomes.

CHAPTER 3

METHODOLOGY

The questions addressed in a study should determine the selection of methods (Savenye & Robinson, 2004, p. 1047). This study attempted to answer the following questions:

1. Do differences exist in student outcomes, as measured by pre- and posttest scores on the standards-driven assessment, between students who use PrimaryAccess vs. those who generate a PowerPoint slideshow?
2. If differences exist, does the effect on student outcomes vary by achievement-level grouping?
3. What environmental factors – such as the curricular context, teacher behaviors, student behaviors, or technological infrastructure – appear to inhibit or promote this effect?

The first two questions implied a proposition, that students using two different technologies for learning would display different learning outcomes on the teacher-designed, standards-driven assessment. This proposition required validation or disconfirmation. Quantitative methods of analysis are best suited for validation (Krathwohl, 1998, p. 25), allowing the researcher to detect differences in outcomes within and between groups. The third question was descriptive, attempting to capture the thoughts and behaviors of the teacher and students while they engaged the subject matter with the technological tools in the context of the curriculum. Description is best conducted with qualitative methods (Ibid.), such as observation, interview, and document analysis. Because the research questions called for both quantitative and qualitative methods, and because intact student groups were placed in control and experimental conditions, this study is best termed a *mixed-method quasi-experimental design*.

Organizing Framework for Technology in History Education

As noted in Chapter Two, the literature on technology in history education has not focused on student learning outcomes. The limited body of research on history education, technology, and students' learning outcomes (Brush & Saye, 2002; Kingsley, 2005; Lee & Molebash, 2004) suggests three categories of technologies for history education: *media*, *scaffolds*, and *canvases*. These categories are not exhaustive, but they do reflect activities that history teachers engage in frequently.

- The *media* category includes both time-honored, physical media (jackdaw kits, textbooks, maps, films, and overhead transparencies) and the digital media (video, websites, electronic copies of images and documents) that have mushroomed in the past decade. In the past, social studies educators equated the term *technology* with *media* (Martorella, 1997, p. 511). Media comparison studies, seeking to discover whether instruction delivered using one medium delivers superior learning results to another, is the most prevalent type of research on media in education (Surry & Ensminger, 2001).

- *Scaffolds* are systems and strategies that provide a guiding structure to a learner as he or she attempts an activity (Collins, 2006, p. 55). Scaffolds provide a mechanism to capitalize upon Vygotsky's concept of the zone of proximal development (ZPD). The ZPD describes the levels of task difficulty between those tasks that the learner can execute un-aided and those tasks the learner can execute only with guidance and help (Hedegaard, 1990, p. 349). Scaffolds can provide the student with this support while he or she undertakes a relatively complex task. Examples of scaffolds in history instruction include sourcing heuristics, such as MAPER, APPARTS, SOAPS and SCIM-C, or fully-developed tutorials in historical inquiry and the construction of historical accounts, such as Historical Scene Investigation and The Mystery of Sam Smiley. They can be as simple as a handout that asks guiding questions or as complex as a computer application guided by some form of artificial intelligence.
- *Canvases*, also called authorware, are tools that allow students and teachers to create new media. Accordingly, these tools can be paper-based, as when students write an essay or create a collage, or they can be electronic, as when teachers or students compose a webpage, a PowerPoint slideshow or a digital movie. Products created by teachers commonly become media, shown to students as part of an instructional presentation. Products created by students, on the other hand, typically have a different purpose. Teachers assign student work on canvases not simply to produce educational media, but to allow students to learn by doing, to "engage students as active participants in their

> learning as authors who need to think critically about what they read, integrate information from difference sources with their own knowledge, and structure their work" (Greene, 1994, p. 166).

These capabilities can also exist in combination, such as an online archive of documents and audio (media) that provides a structured sourcing heuristic (scaffold), or a tool that integrates historically-relevant images (media) with a slideshow-creator or video editor (canvas). The combination can take place through a teacher using a single, integrated application or through a teaching strategy that integrates multiple tools.

Kingsley (2005) assessed the outcomes from students' use of a media package. Saye and Brush (2002) and Lee and Molebash (2004) explored strategies for integrating media and scaffolds. This study focused on student use of an integrated media-scaffold-canvas tool. The study employed a quasi-experimental design in which one cohort of students used the integrated media-scaffold-canvas tool (PrimaryAccess) while another used a media collection (a teacher-provided list of internet images) and a canvas (PowerPoint).

- PrimaryAccess is a web-based digital video editing environment designed specifically for history instruction (Ferster, Hammond, & Bull, 2006). Using PrimaryAccess, a teacher can pre-select a list of archival images, maps, and documents (i.e., media) that relate to the topic being studied. Each item comes with relevant contextual data and links to further information and can be viewed on a timeline (hard scaffolds). These documents are available to students after they log into the application. Students then compose a script, select images, associate the images with the script, and record a voice-over narration (canvas). The final products are available for playback over the internet. Throughout the process of constructing a PrimaryAccess movie, the teacher is able to view student work and leave notes for the student, thus providing formative feedback (soft scaffold).

- PowerPoint is a common if not ubiquitous tool for technology-mediated visual communication. Users can create a collection of slides, each containing text, images, or other media (canvas). Objects on a slide can link to other slides or to external data. Each slide, each object on a slide, and the transitions between slides can be modified in many ways, allowing the user to select colors, fonts, layouts, backgrounds, animations, and even sound effects. A common educational practice by technologically-savvy teachers is to require students to create a presentation illustrating their own content understandings, drawing upon relevant images from the internet (media).

Every technology has affordances and constraints that shape its use and utility (Norman, 1990). Affordances are potential actions or capabilities, and constraints are limitations or incapacities. A ballpoint pen, for example, affords grasping (due to its shape) and writing (assuming it has ink), but has the constraint of writing only on certain materials (i.e., paper and not metal) and indelibility (it cannot be erased). PrimaryAccess integrates three affordances: media, scaffolds, and canvases. PrimaryAccess has the constraint of being web-based, and is therefore vulnerable to broken links, disruptions in internet service, or mis-parsed data. PowerPoint is a canvas that can be easily integrated with media, and it, too, has constraints. First, it uses local data that is easily lost or destroyed. Second, PowerPoint presents the author with a host of distractors—colors, fonts, layouts, backgrounds, transitions, and sound effects—that, in the end, may impede rather than improve communication (Tufte, 2006).

For the purposes of acquiring content knowledge, the affordances of the relative canvases may be critical: PrimaryAccess encourages students to write iteratively and in full statements “as a means for helping them to explore their ideas and for acquiring new knowledge” (Greene, 1994, p. 166). PowerPoint, on the other hand, discourages iterative refinement of the message and encourages the use of bulleted lists for conveying information. Tufte argues that such lists are inferior analytic tools compared to “*sentences*, with subjects and predicates, nouns and verbs, which then combine sequentially to form *paragraphs*” (2006, p. 158). The proposition is that one of the two tools, in conjunction with a teacher’s instruction in the classroom environment, may allow students to achieve greater gains in historical content knowledge due to its affordances and constraints. Quasi-experimentation using mixed methods allows the researcher to infer causal validity and describe the contextual factors of these findings in the authentic environment of intact classrooms engaged in common curricular tasks.

Methodological Approaches and Assumptions

A fundamental activity in education, following teaching and learning, is evaluation. This study was itself a form of evaluation, determining the relative merits of two competing technological tools and their accompanying instructional strategies. Ideally, evaluation should take place in the natural setting of the object under study. The common setting for American K-12 education is the classroom, particularly the public school classroom (Gutmann, 1987). Research on students' use of technology to study the K-12 curriculum, therefore, is best conducted in a school setting. Laboratory-based research lacks ecological validity as it omits the contextual details that make the school classroom a demanding environment: interruptions from visitors or announcements, limited technical support, administrative overhead, variable student behavior, teacher burnout, etc. (Brouwers & Tomic, 2000; Hastings & Bham, 2003). Some public school classrooms, as compared to private school settings, are particularly challenging: student motivation and attendance can vary widely (Coleman, Hoffer, & Kilgore, 1982), and teachers and students face greater pressure from universal high-stakes testing. Classroom-based research, therefore, provided the best opportunity to generate findings that were an authentic evaluation of what these tools (PrimaryAccess and PowerPoint) and teaching practices (end-of-unit, student-generated products) could do within the demanding context of the K-12 history curriculum. Findings from classroom-based research are more valid and generalizable to other instances of K-12 history education, especially when accompanied with enough descriptive detail to inform the application of these findings to other settings.

Traditionally, evaluation studies have a goal or objective in mind (Krathwohl, 1998, p. 594). In this study, this objective was defined by the context of classroom instruction: the goal for students in both conditions (media-scaffold-canvas and media-canvas) was to achieve gains on an end-of-unit, teacher-designed, standards-based posttest and maintain these understandings over time, in preparation for the end-of-year, high-stakes summative examination. Roblyer and Knezek (2003) point out the need for educational technology research to use "established measures of education quality" (p.

65) to measure outcomes, such as this end-of-unit assessment. To provide this basis for relative judgment, students took multiple pretests and posttests.

As previously noted, standardized assessments of historical content knowledge are flawed instruments. The students' responses do not provide a true image of their comprehension, the results are unstable, and yet the measurement is taken as final (Amrein & Berliner, 2002; Nuthall & Alton-Lee, 1995; Wineburg, 2004). In contrast, deeper understandings of students' historical content knowledge emerge through interviews, often combined with a performance task (Brophy & Alleman, 2000; Wineburg, 2001, Ch. 5). This study therefore supplemented and challenged the test-based outcome measures through document analysis of students' products (movies and presentations). The final decision concerning students' historical content knowledge was based on all sources of data, and not test scores alone.

The mixed-method approach allowed for a complementary relationship between the strengths and weaknesses of the quantitative and qualitative techniques utilized. The quantitative component allowed for verification or disconfirmation of statistical and practical differences in student outcomes. Furthermore, by using the teacher-assigned, standards-driven assessment as the instrument, the quantitative aspect was representative of the interests and information needs of the stakeholders in this environment: the students, teacher, administrators, and parents. The qualitative component strengthened the explanatory power of the quantitative results through observation and interviews. Because observation may alter participant behavior (Savenye & Robinson, 2004, p. 1055) and interviewer characteristics may alter responses (Krathwohl, 1998, p. 296), document analysis and performance on the pretests and posttests provided more naturalistic data on student performance.

The intended end product was a thorough evaluation of the students' relative learning gains, as measured by the end-of-unit test scores and verified by document analysis. Extensive classroom observation and teacher interviews provided a detailed description of the instructional methods used, thus satisfying Roblyer's "comprehensive reporting criterion" (2005, p. 196) for high-quality educational technology research.

Participants, Site, and Curricular Context

This study focused on one teacher and his students at a middle school in a mid-size city of the eastern United States. At the time of the study, per-pupil spending in the district was $12,900. The student body at the middle school was comprised of more than 600 students in grades 7 and 8. The ethnic composition was very similar to that of the district as a whole: approximately half of the students were African-American, and the next-largest ethnic group was white. A very small percentage of the student body was identified as Hispanic, Asian, or Other. More than half of the student body received free or reduced lunch. Five percent of the students were English as a Second Language learners. The school administration identified students' reading achievement as a major key to success on standardized tests, noting at a 2003 board meeting that "approximately 50% of…students do not read at a level allowing them to consistently pass standardized tests." The school also struggled with disciplinary challenges in the student body. Over the two years preceding the study, incidents ranged from student assaults on one another (including a stabbing and a shooting), attacks on teachers, and even the theft of a teacher’s car by a student. In response, the school added a dean of students and instituted restrictions on gang-affiliated clothing (Barry, January, 2006; March 2006).

Participants

The participating teacher in this study, Mr. Smith, was a young, white male beginning his second year as a classroom teacher. His teaching placement was in seventh-grade history, covering American history from 1877 to the present day. Prior to becoming an educator, he worked in the financial industry for three years. He completed an undergraduate degree in Economics from a flagship public university in a neighboring state, and he completed a Master’s of Teaching at a nearby liberal-arts college while completing his first year of teaching. He was certified to teach history and language arts for grades 6 through 12. He coached boys’ basketball, and sports images and clippings decorated his classroom walls.

Mr. Smith was a fluent user of technology; when a laptop battery died in the middle of a class presentation, he was able to trade laptops, move the data, and re-start the presentation without halting the class discussion. He was capable of teaching himself new concepts in educational technology, such as a slow-reveal process for images integrated into a PowerPoint presentation. He had a data projector in his room available for use during almost every class period. He routinely created PowerPoint presentations to support his classroom instruction, and one of his end-of-unit projects required students to draw upon designated internet image collections to create their own presentations. He used PrimaryAccess once prior to participating in the study, but his students did not produce full products.

The two other members of the seventh-grade history team had fifteen or more years' experience, making Mr. Smith the junior member of the group. The team shared a common planning period at the end of the day, allowing them to work closely together to produce instructional materials, plan lessons, and design assessments. All three teachers used the same materials and assessments; a PowerPoint presentation or quiz created by one was used by all. The team frequently engaged in collaborative teaching. In some instances the students moved from room to room during the period, engaging in different instructional activities with each of the teachers on the team; on other occasions the classes combined in the auditorium to watch a movie or engage in a poster-based exercise. On occasion, one teacher would be called away on administrative duties and the other two would cover her class either by absorbing her students or surrendering their planning period. All three teachers moved through the curriculum at the same pace, differing by no more than one or two days of instruction. Because their classrooms were contiguous, the teachers interacted frequently during the day, stepping into one another's rooms to with questions or reminders, or simply to socialize.

The students participating in the study were divided into six sections. The sections were tracked according to end-of-year achievement test results and recommendations by teachers. The six sections were at three track levels, with two sections at each of the three levels: high-achieving (Honors), middle (Talent Development), and lower-achieving (General). The class sizes ranged from 10 to 25 students. (See Table 3.) Several students

Table 3

Class Periods, Times, Tracks, and Sizes

Period	Time of day	Track (coding)	Num students (on average)	Num ESL students
1	7:55 - 8:40	Talent Development (TD1)	18	
2	8:45 - 9:30	Talent Development (TD2)	14	
3	9:35 - 10:20	General (Gen1)	10	1
5	11:15 - 12:00	Honors (Hon1)	13	1
6	12:05 - 12:50	General (Gen2)	11	
7	12:55 - 1:40	Honors (Hon2)	24	2

Table 3: Student groupings for the participating teacher. During the study, the exact number of students in each class varied slightly as students entered or left the teacher's roster or were switched between one section and another.

were classified as English as a Second Language (ESL) learners. Most of the students were familiar with PowerPoint, and some had previously used PrimaryAccess.

The school day was divided into eight periods. Each period was 45 minutes long, and interruptions—from announcements, late-arriving students, questions or requests for students over the intercom, or other teachers dropping in—occurred every day, and almost every period. Mr. Smith began most class sessions with a brief Do Now prompt written on the board. These Do Now prompts immediately engaged students in a content-related task, such as "Read p. 13 in the folder and answer questions." Homework assignments for the week were written in advance on a board at the back of the room. Loose leaf paper and pencils were available for students who needed them. Study materials for each unit of study were provided to students in a folder, stored in the classroom. A classroom set of textbooks was available at all times. High-track and middle-track students were provided with a textbook to keep at home; lower-track students used only the classroom set.

The Curriculum and Its Assessment

The curricular context for instruction and assessment was the Virginia Standards of Learning (SOLs) and the end-of-year, high-stakes test (SOL tests). The Virginia Department of Education first published these standards in 1995 and has revised and expanded them over the ensuing decade. The current version is the *History and Social Science Standards of Learning for Virginia Public Schools* (2001b). The standards and their attendant assessments "form the core of the Virginia Board of Education's efforts to strengthen public education across the Commonwealth and to raise the level of academic achievement of all Virginia students" (preface). The 50-page *Standards* is supplemented by a 580-page *Curriculum framework* (2001a) that "defines the content knowledge, skills, and understandings" (preface) that students are to acquire throughout their history education in Virginia public schools, from Kindergarten until graduation from high school. Despite the fact that the standards "do not prescribe the grade level at which the standards must be taught or a scope and sequence within a grade level" (2001b, preface), the Virginia Department of Education presents enhanced scope-and-sequence guidelines for each subject. The scope-and-sequence guideline for the course in this study, United States history: 1877 to present (USII) (2004), contains resource lists, instructional materials and activities, and even sample assessment items.

Since 1998, the Commonwealth of Virginia has assessed every student in some form of History/Social Science at grades 3, 5, 8, and in high school. Like many other standardized history assessments, the tests items are multiple-choice and emphasize factual recall (van Hover, 2006). The content of the test adheres to a blueprint that specifies the standards to be tested and the number of items that address each standard. According to the blueprint (Virginia Department of Education, 2002), the SOL test for USII consists of 50 multiple-choice items. The blueprint flags one standard that is excluded from the assessment, USII.1e: "The student will demonstrate skills for historical and geographical analysis, including the ability to evaluate and debate issues orally and in writing" (p. 9).

No close scrutiny of these SOL tests can be made. While the Virginia Department of Education has released test items from other subjects, it has not released items from the test for USII. The publicly-available data is limited to pass rates, which have demonstrated near-constant improvement in scores. (See Table 4.)

Table 4

Pass Rates on Virginia Standards of Learning Tests of History/Social Science, 1998-2005

Grade level / subject	1998	1999	2000	2001	2002	2003	2004	2005
3rd grade History / Social Science	49	62	65	72	76	82	87	89
5th grade History / Social Science	33	46	51	63	72	79	87	85
8th grade History / Social Science	35	40	50	56	78	80	83	82
High school History / Social Science	N/A	N/A	N/A	N/A	N/A	81	83	86
High school US History	30	32	39	47	72	75	N/A	N/A
High school World History I	62	68	75	83	86	86	N/A	N/A
High school World History II	41	47	60	65	79	82	N/A	N/A

Table 4: Annual pass rates on Virginia History/Social Science assessments. Until 2003, high school students were assessed in subject-specific tests only.
Sources: Virginia Department of Education *1998-2003 statewide standards of learning spring assessment results* (n.d.) and *State report card* (2006).

Classroom Instruction

This study spanned four units of instruction conducted during the fall of 2006. The topics for each unit were drawn from the *History and Social Science Standards of Learning curriculum framework* (Virginia Department of Education, 2001a). The topics and relevant SOLs for the four units of instruction are listed in Appendix A and are summarized in Table 5.

Table 5

Summary of Topics and Standards of Learning Across Units of Instruction

Unit A	Rise of Jim Crow laws and African-American responses to discrimination	USII.3c
	Migration to the Great Plains	USII.2a, 3a
	Cultural conflicts on the Great Plains	USII.3b (selected sections)
Unit B	Immigration	USII.3b
	Industrialization	USII.3d
Unit C	Spanish-American War	USII.4a
	Life in early 20th century America	USII.5a
	Great Migration	USII.5b (selected sections)
Unit D	World War I	USII.4b
	Progressivism	USII.5b
	Harlem Renaissance	USII.5c

Each unit consisted of approximately three weeks of instruction, followed by project work (for those units that included projects, namely Units A and C) and the end-of-unit assessment. During the instructional portion of each unit, the teacher followed a consistent pattern. At the commencement of the unit, students received two packets of teacher-created materials:

- A study guide, containing an overview of the content for the unit followed by blended textual passages, maps, lists of terms, and questions and activities for each topic within the unit. This study guide presented the same information as the curriculum guide in a student-comprehensible format, with some teacher-selected elaboration.
- A classwork packet, containing a series of readings and worksheets to be completed during the course of the unit.

As the unit progressed and new topics were introduced, students also received weekly homework packets that provided questions to be answered and activities to be completed (e.g., word searches). The classwork packet was placed in teacher-provided folders, color-coded by class period. These folders remained in the classroom and were not taken home by the students. The study guide and homework packets were hole-punched and placed in students' binders.

During the unit, the content instruction progressed sequentially through each of the designated sub-topics. (See Table 6.) Unit A, for example, covered Jim Crow laws, then migration to the Great Plains, and finally cultural conflicts on the Great Plains. Later sub-topics built upon earlier sub-topics as appropriate: for example, the discussion of migration to the Great Plains included mention of the Exodusters, African-Americans who moved west to escape discrimination and Jim Crows described earlier in the unit.

The pattern for a typical class session was for students to complete a Do Now prompt (e.g., "On a piece of paper, tell me what has been your favorite topic in this class

Table 6

Pattern of Instruction Across Observed Units

Unit A *(Four weeks: Jim Crow, Great Plains, cultural conflicts)*				Unit B *(Three weeks: Immigration, industrialization)*		Unit C *(Three weeks: Spanish-American War, early 20th century American life, Great Migration)*			
Topic A1	Topic A2	Topic A3		Topic B1	Topic B2	Topic C1	Topic C2	Topic C3	
Complete packet section about Jim Crow laws	Complete packet section about Great Plains	Complete packet section about cultural conflicts on the Great Plains	Complete project (create either a movie or a presentation about self-selected topic)	Complete packet section about immigration	Complete packet section about industrialization	Complete packet section about Spanish-American War	Complete packet section about Great Migration	Complete packet section about early 20th-century life	Complete project (create either a movie or a presentation about self-selected topic)
Multimedia presentation about Jim Crow era	Slideshow about settlement of the Great Plains	Video excerpts about Chief Joseph and Crazy Horse			Photograph analysis activity about child labor	Atlas activity about locations and events of Spanish-American War	Slideshow about Jacob Lawrence		
Internet archive activity about segregation in local community history	Video about building transcontinental railroad								
	Enactive lesson about Great Plains inventions								

Table 6: The researcher observed three units of study. Each unit contained two or three topics; the first and third units included project work. The staple instructional activity was completing a teacher-designed packet of classwork materials. Work on this packet was supplemented by the indicated additional instructional activities.

and why"), followed by an instructional activity. The staple activity for most of the observed instructional days was a teacher-led discussion of the material in the study guide. Students were called upon to read passages from the study guide or textbook, followed by the teacher either summarizing the passage or eliciting students' summaries. The students then answered the study questions in the packet, and the teacher verified that they had the correct answer written down. Each class session covered up to four pages in the classwork packet.

In addition, as indicated by Table 6, the instruction for most topics included activities beyond working on the study guide. For example, during Unit A, the instructional activities displayed considerable variety, as students

- Watched a teacher-created multimedia presentation;
- Browsed material in an internet archive;
- Discussed a teacher-created slideshow;
- Participated in outdoor enactive lessons, including building a sod house and engaging in train-vs.-wagon races; and
- Viewed a custom-created movie consisting of video clips excerpted from films and documentaries.

The instructional activities employed in Units B and C displayed less variation; the teacher adhered more closely to completing the study packet and included fewer supplemental activities. In contrast to the variety of activities used in Unit A, the supplemental activities in Units B and C were paper-based (i.e., examining photocopied images of child labor and looking up locations in atlases) or did not involve student use of technology (i.e., students viewed a teacher-created slideshow of paintings). As might be expected, Unit A received more instructional time (four weeks) compared to Units B and C (three weeks each).

Each of the three track levels received almost identical instruction. From one period to the next, the same study guide was used, the same presentations and videos were shown, and the same questions were posed. The only clear, consistent differentiation among tracks took place in the homework packet: the packets given to the General students contained different worksheets than those given to the Talent Development and Honors students. Typically, the readings in the General homework packet were shorter (e.g., rather than read a four-page article about Booker T. Washington and W.E.B. DuBois, students would read a one-page article) and asked more concrete questions. The questions asked in the homework packet for the Talent Development and Honors students were typically more open-ended: “Do you agree more with Washington or DuBois’ approach to gaining equality for African-Americans? Give reasons for your answer.” The corresponding question in the General students’ homework packet was a multiple-choice question asking, “Who do you think is right?” (Washington, DuBois, or neither), followed by “Why did you decide that way?”

During the teacher interview, Mr. Smith reported two influences on his choices of classroom instructional activities: the state-mandated curriculum and the needs of his students. The described pattern of instructional activities (i.e., using study guides and classwork and homework packets to frame the content and integrating supplemental activities to provide variety) emerges from the teacher’s synthesis of these influences.

According to Mr. Smith, “The SOLs tell me what I need to teach.”

> This [the curriculum guide] is the basics of what they [the students] need to know, so [I go] through, making sure, "OK, I need to cover this, do these kids know this,

> this, this." Before I move on, I need to make sure there's essential questions they need to be able to answer, essential knowledge...those basic things. If they know more, that's great, and a lot of them do know more, but for some of the students that haven't passed SOLs before, that's the focus. You need to know this, and let's get that down before we move on (Mr. Smith, interview, January 25, 2007).

Mr. Smith felt that the current administrative climate in the middle school reinforced strict adherence to the curriculum guide: “We are under the microscope right now, as far as having the state come in and take over. Yeah, that's pushed on us every day” (Mr. Smith, interview, January 25, 2007). The seventh-grade social studies team had a standing Friday-afternoon meeting with the building administration, sometimes attended by district or state personnel to review instruction and ensure that the team was maintaining the necessary pace to cover the entire curriculum by the end-of-year state assessment. “Right now we're about a week off. And they're letting us know” (Mr. Smith, interview, February 5, 2007).

For his students to succeed in mastering the content required by the curriculum, Mr. Smith felt that his instruction should include “repetition after repetition after repetition.” The study guides and homework packets provided the structure for this repetition, and they reinforced the need to progress from one curricular topic to the next. However, this pattern ran the risk of alienating the students. “I've seen some people come up here and just talk at these kids. They're not going to get it. You need to be…entertaining for them” (Mr. Smith, interview, February 5, 2007).

Mr. Smith looked to the supplemental activities identified in Table 6 to provide variety, build interest, and provide strong visual images to associate with the concepts specified in the curriculum. By integrating supplemental activities, Mr. Smith was able to move among different instructional strategies: “We move from discussion to lecture to discussion to multimedia types of things” (Mr. Smith, interview, January 25, 2007). The technology-based supplemental activities, in particular, supported the integration of visuals.

> This year we're trying to do a lot of visual stuff, we have a lot of visual learners. [We show] a short PowerPoint where they take notes off of it. We take all the history classes into the auditorium. We're talking about the event…, taking quick notes on it, then they're seeing a clip, from old newsreels or Hollywood movies. So this PowerPoint incorporates still images but also video clips, audio clips.... [It's] a media-rich PowerPoint (Mr. Smith, interview, January 25, 2007).

This use of technology was personally satisfying to Mr. Smith ("I enjoy the use of technology"), and he felt it was a powerful pedagogical tool: "I think the students like seeing pictures when talking about something. They try to tie in [concepts]" (Mr. Smith, interview, January 25, 2007). Through the use of tools such as classroom buzzers and a PowerPoint-based Jeopardy template, Mr. Smith incorporated compelling games into his instruction. In the spring semester, he taught himself a technique in PowerPoint to animate squares over an image and thereby create a slow-reveal process to heighten suspense. Allowing the students to use technology, according to Mr. Smith, elicited greater effort on their part.

