Improving Student Science and English Language Skills:

A Study of the Effectiveness of BrainPOP

2009

SEG RESEARCH

This study was conducted by SEG Research, an independent educational research firm located in New Hope, Pennsylvania. SEG Research provides research, evaluation, and assessment services to educational publishers, educational technology providers, assessment service providers and government agencies. SEG has been meeting the research and assessment needs of organizations since 1979. This research was supported by a grant from BrainPOP.

Executive Summary

Background and Purpose

During the 2008-2009 school year, SEG Research conducted a multi-site study of students in grades 3, 5 and 8 to evaluate the effectiveness of BrainPOP, a web-based animated instructional tool designed to support educators and engage students. BrainPOP is intended for use in both group and one-on-one settings and can be used in numerous ways, from introducing a new lesson or topic to illustrating complex subject matter to reviewing before a test. Content is aligned to state standards and searchable.

The goal of this study was to evaluate the effectiveness of BrainPOP. The findings indicate that students in classes using BrainPOP made significant improvement in Science, Reading Comprehension, Language, and Vocabulary skills during one school semester (January through June) as compared to students in classes that did not use BrainPOP.

Study Design

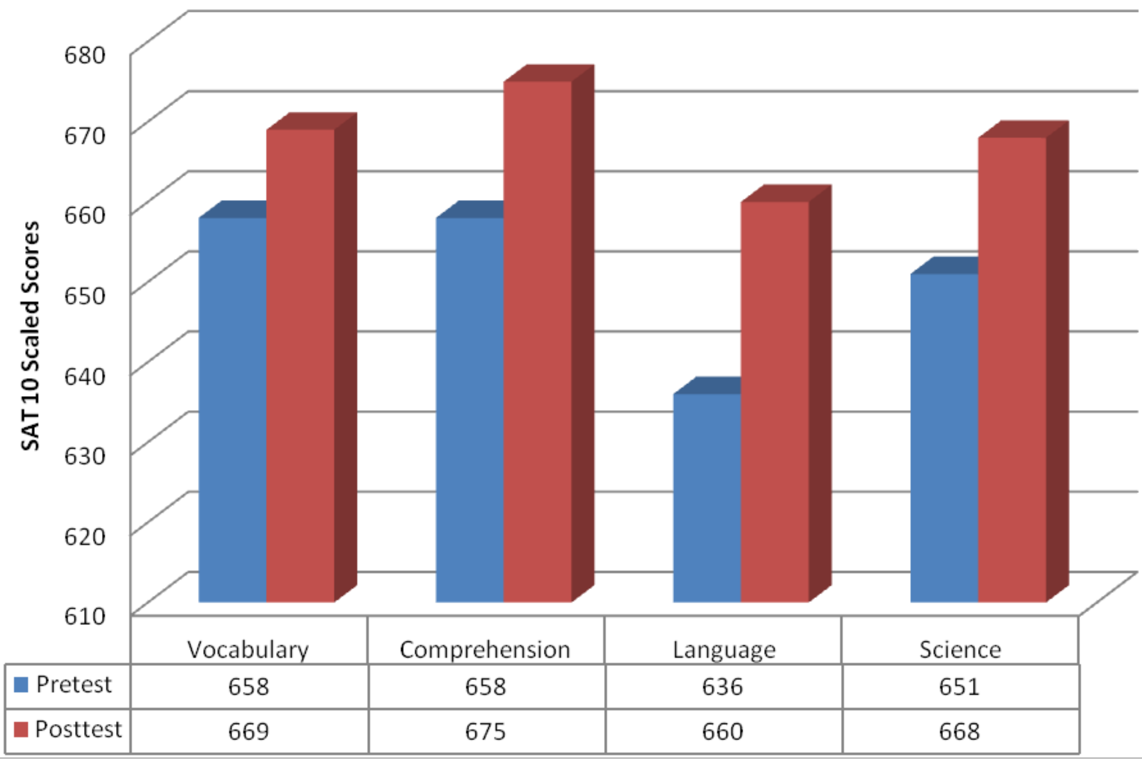
Between January 2009 and June 2009, approximately 1,100 students in 46 classrooms in Palm Beach County, Florida and New York City, New York participated in a controlled study of BrainPOP effectiveness. Using a quasi-experimental, pre-post design, this study compared the growth in Science, Reading Comprehension, Language and Vocabulary skills between students in classes using BrainPOP (Treatment Group) and a comparable group of students in classes that did not use BrainPOP (Control Group). Growth in Reading Comprehension, Vocabulary, Language and Science was measured by comparing scores on the Stanford 10 Achievement Test™ Abbreviated Battery (SAT 10) at the beginning of the second semester of the school year and at the end of the school year. Students in both the Treatment (BrainPOP Users) and Control Groups took a pretest (SAT 10) at the beginning of the second semester of the school year to obtain a baseline measure of student Science, Reading Comprehension, Language and Vocabulary skills. Students in the Treatment Group then received instruction that typically included two to three hours of BrainPOP weekly, while those in the Control Group received instruction without the benefit of BrainPOP. At the end of the school year, students in both the BrainPOP Users Group and Control Groups took a posttest (SAT 10). The results from the pretest and posttest were compared statistically to determine the level of growth in Science, Language, Reading Comprehension, and Vocabulary skills. Students in the BrainPOP Users and the Control Group were well-matched in ability and demographically. Any initial differences in the Science, Reading Comprehension, Language and Vocabulary skills of students in the Treatment and Control Groups were statistically controlled during analysis.

Pre- Post Results for BrainPOP Users

Students who were in classes that used BrainPOP showed substantial growth in Science, Language, and Reading Comprehension and more moderate gains in Vocabulary, during the course of the study. Students in classes using BrainPOP increased their SAT 10 Language scale-scores by 24 points, their Reading Comprehension scores by 17 points, their Science Scores by 17 points, and their Vocabulary scores by 11 points (see Figure 1). Students received approximately 16-20 weeks of instruction using BrainPOP, yet the amount of growth achieved is equivalent to between one and two grade levels of growth when compared to the national sample of students included in the Stanford 10 norm group (Harcourt Assessment, 2002).

To better understand the magnitude of growth for students in BrainPOP classrooms, we looked at the “effect size”, a common metric that can be used to evaluate the amount of growth across studies, even when different measures are used. We found effect sizes of +.47 for Language, +.37 for Reading, +.36 for Science, and+.19 for Vocabulary (unadjusted for Control Group differences). This indicates that students in classes that used BrainPOP showed substantial growth in Reading, Language, and Science and moderate growth in Vocabulary from the beginning to the end of the study.

Figure 1
Comparison of Pretest and Post Test Scores for Students in Classes Using BrainPOP



	Vocabulary	Comprehension	Language	Science
Pretest	658	658	636	651
Posttest	669	675	660	668

While the growth achieved by students using BrainPOP is an important indicator of the effectiveness of BrainPOP, a more complete way to assess growth is to compare the growth achieved by students in classes using BrainPOP (Treatment) to students in classes that did not use BrainPOP (Control).

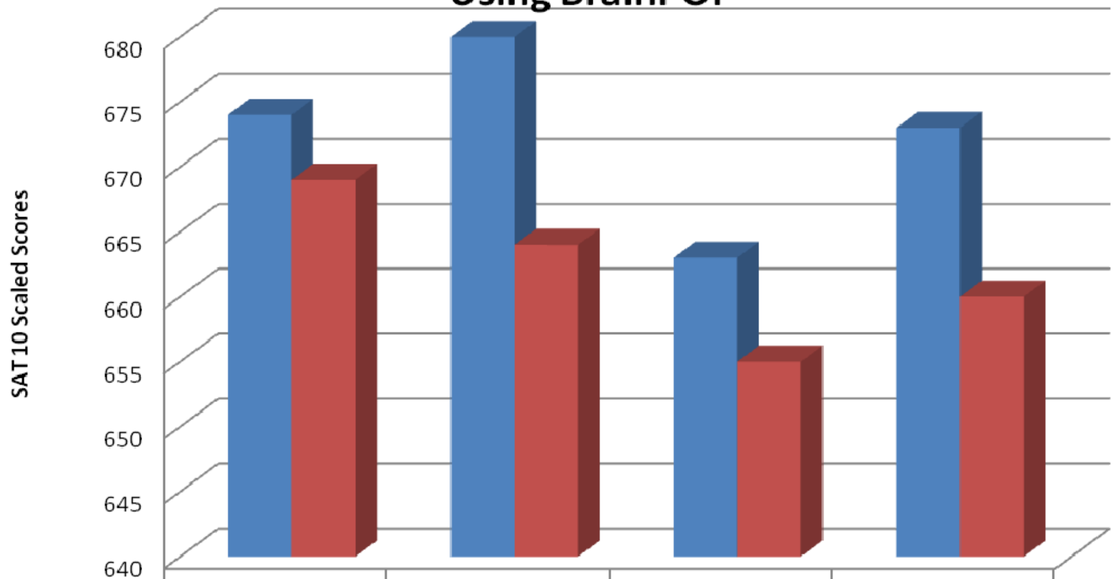
Comparison of BrainPOP Users to the Control Group

We compared the gains made by students in classes using BrainPOP to those of a Control group of students in classes that did not use BrainPOP, controlling for any initial differences in the Science, Reading Comprehension, Language and Vocabulary skills of students in the two groups. We used a statistical procedure known as analysis of covariance (ANCOVA), to provide a more accurate comparison of growth between groups. This analysis compares differences as if the two groups (BrainPOP Users and Control) were identically matched in initial Science, Reading Comprehension, Language and Vocabulary skills. The students using BrainPOP showed statistically greater gains in Science, Language, Vocabulary, and Reading Comprehension than those students in the Control Group who were not using BrainPOP. Gains by the BrainPOP users were statistically significant at the $p<.05$ level.

BrainPOP users showed substantially greater gains in Reading Comprehension, Science and Language and moderately greater gains in Vocabulary than students in classes that did not use BrainPOP. When controlling for students' initial ability, BrainPOP users finished the year with scores that were 16 scale-score points higher in Reading Comprehension, 13 scale-score points higher in Science, 8 scale-score points higher in Language, and 5 points higher in Vocabulary than the Control Group on the SAT 10 assessments (see Figure 2).

To better understand the magnitude of the difference between students in BrainPOP classes and those in classes that did not use BrainPOP, we again looked at the "effect size." We compared the average (mean) score for the BrainPOP users Group to the average (mean) score for the Control Group (adjusted for any initial differences in student ability).

Figure 2
Comparison of Matched Posttest Scores for Students Using BrainPOP and Control-Group Students Not Using BrainPOP



	Vocabulary	Comprehension	Language	Science
BrainPOP Users	674	680	663	673
Control Group	669	664	655	660

We found effect sizes of +.34 for Reading Comprehension, +.29 for Science, +.17 for Language, and + .09 for Vocabulary. The effects for Reading Comprehension, Science and Language are quite substantial, indicating that the students in classes that used BrainPOP performed well above the students in classes that did not use BrainPOP in these skill areas. For comparison, the effect sizes for Reading Comprehension and Science are above the typical effect sizes seen in other studies of instructional programs. Language and Vocabulary were more typical of the comparative gains seen in other studies of instructional programs. (For example, Slavin (2008) in his comprehensive synthesis of middle and high school Reading program research studies reports a mean effect size for instructional-process Reading programs of +.21.)

