Supporting Social Studies Reading Comprehension with an Electronic Pop-up Dictionary

Sara Winstead Fry
Bucknell University

Ross Gosky
Appalachian State University

Abstract

This study investigated how middle school students' comprehension was impacted by reading social studies texts online with a pop-up dictionary function for every word in the text. A quantitative counterbalance design was used to determine how 129 middle school students' reading comprehension test scores for the pop-up dictionary reading differed from test scores for reading hard-copy texts or an online text without the dictionary. The pop-up dictionary reading was shown to be a statistically effective method for improving student test scores. The results suggest pop-up dictionaries may provide a helpful intervention for increasing middle-level learners' reading comprehension. (Keywords: reading comprehension, electronic dictionaries, content area reading.)

Although America's schools face strong pressures to improve standardized test scores (Lewis, 2005), expectations for social studies education were excluded from *No Child Left Behind* legislation. Thus, as school districts endeavor to improve scores in reading and mathematics, they often reduce or eliminate time previously allocated to social studies in elementary schools (Haskvitz, 2006; Risinger, 2006). Since middle school students face increasingly demanding reading as they encounter sophisticated content area texts (Beers, 2002; Ciardiello, 2002), a lack of experience with social studies texts may compound the challenges faced by struggling middle school readers. The stakes are high; large percentages of students who struggle with reading at the end of eighth grade do not complete high school (Papalewis, 2004).

The present study was designed in response to these pressures. Because of the potential for technology-based interventions to offer effective ways to improve students' literacy skills (MacArthur, Ferretti, Okolo, & Cavalier, 2001) and the "need for more and better research on education technology" (Schneiderman, 2004, p. 30), this investigation determined how an electronic pop-up dictionary with middle school level definitions for every word in the text impacted student reading comprehension of social studies texts. One text was read in a hard-copy text format, one was read online, and the third was read online with a pop-up dictionary function for every word in the text. We hypothesized that comprehension would be higher for the pop-up dictionary reading because students could easily check the meanings of words they did not know.

LITERATURE REVIEW

Social Studies, Literacy, and Educational Technology

Social studies textbooks have long been recognized as hard to read because of high reading levels, long sentences, technical vocabulary, too much content, and the use of primary source material (Johnson, 1977). Textbooks rely on expository writing and are particularly difficult for poor readers in the middle grades. Authors of the textbooks make strong cognitive demands on middle school readers—who are in the process of adjusting to the unfamiliar expository format since narrative texts are more common in elementary school. When middle school social studies teachers do not provide reading comprehension as part of their instruction, poor readers are left on their own to make sense of complicated textbooks (Ciardiello, 2002). In response to the reading challenges posed by middle and secondary textbooks, the International Reading Association recommended reading comprehension instruction in all content areas (Moore, Bean, Birdyshaw, & Rycik, 2000).

As a result of challenges with social studies texts, literature describing instructional methods to help improve reading comprehension for middle school social studies students is abundant. For example, Katims and Harmon (2000) used a cognitive strategy called PEP (Person, Event, Place) to support low-achieving and learning disabled middle school students' when reading social studies content. Pre- and post-test scores on reading comprehension tests suggested the PEP technique contributed to significant gains in comprehending social studies texts. Students also took more organized, more focused, and more detailed notes about the text when they used the PEP technique. Massey and Haefner (2004) described how middle school social studies teachers can scaffold adolescent literacy in social studies by helping students 1) establish a purpose for reading, 2) make connections to prior knowledge, 3) use graphic organizers during reading to make sense of complicated text arrangements, 4) make connections between texts, 5) use post-reading questions as a way to further comprehension, and 6) synthesize connections between texts. Ciardiello (2002) developed and recommended a visual learning tool called "Question Networks" that helps middle school students understand cause and effect relationships. Like Massey and Haefner and Katims and Harmon, Ciardiello sought to provide a scaffold to help students understand complicated middle and secondary level textbooks.