> They worked harder on that [making a presentation or a movie] than they normally do on most projects. Putting a computer in front of them is extremely helpful. They enjoy working on the computers. They liked having something that they create (Mr. Smith, interview, February 5, 2007).

Project Work

The experimentation took place in the variation of student end-of-unit project work. Students completed two projects: one movie and one presentation. On Unit A, students in three sections, one of each track level, created movies while students in the remaining three sections created presentations; on Unit C these conditions were reversed. On both Units A and C, the projects took one week (five instructional days) to complete. Both the movie-making and presentation-making projects were completed simultaneously, and no extra instructional time was provided for either group.

Most students worked in pairs, but in most classes at least one student worked alone. The students were already familiar with PowerPoint, having used it to complete projects for other teachers. (During Unit A, a student seeing PrimaryAcess for the first time assumed it was PowerPoint and asked if she could "push the button.") Many students had heard of PrimaryAccess before, and a few Honors students had used it in the previous year as a part of a Gifted and Talented class.

For each unit, the projects addressed one of three sub-topics contained within the unit (see Table 7). Mr. Smith allowed students to select their topics; therefore, the

Table 7

Distribution of Students' Project Topics Within Condition

		Cohort 1 (movies)	Cohort 2 (presentations)
Unit A	Rise of Jim Crow laws	13 (7)	19 (11)
	Migration to the Great Plains	17 (10)	13 (7)
	Cultural conflicts on the Great Plains	10 (5)	14 (9)
	Total	40 (22)	46 (27)
		Cohort 1 (presentations)	Cohort 2 (movies)
Unit C	Spanish-American War	13 (7)	17 (9)
	Life in early 20th century America	13 (8)	14 (7)
	Great Migration	12 (6)	16 (8)
	Total	38 (21)	47 (24)

Table 7: Distribution of students' project topics within condition. The differences in *n* between are due to students entering/leaving the teacher's class roster.

distribution was not equal across topics. During Unit A, for example, the most frequent topic for students making movies was migration to the Great Plains, but the most frequent topic for students making presentations was the rise of Jim Crow laws. During Unit C, the distribution across topics was roughly equal.

In both conditions (movie-making and presentation-making), the task given to the students was identical: for the assigned topic, students were to create a product (a movie or a presentation) that

1. Reviewed the "basic information" (i.e., the information specified in the curriculum guide and therefore highlighted in the study guide),
2. Included "any extra information" that the students felt was interesting or important, and
3. Explained "why is it [the topic] important to us today."

This task (basic information, extra information, and significance) reflected the teacher's decisions regarding instruction. The attention to the information specified in the curriculum guide demonstrated his fidelity to the SOLs ("The SOLs tell me what I need to teach"). When students included this information in their projects, their action provided yet another cycle in the "repetition after repetition after repetition" that Mr. Smith felt that his students needed to internalize concepts. The additional information–gleaned from the study guide, the textbook, or supplemental sources–provided necessary elaboration to help cement the more basic information. For example, the Homestead Act of 1862 was not specified in the USII curriculum guide's listing of reasons for westward expansion. However, when discussing the motive of "Opportunities for land ownership," Mr. Smith identified the Homestead Act and included many details, such as the size of the plots available (160 acres), the required filing fee for a claim ($18), and the requirements to receive ownership (occupation and use for 5 years). This additional information, in Mr. Smith's opinion, reinforced the concept specified by the curriculum guide, namely, that "Opportunities for land ownership" was a reason for westward expansion. The details about the Homestead Act helped students learn and retain the information: "If there's some kind of tie to a basic piece of knowledge, then I try to hit that, so they get that.… If for them to understand it better, they need to know this other piece of information—give it to them. That's what we do as teachers" (Mr. Smith, interview, January 25, 2007). The third requirement, that students explain why a topic was significant and important to the present day, provided an additional opportunity for elaboration and encoding that would assist in students' retention: "They need to know those basic facts that the SOL wants them to know, and when they make those connections [between past and present], they'll remember the facts.… When they do that [make connections], they get those basics. They understand the basic content knowledge" (Mr. Smith, interview, January 25, 2007). The assigned projects, therefore, required students to explicitly meet *and* exceed the content knowledge expectations of the curriculum guide, with the intended outcome being greater understanding and retention of the content for the end-of-year state assessment.

While the statement of the task given to the students was identical for both conditions, the production process that emerged within each condition was very different. Mr. Smith introduced the students to the movie-making tool (PrimaryAccess) and then directed them through an iterative process of writing and selecting images.

> I set goals every day for what they [the students] had to accomplish: "Have your introduction done by the first day, 2-3 paragraphs done the second day," and so

> on. [It was] just a kind of work-in-progress there (Mr. Smith, interview, February 5, 2007).

Students saved their work in different versions as the project progressed: at the end of the first day, the existing script and images were saved as Version 1, at the end of the second day the revised script and images were saved as Version 2, and so forth. (See Figure 2.)

Figure 2: Screenshot of Student Work in PrimaryAccess During Unit A. This image shows students' second draft of three written while creating their script and selecting images. Note the presence of embedded text notes left by the teacher outside of class time. One of these notes has been opened to reveal the teacher's comment; clicking the "x" closes the note and collapses it into an "N" appearing on the vertical bar. On the left are images selected and sequenced to the script. On the right is a viewer for adding motion and titles. Below the script and viewer panels is a recorder for capturing the student narration.

In contrast, the presentation-making groups worked in a linear fashion, progressing from an outline written on paper to making slides in PowerPoint.

> They had to write it out on paper before they got the computer. *These* bullets, *these* slides are going to have *this*. So it was kind of quick. So mostly it was kind

> of putting it into the computer. Totally different processes between the two [PrimaryAccess and PowerPoint] (Mr. Smith, interview, January 25, 2007).

The presentation-making process was further structured by PowerPoint's slide metaphor. (See Figure 3.) As students created their presentations, the first slide was reserved for a title, the following slide addressed basic facts, followed by slides of additional

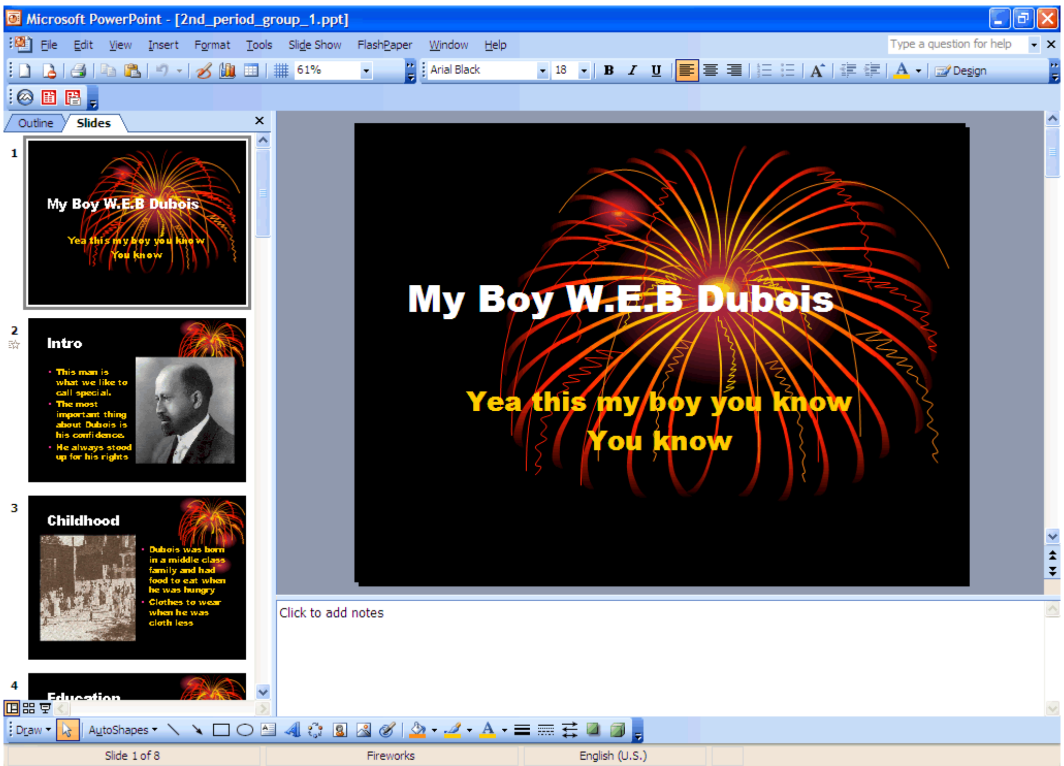


Figure 3: Screenshot of Student Work in PowerPoint During Unit A. This image shows the title slide of the students' final draft. The slideshow consists of eight total slides; each contains text and images discussing the topic.

information, and then a slide explaining the significance of the topic. The final slide listed resources consulted, such as the textbook or websites.

During the project work, students in both conditions (movie-making and presentation-making) received feedback from the teacher. Students using PrimaryAccess received feedback in two forms: verbal feedback and embedded notes left by the teacher. (See Figure 2.) These notes were added outside of class time; each evening, the teacher logged into movie-making students' accounts and reviewed their work.

> Because they were in groups and it didn't seem like it was too much, I actually went in that night and put in notes for all the groups. It normally took about an hour to do that. For three classes, for all three classes together, the total would be about an hour. A quick note, sometimes just to give them a little direction if they were lost or try to clean it up a little bit, give them thoughts if they were stuck. It was a nightly thing (Mr. Smith, interview, February 5, 2007).

These notes were then supplemented by verbal feedback given during class. In contrast, PowerPoint has no equivalent feedback feature. The "Notes" section of the PowerPoint slide (see Figure 3) can be used for this purpose, but it would require the teacher to centralize and transport the files off the laptops to provide this feedback outside of class time. Because there was no convenient way to provide asynchronous feedback in PowerPoint, the presentation-making groups received only verbal feedback from the teacher.

Classroom Assessment

At the end of each unit, student learning was assessed using a teacher-designed test. These tests consisted of 20-40 multiple choice items; the tests for Units A and C also included 3-5 open-ended prompts. (Open-ended prompts had been planned for Units B and D, but the teacher chose not to implement them.) In constructing these tests, the teacher drew upon the relevant sections of the curriculum guide (Virginia Department of Education, 2001a). The multiple choice items were based upon the Essential Knowledge items specified in each unit (e.g., "'Jim Crow' laws were passed to discriminate against African Americans," [p. 162]), and the open-ended items were selected from the Essential Questions provided for each unit (e.g., "How did African Americans respond to discrimination and 'Jim Crow'?" [Ibid.]).

Before the end-of-unit posttests, the teacher prepared for the test by leading the class in a review session. Review sessions lasted one day and were typically done in a game format. During the Unit A review session, the students were placed in small groups and played a team-format Jeopardy game created by the teacher. The teacher had downloaded a PowerPoint template from the internet and stocked it with questions from the unit content. Teams were given signaling buzzers. During the game, an assigned student kept score on the board. The students were highly motivated to ring in and answer questions and beat the other teams. During the review for the Unit C test, the teacher similarly placed students in groups and played a relay game called “Four Minute Fury,” in which students passed a sheet of paper around the group and answered questions in sequence. The group able to answer the most questions won the game. Following Four Minute Fury, the same groups were used to play a Jeopardy game. According to Mr. Smith, these review games are useful because “They're answering, they're reviewing questions, but they have fun with it. It's hitting all those things that I want them to know, but it's also linking directly to what I need them to know on the test” (Mr. Smith, interview, February 5, 2007).

On the day of the end-of-unit tests, the teacher first led the class through a final review. On the Unit B test, for example, the teacher began the class by projecting a copy of the unit study guide with all of the answers filled in. Students were invited to check their work against the completed study guide while the teacher visited students individually to record their homework grades. Then, after completing the homework check, the teacher switched off the overhead projector and called upon students:

> Who are the captains of industry?
> What region was known for textiles?

> Who founded Hull House? … What was Hull House?
> What was the AFC? … What's that stand for? … What's a union?
> Name three big businesses. (Classroom observation, October 17, 2006).

The teacher commented that this practice of reviewing test content immediately before distributing the test was an important technique to allow the lower-achieving students to do well: "They normally don't study the night before. They don't understand how to study for a test. [Before a test,] They'll have their study guides out, and I'll ask questions out loud" (Mr. Smith, interview, February 5, 2007). However, he did not engage in the same review for all groups; the review for the General students was extensive, lasting 10-15 minutes. The Talent Development groups received a shorter review (approximately 5 minutes), and the Honors classes in some instances received no review at all. "My high-level classes I wouldn't [review immediately before distributing the test]: 'You studied last night; OK, here's your test.'"

Students were given 25-35 minutes to complete the multiple choice items and the open-ended prompts. Some students finished the multiple choice section of the test in as little as five minutes, while others took all of the allowed time. After turning in the multiple choice answers, students responded to the open-ended prompts, which were either written on the board or projected on an overhead transparency. Students who finished both the multiple choice and the open-ended responses early were given a word search or a worksheet to complete.

At the end of the semester, students took a 50-item multiple-choice examination. The semester test addressed the same content as the end-of-unit tests, but did not repeat items from these tests. The semester posttest was administered on the first day back from winter break. Prior to the two-week recess, the teacher conducted a review of the material from the semester, again using a custom-created packet of materials. He spent three days reviewing material from Units A through D. No emphasis or privilege was given to any one section of content. On the day of the semester posttest, the teacher did not conduct the same question-and-answer review that he had used for the end-of-unit tests. “I did not do it [review immediately before distributing the test] for the midterm because I wanted to keep that as close to the SOLs as possible, you know, without doing any review other than three days...or two days before” (Mr. Smith, interview, February 5, 2007). Again, the end-of-semester test items were constructed from the curriculum guide and closely modeled the SOL tests.

In addition to these assessments of end-of-unit and end-of-semester knowledge, students took several assessments of previous knowledge. At the beginning of the semester, Mr. Smith administered the 50-item semester exam as a pretest. This semester pretest was administered on the second day of classes. Mr. Smith explained to the students that “I want to see what you know, and I want to focus on what you don’t know. …I don’t expect you to know this” (Classroom observation, August 22, 2006). Classes took between 20 and 30 minutes to complete the 50-item test. During the administration of the test, students appeared to be focused: the room was quiet, the students were looking at their papers, and once the first test was turned in, there was no rush to turn the papers in.

During Units A and C, students’ previous knowledge was also assessed using open-ended prompts. These pretests were administered throughout the unit of instruction; as new topics were introduced, students were pretested. The prompts were selected from the Essential Questions identified for the topic in the curriculum framework. (See Appendix A for examples.) The pretests were given as Do Now activities immediately prior to discussing the topic in question. For example, the Do Now before instruction about migration to the Great Plains was, “How did people’s perceptions and use of the Great Plains change after the Civil War?” However, due to logistical problems during the Spanish-American War segment of Unit C, the pretest for that topic was administered *after* the instruction had taken place. The answers given by the students reflected this fact, as they were far more complete and accurate than any other pretest responses.

Technological Infrastructure

The middle school provided Mr. Smith with the technology necessary for his teaching style and the end-of-unit projects described. The school provided a desktop Macintosh computer for inputting grades, preparing classroom materials, and accessing the district email server. He shared a data projector with the other two teachers on the seventh-grade history team. The school owned two mobile laptop carts, each stocked with 12 Dell Inspiron notebooks. These carts were available for checkout to the teachers, and Mr. Smith used one cart for each implementation of the movie-making and presentation-making projects. The laptops had wireless network adapters, and the laptop carts each contained a wireless hub that plugged into the school’s hard-wired network. The wireless connection was secure; unregistered computers could not use the network.

Students used these laptops to complete their projects. When making presentations, the students used the local data and applications and only accessed the internet to import images from the teacher-selected list available on a web page or to look up information on the internet. When making movies, students used the internet more intensively. Students accessed the movie-making application (PrimaryAccess) via a web browser, and all data for each successive version of the project was stored on a university database.

During the end-of-unit project for Unit C, the school's internet service provider (ISP) made changes to its internal configuration that resulted in the disruption of all internet traffic coming to the city school district from the university server on which the PrimaryAccess database resided (M. Leach, personal communication, November 2, 2006). School district servers were accessible from the university network, and university servers were accessible from any location outside the school district network. From inside the school district network, no university services (PrimaryAccess, web-based email, online card catalogs, etc.) responded, but all non-university web sites (e.g., ask.com) were accessible. During this interruption, the researcher worked with school district technology support personnel to diagnose and resolve the problem.

Data Collection

Property surveyors can establish a boundary with a single measurement from an established point. However, surveyors' standard practice is to *triangulate*, measuring from two points along a baseline, and using these two points to locate a third. This process provides a more accurate location than taking a single measurement (Krathwohl, 1998, p. 275).

The surveyors' practice of triangulation is now a standard technique in social science research. Researchers employ triangulation across data sets, investigators, and methods (Ibid.). This study triangulated across data sets and methods (see Table 8).

- Data sets: The study took place across four consecutive units of instruction. Prior to the units, the students answered a 50-item, multiple-choice pretest over the semester content. During the semester, data from each unit included open-ended and standardized pretest and posttest items, as well as student products (either a movie or presentation). The semester concluded with the re-administration of the same 50-item, multiple-choice test used as the pretest.
- Methods: The study used several methods, including quantitative analysis of scores on multiple-choice items, coding of open-ended responses, line-by-line document analysis of student products, and teacher interviews.

Assessments

The assessments consisted of both multiple choice and open-ended response items conducted throughout the semester. Students took a semester pretest in August and a semester posttest (exam) in January. These semester pre- and posttests consisted of 50 multiple choice items. During the semester, students took four end-of-unit tests. Each end-of-unit test contained 20-40 multiple choice items, and the tests for the intervention units (Units A and C) also included open-ended prompts.

School district personnel scored the multiple choice items, blinded the data, and provided it to the researcher. The teacher collected the open-ended responses and, after conducting his own assessment, provided them to the researcher. The researcher then coded the responses and shared the results with the teacher.

Table 8: Triangulation of Data and Methods

class grouping	sem pre	Unit A *Jim Crow, Great Plains, cultural conflicts*			Unit B *Immigration & industrialization*	Unit C *Spanish-American War, early 20th century American life, Great Migration*			Unit D *WW I, Progressivism, Harlem Renaissance*	sem post	interview
		pre	product	posts	post	pre	product	posts	post		
General track #1 (n=9)	50-item mult. choice	3-4 open-ended	movie doc. analysis	20-item mult. choice repeat open-ended	33-item mult. choice	3-4 open-ended	presentation doc. analysis	33-item mult. choice repeat open-ended	40-item mult. choice	50-item mult. choice (repeats sem pre)	Teacher interview
Talent Development track #1 (n=18)			movie doc. analysis				presentation doc. analysis				
Honors track #1 (n=12)			movie doc. analysis				presentation doc. analysis				
General track #2 (n=11)			presentation doc. analysis				movie doc. analysis				
Talent Development track #2 (n=14)			presentation doc. analysis				movie doc. analysis				
Honors track #2 (n=22)			presentation doc. analysis				movie doc. analysis				

Table 8: This study triangulated in both *data sets* (collecting data over four units of instruction, bookended by two examinations) and *methods* (standardized items, open-ended items, document analysis, and interview)

Projects

Students completed projects (either a movie or a presentation) at the end of Unit A and Unit C. These projects were collected by the researcher and the teacher. Because the movie-making application, PrimaryAccess, is publicly available on the internet, the researcher was able log into students' movies and observe the different versions as they were constructed. Because the presentation-making application, PowerPoint, is a desktop-based application, the teacher collected students' finished products and provided them to the researcher on a removable drive.

Observations

The researcher observed classroom instruction for three of the four units (Units A, B, and C), sampling instruction within each track and for each class period. During the 11 weeks of instruction covering the 3 units, the researcher observed part or all of 24 days, including at least 1 day of each week. The researcher observed multiple class periods each day, totaling 80 classes throughout the 11 weeks. During the end-of-unit projects, the researcher was present for 2 or 3 periods of each group's work with PowerPoint or PrimaryAccess to carefully observe the teacher's and students' interactions with these applications. For every instructional day observed, the researcher collected copies of all handouts and instructional materials.

The researcher followed a passive participation protocol, in which the observer "is present at the scene of action but does not participate or interact with other people to any great extent" (Spradley, 1980, p. 59). Passive participation maintains, to the greatest degree possible, the natural working environment of the classroom. Seated at the back of the room, the researcher listened, watched, and took extensive notes on the teacher and student actions. The researcher recorded all writing on the board, and conducted quick reviews of students' papers at the end of the day. The researcher became an active participant only when requested by the teacher (e.g., resolving the internet access problem described above) or students (e.g., students asked the researcher whether the web address for PrimaryAccess was .org or .com.). The researcher's field notes were transcribed and provided to the classroom teacher within 48 hours of each observation for member-checking.

Because one of the technological tools (PrimaryAccess) is web-based, the researcher's observation of the classroom environment included monitoring of the server on which student data is stored. When the teacher left formative feedback to the students, or when students logged into their project from home, the researcher was able to note this activity via the timestamps on the students' data files.

Interview

Finally, the researcher interviewed the teacher, focusing on the Virginia Standards of Learning, historical content knowledge, and the use of technology, specifically PowerPoint and PrimaryAccess. (See Appendix B for protocol.) The researcher recorded the interviews and took notes during the conversation. The researcher then transcribed the interview and conducted a line-by-line analysis to identify trends and align the teacher's responses with observed behaviors from the classroom observations and document analyses.

Data Analysis

The process of data analysis was shaped by the questions: first, the students' responses on the teacher-designed pretests and posttests were assessed for patterns, using quantitative techniques for the multiple-choice scores and qualitative techniques for the open-ended responses. Once this data analysis was completed, the researcher articulated findings regarding students' short-term (end-of-unit) and long-term (end-of-semester) learning outcomes as measured by the teacher-designed tests. Next, the environmental data (classroom observation notes, student products, and teacher interview transcripts) were analyzed to identify themes and provide context for the previous findings.

Assessments

The multiple choice items from the semester pre- and posttests and from the end-of-unit tests were first sorted by content sampled. For this purpose, the 50-item beginning-of-semester pretest and 50-item end-of-semester posttest was broken into subscales by unit, providing scores for each of the four units. (The total of items from each subscale does not add up to 50; additional items covered content not addressed in Units A-D, such as geography.) Each end-of-unit test was treated similarly: each item was coded by content, and items that addressed content outside of the unit in question were excluded. The construct validity of the resulting subscales was confirmed by an item-by-item coding by a panel of three experts. These experts were selected based on their training and experience in social studies education and familiarity with the Virginia Standards of Learning: all were social studies educators with advanced degrees, and all had taught in the Commonwealth of Virginia. Items that the experts judged to be outside of the content called for by the curriculum guide or poorly written were excluded. The resulting scales underwent a reliability analysis following Cronbach's alpha. (See Table 9.)

Cohort and class performances on each scale were first described by calculating means, standard deviations, and effect sizes using Cohen's *d*. The four students identified as ESL (see Table 3) were consistent outliers on these scales. Because these students were not evenly distributed across sections and track levels, their work (on end-of-unit tests, semester subscales, and project work) was excluded from all further analysis. Next, these statistics were analyzed for significant differences between students' end-of-unit tests and students' semester pretest and posttest subscales for each unit. The means for each end-of-unit test were analyzed, first by condition (i.e., movie-making or presentation-making cohorts) and then by condition within track level (General, Talent

Table 9

Distribution of Items and Coefficients of Reliability for Multiple-Choice Tests

	Number of test items (coefficient of reliability)					
Unit (topics)	sem pre	unit A test	unit B test	unit C test	unit D test	sem post
Unit A (Jim Crow, Great Plains, cultural conflicts)	8 (.24)	15 (.44)				8 (.33)
Unit B (Immigration & industrialization)	12 (.34)		21 (.72)			12 (.38)
Unit C (Spanish-American War, early 20th-century American life, Great Migration)	9 (.58)			24 (.67)		9 (.43)
Unit D (World War I, Progressivism, Harlem Renaissance)	10 (.55)				36 (.78)	10 (.32)

Table 9: Distribution of items and Cronbach's alpha scores for multiple-choice tests. The semester pre- and posttests repeated items; the end-of-unit tests did not. Note that the semester pre- and posttest examinations consisted of 50 items, from which subscales were formed for each unit included in the study. The total number of items in the subscales does not add up to 50 due to inclusion of additional material (e.g., geography) not included in Units A-D.

Development, or Honors classes compared against one another) using an independent samples *t*-test. The analysis for each unit was independent, since each assessment consisted of unique items (no items were repeated from one test to the next) and reflected learning of different content after similar (but not identical) instructional techniques. Because the semester pre- and posttests did not share items with the end-of-unit tests, comparisons were not made across the semester pre- and posttest scales and the end-of-unit tests. Each comparison of semester pre- and posttest subscales was analyzed as a repeated measure, and in the interest of consistency effect sizes were calculated using Cohen's *d*.

The open-ended response items were coded using the constant-comparative technique. The curriculum guide provided the initial themes, but as additional themes emerged, the researcher re-coded the previous data for the theme. All pretests and posttests responses (totaling slightly more than 1,000 responses) were coded. The pretests displayed an overwhelming pattern of "I don't know" (or "I'm not sure," "I forget," etc.) or non-responses (skipped items): more than 50% of responses were coded as "don't know" or non-responses. The remainder of the responses offered some insight into previous instruction (e.g., two students' responses to "Why did westward expansion occur?" cited Manifest Destiny, which is a topic specified in the sixth-grade curriculum guide) but suggested no mastery of the material specified in the USII curriculum guide (i.e., opportunities for land ownership, discoveries of silver and gold, building the transcontinental railroad, etc.—see Appendix A). Consequently, the open-ended pretests were not subjected to further analysis. However, responses on the pretests were used to shape coding categories for the posttests.

Student responses to the prompts administered on the end-of-unit tests were coded and then reduced into three categories:

- Information specified by the curriculum guide;
- Information not mentioned in the curriculum guide, but relevant to the topic; and
- Incorrect or misleading information.

(See Table 10 for examples of each.) Non-responses were not coded, since the researcher was unable to determine which non-responses reflected a lack of time rather than a lack of information. Because very few responses included misconceptions or misleading information (approximately 5% of the total, distributed evenly by condition), this category was dropped from the group analysis. This reduced data set was then sorted by condition and scrutinized for patterns of differences.

Following this analysis, sub-group differences were explored, comparing students by condition within track level and also by topic of project work: students writing on the same topic as their project were compared with those of students who were writing on topics outside of their project. For example, when examining student responses to "Why did westward expansion occur?," the responses of students who created either a movie or presentation about the Great Plains were compared with the responses of students whose projects focused on other topics from the unit (i.e., Jim Crow laws or cultural conflicts on the Great Plains).