The results by grade level indicated that BrainPOP produced significantly greater increases in a composite of Reading, Language, Vocabulary and Science. The effects were most pronounced at the 5th grade level, where the greatest gains were seen.

BrainPOP was found to be equally effective for boys and girls; for students of different ethnicities; and for both students receiving free or reduced lunch services and those not receiving this service. The interaction between BrainPOP use and gender, ethnicity and free or reduced lunch status was not statistically significant.

Teacher Perceptions of Effectiveness

At the conclusion of the study, participating teachers were surveyed regarding their perceptions of BrainPOP. 90% felt that BrainPOP was good or excellent at improving students' attitudes toward school and learning. Similarly, 90% of teachers felt that BrainPOP was good or excellent at increasing students' cognitive/intellectual growth. All (100%) of the teachers indicated that they would definitely recommend BrainPOP to others.

Summary

Students who were in classes that used BrainPOP showed substantial growth in Language, Reading Comprehension and Science and more moderate gains in Vocabulary during the course of the study. Students in classes using BrainPOP increased their SAT 10 scores between 11 and 24 points. Students received approximately 16-20 weeks of instruction using BrainPOP, yet the amount of growth achieved is equivalent to between one and two grade levels of growth when compared to the national norm group.

Students enrolled in classrooms using BrainPOP achieved substantially greater gains in Science, Language, and Reading comprehension than students enrolled in classes that did not use BrainPOP. More moderate gains were also seen for Vocabulary. When controlling for students' initial ability using analysis of covariance, BrainPOP users showed substantially greater gains in Reading Comprehension, Science and Language and moderately greater gains in Vocabulary than students in classes that did not use BrainPOP. BrainPOP users finished the year with scores that were 16 scale-score points higher in Reading Comprehension, 13 scale-score points higher in Science, 8 scale-score points higher in Language, and 5 points higher in Vocabulary than the Control Group on the SAT 10 assessments.

To understand the magnitude of the difference in growth between BrainPOP users and those who did not use BrainPOP, we looked at effect size. We found effect sizes of +.34 for Reading Comprehension, +.29 for Science, +.17 for Language, and + .09 for Vocabulary. This is a large effect, particularly for Reading Comprehension, Science, and Language, indicating that the students in classes using BrainPOP performed well above the students in classes that did not use BrainPOP.

BrainPOP was found to be effective at all three grade levels included in the study (grades 3, 5 and 8). BrainPOP was found to be equally effective for boys and girls; students of different ethnicities; and both students receiving free or reduced lunch services and those not receiving this service.

The findings of this study demonstrate that students in classes using BrainPOP achieve substantially more growth in Science, Reading comprehension and Language skills than students in classes that do not have the benefit of BrainPOP in their classroom.

Improving Student Science and English Language Skills:

A Study of the Effectiveness of BrainPOP

BrainPOP Program Overview

BrainPOP is a web-based animated instructional tool designed to support educators and engage students. BrainPOP is designed for both group and one-on-one settings and can be used in numerous ways, from introducing a new lesson or topic to illustrating complex subject matter to reviewing before a test. Content is aligned to state standards and searchable.

Multimedia-Based

BrainPOP is based on what is often referred to as the *fundamental multimedia principle*: Information is more effective when presented in words and pictures than words alone (Mayer, 2005). Research has shown that the brain processes information using two channels: visual and auditory. The brain can accommodate more information when it is presented both visually and aurally. By taking advantage of this multimodal processing capability, we can dramatically enhance student learning through multimedia instruction.

BrainPOP is a multimedia learning application that delivers instructional content using multiple modes. These include presenting visual and auditory information, which students can then use to construct knowledge.

Information Processing

Our ability to process information is a multi-step process that involves the perception, attention, selection, organization and integration of information (Sweller, 2003). At the center of this process is **long term memory**. Our long term memory stores our accumulated knowledge. Our accumulated knowledge is organized into "chunks" of information in what are known as **schema**. Schemas allow us to organize information in meaningful ways and help us integrate and organize new information (Chi, Glaser, and Rees, 1982). In short, our long term memory is where what we know is stored and where we integrate new information. If information does not find its way into long term memory, it is lost. Learning can be thought of as change in our long term memory.

The limitations of working memory. Before information can be integrated into long term memory it must be received and processed by our working memory. Working memory is very limited; it can only handle small amounts of information before it has to be integrated into our long term memory or lost. George Miller (1956) suggested that we can only process about seven pieces of information at one time. And, we must do so quickly, as working memory can only keep information for about 20 seconds.

Multiple channels for information processing. It is widely believed that there are multiple channels in working memory. Baddeley (1992) proposes an auditory and a visual channel. The auditory channel handles information that is heard, while the visual channel processes information that is seen. Text seems to have unique processing requirements, with words initially captured by the visual channel and then converted to sounds in the auditory channel (Mayer, 2005).

Research suggests that the visual channel handles less information than the auditory channel (Miler, 2005). However, when information is presented using both the visual and auditory channels, working memory can handle more information overall.

Using multiple channels can increase the amount of information that the brain can process (Sweller, 2005). But, there is still the risk of cognitive overload. Too much information delivered in an ineffective manner can interfere with the brain's ability to successfully integrate information into long term memory.

Organizing information using schema. The information in working memory is integrated into long term memory using existing schema (Sweller, 2003). If there are no existing schema in which to "fit" the information, new schema need to be created and working memory may need to do some extra work to help organize the information (Baddeley 1999). If information is poorly organized, or if it is difficult to relate newly presented information to existing schema, working memory can handle even less information. This can be prevented somewhat by presenting organizing information along with the information to be learned.

Brain Processing and Multimedia Learning. So what do we know about brain processing that is relevant to multimedia learning? We know that:

1. Effective multimedia recognizes that working memory has a limited capacity to process information.

2. Effective multimedia presentations take advantage of both the auditory and visual channels in working memory to deliver content. Using multiple channels increases the overall amount of information the brain can process.

3. Effective multimedia understands that text may be particularly challenging to process, with involvement from both the visual and auditory channels required.

4. Effective multimedia presentations recognize that long-term memory organizes information into meaningful chunks called schema. Presenting information in a way that makes use of existing organizing structures

(schema) or that helps students organize the information can greatly assist the learner in incorporating information into Long Term Memory.

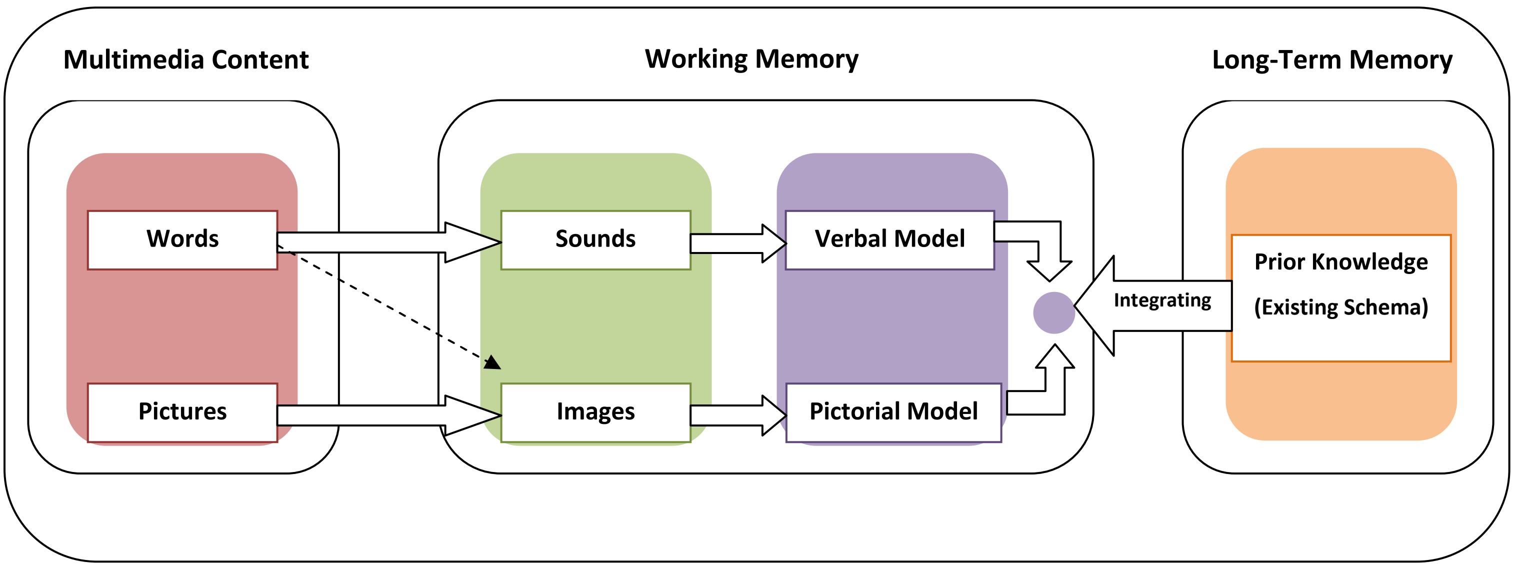


Figure 3: Information Processing Model based on Mayer (2005)

Effectiveness Study Goals and Overview

This report describes a study conducted during the 2008-2009 school year to evaluate the effectiveness of BrainPOP. The study compares the growth in Science, Reading Comprehension, Language and Vocabulary skills of students in grades 3, 5 and 8 who received instruction using BrainPOP (Treatment Group) to those receiving instruction that did not include BrainPOP (Control Group). The study compared student growth in the Treatment and Control Groups. We compared the growth in Language, Reading Comprehension, Vocabulary and Science performance attained by students in the Treatment Group and Control Group between January and June 2009, as measured by the growth in Stanford 10 Achievement Test™ Abbreviated Battery (SAT 10) scores from the pretest administered in January 2009 to the posttest administered at the end of May and in June 2009.

Research Questions

This study investigated the following questions:

1. Do students in grades 3,5 and 8 using BrainPOP as part of their instruction show larger gains in Language, Reading Comprehension, Vocabulary and Science skills than a comparable group of students who do not use BrainPOP as part of their instruction?
2. Do boys and girls using BrainPOP as part of their instruction show larger gains in Language, Reading Comprehension, Vocabulary and Science skills than a comparable group of boys and girls who did not use BrainPOP as part of their instruction?
3. Do students of different ethnicities using BrainPOP as part of their instruction show larger gains in Language, Reading Comprehension, Vocabulary and Science skills than a comparable group of students who did not use BrainPOP as part of their instruction?
4. Do students receiving free or reduced lunch and those not receiving this service using BrainPOP as part of their instruction show larger gains in Language, Reading Comprehension, Vocabulary and Science skills than a comparable group of students who did not use BrainPOP as part of their instruction?