Although the response to Moore et al.'s (2000) call for reading comprehension instruction in middle and secondary level content areas includes specific instructional methods for use in social studies, literature about the use of educational technology to support these instructional methods appears underdeveloped. However, the use of education technology to support other social studies learning in K–12 classrooms is widely represented in the literature. For example, Tally and Goldenberg (2005) effectively used digital primary sources to help middle and high school students develop historical literacy skills, specifically document analysis, observation, making inferences, and corroboration. In a study of fourth grade students using Web sites and PowerPoint as instructional tools in social studies, Tancock (2002) found that reading on the Internet exacerbated the challenges struggling readers already face. Font sizes and colors,

spacing between words, background graphics and colors, and animated graphics were among the characteristics of online texts that presented obstacles for struggling readers. Tancock also found that proficient readers were engaged by the same text characteristics that challenged their less proficient classmates. Many of the struggling readers were so motivated by the chance to use technology that they persevered in order to make sense of the text and complete their assignment.

Dictionary Use, Literacy, and Educational Technology

Educational technology that includes dictionary components has been used successfully to promote literacy skill development among elementary school students. Doty, Popplewell, and Byers (2001) studied how using an interactive CD-ROM storybook impacted second grade students' reading comprehension. The CD-ROM program allowed students to 1) click on words to obtain pronunciation and definitions, and 2) click on pictures to obtain labels and pronunciations. Compared to a control group who read the same story in conventional print format, the second graders who read the CD-ROM storybook scored significantly higher on comprehension questions. Doty et al. suggested that CD-ROM storybooks have a positive impact on reading comprehension because they reduce decoding challenges and allow students to obtain help as often as needed without having to wait for the teacher.

Literature about dictionary use to promote reading comprehension also includes English Language Learners' (ELLs) use of hard-copy dictionaries during tests. Nesi and Meara (1991) reported that such use did not significantly affect adult ELLs' test scores, and students who had access to dictionaries took significantly longer to complete the test. Albus, Thurlow, Liu, and Bielinski (2005) reported more recent research about dictionary use among middle school ELLs on a test that replicated the format of a state standardized reading comprehension test. The effects of access to a dictionary were not significant; however, not all students reported using the provided dictionary. Students who reported using the dictionary had statistically significantly higher test scores, suggesting potential benefits of dictionary use during tests.

Although research about the use of educational technology to support literacy is growing, we know of no studies that address how electronic pop-up dictionaries impact reading comprehension in content areas. This pilot study was designed to fill this gap in the research literature. The research question was:

How does students' reading comprehension differ for texts that are read a) in a hard-copy text format, b) online, and c) online with a pop-up dictionary function for every word in the text?

METHODS

Participants

One-hundred twenty-nine middle school students from a rural, working class town participated in this investigation. The town's economic infrastructure was based on a now waning industry and as a result the school district is experiencing declining enrollment despite an influx of ELL learners from Russian

and Spanish speaking countries. The district is seeking to provide increasingly diverse learning programs with reduced state funding. More than 37% of the school district's population comes from low-income homes. The district faces challenges that are similar to those faced by inner city schools. Improving students' reading comprehension scores in the non-fiction category of the state reading examination is one of the district's targeted goals. Thus, the district's administrators agreed to host the pop-up dictionary investigation in hopes of identifying a technique that scaffolds students who struggle to read content-area text. A total of 129 students participated. Informed consent was obtained from parents and assent was obtained from the students who participated in the study. Spanish language copies of the informed consent letter were made available.

Four class sections each of sixth, seventh, and eighth grade were selected for the study. We followed Fletcher and Hunter's (2003) recommendations for obtaining parental consent for children to participate in research. First we sent a letter of informed consent home with each child in the selected classes; 76 sixth graders, 74 seventh graders, and 71 eighth graders received letters. After a week, we sent a second copy home with the children who had not yet returned the original informed consent letter. A week later, we sent personalized copies home with the name of each child and their parents or guardians hand-written on the letter. Spanish language copies of the letter were also made available. Students who received informed consent from their parents were asked to provide their informed assent on each day of testing. After excluding those who had consent but were absent for one or more day of testing, 37 sixth graders, 33 seventh graders, and 59 eighth graders participated in the study. The higher participation rate among eighth graders seemed to reflect the more extensive reminders these students received from their teachers. Participation was voluntary and no incentives were provided.