Groups' performances on the open-ended prompts were then characterized to describe the coverage of topics within and outside the curriculum guide. If the average response for a group addressed a single topic, the group's performance was described as

Table 10

Examples of Student Responses and Coding Categories for Open-Ended Prompts

Prompt: Why did westward expansion occur? (Essential Question for Virginia SOL USII.3a)	Within-CG	External to CG	Misconception or misleading
"Westward expansion occured from the five reasons for moving west. Adventure, new beginning for former slaves, technological advances including the Trancontinental Railroad, wealth possibility from found gold and silver, and cheap land ownership."	✓		
Explanation	This answer provides exactly the information specified by the curriculum guide (and features almost the exact phrasing) and nothing more.		
"because of manifest destiny"		✓	
Explanation	Manifest Destiny is relevant to the prompt, but is not mentioned in the USII curriculum guide.		
"Because the east got over populated. Looking for gold."	✓		✓
Explanation	The search for gold is identified as Essential Knowledge for USII.3a, but the East was not "over-populated," and this explanation is a distractor from the motives outlined by the curriculum guide.		
"Because of gold and silver, and the Homestead Act."	✓	✓	
Explanation	The search for gold and silver is specified in the curriculum guide; the Homestead Act is not mentioned in the USII curriculum guide		

Table 10: Example of an open-ended prompt and coding categories, drawing upon *History and social science Standards of Learning curriculum framework: Essential knowledge, skills, and understandings*, p. 159. "Within-curriculum guide" responses cite the information listed on that page as "Essential Knowledge". "External to curriculum guide" responses cite information that is not listed but is relevant to the prompt. "Misconception or misleading" responses either provide incorrect information or suggest a line of thinking far removed from the information provided by the curriculum guide and the textbook.

Medium. If the group addressed, on average, more than one point, the group's performance was described as High; if the average included several points, it was described as Very High. Conversely, a group that scored on average less than one point was described as Low or – if the average approached zero – None.

Projects

Following the analysis of students' outcomes on the end-of-unit and semester assessments, the researcher examined the students' products. The researcher first sorted the projects by topic (Jim Crow laws, Great Plains, or cultural conflicts; Spanish-American War, early 20th century American life, or Great Migration). To place both formats (movies and presentations) on a common metric and establish a rough measure of output, the researcher conducted a simple word count and image count for each project. For the purposes of the word count, the researcher excluded non-content text, such as the students' names. Each image used in a movie or presentation was coded as being teacher-selected (i.e., included on the list of images provided to students during each project) or non-teacher selected.

To describe the quality of students' work relative to instruction and assessment, the researcher conducted two rounds of coding. First, projects were scrutinized in a line-by-line analysis for factual errors or lack of conformity to instruction, defined as disagreement with the information presented in the curriculum guide, textbook, teacher-generated handouts, or during observed classroom instruction. For the purposes of this coding, the researcher ignored spelling mistakes ("The assembly line helped make cars aforabol") or simple word substitutions ("They think that the Maine blew up because of minds") and noted only misconceptions ("Well a white man got on and wanted the seat that Plessy was in, so he told him to get up out of his seat"—Plessy in fact informed the conductor, upon surrendering his ticket, that he was breaking the law). Next, students' work was again subjected to a line-by-line analysis using the semester exam as a coding frame: each project was matched to exam questions that corresponded to its topic, and each question was unpacked to generate codes. The codes addressed key information (linking the question stem to the correct answer), stem information (additional information in the question stem), and distractor information (allowing the student to eliminate incorrect answers). (See Table 11 for examples.)

Table 11

Sample Scoring of Student Projects Using Exam Items as Coding Frame

Item	Sample presentation text (Honors track)	Sample movie script (General track)
(USII2.a) 2. Dust storms + ______ + eroded land = Great Plains a. elevated land b. hardwood trees c. low rainfall d. little wind (correct answer = "c. low rainfall")	"Before common inventions were made, the West was seen as a 'treeless wasteland'. Knowing this the settlers were forced to adapt to it's harsh climate, little water supply, and lack of trees."	"The great plains was seen as a treeless wasteland.It also had frequent dust storms,low rainfall,and land eroded by wind and water."
Codes KI = key information (i.e., associates "low rainfall" and "Great Plains") SI = stem information ("dust storms"; "eroded land") DI = distractor information	KI, DI Includes key information for item ("little water supply" ≈ low rainfall) and some distractor information (lack of trees ≠ hardwood trees)	KI, SI, DI Includes key information, complete stem information (dust storms, eroded land) and includes almost complete distractor information (treeless ≠ hardwood trees; little wind ≠ eroded by wind)

Finally, the movie files were examined for evidence of student-teacher interaction. Because PrimaryAccess allows teachers to leave formative feedback in embedded text notes, the researcher was able to conduct a frequency count of feedback left within students' movies. Each instance of feedback was coded as:

- Direction regarding style or usage (e.g., "Make sure this reads smoothly. It seems to have some typos and it jumps around in some places."),
- Non-content encouragement (e.g., "Great start to this paragraph. Keep it up!"),
- Non-specific content encouragement ("Remember to use your study guide."), and/or
- Specific content encouragement ("Make sure you are covering information about electrification as well.").

Because no such notes were left in students' PowerPoint presentations, no equivalent analysis of products could be made for the competing condition. Instead, classroom observation notes (see below) provided the only record of student-teacher interaction.

Observations and Interview

The researcher coded the classroom observation notes and teacher interview transcripts following grounded theory. As themes and trends emerged from the data, they were checked against the themes and trends from the student data (performances on pretests and posttests and scoring of products). Particular emphasis was given to the project work sessions, to note instances of in-class formative feedback from the teacher to students.

The results from the standardized pretests and posttests and open-ended pretests and posttests informed the findings on the first two research questions (i.e., differences in student outcomes). The results from the analysis of student products, classroom observation notes and interview transcripts provided the basis for answering the third research question (contextual factors influencing student outcomes).

Ethical Considerations

The primary mission of any school is to educate its students, not to support research. Instructional time is a precious commodity, and no teacher, parent, or administrator will surrender it lightly. Accordingly, education researchers must approach school-based research situations with the utmost care and ensure that the wishes and best interests of the participants are being respected.

In 2005, the researcher made initial contact with school district central office personnel to discuss possibilities for exploring new forms of technology integration into instruction. The district personnel, who were familiar with PrimaryAccess, focused on that tool and identified a teacher who would be an eager user. The researcher met with the teacher and the school principal to establish guidelines and expectations for the teacher's use of PrimaryAccess and the researcher's role. This pilot experience was a success, culminating in student products being shown at a school board meeting. Following this pilot, the principal and the seventh-grade social studies team decided to adopt the selective use of PrimaryAccess.

The teacher, students, and parents were free to decline to participate in the study. The classroom observation and examination of student work followed a passive assent/consent policy approved by the Institutional Review Board of the University of Virginia (project # 2006-0058-00): students and/or parents were informed of the project and could opt out at any time. No students or parents elected to opt out, allowing the researcher to analyze the full data set.

All data was held securely and confidentially in a locked location in the researcher's home. Objective test data was blinded before coming to the researcher, and the analysis was not available until after the grading period was completed. Open-ended responses, student products, and statements from teacher interviews were stripped of identifiers before being analyzed. Notes on classroom observation and the teacher interviews were not released to anyone other than the respective participant. Participants could leave the study at any time or expunge any section of their respective data upon notifying the researcher. However, no participants made this request.

Validity

A mixed-method design must consider the validity of both its quantitative and qualitative measures, as each domain carries its own unique set of strengths and weaknesses. If a researcher fails to consider validity from each perspective, he or she runs the risk of combining the methods' weaknesses, and not their strengths.

Validity of Quantitative Measures

The primary goal of this study was to evaluate the effectiveness of two competing tools and teaching methods for building students' historical content knowledge. If a significant and practical difference is found, the research must prove that the use of a particular tool was a contributing cause. To make such a causal claim, a study must have internal validity. Relevant threats to the internal validity of this study include the following factors (Krathwohl, 1998, pp. 512-519):

- *Sampling and error chance.* Differences that emerge may be random. Inferential statistics were used to bound the probability of sampling and error chance, with $\alpha = .05$.
- *Testing*. Students' improvements on repeated tests of historical content knowledge may be due to increasing test-taking skill or increased comfort taking test. However, the only repeated items (i.e., those on the semester pre- and posttests) were separated by an entire semester of instruction spanning five months. Accordingly, there was no practical effect from testing.
- *Regression*. When a group or individual is exceptionally above or below the mean, they are likely to score closer to the mean on a subsequent re-testing. Given the repeated measurements taken during this study, any movement toward the mean may be regression, rather than growth or decline. This study minimized the threat of regression by selecting all students within the teacher's classes, rather than seeking to select only high-scorers or low-scorers.
- *Local history*. Events at the school or in the community may affect students' performance on tests. The researcher was alert to the possibility of local history effects by observing the classroom and interviewing the teacher.
- *Mortality*. If participants leave a group during a study, their absence alters the group's mean. The researcher observed for mortality by monitoring

attendance data. Students who left during a unit were removed from the analysis of the unit; students who left during the course of the semester were removed from the semester analysis. The same process was followed for students who entered the class.

- *Maturation*. As the students in this study moved through the instructional units, they were learning not just historical content knowledge but test-taking skills, study habits, and patterns of interaction with the teacher and other students. Seventh grade marks their first year of junior high school, and they inevitably shed some of the habits of elementary school. These changes, rather than a different instructional strategy, may be the cause of differences in performance. This study controlled for maturation by using multiple pretest-posttest measures. If students displayed increasing gaps between their pretest and posttest performance, the posttest may be measuring not just historical content gained but improved test-taking skills. No such gaps were observed. Other changes to maturation, such as boredom, were monitored through classroom observation and teacher interview.
- *Instrument decay*. A measuring device may become less sensitive over time, particularly if the measuring device is a researcher who is progressively less attentive or more habituated to an environment. This study guarded against instrument decay by using multiple measures (e.g., multiple choice tests and open-ended prompts). Furthermore, because the content studied changed from one unit to the next, the coding themes changed for the open-ended items and

document analysis. The rater, therefore, had less opportunity to experience habituation and decay.

- *Selection*. When groups are not randomly assigned, selection bias may be at work, providing a difference that is not the result of a treatment but of some other, pre-existing variable. This study controlled for selection bias by employing switching replication–that is, repeating the experiment, with the conditions reversed. Each group therefore experienced both conditions, controlling for any effect that emerged as a result of the grouping and not the intervention.
- *Interaction*. This study is particularly vulnerable to unintended interaction effects. For example, each unit of history is different than the rest; the carryover from one unit to the next is far smaller than in, say, mathematics or foreign language. Therefore, each unit exerts a content effect–students may find one topic (e.g., the Spanish-American War) more or less difficult than another topic (World War II). This content effect can combine with any other variable (such as tracking) to produce a difference in performance that is falsely attributed to the treatment. Short of repetition, only rigorous consideration of rival hypotheses can adequately address interaction effects.

Validity of Qualitative Measures

Because qualitative data is collected through the filter of human perception and is therefore an ongoing, unstable process of selection, qualitative research requires attention to its own unique set of concerns, mostly centered upon the researcher. Erickson (1986) identifies "five major types of evidentiary inadequacy" (p. 140):

- *Inadequate amount of evidence*. Because qualitative findings are substantiated with illustrative anecdotes and excerpts, and not statistics, a researcher may present a finding with inadequate evidence behind it. Most of the qualitative methods in this study are exhaustive: the researcher examined all student products, for example, and coded all open-ended prompts. The conclusions drawn from data generated via sampling (i.e., classroom observations, interviews) are held more tentatively.
- *Inadequate variety in kinds of evidence*. If a researcher does not triangulate, he or she is vulnerable to basing findings upon an inadequate span of data. Because this study used multiple qualitative and quantitative measures, the researcher was able to base his findings upon a broad evidentiary base.
- *Faulty interpretive status of evidence*. While the researcher is a highly tuned, subtle instrument, he or she observed a complex situation, and the full interplay of the actors may be beyond the researcher's ability to grasp. This study guarded against faulty interpretive status by member-checking: the classroom teacher received copies of the observer's notes within 48 hours of each observation for confirmation and discussion of ambiguous actions.
- *Inadequate disconfirming evidence*. As the researcher observes, he or she may fail to note behaviors that run counter to primary themes. This study guarded against the predilection to find what one is looking for by using the switching

replication in the design (i.e., reversing experimental and control conditions) to specifically observe for disconfirming evidence.

- *Inadequate discrepant case analysis*. A researcher, in pursuit of an effect, may fail to compare disconfirming cases with confirming cases. Because the data collection in this study was exhaustive, including all students' data over all units of study and not just the intervention units, the researcher addressed all cases, and not just those confirming a particular hypothesis.

Researcher as Instrument

The observer's background filters the observation, and researchers are not excepted from the rule (Kuhn, 1962, Ch. 5). While a researcher should be free of "special biases that would distort his view of the phenomena" (Krathwohl, 1998, p. 345), qualitative researchers in particular, should describe their background, influences, and possible biases "so that they can be taken into account when judging the findings" (Ibid.).

I am a former classroom teacher. My undergraduate degree is in history and international studies, and I simultaneously completed a teacher licensure program. Upon graduation I immediately assumed a teaching position at an elite Midwestern all-male, Jesuit day school. I taught Advanced Placement (AP) European History, American History, World History, Government, and English Language and Composition. After six years at this school, I decided to teach overseas. My wife and I spent a total of four years working at independent, American-curriculum schools in Haiti and Saudi Arabia. In these schools, I continued teaching history and English and added new specializations in AP

Economics and computer science. I particularly relished the opportunity to teach AP courses, due to the intellectual challenge of the content, the caliber of the students, and the rigor of the test. Every summer, I looked forward to receiving the report of students' scores through my school administration, and I refined my teaching to try to improve upon those scores in the following year.

After ten years in the classroom, I decided to pursue my growing interest in technologies for teaching and learning by entering a doctoral program in Instructional Technology. I am a graduate fellow at the Center for Technology and Teacher Education at the University of Virginia's Curry School of Education. As a fellow, I teach a pre-service educational technology course for future secondary humanities teachers.

In addition to my coursework and teaching, I have collaborated with Bill Ferster and Glen Bull to develop PrimaryAccess, the online digital video environment to be used in this study. I have conducted previous exploratory studies observing students and teachers using the application in the classroom. I have also delivered presentations and published articles about the application. My interest in PrimaryAccess stems from my beliefs that students learn best when engaged in constructing meaningful whole products and that teachers teach best when their technological tools are designed specifically for the discipline in question.

Because of my role in the development of PrimaryAccess, I am a stakeholder: I have a direct interest in the issues and outcomes of the study. To guard against prejudice against PowerPoint or bias in favor of PrimaryAccess, I have selected a restrictive study design: the outcome measures are teacher-designed, not researcher-designed, and all qualitative analysis flows from the framework of the curriculum and classroom instruction. The strongest safe-guard against over-reaching is the use of switching replication (reversing conditions to observe a repeated effect): if the same effect is not observed in both implementations, then the result is not conclusive.

Summary

This study was a repeated quasi-experimental, mixed-method design. An early-career history teacher led half of his students through an experimental condition under which they used an integrated media-scaffold-canvas tool (PrimaryAccess) to produce online digital documentaries about topics in the curriculum. The other cohort executed a similar project using PowerPoint, a common classroom tool, in an instructional strategy that integrated online media and the canvas capabilities of PowerPoint. On a subsequent unit, these cohorts traded conditions, allowing each group to become its own control.

The researcher collected both outcomes data and observational data. The students took multiple pretests and posttests, responding to both objective and open-ended prompts. These pretests and posttests were dispersed throughout four units of instruction. Additionally, the researcher conducted a document analysis of student products and conducted interviews with the teacher. Conclusions regarding students' growth in historical content knowledge were based on all forms of evidence, and not just upon the standardized test items.

CHAPTER 4

RESULTS

Introduction

This study examined the content knowledge outcomes, as determined by performance on teacher-designed tests, of six classes of seventh-grade history students following teacher instruction and project work using two competing technologies: a purpose-built application for online digital movie-making using primary source images (PrimaryAccess) and a ubiquitous slideware tool for making presentations (PowerPoint). As discussed in Chapter Two, social studies educators have indicated the possibility of improved student outcomes from effective integration of technology into social studies classroom instruction. However, only a handful of studies have examined the results of students' use of technology in social studies education. This study provided an opportunity to observe student learning outcomes and contextualize them within the teacher's instruction and assessment, the curricular framework, the technological infrastructure, and students' own behaviors during the instruction, project work, and assessment. Because the design incorporated switching replication, in which the experimental and control conditions are reversed on a following unit of instruction, any hypothesized effect from the first intervention could be tested against the results from the second intervention. The repeated interventions also provided an opportunity to weigh which contextual factors appear to affect the outcomes.

Sampling an entire semester of history instruction, collecting student products and assessments of student content knowledge, and conducting teacher interviews results in a large quantity of data. This data can be analyzed at many grouping levels, from individual to cohort, and the analysis can be informed by many frameworks and operate at high or low levels of inference. To hew the surest path through this thicket of decisions, the researcher must be guided by the research questions. The questions addressed in this study were:

1. Do differences exist in student outcomes, as measured by pre- and posttest scores on the standards-driven assessment, between students who use PrimaryAccess vs. those who generate a PowerPoint slideshow?
2. If differences exist, does the effect on student outcomes vary by achievement-level grouping?
3. What environmental factors – such as the curricular context, teacher behaviors, student behaviors, or technological infrastructure – appear to inhibit or promote this effect?

This chapter presents the results of the data collection and analysis conducted to answer these questions, organized around the findings that emerged from this process. The first two questions were addressed simultaneously, as they addressed the same construct (content knowledge) and relied upon the same data sources (student performance on pre- and posttest items). The analysis of the end-of-unit and semester pre- and posttest student assessment data suggested two findings:

1. As measured by teacher-designed tests, students who created movies (using PrimaryAccess) displayed no consistent differences in short-term (end-of-unit) learning outcomes compared to students who created presentations (using PowerPoint).

2. As measured by a teacher-designed test, students who created movies (using PrimaryAccess) appeared to have superior long-term (i.e., over several months) learning outcomes compared to students who created presentations (using PowerPoint).

These two findings then shaped the analysis conducted to address the third research question, examining the surrounding environmental factors. This process yielded three additional findings of observed differences in student project work, teacher behaviors, and the technologies used during the projects:

3. As defined by use of teacher-selected information and resources, students working with the movie-making application (PrimaryAccess) during the first intervention demonstrated a greater alignment with classroom instruction and assessment than students working with the presentation-making application (PowerPoint).
4. During the end-of-unit projects, the teacher was able to use the movie-making application (PrimaryAccess) to scaffold students' work; he was unable to use the presentation-making application (PowerPoint) to the same effect.
5. During the second end-of-unit project (Unit C), technical issues stemming from changes in the school's internet service changed the implementation of the movie-making project and reduced the teacher's and students' use of the movie-making application.

The following sections present these findings, and the analysis of the research questions that led to them, in depth.

Finding 1

As measured by teacher-designed tests, students who created movies (using PrimaryAccess) displayed no consistent differences in short-term (end-of-unit) learning outcomes compared to students who created presentations (using PowerPoint).

Overview

On the end-of-unit tests, no clear differences existed between students by condition. The analysis of both the multiple choice items and the open-ended prompts failed to yield a discernable pattern either across implementations (i.e., on Unit A and Unit C) or within a single implementation (on Unit A or Unit C).

End-of-unit Multiple Choice Scores:
No Consistent Differences

The first step was to observe student performance on the end-of-unit, teacher designed multiple-choice test. After excluding items that addressed content outside of the unit being studied, the resulting scales consisted of 15-36 items. Coefficients of reliability for these scales ranged from .44 to .78. (See Table 9 for specific values.) For each cohort and each class, the means, standard deviations, and effect sizes were observed, and significance was tested through a independent-samples *t*-test. (See Table 12.) A statistically significant difference favoring presentation-making existed for Unit A at the $\alpha = .05$ level of significance, but the effect was not repeated when the conditions were

Table 12

Differences on End-of-Unit Tests by Condition for Intervention Units

	Unit A end-of-unit test (15 items)				Unit C end-of-unit test (21 items)			
	Condition	Mean (σ)	*t*	*d*	Condition	Mean (σ)	*t*	*d*
Cohort 1 (mov/pr)	Movie (n=37)	94.1 (4.04)	-2.38 ($p < .05$)	-0.54	Pres (n=35)	91.4 (8.93)	-0.687 ($p > .4$)	-0.1
Cohort 2 (pr/mov)	Pres (n=47)	97.6 (8.28)			Movie (n=46)	92.3 (9.51)		

Table 12: Analysis of differences of end-of-unit tests by condition. The differences in *n* are due to students taking alternative tests or entering/leaving the teacher's class roster. The differences do not affect the overall conclusions, however. Note that the variance for the groups is not equal for Unit A, and the results reported are for equal variances not assumed.

reversed on the second intervention (Unit C). In neither case were the differences practically significant. (In practical terms, the difference registered on Unit A represents approximately a 0.5-item difference on a 15-item scale.)

As a point of comparison, the same analysis was conducted for the non-intervention units. (See Table 13). The differences shown on the non-intervention units suggested that non-significant differences in end-of-unit test scores can be expected. This observation made the difference noted during the first intervention unit (Unit A) appear

Table 13

Differences on End-of-Unit Tests by Condition for Non-Intervention Units

	Unit B end-of-unit test (24 items)				Unit D end-of-unit test (36 items)			
	Condition	Mean (σ)	*t*	*d*	Condition	Mean (σ)	*t*	*d*
Cohort 1 (mov/pr)	No project (n=36)	92.4 (8.77)	-0.15 ($p > .9$)	0.0	No project (n=36)	90.9 (9.05)	-0.756 ($p > .4$)	-0.17
Cohort 2 (pr/mov)	No project (n=46)	92.4 (8.97)			No project (n=44)	92.4 (8.95)		

Table 13: Analysis of differences of end-of-unit tests for non-intervention units. Differences in *n* are due to students taking alternative tests or entering/leaving the teacher's class roster and do not affect the overall conclusions.

more significant. However, the analysis of the non-intervention units also revealed a pattern in all four end-of-unit tests: Cohort 2 equaled or out-performed Cohort 1. The student selection (i.e., the larger number of Honors-track students; see Table 3) may have caused this effect.

The next step was to examine these same results by track level to see whether the sub-groups displayed the same pattern of results as at the cohort level. (See Table 14.) Again, no pattern of differences emerged within any one track. For each track level, one of the two interventions produced statistically significant differences (either Unit A or

Table 14

Differences on End-of-Unit Tests by Condition Within Track Levels for Intervention Units

		Unit A end-of-unit test (15 items)				Unit C end-of-unit test (21 items)			
		Condition	Mean (σ)	*t*	*d*	Condition	Mean (σ)	*t*	*d*
General	Gen1 (mov/pr)	Movie (n=8)	95.8 (6.11)	-0.242 ($p > .8$)	0.1	Pres (n=8)	86.9 (12.4)	-2.184 ($p < .05$)	0.34
	Gen2 (pr/mov)	Pres (n=11)	95.2 (6.03)			Movie (n=11)	82.7 (12.5)		
TD	TD1 (mov/pr)	Movie (n=17)	90.6 (9.45)	-2.862 ($p < .01$)	-0.9	Pres (n=17)	91.9 (8.87)	-0.691 ($p > .4$)	-0.25
	TD2 (pr/mov)	Pres (n=14)	97.6 (3.31)			Movie (n=14)	93.9 (6.84)		
Honors	Hon1 (mov/pr)	Movie (n=12)	97.8 (5.92)	-0.691 ($p > .4$)	-0.2	Pres (n=10)	94.3 (3.76)	-2.184 ($p < .05$)	-0.84
	Hon2 (pr/mov)	Pres (n=22)	98.8 (2.63)			Movie (n=21)	97.5 (3.87)		

Table 14: Analysis of differences of end-of-unit tests within tracks for intervention units. Again, the differences in *n* are due to students leaving the teacher's class roster and do not affect the overall conclusions. Note that the variances for the Talent Development track and Honors track for Unit A are heterogeneous at the .05 level, and therefore equal variances are not assumed.

Unit C), but the other did not. The largest effect sizes were for the Talent Development track on Unit A and the Honors track on Unit C. However, in the first instance the effect favored the presentation-making condition, and the second favored the movie-making condition.

To further explore the random distribution of differences between class groupings within the track levels, the analysis was repeated for the non-intervention units. (See Table 15.) None of the differences were statistically significant, and the effect sizes fell within the same range as on the intervention units.

After examining the end-of-unit multiple-choice item responses, no clear

Table 15

Differences on End-of-Unit Tests by Condition Within Track Levels for Non-Intervention Units

		Unit B end-of-unit test (24 items)				Unit D end-of-unit test (36 items)			
		Condition	Mean (σ)	*t*	*d*	Condition	Mean (σ)	*t*	*d*
General	Gen1 (mov/pr)	No project (n=8)	89.1 (10.9)	1.20 ($p > .2$)	0.56	No project (n=8)	92.4 (9.00)	1.85 ($p > .05$)	0.89
	Gen2 (pr/mov)	No project (n=11)	83.0 (10.9)			No project (n=10)	83.3 (11.2)		
TD	TD1 (mov/pr)	No project (n=18)	90.7 (8.99)	-1.03 ($p > .3$)	-0.37	No project (n=18)	89.5 (7.95)	-0.868 ($p > .3$)	-0.32
	TD2 (pr/mov)	No project (n=14)	93.8 (7.08)			No project (n=13)	91.9 (6.85)		
Honors	Hon1 (mov/pr)	No project (n=10)	97.9 (2.20)	1.22 ($p > .3$)	0.41	No project (n=10)	92.2 (11.3)	-1.30 ($p > .2$)	-0.56
	Hon2 (pr/mov)	No project (n=21)	96.4 (4.62)			No project (n=21)	97.1 (4.85)		

Table 15: Analysis of differences of end-of-unit tests within tracks for non-intervention units. Again, the differences in *n* are due to students leaving the teacher's class roster and do not affect the overall conclusions. Note that the variances for the Honors track for Units A and C are heterogeneous at the .05 level, and therefore equal variances are not assumed.

pattern emerged to suggest an advantage for the movie-making or presentation-making condition, whether by cohort or within track level. The size of differences (measured by *d*) and the significance of differences (indicated by *p*) displayed no pattern, especially when compared against the differences that existed on the non-intervention units: the largest effect size during the intervention ($d = -0.9$, favoring the presentation-makers in the Talent Development track on Unit A) is the same as the largest effect size measured during the non-intervention units ($d = 0.89$ in the General track on Unit D). Whenever a significant difference was observed by condition on Unit A, it was not repeated when the conditions were reversed on Unit C, suggesting that the difference may have been caused by grouping and not by the intervention. As measured by the students' performance on the end-of-unit multiple-choice items, the teacher seemed to be equally successful under all instructional conditions: at every track level, each section averaged better than 80% (and typically greater than 90%) on the end-of-unit test.

End-of-unit Open-ended Responses: No Consistent Differences

Given the possibility of a ceiling effect on the end-of-unit multiple-choice scores, the next step was to examine students' responses to the open-ended prompts. The teacher used four Essential Questions on the Unit A test and three on the Unit C test. (No open-ended prompts were administered during Units B and D.) Students answered each prompt twice, first as a pretest before instruction on a topic (e.g., before instruction on the Great Migration, students were asked to write a response to the question, "Why did African-

Americans migrate to northern cities?") and again during the end-of-unit test after completing the multiple-choice items.