Student Sample

Between January 2009 and June 2009, approximately 1,100 students in 46 classrooms in Palm Beach County, Florida and New York City, New York participated in a controlled study of BrainPOP effectiveness. Classes either used BrainPOP (Treatment Group) or served as a Control Group during the course of the study. There were approximately 572 students in the Treatment group and approximately 518 students in the Control Group. Table 1 shows the number of students comprising each gender, ethnicity, and grade category as reported by their teachers. (The total number of students listed for each background variable may be different since some schools were unable to provide complete background information.)

Table 1. Demographic Profile of Student Participants

Variable	Number (N) of Students	Percentage of Students
GENDER		
Male	432	51%
Female	409	49%
Total (All Gender)	**841**	
FREE OR REDUCED LUNCH		
Receiving Free or Reduced Lunch	450	58%
Not Receiving Free or Reduced Lunch	320	42%
TOTAL (Receiving and Not Receiving)	**770**	
ETHNICITY		
Caucasian	345	41%
African American	127	15%
Hispanic	207	25%
Asian/Pacific Islander	107	13%
Mixed Race and Other	55	7%
Total (All Ethnicity)	**841**	
GRADE		
Grade 3	285	30%
Grade 5	357	38%
Grade 8	305	32%
Total (All Grades)	**947**	

In some cases, teachers did not provide complete background information for a student or a student did not take one of the tests included in the analyses. Where data was missing, the student's results were eliminated from those analyses.

Comparability of Study Groups

It is very important in a study comparing student growth to establish that the Treatment Group and Control Group are similar in background and ability. Demonstrating baseline equivalence of the sample (treatment and comparison groups) minimizes potential bias from selection in quasi-experimental designs that can alter effect size estimates. If the Treatment Group and the Control Group are not similar, we cannot be sure if the growth we see is due to the treatment (in this case, use of BrainPOP) or the result of some differences in the individuals that existed before we conducted the study.

Ideally, this matching is accomplished by sampling study participants of similar ability and with similar background characteristics. However, any observed differences can be adjusted for statistically using analysis of covariance (ANCOVA). The Treatment Group and Control Group were compared with respect to initial Language, Reading Comprehension, Vocabulary and Science ability, as well as their gender, ethnicity and free and reduced lunch status. The results indicate that the groups were similar in ability (see Table 2) and background (see Tables 3, 4 and 5).

Ability Comparison. The SAT 10 pretest scores were used to compare the initial Language, Reading Comprehension, Vocabulary and Science levels for students in both the Treatment and Control Groups. The mean test scores for students in both Groups are presented in Table 2.

Table 2. Comparison of Initial Language, Reading Comprehension, Vocabulary and Science levels (SAT 10 scores) for BrainPOP Group and Control Group

STUDY GROUP	Vocabulary Mean	Vocabulary Standard Deviation	Reading Comprehension Mean	Reading Comprehension Standard Deviation	Language Mean	Language Standard Deviation	Science Mean	Science Standard Deviation
BrainPOP Group	**658.26**	58.30	**657.91**	46.77	**635.80**	55.22	651.53	43.42
	(N=554)		(N=556)		(N=393)		(N=517)	
Control Group	**653.94**	60.71	**657.28**	44.55	**642.30**	46.41	650.01	45.65
	(N=469)		(N=466)		(N=391)		(N=432)	
Total Group	**656.28**	59.42	**657.62**	45.75	**639.04**	44.43	650.84	44.43
	(N=1023)		(N=1022)		(N=784)		(N=949)	

The Treatment and Control Groups were comparable in ability. There were no statistically significant differences between the Treatment and Control groups for Vocabulary (F=1.34, df=1/1023, p<.25), Reading Comprehension (F=4.05, df=1/1022, p<.83), Language (F=3.18, df=1/784, p<.08), and Science (F=.277, df=1/949, p<.28).

Gender, Ethnicity and Free or Reduced Lunch Status Comparison. The number of female and male students, the number of students in each ethnic group, and the number of students receiving and not receiving free and reduced lunch in both the BrainPOP Users and Control groups were counted. These counts are presented in Tables 3, 4 and 5 below.

Table 3. Comparison of Gender Composition of the BrainPOP Group and Control Group

STUDY GROUP	Gender		
	Female	Male	Total
BrainPOP Group	229	229	458
Control Group	196	171	367
Total	425	400	825

Table 4. Comparison of the Ethnic Composition of the BrainPOP Group and Control Group

STUDY GROUP	Ethnicity					
	Caucasian	African American	Hispanic	Asian/ Pacific Islander	Mixed Race or Other	Total
BrainPOP Group	206	65	127	49	23	470
Control Group	130	59	75	58	31	353
Total	336	124	202	107	54	823

Table 5. Comparison of Free or Reduced Lunch Composition of the BrainPOP Group and Control Group

STUDY GROUP	Free or Reduced Lunch		
	Receiving	Not Receiving	Total
BrainPOP Group	221	229	450
Control Group	126	194	320
Total	347	423	770

A statistical comparison of the two study groups shows that the Treatment/BrainPOP Group and Control Group were very similar with respect to gender, ethnicity and free and reduced lunch status.

There were no statistical differences in the expected and observed frequencies for gender (chi square =.95, df=1, p<.36). There was a statistically significant difference in ethnicity (chi square =16.51, df=4, p<.002) and free and reduced lunch (chi square =7.16, df=1, p<.008), however the correlation (contingency coefficient) was .14 for ethnicity and .10 for free and reduced lunch status, suggesting that there was no meaningful difference between the two groups with regard to ethnicity or free and reduced lunch status.

Description of the Pretest and Posttest

The Science, Language, Reading Comprehension, and Vocabulary skills of students participating in the study were measured using the Stanford Achievement Test™, Tenth Edition (SAT 10), Abbreviated Battery, Form A, 2002. The SAT 10 was used as both the pretest and posttest measure; students took the SAT 10 in January at the beginning of the second semester of the school year and then again at the end of May or in June at the end of the school year.

The Language, Reading Comprehension, and Vocabulary subtests of the SAT 10 were used for this study. The Language subtest measures students' achievement in applying the principles that form effective writing from word- and sentence-level skills to whole composition features. The Reading Comprehension subtest measures students' achievement within the framework of three types of materials or purposes for Reading: literary, informational, and functional text. Within each type of text, questions measure achievement in four modes of comprehension: initial understanding, interpretation, critical analysis, and awareness and usage of Reading strategies. The Vocabulary subtest measures students' achievement in Reading Vocabulary skills through use of synonyms, multiple-meaning words, and context clues items that target appropriate words for each grade level. The Science subtest measures student's achievement in Life, Physical and Earth Sciences as well as scientific process. Each subtest ranges from 20-30 items in length (Stanford Achievement Test Series™, Tenth Edition, Technical manual; Harcourt, 2002).

The SAT 10 measures students' skill levels on a single vertical scale ranging from 200-900. The scale-scores represent equal units; differences between scores at any point in the scale represent the same amount of achievement variation. This allows for an accurate comparison of changes over time. The scale is equivalent across forms and grade levels, to provide an accurate comparison across grade levels; a score at one grade level means that same thing at another grade level.

Teachers participating in the study also provided the student results for the New York State Testing Program as an additional posttest measure. However, the results were not usable for this study. Nearly all of the students participating in the study passed the assessment; there was insufficient statistical power to detect differences.

Reliability and Validity

The reliability of the SAT 10 ranges from .89 to .97 (KR-20 reliability coefficient; Harcourt, 2002). Several validity studies conducted for the SAT 10 have found strong evidence for the validity of SAT 10 scores; for example, content expert review found strong alignment with important Reading skills. Strong relationships were found between the SAT 10 and other measures of Reading ability. For a more complete discussion of the SAT 10 reliability and validity, readers are referred to the SAT 10 Technical Manual (Harcourt, 2002).

Description of the Treatment

The Treatment in this study was BrainPOP use. BrainPOP is a web-based animated instructional tool designed to support educators and engage students. BrainPOP is intended for use in both group and one-on-one settings and can be used in numerous ways, from introducing a new lesson or topic to illustrating complex subject matter to reviewing before a test.

BrainPOP comprises a collection of 3-6 minute animated movies used to introduce a new science concept or for review. The movies are accompanied by several tools that can be used to differentiate instruction and engage students' reading, writing, and communication skills. These include instructional Activity Pages and Vocabulary Pages that can be used to reinforce concepts and define new terminology. BrainPOP includes quizzes that can be used to assess prior knowledge before the movie as well as for post-assessment. BrainPOP also includes graphic organizers that can be used to help scaffold student understanding while interacting with the movie as well as short, high-interest readings designed for guided and independent nonfiction reading

Teachers in the Treatment group were provided with access to BrainPOP Educators, an online resource providing video tutorials, webinars, demos and tools to help educators understand how to best use BrainPOP in the classroom.

BrainPOP can be used in several of ways and we observed teachers taking advantage of the BrainPOP tools in a variety of ways in this study. Typically teachers customized their selection of topics and used BrainPOP to fit their pacing calendars. The teachers spent about 15 minutes using a movie and integrating a class discussion during that time period. Teachers " actively viewed" BrainPOP movies with their students, pausing when new concepts or vocabulary is introduced for class discussion, and using the movies' introductory letters to the characters for "quick writes", journaling, or writing prompts. In addition, teachers utilized the various accompanying features described above to further student comprehension.

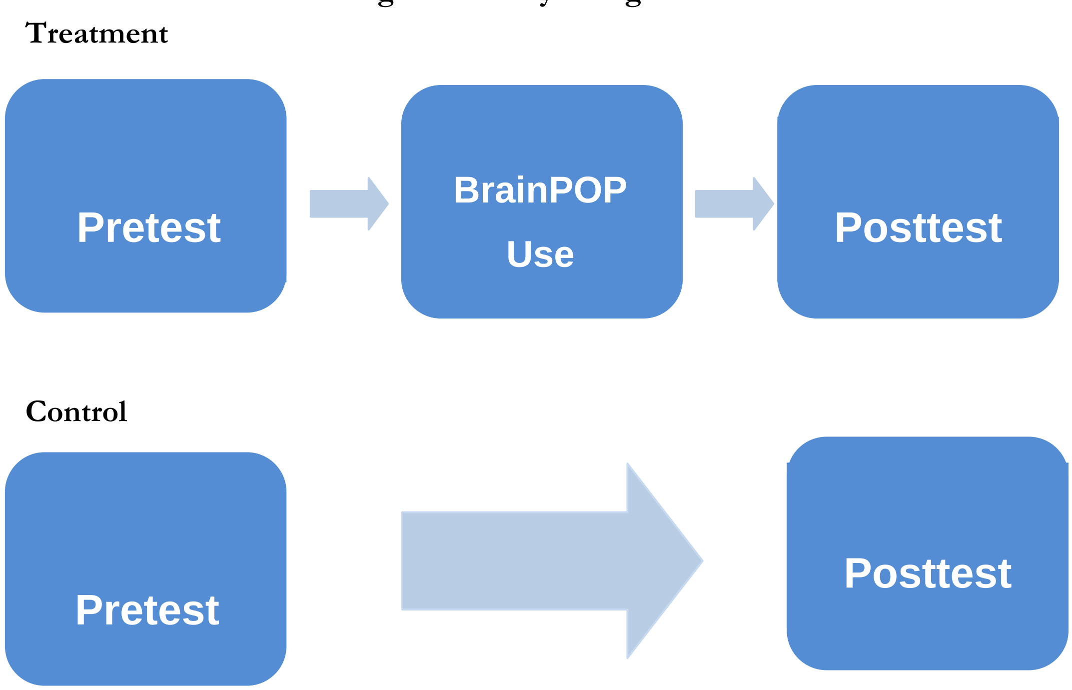
According to the survey of participating teachers conducted at the conclusion of the study, the vast majority (70%) of the teachers used BrainPOP about 2-3 hours per week. Some teachers indicated greater use and 20% reported using BrainPOP about 1 hour or less per week. BrainPOP movies were used every week or every other week by 90% of the teachers and 75% of the teachers reported using the quizzes and activity pages every week or every other week.