Procedure

Students were assigned to one of four sequences of readings. Each student took a test consisting of two parts (referred to hereafter as Part A and Part B) with 10 questions per part after each reading, over a total of two or three testing periods. Part A consisted of comprehension questions presented in multiple-choice format. The questions addressed students' comprehension of main ideas from the chapter; two examples from the first test the eighth grade students took are provided below.

Which of the following was <u>not</u> a result of the Crusades:
      a.   The Pope and Catholic Church grew stronger.
      b.   Many nobles lost their lives.
      c.   European trading cities became more powerful.
      d.   Feudal nobles grew stronger in Europe.

What does the word **reconquer** mean in the following sentence?
Several other Crusades were fought later to **reconquer** the Holy Land.
      a.   trade with   c.   capture again
      b.   capture   d.   destroy

Part B consisted of cloze test items; the passages were taken directly from the texts students read prior to taking the test. An example from the second test the eighth graders took, along with the written directions students were provided, is below.

Directions: Some words have been erased and replaced with blanks. In each blank, write the word that you think should go there.

Town streets _____ narrow, often about 10 _____ wide, and the houses _____ built of wood.

The kinds of questions are commonly used to measure reading comprehension. The tests were completed in pencil and paper for all three readings.

Denoting PU as a pop-up dictionary reading, O as an online reading without the pop-up dictionary, and T as a text-based reading, the sequences for the readings were:

Sequence 1: {T, PU, O}
Sequence 2: {O, T, PU}
Sequence 3: {PU, O, T}
Sequence 4: {T, PU}

Participating students ranged from grades six through eight. In grade six, two sections were given test sequence one, and two sections were given test sequence two. In grade seven, all four sections were given test sequence three. In grade eight, all four sections were given test sequence four. The order and distribution were determined by two concerns. First, groups needed to complete different sequences so we could determine if the order of treatment impacted comprehension. Second, the school schedule and access to computer labs was another issue. By arranging the order of the sequences as listed above, we were able to minimize the impact the testing had on other classes that regularly used the computer labs.

Each grade level teacher selected three sections of their textbook for use in the study. Two readings were converted into electronic format; one of the electronically formatted texts was equipped with the pop-up dictionary. The pop-up dictionary used definitions that were written at the sixth and seventh grade reading level; 100% of the words in the text had pop-up dictionary definitions. The electronically formatted texts, whether or not they were equipped with the pop-up dictionary, were devoid of the complicated and potentially distracting font sizes and colors, background graphics and colors, and animated graphics that Tancock (2002) found likely to exacerbate the challenges faced by struggling readers. The combined length of the converted readings was within copyright guidelines for educational fair use.

The middle school where this study took place had 50-minute class periods. In order to complete each day of testing within this time as well as ensure that the teachers were able to take attendance and check-in with their students, the social studies teachers helped identify three readings they thought their struggling readers could complete in 20 minutes. The comprehension tests were designed to take approximately 10 minutes; in total, data-collection took 30

minutes per test. Students completed the online and pop-up dictionary readings in a computer lab; the text reading was completed in the classroom. By limiting data-collection to 30 minutes, we had 20 extra minutes for the computer lab days. This was a helpful cushion because the students had to walk to the computer lab, login, and occasionally overcome computer or network difficulties. Those who finished with time remaining in the class period participated in a review activity. Each of the readings selected was part of the existing social studies curriculum, so the investigation did not change or increase student work expectations. The testing sequences were scheduled to take place three school days in a row. The start of the testing was staggered by grade level in order to ensure access to the computer lab.

Student use of the pop-up dictionary was measured in two ways. The researchers and their assistants took anecdotal notes about the students' use of the dictionary. The students' initial exploration of the pop-up dictionary is best described as overzealous; they looked up the definition of words such as "is" and "the" because of the novelty of the system. After giving the students a few minutes to explore the reading and the pop-up dictionary functionality freely, students were redirected to the academic focus and instructed to begin reading and using the dictionary when unsure about a word's meaning. As expected, students who were known to be struggling readers used the pop-up dictionary extensively while those who were stronger readers only used it a few times. The stronger readers typically looked up definitions for more complicated social studies terms while struggling readers also looked up less sophisticated words. Sixth and seventh graders used the pop-up dictionary more than the eighth graders. In addition to our anecdotal observations, after completing the pop-up dictionary reading and test students were given a questionnaire on which they indicated the extent to which they used the pop-up function. The questionnaire also included demographic items and items about students' interests in reading. The students self-reports of pop-up use confirmed our anecdotal observations that eighth graders used the function less than the other grades; 20% of sixth and seventh graders reported extensive use of the dictionary while only 13% of eighth graders reported the same. Interestingly, 20% of sixth and seventh graders reported not using the pop-up at all while 28% of eighth graders reported the same. These two methods were used to determine student utilization of the pop-up dictionary because the computer software was not designed to record use. The procedures for this investigation were reviewed and approved by the university's Institutional Review Board (IRB).