The results of the data reduction for each unit and each prompt by condition are displayed in Table 16. (See Table 10 for an example of the coding scheme.) The Unit A posttest displayed considerable variation between groups by condition but featured no clear pattern. The Unit C posttest displayed very little variation by condition. No consistent advantage was evident for either movie-making or presentation-making, either in identifying topics within the curriculum guide or providing relevant information that is not specified within the curriculum guide.

To better explore the subgroup differences, the researcher conducted a second analysis by project topic within the track levels, reasoning that the differences between groups' responses on Unit A may have had more to do with unequal content distribution than the conditions under which the projects were produced: after all, the movie-making group created comparatively more projects (45% of total movies) about the Great Plains than the presentation-making group (26% of total presentations), and three of the four questions asked on the posttest addressed this topic. (See Table 7 for distribution of project topics by unit and by cohort.)

To address this possibility, the General and Talent Development students' responses were re-sorted by project topic of the writer: responses of students writing on the same topic as their project were compared with those of students who were writing on topics outside of their project, generating a total of 26

Table 16

Trends in Open-Ended Posttest Responses

	Topic	Prompt	Group (condition)	Within CG	External to CG
Unit A	Jim Crow	How did African Americans respond to discrimination and “Jim Crow”?	Cohort 1 (movie)	Medium	Low
			Cohort 2 (pres)	Low	Medium
	Great Plains	How did people's perceptions and use of the Great Plains change after the Civil War?	Cohort 1 (movie)	High	High
			Cohort 2 (pres)	High	Low
		Why did westward expansion occur?	Cohort 1 (movie)	High	Low
			Cohort 2 (pres)	Very high	None
		How did people adapt to life in challenging environments?	Cohort 1 (movie)	High	Medium
			Cohort 2 (pres)	Medium	Medium
Unit C	Spanish-American War	What were the results of the Spanish American war?	Cohort 1 (pres)	High	None
			Cohort 2 (mov)	Very high	None
	Early 20th century American life	How was social and economic life in the early twentieth century different from that of the late nineteenth century?	Cohort 1 (pres)	High	Medium
			Cohort 2 (mov)	High	Medium
	Great Migration	Why did African-Americans migrate to northern cities?	Cohort 1 (pres)	Medium	Low
			Cohort 2 (mov)	Medium	Low

Table 16: Display of reduced coding results from open-ended posttest responses. “Misconception or misleading” was dropped as a coding category, as it represented fewer than 5% of the total responses. Note that the responses for the prompts on the Unit A posttest display considerable variation but display no pattern; the responses on the Unit C posttest display very little variation between the groups.

comparisons within a specific track level. First, outstanding individual performances were examined, and no topic effect was observed: out of 26 comparisons, the top-scoring responses were disproportionately students writing *outside* their project topic (12 times) rather than those writing inside their presumed area of expertise (6 times); in the remaining 8 comparisons, the top-scoring writer within a topic was matched by a writer outside the topic. Next, group performance (i.e., all students writing within vs. outside their project topic) was compared. Again, no pattern of differences emerged: over the 26 comparisons, 17 displayed no difference by topic; of the remaining 9 comparisons, only 1 displayed an unambiguous difference in group performance by topic, in which the majority of within-topic writers addressed more topics (both within and external to the curriculum guide) than the outside-of-topic writers. Just as neither condition (movie-making nor presentation-making) created a clear difference in students' performance on the open-ended responses, students' project topic also did not make a clear difference in performance.

Summary

While differences were observed on end-of-unit multiple choice and open-ended responses, no pattern of differences emerged. Mr. Smith's students did well on his end-of-unit multiple choice items regardless of the conditions of instruction, and students' writing of open-ended responses was not impacted by project type (movie or presentation) or topic (e.g., Jim Crow laws, the Great Plains, or cultural conflicts).

Finding 2

As measured by a teacher-designed test, students who created movies (using PrimaryAccess) appeared to have superior long-term (i.e., over several months) learning outcomes compared to students who created presentations (using PowerPoint).

Overview

Analysis of groups' performance on the semester pre- and posttest subscales suggests that students who created a movie on a particular unit displayed greater improvement between the pre- and posttest than students who created presentations. However, the effect was not as pronounced in the second intervention (Unit C). Furthermore, the low coefficients of reliability for each subscale limit the conclusions that can be drawn.

Semester Pre- and Posttest: Consistent Differences Favoring Movie-making

The analysis of student work on the end-of-unit multiple-choice tests and open-ended prompts provided insight into students' short-term learning. Examination of student work on the semester pre- and posttest offers the opportunity to assess long-term learning, as the two tests are separated by more than four months. The 50-question examination was broken into subscales for each unit. The resulting subscales consisted of 8-12 items each; coefficients of reliability ranged from .24 to .58. (See Table 9 for details.) Next, the semester pre- and posttests of the intervention units (A and C) were observed for means and standard deviations. (See Table 17.) Differences between these means were calculated, and effect sizes were observed, using the difference as the numerator and the global standard deviation (σ) as the denominator. On both units, the movie-making group had a lower pretest and a higher posttest score. In practical terms, the difference on the Unit A subscale was approximately 1 item from an 8-item scale; on the Unit C subscale, the difference was slightly larger than 1 item on a 9-item scale.

Table 17

Descriptive Statistics of Semester Pre- and Posttest Subscales for Intervention Units

	Unit A subscale (8 items)					Unit C subscale (9 items)				
	Condition	Pre mean (σ)	Post mean (σ)	Pre-post	*d*	Condition	Pre mean (σ)	Post mean (σ)	Pre-post	*d*
Cohort 1 (mov/pr)	Movie (n=31)	39.9 (19.7)	87.5 (12.1)	47.6	0.76	Pres (n=31)	36.2 (17.9)	91.0 (11.3)	54.8 (21.6)	-0.63
Cohort 2 (pr/mov)	Pres (n=38)	47.4 (16.2)	81.3 (13.2)	33.9		Movie (n=38)	23.1 (21.7)	91.2 (12.7)	68.1 (25.6)	

Table 17: Descriptive statistics of semester pre- and posttest subscales for intervention units. Effect sizes (Cohen's *d*) were calculated using global standard deviation (σ) as the denominator.

For comparison, the same statistics were calculated for the non-intervention units (B and D; see Table 18). The cohorts differed by about 1 item on a 12-item scale for Unit B and by about 0.5 items on a 10-item scale for Unit D. On both subscales, Cohort 2 outperformed Cohort 1. Again, this effect may have been the result of selection (i.e., the overweighting of the Honors track for Cohort 2; see Table 3). Regardless of cause, this pattern stands in contrast to the intervention units, where the movie-making group outperformed the presentation-making groups on both units' subscales.

Table 18

Descriptive Statistics of Semester Pre- and Posttest Subscales for Non-Intervention Units

	Unit B subscale (12 items)					Unit D subscale (10 items)				
	Cond	Pre mean (σ)	Post mean (σ)	Pre-post	*d*	Cond	Pre mean (σ)	Post mean (σ)	Pre-post	*d*
Cohort 1 (mov/pr)	No proj (n=31)	36.0 (17.4)	88.7 (11.7)	52.7	-0.53	No proj (n=31)	29.0 (17.0)	91.8 (7.55)	62.8	-0.30
Cohort 2 (pr/mov)	No proj (n=38)	32.7 (13.2)	93.4 (7.30)	60.7		No proj (n=38)	22.7 (19.6)	91.1 (8.84)	68.4	

Table 18: Descriptive statistics of semester pre- and posttest subscales for non-intervention units. Effect sizes (Cohen's *d*) were calculated using global standard deviation (σ) as the denominator.

The significance of these differences in student performance by condition on the semester pre- and posttests was tested using a split-plot full factorial analysis of variance: the pre- and posttest subscales offered repeated measures of students' content knowledge, separated by more than four months of time. As with the end-of-unit tests, each comparison was treated as independent of the others: items were repeated within but not across subscales, and each subscale assessed similar (but not identical) instruction delivered at different times over different content.

Examining the intervention units first, significant differences by condition were observed on both Unit A and Unit C. (See Table 19.) Specifically, an interaction effect was present, in which the condition (movie-making or presentation-making) did display an effect within the instruction. In other words, the effect sizes noted in Table 17 favoring the movie-making condition were statistically significant. However, as noted on Table 9, the coefficients of reliability for these subscales was low (ranging from .24 to .58), meaning that the differences may not be the result of the intervention but instead stem

Table 19

Source Table for Comparisons by Condition on Intervention Units

	Unit A subscale (8 items)				Unit C subscale (9 items)			
Source	SS	dF	MS	F	SS	dF	MS	F
Between								
Condition *(mov vs. pr)*	12.273	1	12.273	0.039 (p > 0.8)	1423.303	1	1423.303	5.411 (p < .03)
Subjects (condition)	20832.4	67	310.9311		17623.04	67	263.0304	
Within								
Instruction *(pre vs. post)*	56647.2	1	56647.24	331.3 (p > .01)	129076.1	1	129076.1	451.6 (p < .01)
Instruction * Condition	1601.95	1	1601.947	9.368 (p < .01)	1507.692	1	1507.692	5.275 (p < .03)
Instruction * Subjects (condition)	11457.4	67	171.0057		19148.96	67	285.8053	
Total	90551.2	137			168779.1	137		

from unreliability in the teacher-designed test.

To provide context for these apparent differences and control for selection bias, the identical analysis was repeated for the non-intervention units (B and D; see Table 20). In this case, no intervention was present, therefore the *condition* variable was replaced by *grouping*: the students may perform differently on the semester pre- and posttest scales simply by virtue of being consistently grouped throughout the semester. The grouping (Cohort 1 vs. Cohort 2) did display a statistically significant difference within the instruction for Unit B, but not for Unit D. Therefore, a statistically significant level of variation was possible between the two cohorts on any given unit, without the presence of an intervention. However, the intervention units displayed more significant differences

Table 20

Source Table for Comparisons by Condition on Non-Intervention Units

	Unit B subscale (12 items)				Unit D subscale (10 items)			
Source	SS	dF	MS	F	SS	dF	MS	F
Between								
Group *(Cohort 1 vs. Cohort 2)*	15.912	1	15.912	0.0773 (p > .7)	411.79	1	411.79	2.127 (p > .1)
Subjects (Group)	13795.48	67	205.9027		12969.45	67	193.5739	
Within								
Instruction *(pre vs. post)*	109837.6	1	109837.6	943.7 (p < .01)	146887.9	1	146887.9	680.1 (p < .01)
Instruction * Group	554.192	1	554.192	4.761 (p < .04)	273.894	1	273.894	1.268 (p > .2)
Instruction * Subjects (group)	7798.263	67	116.392		14470.39	67	215.9759	
Total	132001.5	137			175013.4	137		

(i.e., lower *p* values and larger *d* values) and consistently favored movie-making (i.e., $d > 0$ on Unit A and $d < 0$ on Unit C) , while the non-intervention units displayed less significant differences and consistently favored the Honors-track-intensive Cohort 2 ($d < 0$ in both cases).

Semester Pre- and Posttest:
Inconsistent Differences by Condition Within Tracks

Next, the semester subscales were broken out by track level to explore whether the observed effect favoring movie-making was consistent across track levels. First, means, standard deviations, and effect sizes were calculated for the subscales of the intervention units. (See Table 21.) Again, effect size was determined using Cohen's *d*, with the numerator being the difference in group means and the denominator being the

Table 21

Descriptive Statistics Within Track Levels of Semester Pre- and Posttest Subscales for Intervention Units

		Unit A subscale (8 items)					Unit C subscale (9 items)				
		Cond	Pre mean (σ)	Post mean (σ)	Pre-post	*d*	Cond	Pre mean (σ)	Post mean (σ)	Pre-post	*d*
General	Gen1 (mov/pr)	Movie (n=7)	46.4 (21.3)	89.3 (15.2)	42.9	1.1	Pres (n=7)	38.1 (23.0)	92.1 (8.40)	54.0	0.72
	Gen2 (pr/mov)	Pres (n=6)	35.4 (14.6)	72.9 (14.6)	37.5		Movie (n=6)	20.4 (20.4)	70.4 (18.1)	50.0	
TD	TD1 (mov/pr)	Movie (n=15)	35.8 (20.5)	86.7 (12.0)	50.9	6.25	Pres (n=15)	31.9 (15.1)	88.1 (13.6)	56.2	-0.72
	TD2 (pr/mov)	Pres (n=13)	50.0 (15.3)	79.8 (14.0)	29.8		Movie (n=13)	32.5 (19.5)	91.5 (6.66)	59.0	
Honors	Hon1 (mov/pr)	Movie (n=9)	41.7 (17.7)	87.5 (10.8)	45.8	2.9	Pres (n=9)	42.0 (18.2)	95.1 (8.07)	53.1	-6.76
	Hon2 (pr/mov)	Pres (n=19)	49.3 (16.4)	84.9 (11.5)	35.6		Movie (n=19)	17.5 (22.3)	97.7 (4.65)	80.2	

Table 21: Descriptive statistics within track levels of semester pre- and posttest subscales for intervention units. Effect sizes (Cohen's *d*) were calculated using the track's global standard deviation (σ) as the denominator.

global standard deviation (σ) within the track level. On Unit A, the effect sizes were all positive, favoring the movie-making groups, and varied from $d > 1$ (for the General track) to $d > 6$ (for the Talent Development track). On Unit C, the results were mixed: two effect sizes were negative, favoring the movie-making groups, but one (the General track) was positive. In other words, within the General track on Unit C, the presentation-making group out-gained the movie-making group on the semester pre- and posttest scales. Again, effect sizes showed considerable variation, ranging from $-1 < d < 1$ (for the General and Talent Development tracks) to $d > 6$ (for the Honors track).

Table 22

Descriptive Statistics Within Track Levels of Semester Pre- and Posttest Subscales for Non-Intervention Units

		Unit B subscale (12 items)					Unit D subscale (10 items)				
		Cond	Pre mean (σ)	Post mean (σ)	Pre-post	*d*	Cond	Pre mean (σ)	Post mean (σ)	Pre-post	*d*
General	Gen1 (mov/pr)	No proj (n=7)	32.1 (18.3)	88.1 (10.6)	56.0	-2.6	No proj (n=7)	29.9 (23.9)	96.1 (4.86)	66.2	3.9
	Gen2 (pr/mov)	No proj (n=6)	20.8 (11.5)	87.5 (12.6)	66.7		No proj (n=6)	33.3 (16.9)	80.3 (13.4)	47.0	
TD	TD1 (mov/pr)	No proj (n=15)	32.2 (13.7)	85.6 (13.5)	53.3	-1.1	No proj (n=15)	28.5 (11.3)	88.5 (8.74)	60.0	-0.04
	TD2 (pr/mov)	No proj (n=13)	35.9 (9.25)	92.3 (5.34)	56.4		No proj (n=13)	30.8 (15.1)	90.9 (6.43)	60.1	
Honors	Hon1 (mov/pr)	No proj (n=9)	45.4 (20.5)	94.4 (7.22)	49.0	-4.3	No proj (n=9)	29.3 (20.7)	93.9 (4.55)	64.6	-4.5
	Hon2 (pr/mov)	No proj (n=19)	34.2 (14.4)	96.1 (5.10)	61.9		No proj (n=19)	13.9 (19.7)	94.7 (5.52)	80.8	

Table 22: Descriptive statistics within track levels of semester pre- and posttest subscales for non-intervention units. Effect sizes (Cohen's *d*) were calculated using the track's global standard deviation (σ) as the denominator.

To provide context for these differences, the semester pre- and posttest subscales were examined for the non-intervention units. (See Table 22.) Consistent with the earlier observations of larger gains for Cohort 2 (see Table 18), the effect sizes for the non-intervention units consistently favored Cohort 2 across all tracks, excepting Unit D at the General track.

To test the statistical significance of these observed differences at the track level, the semester pre- and posttest scores were subjected to a split-plot full factorial analysis of variance, starting with the intervention units (Units A and C). In contrast to the earlier analysis (see Tables 19 and 20), this analysis included two independent variables: the treatment variable (i.e., the movie-making or presentation-making condition) and the blocking variable of achievement-level grouping (track). The inclusion of the blocking variable reduced the error term. (See Table 23.) With the smaller error term, different levels of significance were observed than in the previous analysis. In the previous analysis, both Units A and C displayed a statistically significant interaction effect in which the group's condition changed its performance between the semester pretest and posttest. After adding track levels into the model, only Unit A maintained the statistically significant interaction effect ($p < .02$) by condition; on Unit C, the condition did not alter the group's performance as a whole ($p > .15$).

Table 23

Source Table for Comparisons by Condition Within Track Levels on Intervention Units

	Unit A subscale (8 items)				Unit C subscale (9 items)			
Source	SS	dF	MS	F	SS	dF	MS	F
Between								
Condition *(mov vs. pr)*	183.092	1	183.092	0.6117 (p > .4)	2663.048	1	2663.048	11.58 (p < .01)
Track	433.034	2	216.517	0.7234 (p > .4)	1050.631	2	525.3155	2.284 (p < .01)
Condition * Track	1473.32	2	736.658	2.461 (p > .5)	2349.852	2	1174.926	5.108 (p < .01)
Subjects (condition * track)	18855.6	63	299.2952		14492.3	63	230.0365	
Within								
Instruction *(pre vs. post)*	47610.2	1	47610.21	269.9 (p < .01)	100684.4	1	100684.4	391.1 (p < .01)
Instruction * Condition	1091.13	1	1091.131	6.185 (p < .02)	537.05	1	537.05	2.086 (p > .15)
Instruction * Track	1.337	2	0.6685	0.0038 (p > .99)	1024.877	2	512.4385	1.990 (p > .1)
Instruction * Condition * Track	334.116	2	167.058	0.9470 (p > .3)	1385.609	2	692.8045	2.691 (p > .05)
Subjects (instruction * condition * track)	11113.35	63	176.4023		16220.57	63	257.4693	
Total	81095.18	137			140408.3	137		

Looking within track levels (i.e., Instruction * Condition * Track), Unit A displayed no statistically significant differences by condition within the track level; the interaction effect introduced by the condition was consistent across all three levels. Even with the reduced error term, all three movie-making groups (General, Talent Development, and Honors) observed the same effect. On Unit C, however, the track levels *approached* statistically significant differences ($.06 > p > .05$) by condition; a glance at Table 21 confirms that the Honors track displayed a statistically significant difference by condition, while the others did not.

Finally, by way of adding further context, the non-intervention units were subjected to the same split-plot analysis with two independent variables: the *condition* variable was replaced by *group*, and the blocking variable of *track* was introduced. (See Table 24.) Again, on Unit B, a statistically significant interaction effect was observed between instruction and grouping ($p < .04$), while Unit D did not display this difference ($p > .8$). Within the track levels, Unit B approached statistical significance ($.06 > p > .05$) and Unit D satisfied statistical significance ($p < .03$). The effect of instruction, accordingly, varied by cohort for Unit B and varied by track level within the cohorts for both non-intervention units.

Table 24

Source Table for Comparisons by Condition Within Track Levels on Non-Intervention Units

Source	Unit B subscale (12 items) SS	dF	MS	F	Unit D subscale (10 items) SS	dF	MS	F
Between								
Group *(Coh 1 vs. Coh 2)*	98.597	1	98.597	0.5547 (*p* > .4)	401.24	1	401.24	2.093 (*p* > .15)
Track	2009.114	2	1004.557	5.652 (*p* < .01)	97.51	2	48.755	0.2543 (*p* > .78)
Group * Track	868.214	2	434.107	2.442 (*p* > .09)	691.019	2	345.5095	1.802 (*p* > .17)
Subjects (group * track)	11197.96	63	177.7453		12080.04	63	191.7467	
Within								
Instruction *(pre vs. post)*	95516.86	1	95516.86	806.611 (*p* < .01)	116338.4	1	116338.4	662.8 (*p* < .01)
Instruction * Group	571.765	1	571.765	4.828 (*p* < .04)	6.861	1	6.861	0.03909 (*p* > .8)
Instruction * Track	199.076	2	99.538	0.8406 (*p* > .4)	1504.194	2	752.097	4.285 (*p* < .02)
Instruction * Group * Track	165.319	2	82.6595	0.6980 (*p* > .5)	1358.651	2	679.3255	3.870 (*p* < .03)
Subjects (instruction * group * track)	7460.303	63	118.4175		11058.56	63	175.5327	
Total	118087.2	137			143536.5	137		

Summary

According to a split-plot, full-factorial analysis of students' performance on the semester pre- and posttest subscales, students who made movies on Unit A out-gained students who made presentations on that content. This effect was consistent across all track levels, and in practical terms resulted in a difference of 1 item on an 8-item scale. When the conditions were reversed for Unit C, the initial analysis suggested the same effect, favoring movie-making over presentation-making. However, once differences within the track levels were examined, this pattern fragmented: the movie-making condition decisively benefited only the Honors track, and in the General track, the presentation-making group actually out-gained the movie-making group.

Analysis of the non-intervention units indicated that statistically significant differences can emerge strictly from the grouping of students: on Unit B, Cohort 2 out-gained Cohort 1, and the same pattern was displayed at the track level. However, on Unit D, no such consistent pattern emerged.

Underlying these analyses is a critical caveat: the coefficients of reliability for these semester pre- and posttest scales were low, ranging from .24 to .58. No firm conclusions can be drawn from tests that display this level of sampling error; any comparison, no matter the apparent statistical or practical significance, should be held tentatively.

However, taking all four units as a whole, the dominant pattern was Cohort 2 out-gaining Cohort 1, with the notable exception of Unit A: on this unit, in which Cohort 1 made movies while Cohort 2 made presentations, the movie-makers out-gained the presentation-makers both at the cohort level and within each track. The intervention, in other words, may have allowed the (otherwise) lower-achieving group to outgain the (otherwise) higher-achieving group.

Finding 3

As defined by use of teacher-selected information and resources, students working with the movie-making application (PrimaryAccess) during the first intervention demonstrated a greater alignment with classroom *instruction and assessment than students working with the presentation-making application (PowerPoint).*

Overview

Concurrent with the observed differences in long-term learning, student work during the end-of-unit projects (creating movies or creating presentations) displayed differences in their relative alignment with classroom instruction and assessment. During the first intervention (Unit A), students' movies contained fewer factual errors (i.e., stated information that contradicted the curriculum framework and/or classroom instruction), covered more content assessed on the semester exam, and used more teacher-selected images and resources than when they were in the presentation-making condition. These differences were weaker or non-existent during the second intervention (Unit C).

Differences in Student Products

As discussed in the previous findings, the analysis of the students' multiple choice and open-ended items revealed no consistent differences by condition on the end-of-unit tests but consistent differences by condition on the semester pre- and posttest subscales. The differences on the semester pre- and posttest subscales were statistically and practically significant in Unit A but not in Unit C. To explore these observed differences on the semester exam, the exam items were used as a coding frame (see Table 11 for an example) to assess students' work on each intervention unit (Units A and C). The text and images included in each project were also analyzed to observe errors or inconsistencies with classroom instruction and to note use of teacher-selected images (i.e., from the list provided to students at the start of each project) and teacher-selected resources (e.g., the provided study guides).

Four differences by condition were observed (see Table 25), with the strongest differences emerging on the first intervention (Unit A). Across the Unit A projects, students' presentations demonstrated a far higher error rate than students' movies, and students' movies incorporated more information featured on the semester exam than the corresponding presentations. On Unit C, both the movies and presentations made few errors and included exam-related information at an equal rate. During both Unit A and Unit C, the movies drew from more teacher-selected materials, while presentations did not.

The movies made during Unit A were more accurate and made fewer errors than the presentations. (See Table 26.) Using the curriculum guide, textbook, teacher handouts, and classroom observation notes, all movies were coded for information that contradicted or was not aligned with instruction. Out of 21 movies, 2 contained factual errors, both regarding the relationship between the Nez Perce rebellion (led by Chief Joseph) and the Sioux rebellion (led by Sitting Bull). Out of 27 presentations, 5 contained errors, and these errors were distributed across all three topic areas (Jim Crow laws, migration to the Great Plains, and cultural conflicts on the Great Plains).

Table 25

Differences in Student Products During Intervention Unit Projects

	Unit A		Unit C	
	Condition	Student Behaviors	Condition	Student Behaviors
Cohort 1	Movie	Featured few errors (2 out of 21 projects) Included more information addressed on the semester exam (approx. one item difference) Used only teacher-selected images Used teacher-selected materials	Presentation	Featured few errors (3 out of 19 projects) Included almost all information addressed on the semester exam Used non-teacher-selected images (more than one per project) Used non-teacher-selected materials
Cohort 2	Presentation	Featured several errors (5 out of 27 projects) Included less information addressed in the semester exam (approx. one item difference) Used non-teacher-selected images (more than one per project) Used non-teacher-selected materials	Movie	Featured few errors (2 out of 23 projects) Included almost all information addressed on the semester exam Used only teacher-selected images Used teacher-selected materials

In addition to making fewer errors, students' movies were more comprehensive than the presentation-making cohort's text. Using the semester exam as a coding frame (see Table 11), movie-making students' scripts addressed more than two-thirds (70%) of the relevant items from the Unit A subscale, while the presentation-making students addressed less than half (42%). Furthermore, the movie scripts included twice as much stem and distractor information than the presentations. For example, two of the seven movies addressing Jim Crows laws discussed the topic in the context of Booker T.

Table 26

Errors in Students' Projects

	Unit A			Unit C		
	Condition (total)	Errors ID	Description	Condition (total)	Errors ID	description
Cohort 1	Movie (21 total produced)	1.2	Displays image of Sitting Bull when discussing Chief Joseph	Presentation (19 total produced)	1.6	Displays images of William Randolph Hearst and the *New York Journal* without ever naming Heart, the *Journal*, Yellow Journalism, or the role of the press in the war
		1.6	Displays image of Little Big horn when discussing Chief Joseph		1.8	Displays image of Harlem when discussing the South
					1.9	Displays image from a Napoleonic battle when discussing the Spanish-American War
Cohort 2	Presentation (27 total produced)	2.5	Uses events of Rosa Parks' arrest in 1955 when narrating circumstances of Homer Plessy's arrest in 1892	Movie (23 total produced)	7.5	Confuses end of the Filipino insurrection (1902) with end of Spanish-American War (1898)
		2.8	Locates the Battle of Little Bighorn in the Black Hills		7.6	Asserts existence of a second Great Migration, going from north to south
		6.3	Includes information about English cattle, but the topic is Texas longhorns			
		6.5	Implies that Booker T. Washington founded the NAACP			
		6.7	Suggests that Chief Joseph fought at Battle of Little Bighorn			

Table 26: Summary of errors in student projects. The differences in numbers of movies between this table and Table 7 reflect the exclusion of ESL students' work.