Study Design

The goal of this effectiveness study was to compare the growth in Language, Reading Comprehension, Vocabulary and Science for students in classes receiving instruction that included BrainPOP to a group of students in classes that did not use BrainPOP. Students' growth in Language, Reading Comprehension, Vocabulary, and Science was measured by comparing their proficiency at the beginning of the second semester of the school year and again at the end of the school year after receiving instruction. Students in the BrainPOP Group and the Control Group were administered the SAT 10 test as a pretest at the beginning of the second semester in January of 2009 and at the conclusion of the school year at the end of May or in June 2009. Students received approximately 16-20 weeks of instruction between the pretest and posttest. Students in the BrainPOP Group received instruction using BrainPOP, while those in the Control Group received instruction without the benefit of BrainPOP. The results were then compared statistically (.

The study employed a pre-post, Treatment-Control Group design. Since the students were not randomly assigned to the groups, this is considered a quasi-experimental design. This design is illustrated in Figure 3 below.

Figure 3. Study Design



Data Collection

At the outset of the study, teachers were asked to provide background information about the participating students in order to characterize the sample, compare the differences between the study groups and facilitate the analysis of Language, Reading Comprehension, Vocabulary, and Science gains between the study groups. This information included:

- Student grade level
- Student gender
- Student ethnicity
- Student free or reduced lunch participation
- Study group membership (Treatment/BrainPOP or Control)

Teachers were also asked to provide some additional demographic and instructional information regarding Individual Education Plans (IEP) and disabilities. Due to the unavailability of information and/or privacy concerns, many teachers did not provide this additional information. Therefore, there was insufficient information to provide additional analyses examining these specific variables.

Teachers participating in the study were provided with SAT 10 test booklets and administration manuals for their grade level in January 2009. The teachers then administered the SAT 10 pretest (Language, Reading Comprehension, Vocabulary and Science subtests) according to the administration instructions provided. The completed test booklets and answer sheets were then returned for processing. The answer sheets were scanned and entered into a database. Any questions that the students did not answer were scored as incorrect. Students answering fewer than four questions were removed from the analysis. All data was reviewed and checked for accuracy before scoring and analysis.

At the conclusion of the school year at the end of May or in June 2009, following approximately 16-20 weeks of instruction, teachers administered the SAT 10 posttest (Language, Reading Comprehension, Vocabulary and Science subtests). The SAT 10 pretest and posttest results were compared as a basis for evaluating the growth reported in this study.

Findings

Measuring Growth

The growth in Science, Language, Reading Comprehension, and Vocabulary skills for the Treatment (BrainPOP*)* Group and the Control Group was compared using a statistical procedure known as analysis of covariance (ANCOVA). This approach provides an accurate way to compare growth over time controlling for any potential differences in student skills between the two study groups that may have been present at the beginning of the study. Any differences in skill levels between the BrainPOP Users Group and Control Group that may have existed at the beginning of the study were controlled for to ensure that any differences in subsequent growth were the result of BrainPOP use and not merely the result of differences that existed at the start of the study. Using this method, we were able to compare differences as if the two groups were matched in initial Reading, Language, Vocabulary and Science proficiency. While no procedure can completely eliminate differences that may exist at the outset of a study, ANCOVA is widely recognized as an effective way to control for differences.

Only students for whom matched pretest and posttest results were available were included in the analysis. The analysis looked only at those students who had taken the SAT 10 at the beginning of the second semester of the school year (pretest) and those who had taken the SAT 10 at the end of the school year (posttest). Students who left the class during this period or who joined the class during this period were not included in the growth comparisons.

Pre-Post Growth for BrainPOP users

Students who were in classes that used BrainPOP showed substantial growth from pre- to posttest in Vocabulary, Language, Reading Comprehension and Science. During the course of the study, students in classes using BrainPOP increased their SAT 10 Language scale-scores by 24 points (Mean pretest=636; Mean posttest score=660); their Reading Comprehension scale-scores by 17 points (Mean pretest=658; Mean posttest score=675), their Vocabulary scale-scores by 11 points (Mean pretest=658; Mean posttest score=669) and their Science scale-scores by 17 points (Mean pretest=651; Mean posttest score=668) . This level of growth is about twice the level of one-year growth in Science seen for 3rd, 5th and 8th grade students in the national sample of students included in the Stanford 10 norm group. The level of growth seen in Language and Reading Comprehension was equal to or about one-and-a-half times the annual gain seen in Language and Reading Comprehension among 3rd, 5th and 8th graders in the national sample (Harcourt Assessment, 2002).

To better understand the magnitude of growth for students in BrainPOP classrooms we looked at the "effect size", a common metric that can be used to evaluate the amount of growth across studies, when different measures are used. We found effect sizes of +.37 for Reading, +.47 for Language, +.19 for Vocabulary and +.36 for Science (unadjusted for Control Group differences). This indicates that students in classes that used BrainPOP showed substantial growth in Reading, Language, Vocabulary and Science from the beginning of the second semester of the school year to the end of the school year.

While the growth achieved by students using BrainPOP is an important indicator of the effectiveness of BrainPOP, a more complete way to assess growth is to compare the growth achieved by students in classes using BrainPOP to students in classes that did not use BrainPOP. This allows us to see the unique contribution BrainPOP made to students' growth.

Overall Comparison of Growth

The overall growth in Science, Reading Comprehension, Language and Vocabulary skills as measured by the Language, Reading Comprehension, Vocabulary and Science subtests of the SAT 10 for those students in classes using the BrainPOP program (Treatment Group) were compared to the Language, Reading Comprehension, Vocabulary and Science subtests of the SAT 10 of those students in classes who did not use BrainPOP (Control Group). Multivariate Analysis of Covariance (MANCOVA) was used to evaluate the difference in a composite Science, Reading Comprehension, Language and Vocabulary skill score (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for the initial Reading, Language, Vocabulary and Science levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and Control Group on the same baseline. The comparisons were based on 272 BrainPOP Group students and 275 Control Group students for whom all four pretest measures and all four posttest measures were available.

The results show a significant difference in a composite of the SAT 10 Language, Reading Comprehension, Vocabulary and Science subtest posttest scores between the BrainPOP Group and the Control Group ($df=4/538$; $F=14.81$; $p<.001$) when initial Reading, Language and Science skills are controlled. BrainPOP use accounted for approximately 10% of the variation in a composite of the Language, Reading Comprehension, Vocabulary and Science SAT 10 posttest scores. (Eta squared = .10). This means that about 10% of the growth in overall Reading, Language, Vocabulary and Science skills can be explained by whether or not teachers and their students used BrainPOP. The results, using Pillai's Trace, are summarized in Table 6 below. (The results for Wilks Lamda and Hotelling's T are not reported since with only two groups the results are the same as those shown for Pillai's Trace.)

Table 6. Multivariate Analysis of Covariance Comparison of BrainPOP Group and Control Group Language, Reading Comprehension, Vocabulary and Science Posttest Scores

Effect		Value	F	Hypothesis df	Error df	Significance	Partial Eta Squared
Intercept	Pillai's Trace	.189	31.43[4]	4	538	.001	.19
Language Pretest	Pillai's Trace	.079	11.59	4	538	.001	.08
Reading Pretest	Pillai's Trace	.165	26.6[3]	4	538	.001	.17
Vocabulary Pretest	Pillai's Trace	.268	49.2[1]	4	538	.001	.27
Science Pretest	Pillai's Trace	.234	41.03	4	538	.000	.23
Study Group	Pillai's Trace	.099	14.81	4	538	.001	.10

To provide a more complete understanding of these results for the separate Language, Reading, Vocabulary, and Science skill areas, the individual effects for the Language, Reading Comprehension, Vocabulary, and Science subtests were examined separately using ANCOVA (see Table 7).

Table 7. Analysis of Covariance Comparison of BrainPOP Group and Control Group Language, Reading Comprehension, Vocabulary and Science Posttest Scores

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Signi-ficance	Eta Squared
Corrected Model	Vocabulary Posttest	1.487E6	5	297320.711	345.086	.001	.76
	Reading Posttest	698344.322[b]	5	139668.864	190.513	.001	.64
	Science Posttest	743479.617[c]	5	148695.923	256.376	.001	.70
	Language Posttest	639721.968[d]	5	127944.394	125.248	. .001	.54
Intercept	Vocabulary Posttest	3325.465	1	3325.465	3.860	.050	.01
	Reading Posttest	40021.854	1	40021.854	54.591	.001	.09
	Science Posttest	38328.088	1	38328.088	66.084	.001	.11
	Language Posttest	42566.886	1	42566.886	41.670	.001	.07
Language Pretest	Vocabulary Posttest	11101.560	1	11101.560	12.885	.001	.02
	Reading Posttest	13732.825	1	13732.825	18.732	.001	.03
	Science Posttest	3791.275	1	3791.275	6.537	.001	.01
	Language Posttest	37282.250	1	37282.250	36.497	. .001	.06
Reading Pretest	Vocabulary Posttest	11649.491	1	11649.491	13.521	.001	.02
	Reading Posttest	76325.373	1	76325.373	104.110	.001	.16
	Science Posttest	3257.135	1	3257.135	5.616	.018	.01
	Language Posttest	17234.283	1	17234.283	16.871	.001	.03

Vocabulary Pretest	Vocabulary Posttest	151267.520	1	151267.520	175.569	.001	.25
	Reading Posttest	5820.469	1	5820.469	7.939	.005	.01
	Science Posttest	20907.873	1	20907.873	36.049	.001	.06
	Language Posttest	14791.326	1	14791.326	14.480	.001	.03
Study Group	Vocabulary Posttest	3713.435	1	3713.435	4.310	.038	.01
	Reading Posttest	30786.692	1	30786.692	41.994	.001	.07
	Science Posttest	21141.548	1	21141.548	36.451	.001	.06
	Language Posttest	8094.852	1	8094.852	7.924	.005	.01
Error	Vocabulary Posttest	466116.889	541	861.584			
	Reading Posttest	396617.499	541	733.119			
	Science Posttest	313775.370	541	579.991			
	Language Posttest	552645.491	541	1021.526			
Total	Vocabulary Posttest	2.484E8	547				
	Reading Posttest	2.482E8	547				
	Science Posttest	2.440E8	547				
	Language Posttest	2.389E8	547				
Corrected Total	Vocabulary Posttest	1952720.442	546				
	Reading Posttest	1094961.820	546				
	Science Posttest	1057254.987	546				
	Language Posttest	1192367.459	546				

Table 8. Descriptive Statistics Comparison of BrainPOP Group and Control Group Language, Reading Comprehension, Vocabulary, and Science Posttest Scores (Adjusted for Pretest Covariate)

Dependent Variable	Group	N	Mean SAT 10	Standard Deviation SAT 10
Language Posttest	BrainPOP	272	663.17	49.25
	Control	275	655.26	43.54
	Total	547	659.22	46.73
Reading Comprehension Posttest	BrainPOP	272	679.86	45.08
	Control	275	664.44	42.51
	Total	547	672.15	44.78
Vocabulary Posttest	BrainPOP	272	673.87	58.66
	Control	275	668.52	60.16
	Total	547	671.19	59.80
Science Posttest	BrainPOP	272	672.85	45.02
	Control	275	660.08	41.02
	Total	547	666.46	44.00

Language Growth

The SAT 10 Language subtest scores, for those students in classes using BrainPOP (Treatment Group) were compared to the SAT 10 Language subtest scores of those students in classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Language subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results show a significant difference in Language scores between the BrainPOP Group and the Control Group (df=1/547; F=7.92; p<.01) when initial Language proficiency is controlled. The average Language subtest score for students in the BrainPOP Group (Mean= 663.17) was significantly greater than the average Language subtest score achieved by students in the Control Group (Mean= 655.26). This represents an effect size of +.17. The results are summarized in Table 7 and 8 (see above).