Data Analysis

Crossover analyses allow each subject to receive each treatment (here, Reading Method) in succession. Because each subject is tested under multiple reading methods, the study requires far fewer subjects than if each subject was tested under one reading method. Using the crossover approach, each subject serves as their own control, screening out a substantial amount of experimental error. Relatively strong students are expected to score well overall regardless of Reading Method, for instance, and this model accounts for this variation, which

ultimately reduces the amount of unexplained variation in the response variable substantially. Because each student takes tests in succession, this modeling approach also allows for the possibility of what is termed a carryover effect, in which a reading method potentially affects a test score in the testing period after it is administered. In general, crossover ANOVA analyses are common in clinical trials, and typically test whether a response variable is a function of explanatory variables such as treatment type, order of treatments, carryover effects, treatment period, and other categorical variables. All ANOVA models, including this one, can be viewed as special cases of multiple regression models where the explanatory variables are categorical and coded numerically as binary variables indicating presence or absence of an experimental condition during any observation.

Three separate crossover ANOVA analyses were performed to determine whether the reading type (T, PU, or O) had statistically significant effects on test scores. The response variable for Analysis 1 was the number of correct answers for part A of the test. The response variable for Analysis 2 was the number of correct answers for part B of the test. Finally, the response variable for Analysis 3 was the combined score on both parts of the test.

Potential explanatory variables for each model were Grade Level, Class Section (nested within Grade Level), Student's PSSA score in reading (PSSA stands for Pennsylvania System of School Assessment, the high-stakes standardized testing system used in Pennsylvania), Reading Method (T, PU, or O, corresponding to the previous designations), Lag Variables (indicator variables to represent which exam was taken in the previous period), Gender, Testing Period (1, 2, or 3), Sequence (1, 2, 3, or 4 corresponding to sequences previously defined), and Student (nested within Sequence). We also looked for the possibility of an interaction effect between Reading Method and Sequence. We removed Grade Level and Class Section from consideration because effects of these variables are confounded with the Sequence variable. Strong practical considerations led to students within a class section taking tests in the same sequence. This constraint affected ability to simultaneously estimate grade, section, and sequence effects, but it still allowed the main research questions to be answered. The Lag variables in the model allowed for the previously discussed possibility of a carryover effect from the previous reading period. The Lag variables in the analysis allowed for simultaneous estimation of the carryover effect and direct effect of each reading method. For this analysis, two lag variables were necessary; LagPU denotes the lag variable for the pop-up dictionary and LagO as the lag variable for the online reading.

Some limitations of our approach are worthy of discussion. While estimation of direct Reading Method and carryover Reading Method effects are possible, in Crossover ANOVA analyses, these two effects are never completely separable because not all combinations of each variable occur together. For instance, we never observe both the carryover effect and direct effect of any Reading Method in the same testing period. Because of this, it is often required to assume other effects (such as Testing Period, perhaps) are negligible to avoid further confounding of the explanatory variables with each other. This confounding reduc-

es the ability to separate each effect's impact on the response from that of the other explanatory variable. Ultimately, we needed to reduce the number of explanatory variables in our model, but we decided to remove the non-significant carryover effects from the model, eliminating the confounding problems caused by estimation of carryover effects. Details regarding this decision are discussed in the Results section that follows. One other concern with crossover ANOVA methods is to minimize subject dropout during the analysis. Heavy dropouts and incomplete observations, when a subject receives only some of the treatments in a sequence, can cause significant practical data problems. Fortunately, we had no significant subject dropout during the study. Another weakness of our approach (previously mentioned) is that Testing Sequence was not varied in each grade and section, which affected our ability to estimate Grade and Class Section effects.