Washington and W.E.B. DuBois; both movies described the two men's actions (founding the Tuskegee Institute and the NAACP, respectively) and contrasting political stances (equality through vocational education and equality via political and legal action). In contrast, of the eleven presentations about Jim Crow laws, six mentioned either Washington or DuBois—but never both. The net result was that students' movies addressed, on average, one additional item on the Unit A subscale of the semester exam when compared to their peers' presentations.

When conditions were reversed on Unit C, the same pattern was not displayed. Instead, students' work in both conditions addressed the material thoroughly and committed fewer errors. Both movies and presentations were comprehensive, covering approximately 90% of relevant items in the semester exam subscale. Both movies and presentations were accurate, with only 2 out of 23 movies and 3 out of 19 presentations displaying incorrect or non-aligned information.

Some of the differences were driven by the structure of the application. For example, students making movies could not use images that the teacher had not approved. The teacher pre-selected the images that were available to students; adding new images required the teacher to pull them in from the database. One movie-making student, browsing external websites, located a picture she wanted to use, but the teacher declined to make it available. She was unable to insert the image into her movie and had to select teacher-approved image from the provided list. Students making presentations, on the other hand, were able to use not only a page of teacher-selected images but any accessible digital image. Students obtained images either through the clipart feature of the presentation-making program or from non-teacher-designated internet sites. Students located these external sites through general-purpose search engines (e.g., ask.com) and education-specific search engines (factmonster.com).

Due to this structural difference, presentation-making students incorporated non-teacher-selected information and resources at a far higher rate. Only one movie-making project demonstrated use of external resources: a student researched his topic at home and copy-pasted 15 pages of text from various internet sites into his movie script. In contrast, on both Units A and C, all presentations averaged at least one non-teacher-selected image, and several included information drawn from non-teacher-selected resources. In at least one instance, a student used the web site of a hotel in Havana to gather information for his presentation about the destruction of the U.S.S. Maine.

In some instances, presentation-making students incorporated relevant material from external sites. (See Figure 4 for an example.) However, students seeking images to include in their presentations sometimes included images that did not relate to the content, such as flowers, abstract designs, sunsets, and so forth. (See Figure 5 for an example.) For many presentation-making students, this attention to design and layout consumed large sections of their class time:

> Boys in front of me have been quiet, attentive…. Have been doing some intensive work with backgrounds, animations. I'm actually mildly impressed with the sophistication of their PowerPoint usage. … I don't recall seeing them add/edit content, but I've seen at least three iterations of design changes (Classroom observation, September 20, 2006).

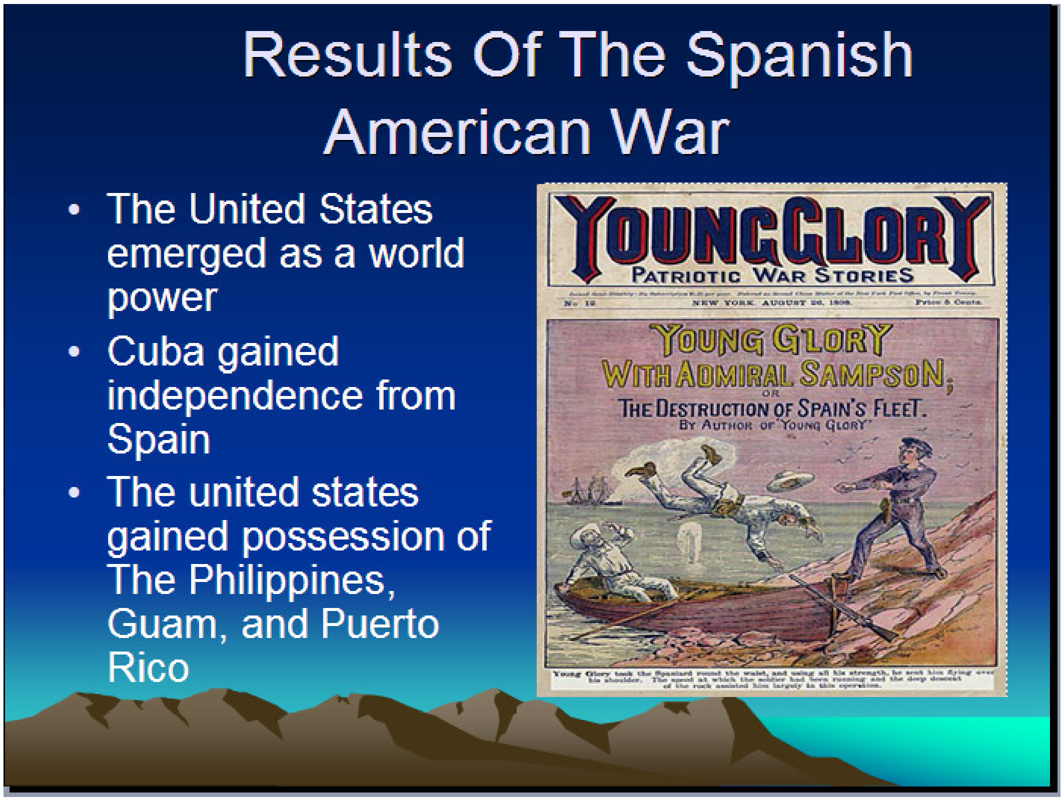


Figure 4: Slide from Presentation Project During Unit C Illustrating Use of Non-Teacher-Selected Images. The image used is not from the original list of teacher-selected images, but it does pertain to the topic (the Spanish-American War), presents information about an event in the conflict (the US naval siege of Santiago, Cuba), and is a primary source illustrating American culture at the time (dime novels, jingoism, views of masculinity, etc.).

In some instances, students' use of non-teacher-selected visuals became the basis for a disruption in class:

> Two boys in middle of room have some clip art…, other students are trying to call attention to it. Woman eating ice cream?? Topic [of presentation] = Spanish American war. Now they're browsing for more clip art...have...a plate of food?? Now a turkey...two turkeys...now a person holding a duck. They show it around the room to elicit laughs from other students (Classroom observation, November 3, 2006).

Other student products demonstrated that non-teacher-selected images could be not only distracting but misleading. One group inserted an image of a Napoleonic battle scene

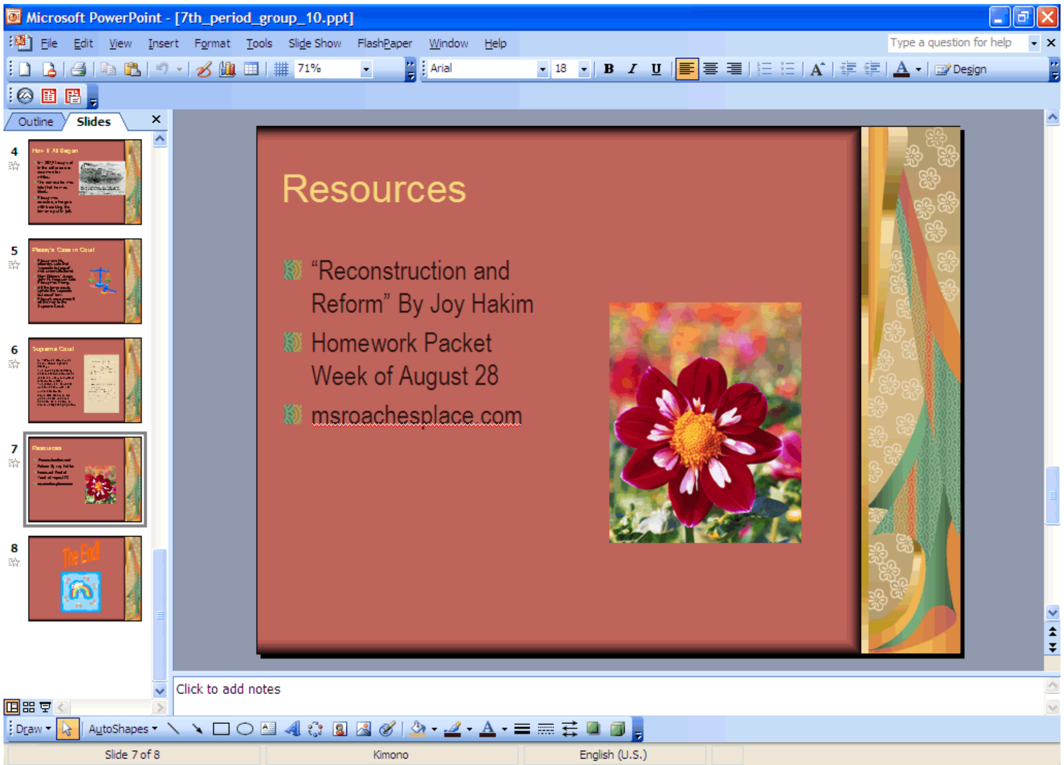


Figure 5: Screenshot of Student Work in PowerPoint During Unit A Illustrating Use of Non-Content Images. Note that the slides visible in the queue on the left demonstrate a mix of content-specific and non-content images.

when discussing the Spanish-American War. (See Figure 6.) Images of the Spanish-American War, and specifically battle scenes from the war, were available in the teacher-selected list, but this group chose to search for a more striking visual. In another instance, students included a non-teacher-selected image of Homer Plessy that contained an error in a critical detail: Plessy was recruited to challenge the Louisiana law requiring segregation because he was 7/8 white and was therefore able to "pass"; the non-teacher-selected image that the students used suggested that Plessy's heritage was predominantly African-American.

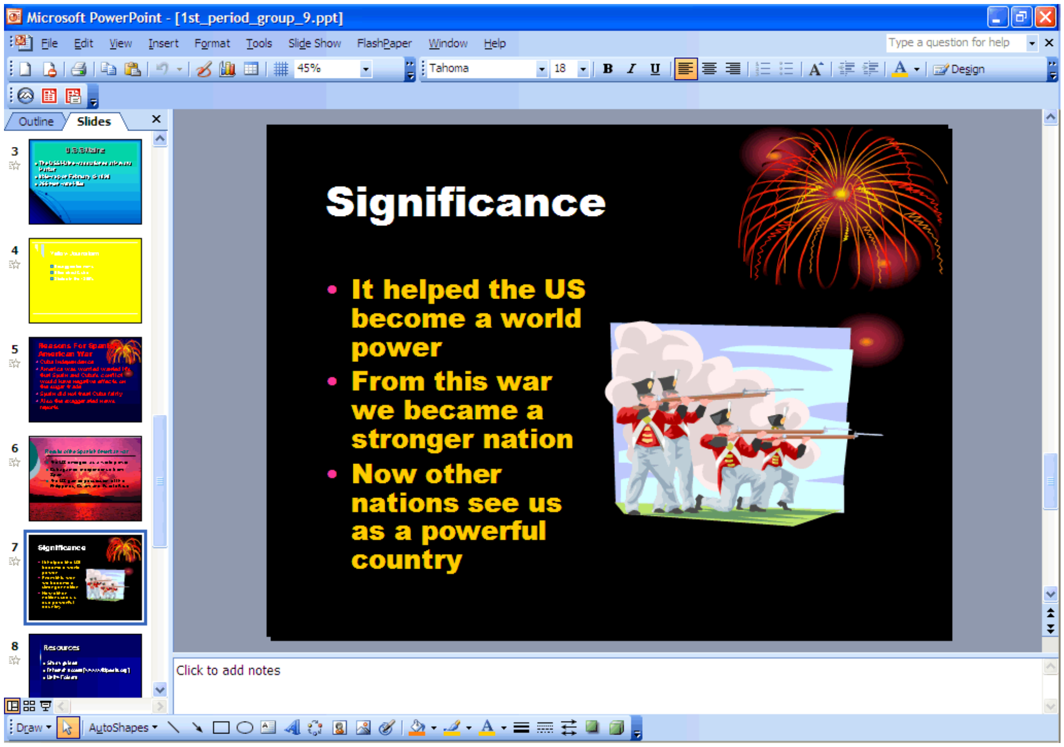


Figure 6: Screenshot of Student Work in PowerPoint During Unit C Illustrating Use of Misleading Images. Note that the slides visible in the queue on the left demonstrate additional use of non-content images.

Second, student use of non-teacher-selected materials was evident in presentation-making students' writing. Several presentations included sentences that did not make sense. Some errors were typographical ("Booker T. Washington believed equality could be achieved through vocal education"), but others were incomprehensible ("As a result many blacks as long as whites were killed in the act of violence and 19th century freedom"). Several of these sentences appeared to result from the use of non-teacher selected materials. A review of classroom observation notes indicated that the author of one such presentation had been browsing an external site (wikipedia.org); a content analysis of his text and the relevant text from the external site revealed that the *entirety* of his text was adapted from this site. (See Table 27.) For the purposes of addressing the content specified in the curriculum guide, the information this student adapted from the

Table 27

Content Analysis of Student Writing in Presentation and Related Passages in Non-Teacher Selected Materials

	Student's text in presentation	Relevant passage from wikipedia
Slide 1: Barbedwires	It was invented by Joseph F. Glidden he was from Dekalb Illinois in 1874.	The most successful barbed wire was patented by Joseph F. Glidden of DeKalb, Illinois in 1874.
Slide 2 (no title)	John Warne Gates demonstraded it in military plaza in San Antonio and Texas in 1876.	John Warne Gates demonstrated barbed wire for Washburn and Moen in Military Plaza, San Antonio, Texas in 1876.
	From enclosure laws in England in the early 18th century.	In the American Southwest barbed wire fencing led to disputes known as the range wars between free-range ranchers and farmers in the late 19th century. These were similar to the disputes which resulted from enclosure laws in England in the early 18th century.
	Barbedwires for aguriutral fence in typicllay available and galvan is for along time.	Barbed wire for agricultural fencing is typically available in two varieties—"soft" or mild-steel wire and "high tensile". Both types are galvanised for long life.
Slide 3: Longhorn cattle	They are a traditional long honed brown and white breed of cattle or in ginating form craven. From the north of England they have that turn down.	Longhorn cattle are a traditional long-horned brown and white breed of cattle originating from Craven in the north of England. They have long horns that turn down, often almost achieving their noses. They have a white patch along the line of their spine and under their bellies.
	There are a breed of cattels in the craven in the north of England.	They are not to be confused with the Texas longhorn breed, which is often called simply "Longhorn cattle".

Table 27: Content analysis of student presentation and wikipedia entries the student used as a source ("Barbed wire" and "Longhorn cattle"). The highlighted text in the last cell is relevant information that the student overlooked: wikipedia associates the term "longhorn cattle" with the English breed and reserves the term "Texas longhorn" to discuss the American breed addressed by the curriculum guide.

non-teacher-selected source was irrelevant and even misleading: the student completely overlooked the fact that the information he sought (i.e., descriptions of American longhorn cattle) was located on a separate page. While this instance was the most striking example of student misuse of non-teacher-selected websites, other presentations demonstrated the same pattern of including information not addressed in the curriculum guide, such as identifying breeds of beef cattle, providing details about the education and military service of George A. Custer, or noting that Native Americans referred to barbed wire as “devil’s rope.” In contrast, only one movie script indicated use of non-teacher-selected materials.

The teacher interview confirmed the difference in student behaviors between the two conditions. Students found the task of creating a movie more challenging than creating a presentation:

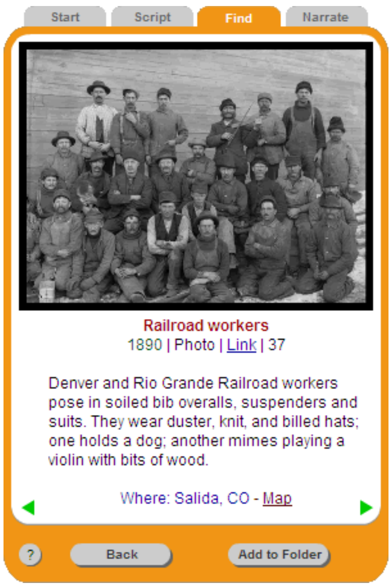
> I had to work harder with PrimaryAccess, that's for sure. [Adopting a student’s voice:] “I don't get what we're supposed to do. I don't understand.” They didn't want to do the extra thinking. It's harder for them. PowerPoint, they can understand the framework. There's three bullets, bam—I can write three short things, and I'm onto the next slide. In PrimaryAccess, they actually had to think and make a smooth transitions, some kind of connections, and also tie the picture in with the movement [pan or zoom]. It's hard for them. I thought, for a lot of the kids, they pushed themselves to do that (Mr. Smith, interview, January 25, 2007).

However, according to Mr. Smith, while movie-making was a more difficult task for the students, they used the structure of the application to spur additional work.

> A lot of times in PrimaryAccess if they got stuck, they would go right to the pictures and say, "What picture do I have that I can use." And they can look at a picture and describe it, and that seemed to trigger what they knew about the topic (Mr. Smith, interview, January 25, 2007).

One example of a movie-making group drawing upon the information provided within the application occurred during Unit A, when students used the contextual information provided with an image in the movie-making application (see Figure 7) to expand their script. The first version of the script merely listed inventions used on the Great Plains. The second version included copy-pasted text from an image of railroad workers. On a later version, the students developed this copy-pasted text into a full thought, providing a transition and putting the concept (exhausting railroad work) into their own words.

Teacher-selected image and contextual information in PrimaryAccess



Successive versions of students' script

Version 1

The great plains was seen as a treeless wasteland.It also had frequent dust storms,low rainfall,and land eroded by wind and water.however because of new technologies people saw the great plains as a vast area to be settled,with many Buffalo

Some of the inventions were,barb wire,steel plows,sod houses,windmills,and railroads.

Version 2

The great plains was seen as a treeless wasteland.It also had frequent dust storms,low rainfall,and land eroded by wind and water.however because of new technologies people saw the great plains as a vast area to be settled,with many Buffalo

Some of the inventions were,barb wire,steel plows,sod houses,windmills,and railroads.
denver and rio grande railroad workers pose in soiled bib overalls

Version 3

The great plains was seen as a treeless wasteland.It also had frequent dust storms,low rainfall,and land eroded by wind and water.however because of new technologies people saw the great plains as a vast area to be settled,with many Buffalo The great plains was in the states of North Dakota,South Dakota,Kansas,Nebraska,Texas

Some of the inventions were, barbed wire which was used to keep people out and animals in. Steel plows were used to rip through the hard ground to plant crops. Sod houses were built with sod because they had no timbers. Windmills were used to generate power to get water from the ground, Railroads were used to bring goods from the east coast to the west coast.

Denver and the Rio Grande were to major railroad companies .Railroad workers had worked for long labor and long periods of time with no food orwater.the great plains was a area that changed the united states today.

Figure 7: Growth of a Movie Script. On the left is an image displayed in the PrimaryAccess resource list for Unit A. Note the contextual information provided beneath the image (title, date, link to additional information, caption, geographical location, and link to a map). To the right are successive versions of the students' script. The first version does not include the contextual information; the second version does include it (highlighted); the third version improves upon the previous version by adding transitions and making appropriate use of the information.

Summary

An analysis of students' projects (movies or presentations) revealed differences in alignment with classroom instruction and assessment for Unit A. Students' movies made fewer errors and covered more material addressed on the exam than corresponding presentations. These differences were not as strong for Unit C, in which both students' movies and presentations made few errors and addressed almost all relevant exam material. On both intervention units, students' movies used teacher-selected images and resources more frequently than corresponding presentations. Examination of classroom observation notes and teacher interview transcripts reinforced these observations and provided context for discussing specific student products.

Finding 4

During the end-of-unit projects, the teacher was able to use the movie-making application (PrimaryAccess) to scaffold students' work; he was unable to use the presentation-making application (PowerPoint) to the same effect.

Overview

The participating teacher displayed different patterns of behavior by condition during the two end-of-unit projects. With the movie-making groups, the teacher used the application (PrimaryAccess) to create a process of iterative refinement, spurred by both synchronous and asynchronous feedback. With the presentation-making groups, he used the application (PowerPoint) to guide students through a linear production process and provided only synchronous feedback. These differences connected to the observed differences in students' end-of-unit project work and provided a portrait of reflective and dynamic use of the movie-making application that changed over the course of the two intervention units. Mr. Smith was able to use the movie-making tool to apply a mix of hard scaffolds (the affordances built into the software itself) and soft scaffolds (teacher feedback) to support and extend students' thinking as they developed their products.

Differences in Teacher's Use of Scaffolds

Following the observation of differences between students by condition (movie-making versus presentation-making) and the differences across implementations (Unit A versus Unit C), the next step was to examine the environmental data to determine what teacher behaviors were connected to these patterns in student behaviors. Again, student products, classroom observation notes, and transcripts of the teacher interviews were reviewed. This analysis provided a clear pattern of differences both between conditions and between implementations of the movie-making project. (See Table 28.)

The teacher used the movie-making application to provide students with a mix of hard and soft scaffolds. The soft scaffolding came in the form of asynchronous (outside of class time) and synchronous (in-class, face-to-face) feedback. Mr. Smith's asynchronous feedback is summarized on Table 28: during Unit A, the teacher left

Table 28

Differences in Teacher Scaffolding of Student Work During Intervention Unit Projects

	Unit A		Unit C	
	Condition of students	Scaffolding provided to students	Condition of students	Scaffolding provided to students
Work with Cohort 1	Movie	Guided students through 3 iterations of script development Provided synchronous and asynchronous feedback (average > than 2.5 notes per project) Provided more content-specific asynchronous feedback (approx. 40% of notes)	Presentation	Guided students through a linear process of adding slides Provided synchronous feedback to students
Work with Cohort 2	Presentation	Guided students through a linear process of adding slides Provided synchronous feedback to students	Movie	Guided students through 2 iterations of script development Provided synchronous and asynchronous feedback (average ≈ 1 note per project) Provided less content-specific asynchronous feedback (approx. 25% of notes) Revised technique for introducing students to application

Table 28: Display of patterns in teacher scaffolding of student work in both conditions of during intervention unit projects.

several notes (mean = 2.78) for each movie; during Unit C, he left just one note (mean = 1.16) for each movie. The notes left during Unit A contained a higher frequency of specific content direction codes. For example, the teacher re-directed student misconceptions (e.g., "Did the railroad only go as far east as Omaha? Look in the book for more info. Pg. 162 of the textbook."), elicited further elaboration of concepts (“What happened at the battle of little bighorn? You are telling me about Crazy horse but dont forget the details about the battle"), and ensured that students addressed the significance of their topic to the present day ("Why are knowing Jim Crow Laws significant to us today? Why are they important?"). Just under 40% of the notes left during the Unit A projects addressed specific content. During Unit C, the rate of content-specific notes was just over 25%. Most of the notes offered generic encouragement ("Good start. Do some research to answer some of your questions") and did not respond to specific points in the students' writing. In addition to providing this asynchronous feedback via embedded text notes, Mr. Smith also provided in-class, face-to-face feedback to students as they worked on their movies.

Mr. Smith commented that logging in and reviewing student work outside of class time was a powerful teaching technique:

> One thing that I thought was terrific about it [the notes feature] was not only could they pull it up and read it and use it as a guide, but it was a help for me, too. As everyone was working, I could go around and say, pull up your note real quick and refresh my memory. There's so many of them [projects], a lot of them didn't have names on them. I didn't have a list of whose group was what [topic]. So I just go through and put the notes and then I can see, OK, pull it open, this is what you need to do. So it was not only good for them to see but it was a refresher for me (Mr. Smith, interview, February 5, 2007).

The written feedback left outside of class was therefore a spur to more structured, more purposeful verbal feedback in class. Furthermore, the asynchronous notes allowed the soft scaffolding (i.e., the teacher feedback) to build upon the hard scaffolds available in the movie-making application, such as the contextual information provided with images. In the case of the student work displayed in Figure 7, the teacher left a note on Version 2 that read, "Do not copy what the information about the picture says. Use the pictures to talk about what you want."

In contrast, the presentation-making students received no hard scaffolding and less soft scaffolding. The presentation-making tool (PowerPoint) was not designed for educational purposes and provided no hard scaffolds; it had no built-in features to support content learners. Presentation-making students had access to the same images as movie-making students, but these images were accompanied by titles alone and not captions, dates, and locations. PowerPoint has no built-in capacity for referencing different versions of a work in progress, so presentation-making students saved all versions of their work into the same file. On more than one occasion, students accidentally over-wrote existing files with blank slides, irretrievably losing their previous work. Finally, PowerPoint has a Notes section than can be used to provide soft scaffolding in the form of teacher feedback. However, the teacher did not have access to the presentation files outside of class, so he had no opportunity to use the Notes area to provide formative feedback. Therefore, the only soft scaffolding presentation-making students received was on-the-spot verbal feedback in class.

The net effect of these differences in teacher actions was to provide more teacher-student interaction about content during movie-making than during presentation-making. The notes placed in students' movies was the clearest example: During both Unit A and Unit C, teacher devoted one hour per day outside of class time reviewing movie-making students' scripts and providing feedback (soft scaffolding) in the form of embedded text notes. (See Figure 2 for an example.) The students making presentations did not receive the benefit of this additional, out-of-class instructional time.

On both intervention units, Mr. Smith used the structure of the movie-making application to guide students through iterative refinement of their work: the task for the first class session was to write one paragraph; the task for the second session was to revise and expand the script to three paragraphs and begin adding images, and so on. Mr. Smith's nightly process of reviewing scripts and adding notes ensured that the second draft was not merely a longer first draft, but a revision of the previous material. (See Figure 7 for an example of successive versions of a movie script.) In contrast, the structure for the presentation-making students during both interventions was to move from slide 1, to slide 2, to slide 3, and so forth. Mr. Smith did not use the structures available in the presentation-making application (slides with sections for notes) to elicit the same process of iterative refinement as with the movie-making application. In some instances, Mr. Smith was able to use in-class feedback to focus students' attention on refining their content:

> Teacher moves around to boys in back. "Whoa-oh-oh-oh! What did I tell you? Oh, you've added one slide ... what are you guys doing?" Teacher now viewing slide show with the students, but drawing their attention to content—points out "Tuskegee", elicits more information:
> Teacher: “What is that?”
> Student: “A college"
> Teacher: “What do they teach you there?”
> Student: "Black stuff?"
> Teacher explains that it's a vocational school. Asks: “What do they teach you there?” No response... Now teacher focuses: “What is a vocational education?”
> Other student: “DuBois learned to be a professor” [Teacher does not react to mention of DuBois during discussion of Tuskegee.]
> Teacher: “I want to see that you've learned something in my class. All right...what did Booker T. Washington do? He founded what?”
> Students: ...institute
> Teacher: “What institute?”
> Student: "Negro institute?"
> Other student: "Tuskegee"
> Teacher: “Did you write that down?” (Classroom observation, September 20, 2006).

Here, the teacher devoted several minutes to clarifying students' understanding and improving their presentation. However, such extended interaction was rare: at any given moment, the teacher typically faced multiple questions from students on a variety of topics:

> Boy in front…has PowerPoint up, but has a question. Boys on far side of room are looking at pics from images list, but have a question. Teacher still talking to students who are not yet on the computer, discussing content, trying to get them organized.... Teacher has to go in [many] directions at once (not to mention put out behavioral fires) (Classroom observation, September 19, 2006).