Reading Comprehension Growth

The SAT 10 Reading Comprehension subtest scores, for those students in classes using BrainPOP (Treatment Group) were compared to the SAT 10 Reading Comprehension subtest scores of those students in classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Reading subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results show a significant difference in Reading Comprehension between the BrainPOP Group and the Control Group (df=1/547; F=41.99; p<.001) when initial Reading proficiency is controlled. The average Reading Comprehension subtest score for students in the BrainPOP Group (Mean= 679.86) was significantly greater than the average Reading Comprehsnion subtest score achieved by students in the Control Group (Mean= 664.44). This represents an effect size of +.34. The results are summarized in Table 7 and 8 (see above).

Vocabulary Growth

The SAT 10 Vocabulary subtest scores, for those students in classes using BrainPOP (Treatment Group) were compared to the SAT 10 Vocabulary subtest scores of those students in classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Vocabulary subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results show a significant difference in Vocabulary achievement between the BrainPOP Group and the Control Group (df=1/547; F=4.31; p<.05) when initial Vocabulary proficiency is controlled. The average Vocabulary subtest score for students in the BrainPOP Group (Mean= 673.87) was significantly greater than the average Vocabulary subtest score achieved by students in the Control Group (Mean= 668.52). This represents an effect size of +.09. The results are summarized in Table 7 and 8 (see above).

Science Growth

The SAT 10 Science subtest scores, for those students in classes using BrainPOP (Treatment Group) were compared to the SAT 10 Science subtest scores of those students in classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Science subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results show a significant difference in Science achievement between the BrainPOP Group and the Control Group (df=1/547; F=36.45; p<.001) when initial Science proficiency is controlled. The average Science subtest score for students in the BrainPOP Group (Mean= 672.85) was significantly greater than the average Science subtest score achieved by students in the Control Group (Mean= 660.08). This represents an effect size of +.29. The results are summarized in Table 7 and 8 (see above).

Results by Grade level

We analyzed the results for each of the three grade levels examined in this study: 3rd, 5th and 8th grades. The results are presented separately for each grade level below.

Grade 3 Overall Comparison of Growth

The overall growth in Science, Reading Comprehension, Language and Vocabulary skills as measured by the Language, Reading Comprehension, Vocabulary and Science subtests of the SAT 10 for those students in grade 3 classes using the BrainPOP program (Treatment Group) were compared to the Language, Reading Comprehension, Vocabulary and Science subtests of the SAT 10 of those students in grade 3 classes who did not use BrainPOP (Control Group). Multivariate

Analysis of Covariance (MANCOVA) was used to evaluate the difference in a composite Science, Reading Comprehension, Language and Vocabulary skill score (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for the initial Reading, Language, Vocabulary and Science levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and Control Group on the same baseline. The comparisons were based on 81 BrainPOP Group students and 76 Control Group students in grade 3 for whom all four pretest measures and all four posttest measures were available.

The results for grade 3 show a significant difference in a composite of the SAT 10 Language, Reading Comprehension, Vocabulary and Science subtest posttest scores between the BrainPOP Group and the Control Group (df=4/148; F=2.52; p<.05) when initial Reading, Language and Science skills are controlled. BrainPOP use accounted for approximately 6% of the variation in a composite of the Language, Reading Comprehension, Vocabulary and Science SAT 10 posttest scores. (Eta squared = .06). This means that about 6% of the growth in overall Reading, Language, Vocabulary and Science skills can be explained by whether or not teachers and their students used BrainPOP. The results, using Pillai's Trace, are summarized in Table 9 below.

Table 9. Multivariate Analysis of Covariance Comparison of BrainPOP Group and Control Group Language, Reading Comprehension, Vocabulary and Science Posttest Scores for Grade 3 Students

Effect		Value	F	Hypothesis df	Error df	Significance	Partial Eta Squared
Intercept	Pillai's Trace	.157	6.89	4.00	148.00	.001	.16
Language Pretest	Pillai's Trace	.095	3.89	4.00	148.00	.005	.10
Reading Pretest	Pillai's Trace	.112	4.67	4.00	148.00	.001	.11
Vocabulary Pretest	Pillai's Trace	.129	5.50	4.00	148.00	.001	.13
Science Pretest	Pillai's Trace	.186	8.46	4.00	148.00	.001	.19
Study Group	Pillai's Trace	.064	2.52	4.00	148.00	.044	.06

To provide a more complete understanding of these results for the separate Language, Reading, Vocabulary, and Science skill areas, the individual effects for the Language, Reading

Comprehension, Vocabulary, and Science subtests were examined separately for grade 3 students using ANCOVA (see Table 10).

Table 10. Analysis of Covariance Comparison of BrainPOP Group and Control Group Language, Reading Comprehension, Vocabulary and Science Posttest Scores for Grade 3 Students

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Signi-ficance	Eta Squared
Corrected Model	Reading Posttest	155527.713[a]	5	31105.543	22.313	.000	.425
	Science Posttest	96429.638[b]	5	19285.928	29.092	.000	.491
	Vocabulary Posttest	154104.934[c]	5	30820.987	34.307	.000	.532
	Language Posttest	122074.980[d]	5	24414.996	25.927	.000	.462
Intercept	Reading Posttest	6952.235	1	6952.235	4.987	.027	.032
	Science Posttest	17124.562	1	17124.562	25.832	.000	.146
	Vocabulary Posttest	5356.831	1	5356.831	5.963	.016	.038
	Language Posttest	7014.878	1	7014.878	7.449	.007	.047
Language Pretest	Reading Posttest	1691.527	1	1691.527	1.213	.272	.008
	Science Posttest	508.761	1	508.761	.767	.382	.005
	Vocabulary Posttest	137.796	1	137.796	.153	.696	.001
	Language Posttest	11458.615	1	11458.615	12.168	.001	.075
Reading Pretest	Reading Posttest	20189.725	1	20189.725	14.483	.000	.088
	Science Posttest	961.303	1	961.303	1.450	.230	.010
	Vocabulary Posttest	2153.018	1	2153.018	2.397	.124	.016
	Language Posttest	9052.441	1	9052.441	9.613	.002	.060
Vocabulary Pretest	Reading Posttest	1035.159	1	1035.159	.743	.390	.005
	Science Posttest	3811.507	1	3811.507	5.749	.018	.037
	Vocabulary Posttest	14432.546	1	14432.546	16.065	.000	.096
	Language Posttest	187.374	1	187.374	.199	.656	.001
Study Group	Reading Posttest	10167.139	1	10167.139	7.293	.008	.046
	Science Posttest	439.814	1	439.814	.663	.417	.004
	Vocabulary Posttest	4533.589	1	4533.589	5.046	.026	.032
	Language Posttest	58.604	1	58.604	.062	.803	.000
Error	Reading Posttest	210505.277	151	1394.075			
	Science Posttest	100102.616	151	662.931			
	Vocabulary Posttest	135656.989	151	898.391			
	Language Posttest	142191.949	151	941.669			
Total	Reading Posttest	6.677E7	157				

	Science Posttest	6.343E7	157				
	Vocabulary Posttest	6.007E7	157				
	Language Posttest	6.391E7	157				
Corrected Total	Reading Posttest	366032.990	156				
	Science Posttest	196532.254	156				
	Vocabulary Posttest	289761.924	156				
	Language Posttest	264266.930	156				

Table 11. Descriptive Statistics Comparison of BrainPOP Group and Control Group Language, Reading Comprehension, Vocabulary, and Science Posttest Scores for Grade 3 Students (Adjusted for Pretest Covariate)

Dependent Variable	Group	N	Mean SAT 10	Standard Deviation SAT 10
Language Posttest	BrainPOP	81	637.32	47.41
	Control	76	636.04	33.33
	Total	157	636.68	41.16
Reading Comprehension Posttest	BrainPOP	81	658.47	52.46
	Control	76	641.71	40.79
	Total	157	650.09	48.44
Vocabulary Posttest	BrainPOP	81	622.50	47.42
	Control	76	611.31	34.87
	Total	157	616.91	43.10
Science Posttest	BrainPOP	81	636.32	39.10
	Control	76	632.83	30.49
	Total	157	634.58	35.49

Language Growth

The SAT 10 Language subtest scores, for those students in grade 3 classes using BrainPOP (Treatment Group) were compared to the SAT 10 Language subtest scores of those students in grade 3 classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Language subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results show that there were no significant difference in Language scores between the BrainPOP Group and the Control Group (df=1/157; F=.062; p>.05) in grade 3 when initial Language proficiency is controlled. However, the mean scores for the Treatment and Control Groups suggest that grade 3 students using BrainPOP (Mean= 637.32) performed slightly better than grade 3 students who did not use BrainPOP (Mean= 636.04). The results are summarized in Table 9 and 10 (see above).

Reading Comprehension Growth

The SAT 10 Reading Comprehension subtest scores, for those students in grade 3 classes using BrainPOP (Treatment Group) were compared to the SAT 10 Reading Comprehension subtest scores of those students in grade 3 classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Reading Comprehension subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results show a significant difference in Reading Comprehension between the BrainPOP Group in grade 3 and the Control Group in grade 3 (df=1/157; F=7.293; p<.01) when initial Reading proficiency is controlled. The average Reading subtest score for grade 3 students in the BrainPOP Group (Mean= 658.47) was significantly greater than the average Reading subtest score achieved by grade 3 students in the Control Group (Mean= 641.71). The results are summarized in Tables 10 and11 (see above).