RESULTS

Table 1 summarizes the overall ANOVA results of each of the three analyses, listing the response variable, explanatory variables with p-values, and the R-squared value for each analysis. The R-squared value for each model is relatively high. In each of the three analyses, the lag variables were not statistically significant and were removed from the model. Retaining non-significant carryover effects in these analyses is sometimes done, and has in fact been recommended (see Abeyasekera & Curnow, 1984). However, it is generally common practice to remove non-significant carryover effects (see Kuehl, 2000). Ultimately, we decided to remove them to allow for much easier interpretation of the remaining model coefficients and because their removal had very little impact on the R-squared value for each analysis. Had we retained the Lag variables in the model, we would have needed to omit the Reading Method and Sequence interaction term as an explanatory variable due to confounding. Additionally, in all three analyses, the explanatory variables for Grade Level, PSSA level, and Gender were removed due to confounding with Testing Sequence and Student. The confounding of Grade Level and Test Sequence has been previously discussed. Once we account for Student variation in this model, variables such as PSSA score and Gender, which are characteristics of an individual student, are not necessary. Finally, the Testing Period explanatory variable was removed from each analysis due to lack of statistical significance.

Table 2 (p. 136) provides a summary of comparisons between the pop-up dictionary reading method and each of the other two reading methods for each sequence in all three analyses. These comparisons are of interest because of the significant interaction between Reading Method and Sequence in each analysis. Because these pairwise comparisons of Reading Methods were planned before the data were gathered, Fisher's F-protected Least Significant Difference (LSD) method was used. Table 2 also lists the estimated student test score increase for the pop-up reading method compared with each of the other two reading methods for each testing sequence, and a p-value for testing the null hypothesis that there was no difference in exam scores for the two reading methods versus an

Table 1: Summary of Statistical Analyses, Listing the Response Variable for Each Analysis and the Explanatory Variables (with p-values), and the R-squared Value for Each Analysis

Analysis	Response Variable	Explanatory Variables and p-values	R-squared
1	Student's Part A test score	Sequence (less than 0.0001) Student (Sequence) (less than 0.0001) Reading Method (0.0009) Test*Sequence interaction (0.0097)	70.8%
2	Student's Part B test score	Sequence (less than 0.0001) Student (Sequence) (less than 0.0001) Reading Method (0.0050) Test*Sequence interaction (less than 0.0001)	82.4%
3	Student's overall test score (Part A score + Part B score)	Sequence (less than 0.0001) Student (Sequence) (less than 0.0001) Reading Method (less than 0.0001) Test*Sequence interaction (0.0066)	83.6%

alternative hypothesis that the difference was nonzero. We used the 5% statistical significance threshold for our comparisons.

From the comparisons in Table 2, we see that in Analysis 1, where the response variable was the student's Part A exam score, the Pop-Up reading method provided a statistically significant benefit in score over the Online reading method for Test Sequences 2 and 3, and was statistically equivalent to the Online reading method for Sequence 1. Similarly, also within Analysis 1, the Pop-Up reading method scores were statistically higher than the Text reading method for Sequences 2 and 3, and statistically equivalent for Sequences 1 and 4.

Also from Table 2, we see that for Analysis 2, where the response variable was the student's Part B exam score, the Pop-Up reading method was only statistically different from the other methods within the same sequence in one case, comparing Testing Sequence 4 Pop-Up scores versus the traditional Text reading method. Note, though, that this difference was of large magnitude (3.19 points). The remaining comparisons within the same testing sequence were statistically equivalent.