While working on slides, students simultaneously generated a wide range of questions, spanning both academic content (what to write, where to find information) and technology-related issues (how to add a new slide, how to insert a picture into a slide). In contrast, while in the movie-making condition, Mr. Smith was able to direct students' attention to one task at a time: writing, selecting images, sequencing images, adding motion, and recording the narration.

Mr. Smith was also able to use his experience from the first intervention unit to modify his use of the movie-making application during the second implementation. During Unit A, Mr. Smith used the entire first class period with the movie-making students to demonstrate all of the steps students would need to complete as they made a movie. During Unit C, Mr. Smith introduced the application's features as needed, revealing only what students needed to know in order to complete the current step. For example, on the first day, he showed students how to log in, start their script, and save their work. On the following day, he demonstrated loading the script and reviewing notes. In later sessions he modeled the use of images, and he introduced narration only after at least one group had completed a script. The total technical instructional time spent between the two interventions (Units A and C) was approximately equal, but in the second round the more structured introduction to the application reinforced students' iterative work with the content.

Summary

The differences observed by condition in students' end-of-unit products were tied to relevant differences in the teacher's instruction during the projects: the teacher provided more feedback (soft scaffolding) to movie-making students than to presentation-making students, and was able to use the structure of the application to structure the task and integrate use of hard and soft scaffolds. During movie-making, the teacher guided students through an iterative development process in which each successive version of the script expanded and refined the previous version. During presentation-making, the teacher guided students through a linear development process of adding slides. The teacher modified his instructional about the movie-making application between the first and second interventions.

Finding 5

During the second end-of-unit project (Unit C), technical issues stemming from changes in the school's internet service changed the implementation of the movie-making project and reduced the teacher's and students' use of the movie-making application.

Overview

During the second intervention (the Unit C end-of-unit project), the internet service provider (ISP) for the school district lost traffic from the PrimaryAccess database on two instructional days. During the impacted class periods, the teacher changed the instructional plan for the movie-making groups. After the project concluded, the teacher reported that the loss of access negatively impacted his work with the movie-making application during the second implementation. Quantitative measures of the movies produced during Unit C compared to Unit A correspond with his observation of the impact of the disruption.

Impact of Loss of Connection to the PrimaryAccess Database During Unit C

As described in Chapter Three, the school district lost the ability to access the university servers during the Unit C project work. This disruption took place across two instructional days. All movie-making classes were affected for one or more sessions. The lack of access to university servers had greater consequences for the movie-making students (Cohort 2) than for the presentation-making students (Cohort 1). The movie-making tool (PrimaryAccess) is entirely web-based and is hosted on university servers, but the presentation-making tool (PowerPoint) uses local data and applications. Comparing the movie-making work during Unit A to the movie-making work in Unit C, several differences in student and teacher behaviors correspond to the disruption in access. During the Unit C project, movie-making students wrote less and used fewer images than their peers during Unit A. The teacher also left fewer notes and fewer content-specific notes during Unit C than during Unit A. (see Table 29).

Both teacher report and quantitative measures of students' movies across the two interventions (Unit A and Unit C) correspond to the impact of the disruption of traffic from the PrimaryAccess server. As noted on Table 29, Mr. Smith offered more notes (mean = 2.78) and more content-specific notes (just under 40% of the total) during Unit A. During Unit C, the average number of notes fell to approximately one per movie (mean = 1.16) and the frequency of content-specific notes also dropped (to just over 25%). The teacher attributed this change in quantity and quality of feedback to the disruption in the connection to the PrimaryAccess server: "The internet goes down here [at school]. Specifically, dealing with something like PrimaryAcess, you can't get on the UVA server, [and] if you can't get on...there's downfalls" (Mr. Smith, interview, February 5, 2007). These changes in teacher behaviors across the two interventions correlate with

Table 29

Differences in Implementation of Movie-Making Across the Intervention Unit Projects

	Unit A (Cohort 1)	Unit C (Cohort 2)
Student behaviors	Wrote more (average word count > 200) Used more images (average ≈ 7)	Wrote less (average word count < 175) Used fewer images (average ≈ 6)
Teacher behaviors	Left more notes on students' projects (average > than 2.5 notes per project) Provided more content-specific asynchronous feedback (approx. 40% of notes)	Left fewer notes on students' projects (average ≈ 1 note per project) Provided less content-specific asynchronous feedback (approx. 25% of notes) Changed the movie-making process to compensate for loss of access to the online movie-making application

changes in the movie-making students' writing and image use. As noted on Table 29, the students in the movie-making condition in Unit C used fewer images (average image count = 6.08) and wrote less (average word count = 189) than the students in the movie-making condition during Unit A (average image count = 6.81; average word count = 206). Comparing groups within track levels, the scripts generated within the General and the Honors tracks were 15% shorter during Unit C than during Unit A (General: 144 on Unit A, 115 on Unit C; Honors: 247 on Unit A, 209 on Unit C). Students in the Talent Development sections produced approximately equal-length scripts on both implementations (Unit A: 209 words; Unit C: 211 words). In contrast to the variability in movie-making students' scripts, the presentation-making students wrote the same amount during both projects (Unit A average word count = 173; Unit C average word count = 172).

The disruption in internet access during Unit C elicited further evidence of Mr. Smith's ability to adapt his use of the movie-making application. During the affected class periods, he adjusted by instructing students to compose their scripts in a word-processing program (Microsoft Word) rather than in PrimaryAccess. When the school's ISP corrected their error and restored connectivity between the school district and the university servers, Mr. Smith directed the students to paste these word-processed scripts into the PrimaryAccess script editor and resume work within the online application. Despite the frustrations caused by the loss of access to the PrimaryAccess database, Mr. Smith underscored the value of integrating technology into his teaching:

> You can find other ways to do it. Like doing PrimaryAccess…, when we couldn't get on the server, they could still write about their topic, give a general idea, get with your partner…. The benefits of technology, even when they go wrong

> occasionally, I think, are much greater than those when you have to come up with some kind of backup plan (Mr. Smith, interview, February 5, 2007).

Even with the drawback of occasional technological failure, Mr. Smith valued the technology-based instruction more than the “backup” model of traditional instruction.

Summary

The teacher and students lost access to the online movie-making application on two instructional days. Accordingly, the movie-making projects during Unit C followed a different process and produced different products than during Unit A. During Unit C, movie-making students wrote less and used fewer images than their peers did on the Unit A movies. These changes in student behaviors were matched by changes in teacher behaviors, with the teacher leaving fewer notes and fewer content-specific notes during Unit C than during Unit A. The teacher sustained the movie-making project by shifting student work to a word-processing program during the impacted class periods.

Chapter summary

The data collected in this study revealed mixed results in student content knowledge as measured by teacher-designed end-of-unit and semester tests. On the end-of-unit tests, no clear pattern of difference by condition (movie-making or presentation-making) emerged, either in the quantitative data (multiple choice scores) in the qualitative data (responses to open-ended prompts). However, the semester pre- and posttests revealed a consistent pattern of differences suggesting that students who made movies on a particular topic demonstrated increased long-term (two or more months) learning outcomes. The effect was clearly visible during the first intervention (Unit A) and present, but not statistically significant, during the second intervention (Unit C). Neither effect is considered conclusive, however, since the coefficients for reliability on the pre- and posttest subscales were unacceptably low (i.e., less than .7).

These differences in test scores were mirrored by differences in observed student and teacher behaviors during the end-of-unit projects. During Unit A, movie-making students made fewer errors, addressed more content covered on the semester exam, and hewed to teacher-selected materials. During Unit C, movie-making students and presentation-making students both made fewer errors and covered equal amounts of exam material. However, presentation-making students integrated non-teacher selected images and information during both Unit A and Unit C. The teacher was able to provide movie-making students with a highly-structured, iterative process that integrated hard scaffolds (e.g., contextual information about images) and reflective soft scaffolding (i.e., embedded text notes written outside of class time). The teacher was unable to achieve the same effect with the presentation-making students. They worked in a linear process and received only synchronous, face-to-face feedback.

The movie-making process during Unit C displayed several differences compared the process followed during Unit A. Technical problems centering around loss of access between school district computers and the university servers that hosted the online movie editor (PrimaryAccess) changed the teacher's instruction for all three movie-making classes. Correspondingly, the movie-making students wrote less and used fewer images on Unit C than during Unit A. The teacher also left fewer notes and offered less content-specific feedback during Unit C than during Unit A.

The last chapter in this study includes a discussion of these findings in light of the previous literature. It also includes implications for future research, for history instruction, and for the training and preparation of history teachers.

CHAPTER 5

CONCLUSIONS AND IMPLICATIONS

Introduction

Very few studies have investigated the student outcomes of technology-enriched history instruction. This study investigated the outcomes in one seventh-grade history teacher's six classes as he integrated two competing technological tools: a ubiquitous productivity tool (PowerPoint) and an online digital movie-making application designed for history instruction (PrimaryAccess). Each application was used by half the students during one end-of-unit project; on a subsequent unit the conditions were reversed. After examining students' performance on teacher-designed tests, field notes, document analysis, and teacher interview, the researcher formulated five findings that indicated possible differences by condition (movie-making or presentation-making) in learning and more definite differences by condition in student behaviors, teacher behaviors, and technological infrastructure.

Review of the Findings

The first two findings addressed students' content knowledge outcomes. The analysis of short-term outcomes (i.e., on end-of-unit tests) revealed differences between groups, but these differences were not consistent by condition (movie-making or presentation-making) or within tracks. The analysis of both the quantitative data (multiple-choice scores) and qualitative data (open-ended responses) revealed mixed results by condition at the cohort level and within tracks, and the qualitative data failed to show variation even by topic. Accordingly, the analysis concluded that there were no meaningful differences in short-term learning outcomes.

The second finding addressed students' long-term learning outcomes, that is, the differences between a semester pretest and posttest. This analysis revealed differences on the subscales for both intervention units, with movie-making students performing better than presentation-making students each time. On the first intervention unit, this difference was statistically significant and present within all track levels. On the second intervention, the difference was not statistically significant and was inconsistent within track levels. The coefficients of reliability for the content subscales on the semester pre- and posttest were extremely low, so any observed differences are treated as exploratory and not conclusive.

The third finding described the observed differences in student behaviors during the end-of-unit projects. During the first intervention, students who made movies made fewer factual errors and addressed more exam material than students who made presentations. These differences were not present during the second intervention: both movie-making students and presentation-making students made relatively few errors and addressed almost all relevant topics tested on the exam. On both interventions, presentation-making students accessed non-teacher-selected images and resources, while movie-making students did not.

The fourth finding described the teacher's role during the end-of-unit projects. Because the movie-making application contains built-in hard scaffolds and channels for soft scaffolding, the teacher was able to structure students' work as they made movies. During both implementations of the movie-making project, the teacher had students work iteratively, developing two or more versions of the final product. As part of this process, the teacher left embedded text notes on students' scripts outside of class time. For presentation-making students, the teacher conducted a linear production process and provided only synchronous in-class, face-to-face feedback.

The fifth finding described changes in the implementation of the movie-making projects between the first intervention and the second intervention. During the second intervention, a disruption in access to the movie-making database changed the teacher's instructional process for students making movies. Compared to the first intervention, the teacher provided less feedback to students as they developed their movies. Movie-making students also wrote less and used fewer images during the second intervention compared to their peers during the first intervention.

These findings suggest several conclusions regarding these students' learning of historical content knowledge and the affordances of the two applications used, and the teacher's structuring. These observations are exploratory and must therefore be tentatively held, but they may provide the basis for further research.

Conclusions

As Tufte observed, collected, analyzed data presents the researcher with *la rage de vouloir conclure*—the desire to draw firm conclusions and ignore complexity (2006, p. 154). However, a quasi-experiment, and particularly one conducted using teacher-designed measures of the dependent variable, is not a structure that allows the researcher to speak with authority. Accordingly, the process of knitting together the findings into a conclusion must be approached carefully.

The research questions and findings in this study worked outwards from a narrow frame of students' content knowledge outcomes. Given the observed differences in students' performance on the semester pre- and posttest, the next step was to examine the independent variable in the study (use of the movie-making application), starting first with students' behaviors and then moving to the teacher actions and technological issues that interacted with these student behaviors. By looking at all three levels of assertions (assessment outcomes, student behavior differences, and teacher action and technological infrastructure), a considered judgment can be made regarding what did or did not happen during the course of the interventions.

Conclusions Regarding Content Knowledge

The observed differences in students' content knowledge outcomes by condition (movie-making or presentation-making) cannot be viewed as conclusive but must be viewed as exploratory. Significant and/or practical differences may or may not exist in these students' short-term and long-term content learning outcomes depending upon their project work. First, the teacher-designed tests used to measure the outcome variable (historical content knowledge, specifically that information required for the end-of-year high-stakes test) featured extremely low coefficients of reliability, ranging from .24 to .8. Second, as noted in the literature review, students' performances on tests cannot be interpreted to mean that they know or are able to do the same things, regardless of a test's reliability (Nuthall & Alton-Lee, 1995). Third, as noted in Chapter Three, the teacher reviewed test material immediately before administering the end-of-unit tests, and the open-ended prompts did not align with students' project task. Accordingly, the lack of consistent differences by condition on the end-of-unit tests is not definitive; it may say more about the test or the administration of the test than students' actual content knowledge.

The consistent differences by condition on the semester pre- and posttest subscales are suggestive but not conclusive. First, the subscales' low coefficients of reliability (see Table 9) prohibit aggressive interpretation. Second, while the observed differences by condition in factual errors on students' projects (see Table 26) support the difference on the semester exam subscales, a more powerful piece of disconfirming evidence is the lack of difference in students' end-of-unit tests (see Table 12). If the different project processes (movie-making and presentation-making) do produce differences in long-term learning, why was no effect observed on the short-term?

The teacher's observed end-of-unit test administration procedures may be decisive here: his review of unit content immediately before distributing the assessment may have exerted a far stronger impact on groups' test performance than the project work. As a result, observed differences in students' end-of-unit test scores may emerge from the review, not student learning during instruction or during project work. For example, on the end-of-unit test for Unit A, the largest differences are observed within the Talent Development track, favoring the presentation-making group (see Table 14). Immediately before distributing the test, the teacher conducted a review for both the presentation-making group and the movie-making group. The movie-making group received approximately five minutes of review and the presentation-making group received approximately ten minutes of review. These differences in content review immediately before the completing the test may have contributed to the differences between the groups' performance and either over-stated or under-stated differences between the groups' content knowledge as a result of the unit instruction and project work.

The implementation problems in the end-of-unit tests are confirmed by the lack of differences in the open-ended responses by topic. The fact that students who just completed a week-long project on the Great Migration answer a prompt on the topic no differently than students who completed a project on the Spanish-American War is counter-intuitive, to say the least. A different implementation of the end-of-unit tests, including no review of the content and a different strategy for administering the open-ended prompts, may have yielded more examples of concomitant variation.

The lack of concordance among all three measures (end-of-unit tests, semester exam subscales, and analysis of student products) and the lack of a repeated effect during the reversal of conditions on Unit C prohibits the conclusion that differences exist in student learning outcomes by condition. Students who made movies during Unit A (i.e., Cohort 1) or during Unit C (Cohort 2) may have greater mastery of that material than the students who made presentations (Cohorts 2 and 1, respectively). Two sources of data—performance on semester exam subscales and analysis of student products—support this statement, particularly for the Unit A content. However, this effect was not observed on the end-of-unit tests and was not repeated on the Unit C content. The final judgment, therefore, is that while there is some basis for believing that differences in learning outcomes exist by condition (movie-making or presentation-making), this finding is not conclusive.

Conclusions Regarding the Technologies Used

The differences in students' products (movies or presentations) indicate that the two tools used empower very different behaviors by students and teachers. When using PrimaryAccess to make movies, students wrote more and used more teacher-selected images and resources than when they used PowerPoint to make presentations. On the first intervention, movie-making students also made fewer factual errors and addressed more content covered on the semester exam. On both interventions, the teacher provided movie-making students with asynchronous formative feedback eliciting content elaboration, providing guidance on writing, and offering encouragement. Through this formative feedback, the teacher guided students through an iterative development process, rather than a linear mode of adding new slides. The structure of the movie-making application supported these behaviors, as it limited students to working with teacher-selected images, supported successive version control, and provided the teacher with the ability to view students' work and provide reflective feedback as the project unfolded.

In contrast, PowerPoint did not provide the same support to the teacher's instructional choices. Because PowerPoint accepts a wide variety of digital image formats, students used non-teacher-selected images as well as teacher-selected images. In most cases these images were relevant to their content. In many cases, however, students added images that were not relevant to their content; in other cases, students used images that contradicted the curriculum. Furthermore, PowerPoint does not provide successive version control, encouraging a linear rather than an iterative production process. Finally, because PowerPoint is a local application using local data, the teacher was not able to review students' work outside of class and offer reflective formative feedback as they developed their presentations.

The two tools used provide different spaces for thinking and composition. Both the movie-making and presentation-making projects required students to cluster information: presentation-making students gathered information on a slide, and movie-making students gathered information in a paragraph. Clustering is a technique that aids memory and recall (Miller, 1956), but the act of writing provides a deeper level of encoding: the writer must create structure, and not merely sequence (Greene, 1994). Large-scale studies have repeatedly confirmed a specific connection between writing tasks and student learning outcomes in social studies (Risinger, 1987, 1992; Smith & Niemi, 2001). The use of images provides another avenue for both understanding and recall: students can "read" an image more easily than an abstract representation, such as text. Images are therefore a powerful entry point into a conversation about history (Barton & Levstik, 1996), as well as a mnemonic device. The structure of the movie-making application restricts students to using those images selected by the teacher and therefore focuses students' attention on the highest-priority visuals. The availability of hard scaffolds such as contextual information and timelines assists students' "learning with understanding" (Bransford, Brown, & Cocking, 2000, p. 8) rather than accreting discrete facts. The soft scaffolding provided by teacher feedback provides another support for learning with understanding, as well as an opportunity to respond to students' previous knowledge and elicit student metacognition. In sum, the movie-making application provides a learning environment that is consonant with insights from constructivist models of learning and teaching: the teacher can alter the levels of feedback, access to images, or access to images to allow students to work within their zone of proximal development. For the purposes of history education, PrimaryAccess provides an environment for the "learning-by-making" described by Papert (Harel & Papert, 1991).

The interruption in access to the movie-making database during the second intervention actually provided an opportunity to observe concomitant variation: as students' access to the online movie-making application was reduced, students wrote less and used fewer images. Equivalently, the teacher offered less formative feedback and less content-specific feedback. Despite the disrupted access, movie-making students made few errors in their scripts and addressed the relevant exam content. Because the observed differences on the semester pre- and posttest subscales correlate with the differences in movie-making implementations, the possibility of a relationship between the two events can be entertained. On the first intervention, when the differences in long-term learning outcomes were the strongest, the differences in student products and teacher behaviors were also strongest. On the second intervention, when the differences on the exam subscales were more modest, the differences in student products were also smaller, and the teacher provided less scaffolding to the students during the movie-making process. These parallel variations in student outcomes and implementation of the movie-making projects suggest, but do not prove, that the two may be connected. At the very least, they provide ground for further consideration.

Limitations

This quasi-experiment was not a model of design fidelity. Instead, the deviations from the design (movement of students in and out of the sample, administration of a pretest after instruction, interruption in the experimental treatment during the second intervention) provide confirmation that the classroom is a challenging research environment: complex, dynamic, and ill-structured (Mishra & Koehler, 2006). Because of this fact, classroom-based research captures the authentic problems, such as disruptions in internet access, that teachers and students face on a daily basis. However, any observed or implied result must be confirmed by replication. In hopes of such future replication, this section will highlight the limitations of this study and discuss how the design can be improved upon.

The findings and conclusions from this study cannot be readily applied to other students, other teachers, and other curricula. The curriculum framework used in the Commonwealth of Virginia is highly specific and focuses on content knowledge (Virginia Department of Education, 2001a). The testing practices employed are high-stakes, impacting students and teachers and schools alike, and focus exclusively on content knowledge (Virginia Department of Education, 2002, 2006). The school district in which this study took place faces challenges to improve their performance on these end-of-year high-stakes tests and emphasizes adherence to the curriculum framework and pacing guide. The participating teacher is a capable user of technology and enthusiastically adopts new technological tools. The students participating in this study are tracked into achievement-level groupings. While this curriculum, this school, this teacher, and these students are representative of broader trends in American K-12 education, such as pressure from high-stakes tests, the patterns displayed will not be repeated elsewhere. Instead, only selected aspects of the pattern, such as the role of formative feedback, may be applicable to other curricula, other teachers, and other classrooms.

Generalization is moot until causal validity is established (Krathwohl, 1998). The design of this study can be improved in several ways to strengthen causal inference. First, while the teacher-designed tests did have high content validity, they lacked reliability. Any attempt at replicating this design with teacher-designed tests should use more reliable measures. One way to improve the reliability of the subscales may be as simple conducting an initial review of the test items to eliminate or improve weak items. In the current study, the panel of experts examining content validity flagged one item on the teacher-designed test as being poorly constructed. This item was therefore excluded from the final analysis of groups' performances on the subscale. The impact of excluding this poorly-written item was to raise the subscale's coefficient of reliability from .17 to .31. While .31 is by no means a desirable level of reliability, that single weak item reduced the reliability even further. A review of the construction of test items before administering the assessment may help provide more stable and reliable measures of students' content knowledge.

Another possible improvement in the design is to alter the teacher's test administration procedures. On the end-of-unit multiple choice items, if a review of the unit content is not conducted immediately before distributing the test, differences may be observed by condition that were not discernable in the present study. Additionally, the open-ended prompts can be administered in a manner that elicits more extended student responses.

- The open-ended items can be given before the multiple-choice items, ensuring that students have adequate time to demonstrate what they know.
- The prompts can be printed out and provided to students on paper rather than being projected on an overhead transparency or written on the board.
- The teacher can clarify the questions by discussing them with students, ensuring that those who did not understanding the wording of a question (e.g., "How did people adapt to life in challenging environments?") have the opportunity to grasp what sections of the unit content the question is addressing (i.e., the challenging environment of the Great Plains).

Alternatively, the framing of the task during project work can be modified to align with the open-ended prompts: rather than ask students to present information about the Great Plains, the assigned task can be to answer the question, "How did people adapt to life in challenging environments?" This modification would ensure that students answering a question about their project topic understand the connection between the prompt and specific sections of the unit content.

In addition to being improved, this design can be granulated, providing variation to determine where the greatest teaching and learning benefits lie. For example, the number of images and other resources available to students as they work can also be varied, with some sections drawing upon a large or unrestricted image pool and another group working from a much smaller, constrained list. Additionally, a teacher can provide more formative feedback in one section than another, or more content-specific feedback in one section and more generic feedback in another section. Alternative strategies of structuring the students' writing process can be implemented, such as drawing attention to the construction of storygrams (Polman, 2006) or using the Calibrated Peer Review system (Chapman, 2001). Finally, student interaction with PrimaryAccess outside of class time can be varied. Although very few students in the current study took advantage of the opportunity to log into PrimaryAccess from outside of school and continue work on their projects, other students may do so, especially if encouraged by their teacher. If a teacher were to highlight this possibility and encourage students to work from home or an internet-accessible site in the community, students could extend the amount of time spent working on the project. Their behaviors and outcomes could be compared with students who use of PrimaryAccess was limited to class time only.

The current design can be re-purposed to explore other goals of history education. The current study, responding to the curricular framework and instructional and assessment choices of the participating teacher, focused on students' acquisition of historical content knowledge. However, student creation of multimedia can be directed towards a variety of goals in history education, including but not limited to the development of disciplinary thinking skills, such as those outlined by the National Center for History in the Schools (1994) and the National Council for the Social Studies (1996) or even those specified (but not assessed) by the Virginia Department of Education

(2002). In addition to disciplinary thinking skills, the design of the current study can be applied toward the development of student writing skills, vocabulary acquisition, oral fluency, attitudes regarding education or the discipline of history, or attitudes regarding topics in the history curriculum.

Finally, the projected sequence of classroom-based studies can be supplemented by more carefully controlled laboratory studies to explore the possible causal mechanisms. Assuming that there is a connection between students' movie-making and their learning outcomes, what is the mechanism of the effect? By what channel does the writing, use of images, teacher feedback, and iterative development of the script become an observably different outcome in learning? What implications for transfer of content knowledge lie within the various strategies for implementing presentation-making or movie-making projects? For example, can students whose project required them to analyze primary source images also analyze primary source documents? Do students who are asked to reflect on the epistemological underpinnings of their own historical account recognize the same issues in others' accounts or historically-based arguments?

Relationship of Current Study to Previous Research

This study adds to the small but growing base of research on technology in social studies education, particularly studies that examine student content learning outcomes as a dependent variable. Its findings must therefore be considered in light of the observations of previous researchers.

Kingsley (2005) demonstrated that a multimedia tool (i.e., the Ignite!Learning package of 15 units of media and assessments) can improve students' content knowledge: over seven months, students using the tool raised their scores on a test of content knowledge. However, two qualifying considerations emerge from examining the study. First, the impact on students' knowledge was minimal: in the experimental group, students' scores rose by 12.2%, and the control group students' scores rose by 6.1%. Neither increase is dramatic, or can be described as a desirable outcome following seven months of instruction. In contrast, both the experimental and control groups in this study demonstrated far larger gains across all subscales, from 29.8% to 80.8% (see Table 17). Second, the Ignite!Learning tool was used as a closed environment: at no point did the researcher describe teachers modifying the media and assessments provided within each unit to create a tighter alignment with their own instruction. Teachers also could not interact with the students through the application; interaction was between the students and the multimedia package only. A tool that allows teachers to adapt the content used and provide channels for teacher-student interaction may provide more significant results. For example, both the movie-making (PrimaryAccess) and presentation-making (PowerPoint) tools used in this study allowed the teacher to adapt the project work to the exact content of the unit. The movie-making tool also provided opportunities for the teacher to review student work and provide feedback as the project unfolded. This combination of adaptable content and purposeful interaction between the students and the teacher during project work, combined with the during-unit instruction, may have provided the larger learning effects.

The findings of the current study support the conclusions of Brush and Saye (2002) and Lee and Molebash (2004): the quantity and quality of the scaffolding provided to students during a task may have an impact on the learning outcomes. As Brush and Saye increased the scaffolding (including teacher feedback) available to students during a task, student learning outcomes improved. Lee and Molebash found that certain combinations of scaffolds (i.e., the initial selection of documents and a sourcing heuristic) produced discernibly different results in long-term learning. Similarly, this study found that students with greater access to scaffolds appeared to have greater long-term learning, especially when they received content-specific feedback from the teacher.

Taken together, the current study and those that preceded it (Brush & Saye, 2002; Kingsley, 2005; Lee & Molebash, 2004) provide insights for the use of technology in history education and suggest certain strategies in the preparation and training of history educators.