Vocabulary Growth

The SAT 10 Vocabulary subtest scores, for those students in grade 3 classes using BrainPOP (Treatment Group) were compared to the SAT 10 Vocabulary subtest scores of those students in grade 3 classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Vocabulary subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results for grade 3 show a significant difference in Vocabulary achievement between the *BrainPOP* Group and the Control Group (df=1/157; F=5.05; p<.05) when initial Vocabulary proficiency is controlled. The average Vocabulary subtest score for grade 3 students in the BrainPOP Group (Mean= 622.50) was significantly greater than the average Vocabulary subtest score achieved by grade 3 students in the Control Group (Mean= 611.31). The results are summarized in Tables 10 and 11 (see above).

Science Growth

The SAT 10 Science subtest scores, for those students in grade 3 classes using BrainPOP (Treatment Group) were compared to the SAT 10 Science subtest scores of those students in grade 3 classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Science subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the *BrainPOP* Group and the Control Group on the same baseline.

The results for grade 3 show no significant differences in Science achievement between the BrainPOP Group and the Control Group (df=1/167; F=.663; p>.05) when initial Science proficiency is controlled. However, the average Science subtest score for grade 3 students in the BrainPOP Group (Mean= 636.32) was slightly greater than the average Science subtest score achieved by grade 3students in the Control Group (Mean= 632.83). The results are summarized in Tables 10 and 11 (see above).

Grade 5 Overall Comparison of Growth

The overall growth in Science, Reading Comprehension, Language and Vocabulary skills as measured by the Language, Reading Comprehension, Vocabulary and Science subtests of the SAT 10 for those students in grade 5 classes using the BrainPOP program (Treatment Group) were compared to the Language, Reading Comprehension, Vocabulary and Science subtests of the SAT 10 of those students in grade 5 classes who did not use BrainPOP (Control Group). Multivariate Analysis of Covariance (MANCOVA) was used to evaluate the difference in a composite Science, Reading Comprehension, Language and Vocabulary skill score (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for the initial Reading, Language, Vocabulary and Science levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and Control Group on the same baseline. The comparisons were based on 105 BrainPOP Group students and 109 Control Group students in grade 5 for whom all four pretest measures and all four posttest measures were available.

The results for grade 5 show a significant difference in a composite of the SAT 10 Language, Reading Comprehension, Vocabulary and Science subtest posttest scores between the BrainPOP Group and the Control Group ($df=4/205$; $F=12.01$; $p<.001$) when initial Reading, Language and Science skills are controlled. BrainPOP use accounted for approximately 19% of the variation in a composite of the Language, Reading Comprehension, Vocabulary and Science SAT 10 posttest scores. (Eta squared = .19). This means that about 19% of the growth in overall Reading, Language, Vocabulary and Science skills can be explained by whether or not teachers and their students used BrainPOP. The results, using Pillai's Trace, are summarized in Table 12 below.

Table 12. Multivariate Analysis of Covariance Comparison of BrainPOP Group and Control Group Language, Reading Comprehension, Vocabulary and Science Posttest Scores for Grade 5 Students

Effect		Value	F	Hypothesis df	Error df	Significance	Partial Eta Squared
Intercept	Pillai's Trace	.163	9.95[a]	4.00	205.00	.001	.16
Language Pretest	Pillai's Trace	.146	8.74[a]	4.00	205.00	.001	.15
Reading Pretest	Pillai's Trace	.170	10.49[a]	4.00	205.00	.001	.17
Vocabulary Pretest	Pillai's Trace	.183	11.48[a]	4.00	205.00	.001	.18
Science Pretest	Pillai's Trace	.222	14.64[a]	4.00	205.00	.001	.22
Study Group	Pillai's Trace	.190	12.01[a]	4.00	205.00	.001	.19

To provide a more complete understanding of these results for the separate Language, Reading, Vocabulary, and Science skill areas, the individual effects for the Language, Reading Comprehension, Vocabulary, and Science subtests were examined separately for grade 5 students using ANCOVA (see Table 13).

Table 13. Analysis of Covariance Comparison of BrainPOP Group and Control Group Language, Reading Comprehension, Vocabulary and Science Posttest Scores for Grade 5 Students

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Signi-ficance	Eta Squared
Corrected Model	Reading Posttest	115195.940[a]	5	23039.188	71.571	.000	.632
	Science Posttest	233810.810[b]	5	46762.162	77.859	.000	.652
	Vocabulary Posttest	232398.902[c]	5	46479.780	62.920	.000	.602
	Language Posttest	214837.027[d]	5	42967.405	54.900	.000	.569
Intercept	Reading Posttest	10051.146	1	10051.146	31.224	.000	.131
	Science Posttest	616.562	1	616.562	1.027	.312	.005
	Vocabulary Posttest	2.322	1	2.322	.003	.955	.000
	Language Posttest	293.227	1	293.227	.375	.541	.002
Language Pretest	Reading Posttest	1885.187	1	1885.187	5.856	.016	.027
	Science Posttest	4812.712	1	4812.712	8.013	.005	.037
	Vocabulary Posttest	16817.211	1	16817.211	22.766	.000	.099
	Language Posttest	15245.772	1	15245.772	19.480	.000	.086
Reading Pretest	Reading Posttest	12231.510	1	12231.510	37.997	.000	.154
	Science Posttest	2838.669	1	2838.669	4.726	.031	.022
	Vocabulary Posttest	3166.778	1	3166.778	4.287	.040	.020
	Language Posttest	13531.887	1	13531.887	17.290	.000	.077

Vocabulary Pretest	Reading Posttest	5390.112	1	5390.112	16.744	.000	.075
	Science Posttest	4728.024	1	4728.024	7.872	.005	.036
	Vocabulary Posttest	25500.265	1	25500.265	34.520	.000	.142
	Language Posttest	14725.585	1	14725.585	18.815	.000	.083
Study Group	Reading Posttest	7278.771	1	7278.771	22.611	.001	.098
	Science Posttest	22628.386	1	22628.386	37.676	.001	.153
	Vocabulary Posttest	7559.345	1	7559.345	10.233	.002	.047
	Language Posttest	6088.326	1	6088.326	7.779	.006	.036
Error	Reading Posttest	66956.417	208	321.906			
	Science Posttest	124925.383	208	600.603			
	Vocabulary Posttest	153652.392	208	738.713			
	Language Posttest	162790.001	208	782.644			
Total	Reading Posttest	9.462E7	214				
	Science Posttest	9.642E7	214				
	Vocabulary Posttest	9.750E7	214				
	Language Posttest	9.519E7	214				
Corrected Total	Reading Posttest	182152.357	213				
	Science Posttest	358736.193	213				
	Vocabulary Posttest	386051.294	213				
	Language Posttest	377627.028	213				

Table 14. Descriptive Statistics Comparison of BrainPOP Group and Control Group Language, Reading Comprehension, Vocabulary, and Science Posttest Scores for Grade 5 Students (Adjusted for Pretest Covariate)

Dependent Variable	Group	N	Mean SAT 10	Standard Deviation SAT 10
Language Posttest	BrainPOP	105	671.57	39.50
	Control	109	659.86	43.41
	Total	214	665.72	42.11
Reading Comprehension Posttest	BrainPOP	105	670.84	26.13
	Control	109	658.04	29.70
	Total	214	664.44	29.24
Vocabulary Posttest	BrainPOP	105	680.28	40.58
	Control	109	667.24	43.04
	Total	214	673.76	42.57
Science Posttest	BrainPOP	105	681.48	37.13
	Control	109	658.92	40.51
	Total	214	670.20	41.04

Language Growth

The SAT 10 Language subtest scores, for those students in grade 5 classes using BrainPOP (Treatment Group) were compared to the SAT 10 Language subtest scores of those students in Grade 5 classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Language subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results show that there was a significant difference in Language scores between the BrainPOP Group and the Control Group (df=1/213; F=7.78; p>.01) in grade 5 when initial Language proficiency is controlled. The average Language subtest score for grade 5 students in the BrainPOP Group (Mean= 671.57) was significantly greater than the average Language subtest score achieved by grade 5 students in the Control Group (Mean= 659.86). The results are summarized in Tables 13 and 14 (see above).

Reading Comprehension Growth

The SAT 10 Reading Comprehension subtest scores, for those students in grade 5 classes using BrainPOP (Treatment Group) were compared to the SAT 10 Reading Comprehension subtest scores of those students in grade 5 classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Reading subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results show a significant difference in Reading Comprehension between the BrainPOP Group in grade 5 and the Control Group in grade 5 (df=1/213; F=22.61; p<.001) when initial Reading proficiency is controlled. The average Reading subtest score for grade 5 students in the BrainPOP Group (Mean= 670.84) was significantly greater than the average Reading subtest score achieved by grade 5 students in the Control Group (Mean= 658.04). The results are summarized in Tables 13 and 14 (see above).

Vocabulary Growth

The SAT 10 Vocabulary subtest scores, for those students in grade 5 classes using BrainPOP (Treatment Group) were compared to the SAT 10 Language subtest scores of those students in grade 5 classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Vocabulary subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results for grade 5 show a significant difference in Vocabulary achievement between the BrainPOP Group and the Control Group (df=1/213; F=10.23; p<.05) when initial Vocabulary proficiency is controlled. The average Vocabulary subtest score for grade 5 students in the BrainPOP Group (Mean= 680.28) was significantly greater than the average Vocabulary subtest score achieved by grade 5 students in the Control Group (Mean= 667.24).
The results are summarized in Tables 12 and 13 (see above).

Science Growth

The SAT 10 Science subtest scores, for those students in Grade 5 classes using BrainPOP (Treatment Group) were compared to the SAT 10 Science subtest scores of those students in grade 5 classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Science subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results for grade 5 show no significant differences in Science achievement between the BrainPOP Group and the Control Group (df=1/213; F=37.68; p>.001) when initial Science proficiency is controlled. However, the average Science subtest score for grade 5 students in the BrainPOP Group (Mean= 636.32) was slightly greater than the average Science subtest score achieved by grade 5 students in the Control Group (Mean= 632.83). The results are summarized in Tables 13 and 14 (see above).

Grade 8 Overall Comparison of Growth

The overall growth in Science, Reading Comprehension, Language and Vocabulary skills as measured by the Language, Reading Comprehension, Vocabulary and Science subtests of the SAT 10 for those students in grade 8 classes using the BrainPOP program (Treatment Group) were compared to the Language, Reading Comprehension, Vocabulary and Science subtests of the SAT 10 of those students in grade 8 classes who did not use BrainPOP (Control Group). Multivariate Analysis of Covariance (MANCOVA) was used to evaluate the difference in a composite Science, Reading Comprehension, Language and Vocabulary skill score (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for the initial Reading, Language, Vocabulary and Science levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and Control Group on the same baseline. The comparisons were based on 59 BrainPOP Group students and 69 Control Group students in grade 8 for whom all four pretest measures and all four posttest measures were available.

The results for grade 8 show a significant difference in a composite of the SAT 10 Language, Reading Comprehension, Vocabulary and Science subtest posttest scores between the BrainPOP Group and the Control Group (df=4/119; F=4.73; p<.001) when initial Reading, Language and Science skills are controlled. BrainPOP use accounted for approximately 14% of the variation in a composite of the Language, Reading Comprehension, Vocabulary and Science SAT 10 posttest scores. (Eta squared = .14). This means that about 14% of the growth in overall Reading, Language, Vocabulary and Science skills can be explained by whether or not teachers and their students used BrainPOP. The results, using Pillai's Trace, are summarized in Table 15 below.