Finally, from Table 2, we see that for Analysis 3, where the response variable was the student's overall exam score, the Pop-Up method was significantly better than both the online and text reading methods for Test Sequences 2 and 3 (and also for Sequence 4 for the comparison between Pop-Up and Text reading methods). The Pop-Up reading method was statistically equivalent to the other reading methods for Sequence 1. Because Sequences 1 and 2 were assigned only within the sixth grade, it is possible that the increase in score from the Pop-Up method had more to do with student age than testing sequence. However, Sequence 2 showed the Pop-Up reading method to be statistically better than the

Table 2: Summary of Fisher's F-protected Least Significant Difference Comparisons Between the Pop-up Reading Method and Each of the Other Reading Methods, Using Least Squares Means for Each Comparison (Statistically significant comparisons in bold font)

Analysis	Testing Sequence	Comparison	Estimated Effect	p-value
1	1	Pop-Up vs. Online	0.36 point increase for Pop-Up	0.547
	2		**1.20 point increase for Pop-Up**	**0.0033**
	3		**1.41 point increase for Pop-Up**	**Less than 0.0001**
1	1	Pop-Up vs. Text	0.09 point increase for Pop-Up	0.8803
	2		**1.06 point increase for Pop-Up**	**0.0082**
	3		**1.50 point increase for Pop-Up**	**Less than 0.0001**
	4		0.11 point decrease for Pop-Up	0.6834
2	1	Pop-Up vs. Online	0.09 point decrease for Pop-Up	0.4991
	2		0.45 point increase for Pop-Up	0.3406
	3		0.56 point increase for Pop-Up	0.1841
2	1	Pop-Up vs. Text	1 point decrease for Pop-Up	0.1556
	2		0.71 point increase for Pop-Up	0.1244
	3		0.16 point increase for Pop-Up	0.7065
	4		**3.19 point increase for Pop-Up**	**Less than 0.0001**
3	1	Pop-Up vs. Online	0.27 point increase for Pop-Up	0.774
	2		**1.65 point increase for Pop-Up**	**0.0101**
	3		**1.9 point increase for Pop-Up**	**0.0011**
3	1	Pop-Up vs. Text	0.91 point decrease for Pop-Up	0.3391
	2		**1.78 point increase for Pop-Up**	**0.0051**
	3		**1.70 point increase for Pop-Up**	**0.0038**
	4		**3.13 point increase for Pop-Up**	**Less than 0.0001**

other reading methods so the sixth grade students did show a benefit from the Pop-Up reading method, but only in Sequence 2, where the Pop-Up method was the last test taken in the sequence. More data are needed to determine whether a clear interaction effect exists between grade level and reading method.

Overall, the pop-up dictionary reading was shown to be a statistically effective method for improving student test scores. In the analysis of overall student scores, it proved to be clearly superior to a traditional text reading method for three of the four testing sequences, and superior to the online reading method for two of the three testing sequences.

DISCUSSION

The results of this study suggest that pop-up dictionaries provide a helpful intervention for increasing middle-level learners' reading comprehension in social studies. While it is unrealistic to expect a pop-up dictionary (or any one reading intervention) to provide all the remediation necessary to help struggling readers reach grade-level proficiency, the results of this study are promising enough to suggest that pop-up dictionaries offer one path to improving reading comprehension.

Future studies over a longer period of time that use standardized high-stakes state reading assessments to measure effectiveness are needed to demonstrate the long-term feasibility of pop-up dictionaries as an intervention for struggling readers. It may also be worthwhile to determine if pop-up dictionaries support improved reading comprehension at the elementary and high school levels as well as in additional content areas. Regular access to computer labs for reading will be an obstacle for some school districts because of financial, space, and logistical limitations. Therefore, future studies should investigate the use of pop-up dictionaries on hand-held computers because they are less expensive and require less maintenance and space compared to desk or lap top models. Swan, van t' Hooft, Kratcoski, and Unger (2005) found hand-held computer use helped increase student motivation and engagement in learning tasks for students in grades 3, 4, 6 and 7. Thus, future studies can also investigate whether merging pop-up dictionaries and hand-held computer technology increases student motivation and engagement in reading.

In conclusion, because of the pressures of high-stakes standardized tests, the benefits of using reading interventions in social studies and other content areas where standardized assessment is not mandated are twofold. First, the importance of helping struggling readers improve their reading comprehension cannot be understated. Second, using non-tested content areas to help increase student reading comprehension increases the viability of these subjects that are worthwhile in their own right yet are in danger of being pushed out of the public school curriculum because of testing pressures.

Contributors

Sara Winstead Fry is an assistant professor of education at Bucknell University. Her research interests include beginning teacher induction, social studies education, and enhancing preservice teacher and K–12 student learning

through educational technology. (Address: Sara Winstead Fry, Education Department, Bucknell University, 471 Olin Hall, Lewisburg, PA 17837; Phone: 570.577.3459; sfry@bucknell.edu.)