Implications for Practice

The high-stakes testing regimen in place in the Commonwealth of Virginia provided the context for this study and informed the research questions, design, and data analysis. Given the pressures of this environment, van Hover (2006) asks, “How can we continue to encourage beginning teachers to consider and think about ambitious history teaching within a high-stakes context in which the standards and the end-of-year tests pervade all aspects of teaching?” (p. 216). The description of the participating teacher in this study provides a portrait of one such beginning teacher. The insights gained from working with this teacher and his students can inform others' strategies for coping with coverage, control, and high-stakes testing. Assuming that further research provides more robust results and allows the findings in the present study to be viewed as more than exploratory, several considerations emerge regarding history education and the preparation of history educators.

First, the findings suggest that history teachers can usefully employ open-ended technological tools to meet the pressures of the current high-stakes tested, content-focused history curriculum. State assessments such as the SOL tests are administered at the end of the academic year; topics covered on the test include material studied many months before. Because students, teachers, and schools are judged by their performance on these tests (Virginia Department of Education, 2006), students' long-term retention of understanding and knowledge from this material is critical. Accordingly, teachers often spend weeks, if not months, of instructional time during the run-up to the state-mandated tests reviewing previous instruction instead of introducing new material. As a result, teachers are hard-pressed to cover the required curricular content in the foreshortened instructional segment of the year, and students are denied the opportunity to learn new material during the test-review time. In this study, the implied impact of the technology (PrimaryAccess) and its integration into instruction (the teacher's actions during end-of-unit project work) upon student outcomes (long-term retention of content knowledge) suggests that technology can be used to address these challenges. Students who create rich, content-focused products within an environment that supports and extends student-teacher interaction and student access to content may retain their understandings better than students who do not. Martorella (1997) spoke of technology as "a sleeping giant" in social studies education; the findings of this study may help the giant awaken. Students and teachers can use this movie-making tool—and other purpose-built, content-specific applications—as an effective resource for meeting the challenges of a high-stakes testing environment.

Second, history educators need training and preparation that enables them to make informed choices when selecting technologies to integrate into their instruction. The technologies used in this study (PrimaryAccess and PowerPoint) have very different roles to play in the classroom. While PowerPoint is ubiquitous, it does not naturally afford the scaffolding and student-teacher interaction that appear to be the critical difference in the movie-making students' performance on the end-of-semester assessment. This observation extends to other widely-available technologies, such as generic desktop video editing environments (e.g., iMovie and Windows Movie Maker); if the tool does not scaffold students' work with media and allow outside-of-class teacher feedback, it may not be the best choice for building students' long-term content knowledge outcomes. The critical step is that teachers be able to observe not just what a technology does (e.g., make a movie or make a presentation or provide information) but what teacher and student behaviors it supports (iterative development of an idea, focus on content or message rather than decoration, formative feedback, etc.). The technology must be viewed not as product but as process (Heinich, 1995).

Third, the preparation of history educators must provide insight into high-value teaching and learning behaviors, including some that are not currently part of the teacher preparation process for social studies. While the training of social studies educators typically highlights academic learning time as an important variable in shaping student learning outcomes or underscores the role of recitation and questioning during instruction (Berliner, 1990; Wilen & White, 1991), little or no attention may be given to the process of scaffolding student writing. Student writing is infrequent in history education and favors writing short answers rather than extended reports (Beatty, Reese, Persky, & Carr, 1996; Lapp, Grigg, & Tay-Lim, 2002). History teachers' training does not include attention to developing, structuring, and assessing student writing. Furthermore, if issues such as the selection and sequencing of images and the addition of purposeful motion were also critical to the movie-making students' learning outcomes, then these issues are also absent from the preparation of social studies educators.

Finally, the current study provides a concrete example of a flexible approach to integrating constructivist principles into history education. For teachers who struggle to understand constructivism or apply it to their own classrooms (Marlowe & Page, 2005), PrimaryAccess can provide an entry point and a structure for designing instruction that engages students' prior knowledge, introduces information and conceptual frameworks simultaneously, encourages student production of knowledge, and prompts metacognition.

Summary

The findings in this study are exploratory, not conclusive. Based on the observed patterns in student performance in teacher-designed assessments and students' end-of-unit projects, seventh-grade students who made movies using PrimaryAccess appear to have achieved superior long-term learning outcomes when compared to students who made presentations using PowerPoint, at least on the content addressed during the first intervention. However, the teacher-designed tests used to assess these outcomes suffered from extremely low coefficients of reliability. Therefore, actual differences in student content knowledge outcomes may have been overstated or understated.

A stronger pattern of difference emerged in student and teacher behaviors during the end-of-unit movie-making and presentation-making projects. During the movie-making projects, the teacher gave asynchronous formative feedback and guided students through iterative development of their scripts. Students used only teacher-selected images and were able to build upon provided contextual information about the images. While making movies, students wrote more and, on the first implementation of the movie-making projects, made fewer errors on their scripts and covered more relevant exam material than students working making presentations. During both implementations, students making presentations used non-teacher-selected images and resources and worked in a linear process of adding slides. They received only in-class, face-to-face feedback. Accordingly, the movie-making application used (PrimaryAccess) provided a richer environment for scaffolding student work than the presentation-making application used (PowerPoint).

This study contained many limitations due to weaknesses in design (e.g., the open-ended prompts used to triangulate student content knowledge did not align with the project task given to students) and irregularities during implementation (e.g., loss of connection to the PrimaryAccess server over two instructional days on the second implementation). Future research can improve upon the design and address other goals of history instruction, other content areas within social studies, or cross-curricular goals. A series of more controlled experiments can address the causal mechanism that may exist, connecting students' writing and project work to increases in content knowledge.

This study builds upon previous research on technology in social studies. Specifically, the conclusions regarding content knowledge conform with the work of Lee and Molebash (2004) in highlighting the value of teacher-selected materials. The conclusions regarding students' and teachers' patterns of use of the two applications conform with the work of Brush and Saye (2002) in noting the value added by hard (built into the program) and soft (teacher feedback) scaffolding in improving student outcomes.

Taking into account these findings and those of previous researchers, this study has several implications for practitioners. First, open-ended technological tools such as the movie-making application used (PrimaryAccess) can be used to develop students' content knowledge. However, teachers will need training to develop the ability to select the technologies that best support their instructional goals. Furthermore, the value added by this technology hinges upon the teacher's actions in selecting appropriate images, providing formative feedback, and guiding students through an iterative development process. Accordingly, history teachers will need training in these behaviors, as well as exposure to constructivist models of learning and teaching.

REFERENCES

Amrein, A., & Berliner, D. (2002). High-stakes testing, uncertainty, and student learning. *Educational Policy Analysis Archives*, *10*(18). Retrieved March 1, 2007 from http://epaa.asu.edu/epaa/v10n18/

Angulo, M. (2001). Leveraging learning for Generation I. *School Administrator*, 53(3), 28-31.

Ashby, R. (2005). Students' approaches to validating historical claims. In R. Ashby, P. Gordon, & P. Lee (Eds.), *Understanding history: Recent research in History Education*. London: RoutledgeFalmer.

Ayers, E. (1999). The pasts and futures of digital history. Retrieved July 12, 2005 from http://www.vcdh.virginia.edu/PastsFutures.html

Barber, B. (1992). *An aristocracy of everyone: The politics of education and the future of America*. New York: Ballantine Books.

Barry, S. (2006, January 9). Gang policy at schools questioned. *The Daily Progress*. Retrieved February 4, 2007 from http://www.dailyprogress.com/servlet/Satellite?pagename=CDP/MGArticle/CDP_BasicArticle&c=MGArticle&cid=1128769195179&path=

Barry, S. (2006, March 15). Safety fears at Buford. *The Daily Progress*. Retrieved February 4, 2007 from http://www.dailyprogress.com/servlet/Satellite?pagename=CDP/MGArticle/CDP_BasicArticle&c=MGArticle&cid=1137834725120

Barton, K., & Levstik, L. (1996). "Back when God was around and everything": The development of elementary children's understanding of historical time. *American Educational Research Journal*, 33, 419-454.

Barton, K. (1997a). "Bossed around by the Queen": Elementary students' understanding of individuals and institutions in history. *Journal of Curriculum and Supervision*, 12, 290-314. Retrieved 8/31/2005 from hwwilsonweb.com.

Barton, K. (1997b). "I just kinda know": Elementary students' ideas about historical evidence. *Theory and Research in Social Education*, *25*(4), 407-430.

Barton, K., & Levstik, L. (2004). *Teaching history for the common good*. Mahweh, New Jersey: Lawrence Erlbaum Associates, Publishers.

Beatty, A., Reese, C., Persky, H., & Carr, P. (1996). *NAEP 1994 U.S. History Report Card: Findings from the National Assessment of Educational Progress*. (National Center for Education Statistics Rep. No. NCES 96085). Washington, DC: US Department of Education. Office of Educational Research and Improvement.

Bell. J., & McCollum, D. (1917). A study of the attainments of pupils in United States history. *Journal of Educational Psychology*, *8*(5), 257–74.

Berliner, D. (1990). What's all the fuss about instructional time? In M. Ben-Peretz and R. Bromme (Eds.), *The nature of time in schools: Theoretical concepts, practitioner perceptions* (pp. 3-35). New York & London: Teachers College Press.

Berners-Lee, T. (1999). *Weaving the Web: The original design and ultimate destiny of the World Wide Web*. San Francisco: Harper.

Bolick, C., Berson, M., Coutts, C. & W. Heinecke (2003). Technology applications in social studies teacher education: A survey of social studies methods faculty. *Contemporary Issues in Technology and Teacher Education*, *3*(3), 300-309.

Boote, D., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Education Researcher 34*(6), 3-15.

Booth, M. (1987). Ages and concepts: A critique of the Piagetian approach to history teaching. In C. Portal (Ed.), *The History curriculum for teachers* (pp. 22-38). London, England: The Falmer Press.

Booth, M. (1993). Students' historical thinking and the national History curriculum in England. *Theory and Research in Social Education*, *21*(2), 105-127.

Bork, A. (1980) Interactive learning. In R. Taylor (Ed.), *The computer: Tutor, tool, tutee* (pp. 53-66). New York: Teachers College Press.

Bradley Commission on History in the Schools. (1988). *Building a history curriculum: Guidelines for teaching history in schools*. Westlake, OH: National Council for History Education.

Bransford, J., Brown, A., & Cocking, R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academies Press.

Brophy, J., & Alleman, J. (2000). Primary-grade students' knowledge and thinking about Native American and pioneer homes. *Theory and Research in Social Education*, *28*(1), 96-120.

Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education, 16*, 239-253.

Brush, T., & Saye, J. (2001). The use of embedded scaffolds with a hypermedia-supported student-centered learning activity. *Journal of Educational Multimedia and Hypermedia*, *10*(4), 333-356.

Brush, T., & Saye, J. (2002). A summary of research exploring hard and soft scaffolding for teachers and students using a multimedia supported learning environment. The *Journal of Interactive Online Learning*, *1*(2). Retrieved September 9, 2006 from http://www.ncolr.org/jiol/issues/PDF/1.2.3.pdf

Brush, T., & Saye, J. (2005, April). The effects of multimedia-supported problem-based inquiry on student engagement, empathy, and assumptions about history. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

Cannell, J. (2006). "Lake Wobegon," twenty years later. *Third Education Group Review*, *2*(1). Retrieved April 7, 2007 from http://www.thirdeducationgroup.org/Review/Essays/v2n1.pdf

Cantu, D., & Warren, W. (2003). *Teaching history in the digital classroom*. Armonk, NY: M.E. Sharpe, Inc.

Chapman, O. (2001). Calibrated Peer Review: A writing and critical thinking instructional tool. Retrieved December 19, 2006 from http://cpr.molsci.ucla.edu/cpr/resources/documents/misc/CPR_White_Paper.pdf

Cheney, L. The end of history. (1994, October 20). *The Wall Street Journal*. Retrieved August 22, 2006 from http://www.h-net.org/~teach/threads/standard.html

Cheney, L. (2001, October 5). Remarks at the Dallas Institute of Humanities and Culture. Retrieved September 13, 2006 from http://www.whitehouse.gov/mrscheney/news/20011005.html

Clark, R. (1983). Reconsidering research on learning from media. *Review of Educational Research*, *53*(4), 445-459.

Clarke, W., & Lee, J. (2004). The promise of digital history. *The Clearing House*, *78*(2), 84-87.

Cognition and Technology Group at Vanderbilt (CGTV). (1996). Looking at technology in context: A framework for understanding technology and education research. In D.C. Berliner & R.C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 807-840). New York: Macmillan.

Cole, M. & Wertsch, J. V. (1996). Beyond the individual-social antinomy in discussions of Piaget and Vygotsky. *Human Development*, *39*(5), 250-256.

Coleman, J., Hoffer, T., & Kilgore, S. (1982). Cognitive outcomes in public and private schools. *Sociology of Education*, *55*(2/3), 65-76.

Collins, A. (2006). Cognitive apprenticeship. In R. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 47-60). Cambridge: Cambridge University Press.

Crowe, A. (2004). Teaching by example: Integrating technology into Social Studies education courses. *Journal of Computing in Education*, *20*(4), 159-165.

Cuban, L. (1991). History of teaching in social studies. In P. Shaver (Ed.), *Handbook of research on social studies teaching and learning* (pp. 197-209). New York City, NY: Macmillan Publishing Company.

Cuban, L. (2001). *Oversold and underused: Computers in the classroom*. Cambridge, MA: Harvard University Press.

Dawson, K., Bull, G., & Swain, C. (2000). Considerations for the diffusion of technological innovations in social studies teaching and learning. *Theory and Research in Social Education*, *28*, 587-595.

Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. New York: McMillan.

DeWitt, S. (2004). Computer use in high school social studies classrooms: A comparative case study within social class context. Unpublished doctoral dissertation, University of Wisconsin-Madison.

Diem, R. (2000). Can it make a difference? Technology and the social studies. *Theory and Research In Social Education*, *28*(4), 493-501.

Donovan, M., & Bransford, J. (2005). *How students learn*: *History, mathematics, and science in the classroom*. Washington, DC: National Academies Press.

Doolittle, P., & Hicks, D. (2003). Constructivism as a Theoretical Foundation for the use of technology in Social Studies. *Theory and Research in Social Education*, *31*(1), 72-104.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L., Means, B., Murphy, R., Penuel, W., Javitz, H., Emery, D., & Sussex, W. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort*. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences.

Dynarski, M., Honey, M., & Levin, D. (2002). *Designing a study of the effectiveness of education technology: Background material for the first meeting of the Technical Working Group*. Washington, DC: United States Department of Education.

Ely, D., Foley, A., Freeman, W., & Scheel, N. (1992). Trends in educational technology 1991. In D.P. Ely & B.B. Minor (Eds.), *Educational media and technology yearbook*, *18* (pp. 1-30). Englewood, CA: Libraries Unlimited.

Erickson, F. (1986). Qualitative methods in research on teaching. In M. Wittrock (Ed.), *Handbook of research on teaching*. New York: Macmillan.

Ferster, B., Hammond, T., & Bull, G. (2006). PrimaryAccess: Creating digital documentaries in the social studies classroom. *Social Education*, *70*(3), 147-150.

Friedman, A. (2004). Digital primary source use in World History and World Geography. Unpublished doctoral dissertation, University of Virginia.

Friedman, A. (2006). State standards and digital primary sources: A divergence. *Contemporary Issues in Technology and Teacher Education*, *6*(3). Retrieved September 10, 2006 from http://www.citejournal.org/vol6/iss3/socialstudies/article1.cfm

Friedman, A., & Hicks, D. (2006). The state of the field: Technology, social studies, and teacher education. *Contemporary Issues in Technology and Teacher Education*, *6*(2), 246-258. Retrieved September 8, 2006 from http://www.citejournal.org/articles/v6i2socialstudies1.pdf

Gilbert, D. (2006). *Stumbling on Happiness*. New York: A. A. Knopf.

Goodlad, J. (1984). *A place called school: Prospects for the future*. New York: McGraw-Hill.

Goolsbee, A., & Guryan, J. (2006). The impact of internet subsidies in public schools. *The Review of Economics and Statistics*, *88*(2), 336–347.

Grant, S. (2003). *History lessons: Teaching, learning, and testing in U.S. high school classrooms*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Grant, S. (2006). Measuring history through state-level tests. In S.G. Grant (Ed.), *Measuring History: Cases of state-level testing across the United States* (pp. 303-320). Greenwich, CT: Information Age Publishing.

Grant, S., & Horn, C. (2006). The state of state-level history tests. In S.G. Grant (Ed.), *Measuring History: Cases of state-level testing across the United States* (pp. 9-27). Greenwich, CT: Information Age Publishing.

Greene, S. (1994). Students as authors in the study of history. In G. Leinhardt, I. Beck, & K. Stainton (Eds.), *Teaching and Learning in History* (pp. 133-168). Hillsdale, NJ: Lawrence Erlbaum.

Gutmann, A. (1987). *Democratic education*. Princeton, NJ: Princeton University Press.

Guzzetta, C. (1969). Children's knowledge of historically important Americans. In W. Herman (Ed.), *Current research in elementary school social studies* (pp. 392-399). London: The Macmillan Company.

Haney, W. (2000). The myth of the Texas miracle in education. *Educational Policy Analysis Archives*, *8*(41). Retrieved April 1, 2007 from http://epaa.asu.edu/epaa/v8n41/index.html

Hanson, K. & Carlson, B. (2005). Effective access: Teachers' use of digital resources in STEM teaching. Retrieved June 15, 2005 from http://www2.edc.org/GDI/publications_SR/EffectiveAccessReport.pdf

Harel, I., & Papert, S. (1991). *Constructionism: Research Reports and Essays, 1985-1990*. Norwood, NJ: Ablex.

Hastings, R., & Bham, M. (2003). The relationship between student behaviour patterns and teacher burnout. *School Psychology International*, 24(*1*), 115-127.

Heafner, T. (2004). Using technology to motivate students to learn social studies. *Contemporary Issues in Technology and Teacher Education*, *4*(1), 42-53.

Hedegaard, M. (1990). The zone of proximal development as a basis for instruction. In L. Moll (Ed.), *Vygotsky and education: Instructional implications and applications of sociohistorical psychology* (pp. 349-371). Cambridge: Cambridge University Press.

Heinich, R. (1995). The proper study of instructional technology. In G. Anglin (Ed.), *Instructional technology: Past, present and future* (pp. 61-83). Colorado: Libraries Unlimited.

Hickey, D., Moore, A., & Pellegrino, J. (2001) The motivational and academic consequences of elementary mathematics environments: Do constructivist innovations and reforms make a difference? *American Educational Research Journal*, 38(3), 611-652.

Hicks, D., Doolittle, P., & Lee, J. (2004). Social studies teachers' use of classroom-based and web-based historical primary sources. *Theory and Research in Social Education*, *32*(2), 213-247.

Hill, C. (2006, January 1). Social studies. In *Developing educational standards*. Retrieved August 16, 2006 from http://www.edstandards.org/StSu/Social.html

Hirsch, E.D., Jr. (1995). *The dictionary of cultural literacy: What every American needs to know*. Boston: Houghton Mifflin Company.

Hofer, M., & Owings-Swan, K. (2005). Digital moviemaking—the harmonization of technology, pedagogy and content. *International Journal of Technology in Teaching and Learning*, *1*(2), 102-110.

Horn, C. (2006). The technical realities of measuring history. In S. G. Grant (Ed.), *Measuring History: Cases of state-level testing across the United States* (pp. 57-74). Greenwich, CT: Information Age Publishing.

Jefferson, Thomas. (1787). *Notes on the State of Virginia* (W. Peden, Ed.). New York: W.W. Norton & Company.

K-20 Education Code, Florida Statutes Title XLVIII, Chapter 1003. (2006). Retrieved August 27, 2006 from http://www.leg.state.fl.us/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=Ch1003/SEC42.HTM&Title=-%3E2006-%3ECh1003-%3ESection%2042#1003.42

Keiper, T., Harwood, A., & Larson, B. (2000). Preservice teachers' perceptions of infusing computer technology into social studies instruction. *Theory and Research in Social Education*, *28*(4), 566-579.

Kingsley, K. (2005). A quantitative investigation of American History software on middle school student achievement scores. Unpublished doctoral dissertation, University of Nevada, Las Vegas.

Kozma, R. (1994). A reply: Media and methods. *Educational Technology Research & Development*, *42*(3), 11-14.

Krathwohl, D. (1998). *Methods of educational & social science research: An integrated approach* (2nd ed.). New York: Longman.

Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago, IL: The University of Chicago Press.

Kurfman, D. (1991). Testing as context for social education. In P. Shaver (Ed.), *Handbook of research on social studies teaching and learning* (pp. 310-320). New York City, NY: Macmillan Publishing Company.

Lapp, M., Grigg, W., & Tay-Lim, B. (2002). *The nation's report card: U.S. history 2001* (National Center for Education Statistics Rep. No. 2002-483). Washington, DC: US Department of Education. Office of Educational Research and Improvement.

Lee, J. (2002). Digital history in the History/Social Studies classroom. *The History Teacher*, *35*(4), 503-518.

Lee, J., & Molebash, P. (2004). Outcomes of various scaffolding strategies on student teachers' digital historical inquiries. *Journal of Social Studies Research*, *26*(1), 25-35.

Lenhart, A., & Madden, M. (2005, November 2). *Teen content creators and consumers*. Retrieved November 7, 2005, from http://www.pewinternet.org/pdfs/PIP_Teens_Content_Creation.pdf

Levin, D., & Arafeh, S. (2002) *The digital disconnect: The widening gap between internet-savvy students and their schools*. Retrieved June 21, 2006, from http://www.pewinternet.org/pdfs/PIP_Schools_Internet_Report.pdf

Levstik, L., & Barton, K. (1997). *Doing history: Investigating with children in elementary and middle schools*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Littlefield, J., Declos, V., Lever, S., Clayton, K., Bransford, J., & Franks, J. (1988). Learning LOGO: Method of teaching, transfer of general skills, and attitudes toward school and computers. In R. E. Mayer (Ed.), *Teaching and learning computer programming: Multiple research perspectives* (pp. 111-135). Hillsdale, NJ: Lawrence Erlbaum Associates.

Loewen, J. (1995). *Lies my teacher told me: Everything your American History textbook got wrong*. New York: The New Press.

Lowenthal, D. (2000). Dilemmas and delights of learning history. In P. Stearns, P. Seixas, & S. Wineburg (Eds.), *Knowing, teaching, and learning history* (pp. 63-82). New York: New York University Press.

Lowther, D., Ross, S., & Morrison, G. (2003). When each one has one: The influences on teaching strategies and student achievement of using laptops in the classroom. *Educational Technology Research & Development*, *51*(3), 23-44.

Marlowe, B., & Page, M. (2005). *Creating and sustaining the constructivist classroom* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Martorella, P. (1997). Technology and the social studies—or: Which way to the sleeping giant? *Theory and Research in Social Education*, *25*(4), 511-514.

Mason, C. (1998). An analysis of patterns of computer mediated communication within the social studies student teaching experience. Unpublished doctoral dissertation, North Carolina State University.

Mason, C., & Berson, M. (2000). Computer mediated communication in elementary social studies methods: An examination of students' perceptions and perspectives. *Theory and Research in Social Education*, *28*(4), 527-545.

Mason, C., Berson, M., Diem, R., Hicks, D., Lee, J., & Dralle, T. (2000). Guidelines for using technology to prepare social studies teachers. *Contemporary Issues in Technology and Teacher Education*, *1*(1), 107-113.

McMichael, A., Rosenzweig, R., & O'Malley, M. (1996). Historians and the web: A beginner's guide. *Perspectives* (American Historical Association Newsletter), *34*(January 1996). Retrieved July 13, 2006, from http://www.historians.org/perspectives/issues/1996/9601/9601COM3.CFM

Merryfield, M. (2000). Using electronic technologies to promote equity and cultural diversity in social studies and global education. *Theory and Research in Social Education*, *28*(4), 502-526.

Miller, G. (1956). The magic number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97.

Milman, N., & Heinecke, W. (1999). Technology and constructivist teaching in post-secondary instruction: Using the world wide web in an undergraduate history course. Retrieved September 15, 2006 from http://www.vcdh.virginia.edu/milman_heinecke.html

Milman, N. B., & Heinecke, W. F. (2000). Innovative integration of technology in an undergraduate History course. *Theory and Research in Social Education*, *28*(4), 537-556.

Mishra, P., & Koehler, M. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017-1054.

NAEP U.S. History Consensus Project. (2001). *History framework for the 1994 and 2001 national assessment of educational progress*. Washington, DC: National Assessment Governing Board. Retrieved August 16, 2006 from http://www.nagb.org/pubs/hframework2001.pdf

Nash, G., Crabtree, C., & Dunn, R. (1997). *History on trial: Culture wars and the teaching of the past*. New York: Vintage Books.

Nash, G. (2004). Lynne Cheney's attack on the history standards, 10 years later. *History News Network*. Retrieved September 13, 2006 from http://hnn.us/articles/8418.html

National Center for History in the Schools. (1996). *National standards for history* (Basic edition). Los Angeles, CA: National Center for History in the Schools.

National Council for the Social Studies. (1994). *Expectations of excellence: Curriculum standards for social studies*. Washington, DC: National Council for the Social Studies.

National Council for the Social Studies. (1997). Teacher standards Volume I: Disciplinary standards. In *NCSS national standards for social studies teachers*. Retrieved August 23, 2006 from http://www.socialstudies.org/standards/teachers/vol1/disciplinary/

National Council for the Social Studies. (2001). *Creating effective citizens*. Retrieved August 14, 2006 from http://www.socialstudies.org/positions/effectivecitizens/

Norman, D. (1990). *The design of everyday things.* New York: Doubleday.

Nuthall, G. & Alton-Lee, A. (1995). Assessing classroom learning: How students use their knowledge and experience to answer classroom achievement test questions in science and social studies. *American Education Research Journal*, *32*(1), 185-223.

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

O'Reilly, T. (2005). What is Web 2.0. Retrieved March 1, 2007 from http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html

Paige, R. (2005). *Toward a new golden age in American education: How the internet, the law and today's students are revolutionizing expectations*. (National Education Technology Plan 2004). The US Department of Education: Jessup, M.D.

Papert, S. (1980). Teaching children thinking. In R. Taylor (Ed.), *The computer: Tutor, tool, tutee* (pp. 161-176). New York: Teachers College Press.

Papert, S. (1981). *Mindstorms: Children, computers, and powerful ideas*. New York: Basic Books.

Papert, S. (1993). *The children's machine: Rethinking school in the age of the computer*. New York: Basic Books.

Paxton, R. (2003). Don't know much about history – never did. *Phi Delta Kappan*, *85*(4), 264-273.

Polman, J. (2006). True stories, storied truth: Stitching narrative and logico-scientific discourse together in an age of "spin." Paper presented at International Conference of Learning Sciences, Bloomington, IN.