Table 15. Multivariate Analysis of Covariance Comparison of BrainPOP Group and Control Group Language, Reading Comprehension, Vocabulary and Science Posttest Scores for Grade 8 Students

Effect		Value	F	Hypothesis df	Error df	Significance	Partial Eta Squared
Intercept	Pillai's Trace	.172	6.18	4	119.00	.001	.17
Language Pretest	Pillai's Trace	.182	6.62	4	119.00	.001	.18
Reading Pretest	Pillai's Trace	.126	4.28	4	119.00	.003	.13
Vocabulary Pretest	Pillai's Trace	.199	7.38	4	119.00	.001	.20
Science Pretest	Pillai's Trace	.244	9.60	4	119.00	.001	.24
Study Group	Pillai's Trace	.137	4.7[3]	4	119.00	.001	.14

To provide a more complete understanding of these results for the separate Language, Reading, Vocabulary, and Science skill areas, the individual effects for the Language, Reading

Comprehension, Vocabulary, and Science subtests were examined separately for grade 8 students using ANCOVA (see Table 16).

Table 16. Analysis of Covariance Comparison of BrainPOP Group and Control Group Language, Reading Comprehension, Vocabulary and Science Posttest Scores for Grade 8 Students

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Signi-ficance	Eta Squared
Corrected Model	Reading Posttest	132838.599[a]	5	26567.720	42.420	.001	.635
	Science Posttest	109852.920[b]	5	21970.584	51.355	.001	.678
	Vocabulary Posttest	181347.154[c]	5	36269.431	44.316	.001	.645
	Language Posttest	190480.621[d]	5	38096.124	43.284	.001	.640
Intercept	Reading Posttest	4480.409	1	4480.409	7.154	.009	.055
	Science Posttest	8392.344	1	8392.344	19.617	.001	.139
	Vocabulary Posttest	231.886	1	231.886	.283	.595	.002
	Language Posttest	.223	1	.223	.000	.987	.000
Language Pretest	Reading Posttest	4491.227	1	4491.227	7.171	.008	.056
	Science Posttest	651.064	1	651.064	1.522	.220	.012
	Vocabulary Posttest	3969.751	1	3969.751	4.850	.030	.038
	Language Posttest	21229.394	1	21229.394	24.121	.001	.165
Reading Pretest	Reading Posttest	9932.705	1	9932.705	15.859	.001	.115
	Science Posttest	506.454	1	506.454	1.184	.279	.010
	Vocabulary Posttest	2799.664	1	2799.664	3.421	.067	.027
	Language Posttest	4384.905	1	4384.905	4.982	.027	.039
Vocabulary Pretest	Reading Posttest	953.638	1	953.638	1.523	.220	.012
	Science Posttest	4021.346	1	4021.346	9.400	.003	.072
	Vocabulary Posttest	19796.301	1	19796.301	24.188	.001	.165
	Language Posttest	2775.459	1	2775.459	3.153	.078	.025
Study Group	Reading Posttest	5609.097	1	5609.097	8.956	.003	.068
	Science Posttest	6429.971	1	6429.971	15.030	.001	.110
	Vocabulary Posttest	133.879	1	133.879	.164	.687	.001
	Language Posttest	1299.120	1	1299.120	1.476	.227	.012
Error	Reading Posttest	76409.265	122	626.305			
	Science Posttest	52193.985	122	427.820			
	Vocabulary Posttest	99848.346	122	818.429			
	Language Posttest	107376.848	122	880.138			

Total	Reading Posttest	6.472E7	128				
	Science Posttest	6.238E7	128				
	Vocabulary Posttest	6.794E7	128				
	Language Posttest	5.968E7	128				
Corrected Total	Reading Posttest	209247.864	127				
	Science Posttest	162046.905	127				
	Vocabulary Posttest	281195.500	127				
	Language Posttest	297857.469	127				

Table 17. Descriptive Statistics Comparison of BrainPOP Group and Control Group Language, Reading Comprehension, Vocabulary, and Science Posttest Scores for Grade 8 Students (Adjusted for Pretest Covariate)

Dependent Variable	Group	N	Mean SAT 10	Standard Deviation SAT 10
Language Posttest	BrainPOP	59	684.68	49.93
	Control	69	678.12	45.20
	Total	128	681.40	48.43
Reading Comprehension Posttest	BrainPOP	59	717.25	37.60
	Control	69	703.62	39.51
	Total	128	710.43	40.59
Vocabulary Posttest	BrainPOP	59	728.20	42.81
	Control	69	726.09	49.62
	Total	128	727.15	47.06
Science Posttest	BrainPOP	59	705.03	31.66
	Control	69	690.44	35.11
	Total	128	697.74	35.72

Language Growth

The SAT 10 Language subtest scores, for those students in grade 8 classes using BrainPOP (Treatment Group) were compared to the SAT 10 Language subtest scores of those students in grade 8 classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Language subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results show that there was no significant difference in Language scores between the BrainPOP Group and the Control Group (df=1/127; F=1.48; $p>.05$) in grade 8 when initial Language proficiency is controlled. However, the average Language subtest score for grade 8 students in the BrainPOP Group (Mean= 684.68) was somewhat greater than the average Language subtest score achieved by grade 8 students in the Control Group (Mean= 678.12). The results are summarized in Tables 16 and 17 (see above).

Reading Comprehension Growth

The SAT 10 Reading Comprehension subtest scores, for those students in grade 8 classes using BrainPOP (Treatment Group) were compared to the SAT 10 Reading Comprehension subtest scores of those students in grade 8 classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Reading subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results show a significant difference in Reading achievement between the BrainPOP Group in grade 8 and the Control Group in grade 8 (df=1/127; F=8.96; $p<.01$) when initial Reading proficiency is controlled. The average Reading subtest score for grade 8 students in the BrainPOP Group (Mean= 703.62) was significantly greater than the average Reading subtest score achieved by grade 8 students in the Control Group (Mean= 710.43). The results are summarized in Tables 16 and 17 (see above).

Vocabulary Growth

The SAT 10 Vocabulary subtest scores, for those students in grade 8 classes using BrainPOP (Treatment Group) were compared to the SAT 10 Vocabulary subtest scores of those students in grade 8 classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Vocabulary subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results for grade 8 show no significant differences in Vocabulary achievement between the BrainPOP Group and the Control Group (df=1/21; F=.16; p>.05) when initial Vocabulary proficiency is controlled. The average Vocabulary subtest score for grade 8 students in the BrainPOP Group and the Control Group were nearly identical. The results are summarized in Tables 16 and 17 (see above).

Science Growth

The SAT 10 Science subtest scores, for those students in grade 8 classes using BrainPOP (Treatment Group) were compared to the SAT 10 Language subtest scores of those students in grade 8 classes that did not use BrainPOP (Control Group). ANCOVA was used to evaluate the difference in Science subtest scores (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) controlling for initial proficiency levels of the students (covariate). The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The results for grade 8 show a significant difference in Science achievement between the BrainPOP Group and the Control Group (df=1/127; F=15.03; p>.001), when initial Science proficiency is controlled. The average Science subtest score for grade 8 students in the BrainPOP Group (Mean= 705.03) was significantly greater than the average Science subtest score achieved by grade 8 students in the Control Group (Mean= 690.44). The results are summarized in Tables 16 and 17 (see above).

Gender and Ethnicity Analysis

We examined whether there were any differences in growth between male and female students (main effects) and whether or not there were differences in effectiveness for male and female students in classes using BrainPOP as compared to male and female students in classes not using BrainPOP (interaction effects). We also examined the (main and interaction) effects of BrainPOP usage among students in different ethnic groups

To this end, the overall growth in Reading, Language, Vocabulary and Science skills for students in classes using BrainPOP (Treatment Group) as measured by the SAT 10 subtests was compared to the overall growth in Reading, Language, Vocabulary and Science skills for students in classes that did not use BrainPOP (Control Group) as measured by the SAT 10 considering both the gender and ethnicity of the students. MANCOVA was used to evaluate the difference in a composite score (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) of different gender and ethnic backgrounds (independent variable) controlling for the initial skill levels of the students (covariate) in each group. The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The gender comparisons were based on 190 female students and 168 male students. The ethnicity comparisons were based on 204 students classified as White/Caucasian; 30 students categorized as Black/African American; 73 students classified as Hispanic; 30 students classified as Asian/Pacific Islander; and 21 students classified as Mixed Race or Other. The counts for gender and ethnicity are presented in Table 18 (see below).

Table 18. Number of Students by Gender and Ethnicity

	Group	N
Gender	Female	190
	Male	168
	Total	358
Ethnicity	White/ Caucasian	204
	Black/African American	30
	Hispanic	73
	Asian/Pacific Islander	30
	Mixed Race or Other	21
	Total	358

The main effect for BrainPOP use (Treatment Group), was again confirmed; there was a significant difference in a composite of the SAT 10 Language, Reading Comprehension, and Vocabulary posttest scores between students in the BrainPOP Group and students in the Control Group when initial Reading and Language proficiency levels are controlled (F=8.49; df=4/331 p<.001). BrainPOP use accounted for approximately 9% of the variation in a composite of the SAT 10 Language, Reading Comprehension, Vocabulary and Science posttest scores overall (Eta squared=.09). There were no significant main effects for either gender (F=2.07; df=4/331; p>.05) or ethnicity (F=.99; df=4/1336; p>.05).

The interaction effects for gender by study group (F=.51 df=4/331; p>.05) and ethnicity by study group (F=.69; df=16/1336; p>.05) were not statistically significant. This indicates that BrainPOP wass equally effective with boys and girls and equally effective among students of different ethnicities. The results, using Pillai's Trace, are summarized in Table 19 (see below).

**Table 19. Multivariate Analysis of Covariance
Comparison of BrainPOP Group and Control Group by Gender and Linguistic Background
Language, Vocabulary, Reading and Science Posttest Scores**

Effect		Value	F	Hypothesis df	Error df	Significance	Partial Eta Squared
Intercept	Pillai's Trace	.170	16.91	4	331	.001	.17
Language Pretest	Pillai's Trace	.099	9.1[1]	4	331	.001	.10
Vocabulary Pretest	Pillai's Trace	.231	24.89	4	331	.001	.23
Reading Pretest	Pillai's Trace	.174	17.43	4	331	.001	.17
Science Pretest	Pillai's Trace	.275	31.38	4	331	.001	.28
Study Group	Pillai's Trace	.093	8.49	4	331	.001	.09
Gender	Pillai's Trace	.024	2.07	4	331	.084	.02
Ethnicity	Pillai's Trace	.047	.99	16	1336	.466	.01
Study Group by Gender	Pillai's Trace	.006	.5[1]	4	331	.729	.01
Study Group by/ Linguistic Background	Pillai's Trace	.033	.69	16	1336.	.803	.01

Free and Reduced Lunch Analysis

We examined whether there were any differences in growth between students receiving free or reduced lunch and those who were not receiving this service (main effects). We also looked at whether or not there were differences in effectiveness for free and reduced lunch recipients in classes using BrainPOP as compared to free and reduced lunch recipients in classes not using BrainPOP (interaction effect).