Ross Gosky is an assistant professor in the Mathematical Sciences Department at Appalachian State University. His research area is statistics. (Address: Ross Gosky, Appalachian State University, Department of Mathematical Sciences, Appalachian State University, 329 Walker Hall, 121 Bodenheimer Dr, Boone, NC 28608; Phone: 828.262.2364; goskyrm@appstate.edu.)

References

Abeyasekera, S. & Curnow, R. N. (1984). The desirability of adjusting for residual effects in a crossover design. *Biometrics, 40*(4), 1071–1078.

Albus, D., Thurlow, M., Liu, K., & Bielinski, J. (2005). Reading test performance of English-language learners using an English dictionary. *The Journal of Educational Research, 98*(4), 245–255.

Beers, K. (2002). *When kids can't read—What teachers can do.* Portsmouth, N.H.: Heinemann.

Ciardiello, A. V. (2002). Helping adolescents understand cause/effect text structure in social studies. *The Social Studies, 93*, 31–36.

Doty, D. E., Popplewell, S., & Byers, G. O. (2001). Interactive CD-ROM storybooks and reading comprehension of young readers. *Journal of Research on Computing in Education, 33*(4), 374–384.

Fletcher, A. C., & Hunter, A. G. (2003). Strategies for obtaining parental consent to participate in research. *Family Relations, 52*(3), 216–21.

Haskvitz, A. (2006, March 22). Disrespecting social studies. *Education News.* Retrieved September 21, 2007, from http://www.ednews.org/articles/547/1/The-Disrespecting-of-Social-Studies/Page1.html

Johnson, R. E. (1977). What, besides the reading level, makes social studies textbooks hard to read? *Social Science Record, 14*(2), 20–22.

Katims, D. S., & Harmon, J. M. (2000). Strategic instruction in middle school social studies: Enhancing academic and literacy outcomes for at-risk students. *Intervention in School and Clinic, 35*(5), 280–289.

Kuehl, R. O. (2000). *Design of Experiments: Statistical Principles of Research Design and Analysis* (2nd ed). Pacific Grove, CA: Duxbury Press.

Lewis, A. C. (2005). An ammunition dump? *Phi Delta Kappan, 87*(4) 260–261.

MacArthur, C. A., Ferretti, R. P., Okolo, C. M., & Cavalier, A. R. (2001). Technology applications for students with literacy problems: A critical review. *The Elementary School Journal, 101*(3), 273–301.

Massey, D. D., & Heafner, T. L. (2004). Promoting reading comprehension in social studies. *Journal of Adolescent & Adult Literacy, 48*(1), 26–41.

Moore, D. W., Bean, T. W., Birdyshaw, D. & Rycik, J. A. (2000). *Adolescent literacy: A position statement for the Adolescent Literacy Commission of the International Reading Association.* Newark, DE: International Reading Association.

Nesi, H., & Meara, P. (1991). How using dictionaries affects performance in multiple-choice EFL tests. *Reading in a Foreign Language, 8*(1), 631–643.

Papalewis, S. (2004). Struggling middle school readers: Successful, accelerating intervention. *Reading Improvement, 41*, 24–37.

Risinger, C. F. (2006). Teaching what we should be teaching using the Internet. *Social Education, 70*(4), 197–198.

Schneiderman, M. (2004). What does SBR mean for education technology? (A closer look at scientifically based research). *T.H.E. Journal (Technological Horizons In Education), 31*(11), 30–35.

Swan, K., van t' Hooft, M., Kratcoski, A., & Unger, D. (2005). Uses and effects of mobile computing devices in K–8 classrooms. *Journal of Research on Technology in Education, 38*(1), 99–112.

Tally, B., & Goldenberg, L. B. (2005). Fostering historical thinking with digitized primary sources. *Journal of Research on Technology in Education, 38*(1), 1–21.

Tancock, S. M. (2002). Reading, writing, and technology: A healthy mix in the social studies curriculum. *Reading Online, 5*(8). Available at: http://www.readingonline.org/articles/art_index.asp?HREF=tancock/index.html