Rainie, L., & Horrigan, J. (2005). *A decade of adoption: How the internet has woven itself into American life*. Retrieved July 13, 2006, from http://www.pewinternet.org/pdfs/Internet_Status_2005.pdf

Ravitch, D. & Finn, C. (1987). *What do our 17-year-olds know? A report on the first national assessment of history and literature*. New York: Harper & Row, Publishers.

Risinger, C.F. (1987). *Improving writing skills through social studies* (ERIC Digest No. 40). Bloomington, IN: ERIC Clearinghouse for Social Studies/Social Science Education. (ERIC Document Reproduction Service No. ED 285829).

Risinger, C.F. (1992). *Current directions in K-12 Social Studies*. Boston, MA: Houghton Mifflin Co. (ERIC Document Reproduction Service No. ED359130)

Roblyer, M.D. (2005). Educational technology research that makes a difference: Series introduction. *Contemporary Issues in Technology and Teacher Education*, 5(2), 192-201.

Roblyer, M.D., & Knezek, G.A. (2003). New millennium research for educational technology: A call for a national research agenda. *Journal of Research on Technology in Education*, *36*(1), 60-71.

Rogers, E. (2003). *The diffusion of innovations* (5th ed.). New York: Free Press.

Rosenzweig, R., & Thelen, D. (1998). *The presence of the past: Popular uses of history in American life*. New York: Columbia University Press.

Saettler, P. (1990). *The evolution of American educational technology*. Greenwich, CT: Information Age Publishing.

Salinas, C. (2006). Teaching in a high-stakes testing setting: What becomes of teacher knowledge? In S.G. Grant (Ed.), *Measuring History: Cases of state-level testing across the United States* (pp. 177-193). Greenwich, CT: Information Age Publishing.

Sansom, C. (1987). Concepts, skills, and content: A developmental approach to the history syllabus. In C. Portal (Ed.), *The History curriculum for teachers* (pp. 116-141). London, England: The Falmer Press.

Savenye, W., & Robinson, R. (2004). Qualitative research issues and methods: An introduction for educational technologists. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology* (2nd ed.) (pp. 1045-1071). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Saye, J. (1998). Technology in the classroom: The role of dispostions in teacher gatekeeping. *Journal of Curriculum and Supervision*, *13*(3), 210-234.

Saye, J., & Brush, T. (1999). Student engagement with social issues in a multimedia-supported learning environment. *Theory and Research in Social Education*, *27*(4), 472-504.

Saye, J., & Brush, T. (2002). Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments. *Educational Technology Research & Development*, *50*(3), 77-96.

Saye, J., & Brush, T. (2004a). Promoting civic competence through problem-based history learning environments. In G. E. Hamot, and J. J. Patrick, (Eds.), *Civic learning in teacher education* (Volume 3, pp. 123-145). Bloomington, Indiana: ERIC Clearinghouse for Social Studies/Social Science Education.

Saye, J., & Brush, T. (2004b). Scaffolding problem-centered teaching in traditional social studies classrooms. *Theory and Research in Social Education*, *32*(3), 349-378.

Saye, J., & Brush, T. (2006). Comparing teachers' strategies for supporting student inquiry in a problem-based multimedia-enhanced history unit. *Theory and Research in Social Education*, *34*(2), 183-212.

Schramm, L. (1977). *Big media, little media: Tools and technologies for instruction*. Thousand Oaks, CA: Sage Publications.

Seixas, P. (2000). Schweigen! die Kinder! In P. Stearns, P. Seixas, & S. Wineburg (Eds.), *Knowing, teaching, and learning history* (pp. 19-27). New York: New York University Press.

Shadish, W., Cook, T., & Campbell, D. (2002). *Quasi-experimentation: Design & analysis issues for field settings*. Boston, MA: Houghton Mifflin Company.

Shemilt, D. (1987). Adolescent ideas about evidence and methodology in History. In C. Portal (Ed.), *The History curriculum for teachers*. London, England: The Falmer Press.

Shrock, S. (1995). A brief history of instructional development. In G. Anglin (Ed.), *Instructional technology: Past, present and future* (pp. 11-19). Colorado: Libraries Unlimited.

Smith, J., & Niemi, R. (2001). Learning history in school: The impact of course work and instructional practices on achievement. *Theory and Research in Social Education*, *29*(1), 18-42.

Spradley, J. (1980). *Participant observation*. Fort Worth, TX: Harcourt Brace College Publishers.

Staley, D. (2003). *Computers, visualization, and history: How new technology will transform our understanding of the past*. Armonk, NY: M.E. Sharpe, Inc.

Surry, D. W. & Ensminger, D. (2001). What's wrong with media comparison studies? *Educational Technology*, *31*(4), 32-35.

Swan, K. (2004). Examining the use of technology in supporting historical thinking practices in three American History classrooms. Unpublished doctoral dissertation, University of Virginia.

Symcox, L. (2002). *Whose history? The struggle for national standards in American classrooms*. New York: Teachers College Press.

Thomas, W., & Ayers, E. (2003). The differences slavery made: A close analysis of two American communities. *The American Historical Review*, *108*(5), 1299-1307. Retrieved 8 September, 2006, from http://www.vcdh.virginia.edu/AHR

Tufte, E. (2006). *Beautiful evidence*. Cheshire, CT: Graphics Press.

United States Department of Education (n.d.) *Executive Summary*. Retrieved March 1, 2007 from http://www.ed.gov/nclb/overview/intro/execsumm.pdf

van Hover, S. (2006). Teaching history in the Old Dominion: The impact of Virginia's accountability reform on seven secondary beginning History teachers. In S.G. Grant (Ed.), *Measuring History: Cases of state-level testing across the United States* (pp. 195-219). Greenwich, CT: Information Age Publishing.

van Hover, S., & Heinecke, W. (2005). The impact of accountability reform on the "wise practice" of secondary history teachers: The Virginia experience. In E.A. Davis & O.L. Davis, Jr. (Eds.), *Wise social studies teaching in an age of high-stakes testing: Essays on classroom practices and possibilities* (pp. 89-115). Greenwich, CT: Information Age Publishing.

VanSledright, B. (2002). *In search of America's past: Learning to read history in elementary school*. New York: Teachers College Press.

VanSledright, B. (2004). What does it mean to think historically…and how do you teach it? *Social Education*, *68*(3), 230-233.

VanSledright, B., & Limon, M. (2006). Learning and teaching social studies: A review of the cognitive research in history and geography. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (pp. 545-570). Mahwah, NJ: Lawrence Earlbaum Associates, Publishers.

Virginia Department of Education. (n.d.). *1998-2003 statewide Standards of Learning spring assessment results*. Retrieved September 3, 2006 from http://www.pen.k12.va.us/VDOE/src/SOLresults98-03.pdf

Virginia Department of Education. (n.d.). School accreditation ratings description. Retrieved February 4, 2007 from http://www.pen.k12.va.us/VDOE/src/accred-descriptions.shtml

Virginia Department of Education. (2001a). *History and Social Science Standards of Learning curriculum framework: Essential knowledge, skills, and understandings*. Retrieved July 20, 2006 from http://www.pen.k12.va.us/VDOE/Instruction/History/histframework2001.pdf

Virginia Department of Education. (2001b). *History and Social Science Standards of Learning for Virginia public schools*. Retrieved July 20, 2006 from http://www.pen.k12.va.us/VDOE/Superintendent/Sols/historysol2001.pdf

Virginia Department of Education. (2002). *Virginia Standards of Learning assessments for the 2001 history and social science Standards of Learning*. Retrieved February 4, 2007 from http://www.pen.k12.va.us/VDOE/Assessment/HistoryBlueprints03/2002Blueprint4USII.pdf

Virginia Department of Education. (2004). *History and social science Standards of Learning enhanced scope and sequence: United States history: 1877 to present*. Retrieved October 3, 2006 from http://www.pen.k12.va.us/VDOE/EnhancedSandS/histUSII.pdf

Virginia Department of Education. (2006). *State report card*. Retrieved July 30, 2006 from http://pen2.vak12ed.edu/sasweb/rqa/pdf344047.pdf

Wells, J., and Lewis, L. (2006). *Internet access in U.S. public schools and classrooms: 1994–2005* (NCES 2007-020). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Waring, S. (2003). Fourth grade students' experiences during the creation of a technology enriched community history project. Unpublished doctoral dissertation, University of Virginia.

Wertsch, J. (1998). *Mind as action*. New York: Oxford University Press.

Wilen, W., & White, J. (1991). Interaction and discourse in social studies classrooms. In P. Shaver (Ed.), *Handbook of research on social studies teaching and learning* (pp. 483-495). New York City, NY: Macmillan Publishing Company.

Whitaker, S. (2003). The impact of digital images and visual narratives on the ability of fourth graders to engage in historical thinking. Unpublished doctoral dissertation, University of Virginia.

White, C. (1997). Information technology in U.S. Social Studies classrooms: Potentials and realities. Korea-U.S. Symposium, August 8, 1997, EDRS, ED 421415. Retrieved Spetember 9, 2006 from http://eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED421415

Whitworth, S., & Berson, M. (2003). Computer technology in the social studies: An examination of the effectiveness literature (1996-2001). *Contemporary Issues in Technology and Teacher Education*, *2*(4), 472-509.

Wineburg, S. (2001). *Historical thinking and other unnatural acts: Charting the future of teaching the past*. Philadelphia: Temple University Press.

Wineburg, S. (2004). Crazy for history. *The Journal of American History*, *90*(4). Retrieved December 13, 2005 from http://www.historycooperative.org/journals/jah/90.4/wineburg.html

Yeager, E.A. (2005). Introduction: The "wisdom of practice" in the challenging context of standards and high-stakes testing. In E.A. Davis & O.L. Davis, Jr. (Eds.), *Wise social studies teaching in an age of high-stakes testing: Essays on classroom practices and possibilities* (pp. 1-9). Greenwich, CT: Information Age Publishing.

Yeager, E.A., & van Hover, S. (2006). Virginia vs. Florida: Two beginning history teachers' perceptions of the influence of high-stakes tests on their instructional decision-making. *Social Studies Research and Practice*, *1*(3), 340-358.

Zimmerman, J. (2002). Whose America? Culture wars in the public schools. Cambridge, MA: Harvard University Press.

APPENDIX A

INSTRUCTION UNITS, TOPICS, AND RELATED STANDARDS OF LEARNING

Unit A — Rise of Jim Crow laws and African-American responses to discrimination

STANDARD USII.3c
The student will demonstrate knowledge of how life changed after the Civil War by describing racial segregation, the rise of "Jim Crow," and other constraints faced by African Americans in the post-Reconstruction South.

Essential Understandings	*Essential Questions*	*Essential Knowledge*	*Essential Skills*
Discrimination against African Americans continued after Reconstruction. "Jim Crow" laws institutionalized a system of legal segregation. African Americans differed in their responses to discrimination and "Jim Crow."	What is racial segregation? How were African Americans discriminated against? How did African Americans respond to discrimination and "Jim Crow"?	Racial segregation • Based upon race • Directed primarily against African Americans, but other groups also were kept segregated "Jim Crow" laws were passed to discriminate against African Americans. "Jim Crow" laws • Made discrimination practices legal in many communities and states • Were characterized by unequal opportunities in housing, work, education, government African American response • Booker T. Washington—Believed equality could be achieved through vocational education; accepted social separation • W.E.B. Du Bois—Believed in full political, civil, and social rights for African Americans	Analyze and interpret primary and secondary source documents to increase understanding of events and life in United States history. (USII.1a) Make connections between past and present. (USII.1b) Sequence events in United States history. (USII.1c) Interpret ideas and events from different historical perspectives. (USII.1d)

Migration to the Great Plains

STANDARD USII.2a
The student will use maps, globes, photographs, pictures, and tables for explaining how physical features and climate influenced the movement of people westward.

Essential Understandings	*Essential Questions*	*Essential Knowledge*	*Essential Skills*
During the nineteenth century, people's perceptions and use of the Great Plains changed. Technological advances allowed people to live in more challenging environments.	How did people's perceptions and use of the Great Plains change after the Civil War? How did people adapt to life in challenging environments?	Physical features/climate of the Great Plains • Flatlands that rise gradually from east to west • Land eroded by wind and water • Low rainfall • Frequent dust storms Because of new technologies, people saw the Great Plains not as a "treeless wasteland" but as a vast area to be settled. Inventions/adaptations • Barbed wire • Steel plows • Dry farming • Sod houses • Beef cattle raising • Wheat farming • Windmills • Railroads	Analyze and interpret primary and secondary source documents to increase understanding of events and life in United States history. (USII.1a) Analyze and interpret maps that include major physical features. (USII.1f)

STANDARD USII.3a
The student will demonstrate knowledge of how life changed after the Civil War by identifying the reasons for westward expansion.

Essential Understandings	*Essential Questions*	*Essential Knowledge*	*Essential Skills*
New opportunities and technological advances led to westward migration following the Civil War.	Why did westward expansion occur?	Reasons for westward expansion • Opportunities for land ownership • Technological advances, including the Transcontinental Railroad • Possibility of wealth created by the discovery of gold and silver • Adventure • A new beginning for former slaves	Interpret ideas and events from different historical perspectives. (USII.1d)

Cultural conflicts on the Great Plains

(Selected sections of Standard USII.3b–see below)

Unit B

Immigration

STANDARD USII.3b
The student will demonstrate knowledge of how life changed after the Civil War by explaining the reasons for the increase in immigration, growth of cities, new inventions, and challenges arising from this expansion.

Essential Understandings	*Essential Questions*	*Essential Knowledge*	*Essential Skills*
Population changes, growth of cities, and new inventions produced interaction and often conflict between different cultural groups. Population changes, growth of cities, and new inventions produced problems in urban areas. Inventions had both positive and negative effects on society.	Why did immigration increase? Why did cities develop? What inventions created great change and industrial growth in the United States? What challenges faced Americans as a result of those social and technological changes?	Reasons for increased immigration • Hope for better opportunities • Religious freedom • Escape from oppressive governments • Adventure Reasons why cities developed • Specialized industries including steel (Pittsburgh), meat packing (Chicago) • Immigration from other countries • Movement of Americans from rural to urban areas for job opportunities Inventions that contributed to great change and industrial growth • Lighting and mechanical uses of electricity (Thomas Edison) • Telephone service (Alexander Graham Bell) Rapid industrialization and urbanization led to overcrowded immigrant neighborhoods and tenements. Efforts to solve immigration problems • Settlement houses, such as Hull House founded by Jane Addams • Political machines that gained power by attending to the needs of new immigrants (e.g., jobs, housing) Interaction and conflict between different cultural groups • Indian policies and wars – Reservations – Battle of Little Bighorn – Chief Joseph • Discrimination against immigrants: Chinese, Irish Challenges faced by cities • Tenements and ghettos • Political corruption (political machines)	Make connections between past and present. (USII.1b) Sequence events in United States history. (USII.1c) Interpret ideas and events from different historical perspectives. (USII.1d) Analyze and interpret maps that include major physical features. (USII.1f)

Industrialization

STANDARD USII.3d
The student will demonstrate knowledge of how life changed after the Civil War by explaining the rise of big business, the growth of industry, and life on American farms.

Essential Understandings	Essential Questions	Essential Knowledge	Essential Skills
Between the Civil War and World War I, the United States was transformed from an agricultural to an industrial nation.	What created the rise in big business? What factors caused the growth of industry? How did industrialization and the rise in big business influence life on American farms?	Reasons for rise and prosperity of big business • National markets created by transportation advances • Captains of industry (John D. Rockefeller, oil; Andrew Carnegie, steel; Henry Ford, automobile) • Advertising • Lower-cost production Factors resulting in growth of industry • Access to raw materials and energy • Availability of work force • Inventions • Financial resources Examples of big business • Railroads • Oil • Steel Postwar changes in farm and city life • Mechanization (e.g., the reaper) had reduced farm labor needs and increased production. • Industrial development in cities created increased labor needs. • Industrialization provided access to consumer goods (e.g., mail order).	Make connections between past and present. (USII.1b) Sequence events in United States history. (USII.1c) Analyze and interpret maps that include major physical features. (USII.1f)

Unit C Spanish-American War

STANDARD USII.4a
The student will demonstrate knowledge of the changing role of the United States from the late nineteenth century through World War I by explaining the reasons for and results of the Spanish American War.

Essential Understandings	*Essential Questions*	*Essential Knowledge*	*Essential Skills*
The United States emerged as a world power as a result of victory over Spain in the Spanish American War. Economic interests and public opinion often influence U.S. involvement in international affairs.	What were the reasons for the Spanish American War? What were the results of the Spanish American War?	Reasons for the Spanish American War • Protection of American business interests in Cuba • American support of Cuban rebels to gain independence from Spain • Rising tensions as a result of the sinking of the U.S.S. Maine in Havana Harbor • Exaggerated news reports of events (Yellow Journalism) Results of the Spanish American War • The United States emerged as a world power. • Cuba gained independence from Spain. • The United States gained possession of the Philippines, Guam, and Puerto Rico.	Analyze and interpret primary and secondary source documents to increase understanding of events and life in United States history. (USII.1a) Sequence events in United States history. (USII.1c)

Life in early 20th century America

STANDARD USII.5a
The student will demonstrate knowledge of the social, economic, and technological changes of the early twentieth century by explaining how developments in transportation (including the use of the automobile), communication, and electrification changed American life.

Essential Understandings
Technology extended progress into all areas of American life, including neglected rural areas.

Essential Questions
How was social and economic life in the early twentieth century different from that of the late nineteenth century?

Essential Knowledge
Results of improved transportation brought by affordable automobiles
• Greater mobility
• Creation of jobs
• Growth of transportation-related industries (road construction, oil, steel, automobile)
• Movement to suburban areas

Invention of the airplane
• The Wright brothers

Use of the assembly line
• Henry Ford

Communication changes
• Increased availability of telephones
• Development of the radio (role of Guglielmo Marconi) and broadcast industry (role of David Sarnoff)
• Development of the movies

Ways electrification changed American life
• Labor-saving products (e.g., washing machines, electric stoves, water pumps)
• Electric lighting
• Entertainment (e.g., radio)
• Improved communications

Essential Skills
Make connections between past and present. (USII.1b)
Interpret ideas and events. (USII.1d)

Great Migration

(Selected sections of Standard USII.5b–see below)

Unit D World War I

STANDARD USII.4b
The student will demonstrate knowledge of the changing role of the United States from the late nineteenth century through World War I by explaining the reasons for the United States' involvement in World War I and its leadership role at the conclusion of the war.

Essential Understandings
The United States involvement in World War I ended a long tradition of avoiding involvement in European conflicts and set the stage for the United States to emerge as a global superpower later in the 20th century.
There were disagreements about the extent to which the United States should isolate itself from world affairs.

Essential Questions
What were the reasons for the United States becoming involved in World War I?
Who were the Allies?
Who were the Central Powers?
In what ways did the United States provide leadership at the conclusion of the war?

Essential Knowledge
Reasons for U.S. involvement in war
• Inability to remain neutral
• German submarine warfare—sinking of Lusitania
• U.S. economic and political ties to Great Britain

Allies
• Great Britain
• France
• Russia
• Serbia
• Belgium

Central Powers
• Germany
• Austria-Hungary
• Bulgaria
• Ottoman Empire

U.S. leadership as the war ended
• At the end of World War I, President Woodrow Wilson prepared a peace plan that called for the formation of the League of Nations, a peace-keeping organization.
• The United States decided not to join the League of Nations.

Essential Skills
Analyze and interpret primary and secondary source documents to increase understanding of events and life in United States history. (USII.1a)
Sequence events in United States history. (USII.1c)
Interpret ideas and events from different historical perspectives. (USII.1d)

Progressivism

STANDARD USII.5b
The student will demonstrate knowledge of the social, economic, and technological changes of the early twentieth century by describing the social changes that took place, including Prohibition, and the Great Migration north.

Essential Understandings
Reforms in the early twentieth century could not legislate how people behaved.
Economic conditions and violence led to the migration of people.

Essential Questions
What was Prohibition, and how effective was it?
Why did African Americans migrate to northern cities?

Essential Knowledge
Prohibition was imposed by a constitutional amendment that made it illegal to manufacture, transport, and sell alcoholic beverages.

Results of Prohibition
• Speakeasies were created as places for people to drink alcoholic beverages.
• Bootleggers smuggled illegal alcohol and promoted organized crime.

Great Migration north
• Jobs for African Americans in the South were scarce and low paying.
• African Americans faced discrimination and violence in the South.
• African Americans moved to northern cities in search of better employment opportunities.
• African Americans also faced discrimination and violence in the North.

Essential Skills
Interpret ideas and events from different historical perspectives. (USII.1d)
Analyze and interpret maps that include major physical features. (USII.1f)

Harlem Renaissance

STANDARD USII.5c
The student will demonstrate knowledge of the social, economic, and technological changes of the early twentieth century by examining art, literature, and music from the 1920s and 1930s, emphasizing Langston Hughes, Duke Ellington, and Georgia O'Keeffe and including the Harlem Renaissance.

Essential Understandings
The 1920s and 1930s were important decades for American art, literature, and music.
The leaders of the Harlem Renaissance drew upon the heritage of black culture to establish themselves as powerful forces for cultural change.

Essential Questions
Who were the leaders in art, literature, and music?
What were their contributions?
How did the Harlem Renaissance influence American life?

Essential Knowledge
Cultural climate of the 1920s and 1930s
• Art—Georgia O'Keeffe, an artist known for urban scenes and, later, paintings of the Southwest
• Literature—F. Scott Fitzgerald, a novelist who wrote about the Jazz Age of the 1920s; John Steinbeck, a novelist who portrayed the strength of poor migrant workers during the 1930s
• Music—Aaron Copland and George Gershwin, composers who wrote uniquely American music

Harlem Renaissance
African American artists, writers, and musicians based in Harlem revealed the freshness and variety of African American culture.
• Art—Jacob Lawrence, painter who chronicled the experiences of the Great Migration north through art
• Literature—Langston Hughes, poet who combined the experiences of African and American cultural roots
• Music—Duke Ellington and Louis Armstrong, jazz composers; Bessie Smith, blues singer
Popularity of these artists spread to the rest of society.

Essential Skills
Analyze and interpret primary and secondary source documents to increase understanding of events and life in United States history. (USII.1a)
Sequence events in United States history. (USII.1c)
Interpret ideas and events from different historical perspectives. (USII.1d)

APPENDIX B

INTERVIEW PROTOCOL

-Questions for interview with teacher-

General
How long have you been teaching?

How long have you been teaching US history, 1877 to the present?

On an average day, how much planning time do you have?

Professional Training
What is your teacher education background?

What is the highest academic degree that you have been awarded?

What subject is it in?

Teaching Practice
What unit are you working on right now?

Please describe the instructional methods you are using to teach this unit.

How do you assess student learning?

Is this typical of how you teach different units?

Why do you teach in this manner?

SOLs

What type of learning do you think the SOLs represent?

Do you feel any pressure from the SOLs? Please describe.

How does this pressure affect your teaching?

If the SOLs didn't exist, would you teach the same way? If not, please describe how you would teach.

Historical Content Knowledge

How do you define historical content knowledge?

Do you think your students are capable of acquiring historical content knowledge?

If so, what are some methods that you use to encourage your students to acquire historical content knowledge?

Is this idea of encouraging your students to acquire historical content knowledge consistent with the SOLs? Why?

What are some examples of student work that demonstrate historical content knowledge?

Students' PrimaryAccess and PowerPoint use

What are your feelings about using PrimaryAccess and PowerPoint while teaching US history?

In your own words, briefly describe the activity, as if to a fellow US history teacher.

What types of media (newspapers, diaries, images) did you use in the projects?

How did the students respond to these activities?

Name one thing the students did or said during the PrimaryAccess or PowerPoint activities that was a surprise to you, that was unexpected.

What elements would you have liked to have incorporated in the projects but were unable to?

Which project was more effective? Why?

Do you integrate similar projects into your instruction?

> If so, where do these projects come from?
>
> How do you design your implementation of them?
>
> Would you plan more projects if you had the time and materials?

Do you think about how you might use PowerPoint or PrimaryAccess when you plan your lessons?

Are the X minutes (from above) of planning time enough for you to plan lessons/units that include these projects?

Have you noticed a difference in students' test scores (your tests) when you use these projects?

Have you noticed a difference in students' SOL scores when you use these projects?

Are there any limitations or barriers you face in terms of integrating projects into your curriculum?

Are there any factors in place here at ____(MS) that encourage you to integrate projects into your curriculum?

Technology
In the past few years, there has been a push to integrate technology into social studies. How do you feel about this?

Has it affected your teaching? How?

How comfortable do you feel using technology?

Do you think teaching with technology is better than more traditional methods? Why?

Do you think your students can learn from using technology? How?

If a technology user-What type of learning outcomes do you generally see when you teach technology-rich lessons? Is this different?

Is the learning that the SOLs represent (from above) consistent with using technology?

You said you face _____ pressure from the SOLs. Does this pressure affect your use of technology? How?

Given an unlimited budget/time how much technology would you use? Why?
Please describe your classroom in terms of technology.

Do you have a projector?

What could be done to improve your classroom environment in terms of technology?

What would be an ideal classroom?

What is the procedure for using a classroom set of computers?

Is this effective? How could it be improved?

Are there any general limitations or barriers you face in terms of integrating technology into your curriculum?

Are there any factors in place that encourage you to integrate technology into your curriculum?

Possible barriers to technology use
Time

How much time does it take to prepare a lesson that uses technology?

You said you had _____ minutes of planning time. (If there a difference-how do you account for it)?

If you had more time, would you integrate technology more into your teaching?

Does finding appropriate technology and thinking of ways to use it in your instruction take extra planning time?

When do you find this time?

Do you feel that it's worth the time you spend/have spent?

Training
Have you been trained to use technology in your classroom? If so, how?

When was this training offered? By whom?

What was the best (or most useful) thing you learned in this training?

What was the least useful thing?

How would you rate the training overall?

Are there any technology skills that you would like to learn? Which? Why?

Support
Do you have a computer in your classroom? If so, how many? Do they have an Internet connection?

How dependable are the computers and Internet connection at your school?

If your computer wasn't working, what would you do?

Please describe the support you receive regarding technology.

Is there anybody you can consult with that can help you integrate technology into specific lessons/units?

If so, please describe your interactions with them.

If not, how do you think you would make use of this type of person?

Digital media

Do you use web sites as resources in your instruction?

Which sites?

How did you find out about these sites?

Are some sites easier to use than others?

Why? What makes them easier?

How do you use them? (student-centered, teacher-centered, print out handouts, images)

Are they easy to navigate?

What do you think of these sites?

Do they have factually accurate information?

What are some problems students typically encounter?

Finding information?

Understanding words?

Please describe a typical lesson involving a website.
What are students doing?

Do you find students are more engaged when you teach with digital primary sources?
Which type of digital primary source do students seem to like the best?

Do you think they learn more when they use this digital primary source?

Are their test scores improved?

Are their SOL scores improved?

Are there any limitations or barriers you face in terms of integrating digital primary sources into your curriculum?

Are there any factors in place that encourage you to integrate digital primary sources into your curriculum?

Is there anything else you would like to say about media, technology, or digital media?