To this end, the overall growth in Reading, Language, Vocabulary and Science skills for students in classes using BrainPOP (Treatment Group) as measured by the SAT 10 subtests was compared to the overall growth in Reading, Language, Vocabulary and Science skills for students in classes that did not use BrainPOP (Control Group) as measured by the SAT 10 considering whether or not the students were receiving free or reduced lunch. MANCOVA was used to evaluate the difference in a composite score (dependent variable) between BrainPOP users and non-BrainPOP users (independent variable) of students receiving free or reduced lunch and those not receiving this service (independent variable), controlling for the initial skill levels of the students (covariate) in each group. The SAT 10 pretest scores were used as the covariate to place students in the BrainPOP Group and the Control Group on the same baseline.

The free and reduced lunch comparisons were based on 239 students who were not receiving free or reduced lunch and 149 students who were receiving free or reduced lunch.

The main effect for BrainPOP use (Treatment Group), was again confirmed; there was a significant difference in a composite of the SAT 10 Language, Reading Comprehension, and Vocabulary posttest scores between students in the BrainPOP Group and students in the Control Group when initial Reading and Language proficiency levels are controlled ($F=12.83$; $df=4/377$ $p<.001$). BrainPOP use accounted for approximately 12% of the variation in a composite of the SAT 10 Language, Reading Comprehension, Vocabulary and Science posttest scores overall (Eta squared=.12).

The interaction effect for free or reduced lunch status by study group ($F=.71$ $df=4/377$; $p>.05$) was not statistically significant. BrainPOP was equally effective with students receiving free or reduced lunch and those who did not receive free or reduced lunch. The results, using Pillai's Trace, are summarized in Table 20 (see below).

Table 20. Multivariate Analysis of Covariance
Comparison of BrainPOP Group and Control Group by Free or Reduced Lunch Status
Language, Vocabulary, Reading and Science Posttest Scores

Effect		Value	F	Hypothesis df	Error df	Significance	Partial Eta Squared
Intercept	Pillai's Trace	.229	27.94	4	377	.001	.23
Language Pretest	Pillai's Trace	.079	8.1	4	377	.001	.08
Vocabulary Pretest	Pillai's Trace	.247	30.85	4	377	.001	.25
Reading Pretest	Pillai's Trace	.156	17.4	4	377	.001	.16
Science Pretest	Pillai's Trace	.251	31.61	4	377	.001	.25
Study Group	Pillai's Trace	.120	12.83	4	377	.001	.12
Free or Reduced Lunch	Pillai's Trace	.012	1.18	4	377	.318	.01
Study Group by/ Free or Reduced Lunch	Pillai's Trace	.007	.71	4	377	.585	.01

Teacher Survey

At the conclusion of the study, teachers using BrainPOP were asked to complete a 32 question survey to provide background information about themselves and their schools and to find out about their perceptions of BrainPOP effectiveness. The survey was conducted online over a period of three weeks in June 2009. Twenty (of the twenty four participating BrainPOP using teachers) teachers returned completed surveys. Highlights of the results are provided below.

Teacher Profile. Four fifths (80%) of the teachers responding indicated that they were female and one fifth (20%) indicated they were male. More than one third (35%) of the teachers indicated that they were between the ages of 41 and 50, while a quarter (25%) of the group indicated that they were between the ages of 51 and 60 and another fifth (20%) indicated that they were between the age of 31 and 40. About four fifths (85%) of the participating teachers reported their ethnic background was white/Caucasian; the remaining teachers categorized themselves as Asian, Native American or Other.

About half (45%) of the teachers reported that they had 10 or more years of teaching experience, with about a quarter (30% indicating they had 2-4 years of experience and another quarter indicating they had 5-9 years of teaching experience.

School Characteristics. Three quarters (75%) of the teachers responding indicated that they taught in a school with 800 or more students with another fifth (20%) indicating they taught in a school with 600-800 students. About two thirds (65%) reported that they taught in a class with 21-30 students and another quarter (25%) reported teaching in a class with 11-20 students.

Perceptions of Effectiveness. When asked to indicate how likely they were to use BrainPOP in the future, all (100%) of the teachers indicated that they would definitely use BrainPOP in the future. Similarly, all (100%) of the teachers indicated that they would definitely recommend BrainPOP to others, when asked their likelihood of recommending BrainPOP for use by others.

Ninety percent of the teachers felt that BrainPOP was good or excellent in increasing students' cognitive/intellectual growth. Similarly, 90% felt that BrainPOP was good or excellent at improving student's attitudes toward school and learning. When asked how well BrainPOP helps their students improve their Science skills, more than four fifths (85% indicated extremely well or very well, with the remaining respondents indicating that it increased these skills moderately well. When asked how well BrainPOP helps their students improve their Reading and Language arts skills, three fifths (60%) indicated extremely well or very well, with the remaining respondents indicating that it increased these skills moderately well. Only about one quarter (25%) of the respondents felt that BrainPOP worked extremely well or very well in increasing their students mathematics skills, with the majority of teachers (60%) reporting that BrainPOP worked moderately well in improving these skills.

Four fifths (80%) of those teachers responding indicated that BrainPOP was extremely well or very well aligned with the instructional goals and standards that guide their instruction; the remaining respondents felt that BrainPOP was moderately well aligned to those goals and standards.

BrainPOP Use. About a third (35%) of the teachers indicated that they used BrainPOP about 3 hours per week, about a third (35%) indicated that they used BrainPOP about 2 hours per week and another 20% reported using BrainPOP about 1 hour or less per week. One teacher reported using BrainPOP 6 or more hours per week.

BrainPOP movies enjoyed the highest use of any of the BrainPOP components with 90% of the teachers reporting use of the movies either every week or every two weeks. BrainPOP quizzes were also a frequently used component with three quarters (75%) of the teachers indicating that they used them every week or every two weeks. BrainPOP activity pages were also used by three quarters (75%) of the educators every week or every two weeks. BrainPOP Educators was used every week or every two weeks by about two fifths (40%) of the teachers.

SUMMARY AND DISCUSSION

This study examined the effectiveness of BrainPOP, a web-based animated instructional tool that supports educators and engages students. The study examined the effectiveness of BrainPOP by looking at the growth of Science, Reading Comprehension, Language and Vocabulary skills among students in classes using BrainPOP and comparing the level of growth to students in classes that did not use BrainPOP. The study examined differences in growth among the total group of students and between students of different gender and ethnic backgrounds as well as whether or not they received free or reduced lunch.

Study Design

Between January 2009 and June 2009, approximately 1,100 students in 46 classrooms in Palm Beach County, Florida and New York City, New York participated in a controlled study of BrainPOP effectiveness. Classes either used BrainPOP (Treatment Group) or did not use BrainPOP (Control Group) during the course of the study. Students in both the BrainPOP users and the Control Groups were similar in ability and demographically.

Using a quasi-experimental, pre-post design, this study compared the growth in Language, Reading Comprehension, and Vocabulary skills between students in classes using BrainPOP (Treatment Group) and a comparable group of students in classes that did not use BrainPOP (Control Group). Student growth in Language, Reading comprehension, Vocabulary and Science was measured by comparing scores from the Stanford 10 Achievement Test™, Abbreviated Battery (SAT 10), at the beginning of the second semester of the school year and end of the 2008-2009 school year.

Students in both the treatment (BrainPOP users) and the Control Groups took a pretest (SAT 10) at the beginning of the second semester of the school year to obtain a baseline measure of student Science, Reading Comprehension, Language and Vocabulary skills. Students in the Treatment Group then received instruction that included BrainPOP, while those in the Control Group received instruction without the benefit of BrainPOP. At the end of the school year (May and June of 2009), students in both the BrainPOP Group and the Control Group were administered a posttest (SAT 10). The results from the SAT 10 pretest and posttest were compared statistically to determine the level of growth in Language, Reading Comprehension, Vocabulary and Science skills. The study controlled for any initial differences in the Science, Reading Comprehension, Language and Vocabulary skills of students in the treatment and Control Groups during analysis using ANCOVA.

Summary of Results

Pre-Post Growth for BrainPOP users

The students in classes using BrainPOP showed significant gains from pre- to posttest in Language, Reading comprehension, Vocabulary and Science. From January 2009 to June 2009, students in BrainPOP classrooms increased their SAT 10 Language scale-scores by 24 points, their Reading Comprehension scale-scores by 17 points, their Vocabulary scale-scores by 11 points and their Science scores by 17 points .

To better understand the magnitude of growth for BrainPOP users we looked at the "effect size," a common metric that can be used to evaluate the amount of growth across studies, when different measures are used. We found effect sizes of +.37 for Reading, +.47 for Language, +.19 for Vocabulary and +.36 for Science (unadjusted for Control Group differences). This indicates that students in classes that used BrainPOP showed substantial growth in Reading, Language, Vocabulary and Science from the beginning to the end of the study.

Comparison of Growth Between BrainPOP users and the Control Group

The students using BrainPOP showed significantly greater gains in Science, Language, and Reading comprehension than those students in the Control Group who were not using BrainPOP.
This study controlled for any initial differences in the Science, Reading Comprehension, Language and Vocabulary skills of students in the BrainPOP Users and Control Groups using a statistical procedure known as analysis of covariance (ANCOVA). When controlling for students' initial ability

using ANCOVA, the findings of this study are quite significant. BrainPOP users finished the year with scores that were 16 scale-score points higher in Reading Comprehension, 13 scale-score points higher in Science, 8 scale-score points higher in Language, and 5 points higher in Vocabulary than the Control Group on the SAT 10 assessments.

We found effect sizes of +.34 for Reading Comprehension, +.29 for Science, +.17 for Language, and + .09 for Vocabulary. This is a large effect, indicating that the students in classes that used BrainPOP performed well above the students in classes that did not use BrainPOP, particularly in Reading Comprehension, Science and Language.

The results by grade level indicated that BrainPOP produced significantly greater increases in a composite of Reading, Language, Vocabulary and Science. The effects were most dramatic at the 5th grade level, where BrainPOP accounted for 19% of the variance in growth. This means that 19% of the growth in a composite of Reading, Language, Vocabulary and Science can be accounted for by BrainPOP alone.

BrainPOP was found to be equally effective for boys and girls, students of different ethnicities and both students receiving free or reduced lunch services and those not receiving this service. The interaction between BrainPOP use and gender, ethnicity and free or reduced lunch status was not statistically significant.

Teacher Perceptions of Effectiveness

When asked to indicate how likely they were to use *BrainPOP* in the future, all (100%) of the teachers indicated that they would definitely use *BrainPOP* in the future. Similarly, all (100%) of the teachers indicated that they would definitely recommend BrainPOP to others, when asked their likelihood of recommending *BrainPOP* for use by others.

Conclusion

Students enrolled in classes using BrainPOP achieved significantly greater gains in Science, Language, and Reading comprehension than students enrolled in classes that did not use BrainPOP was found to be equally effective for students regardless of gender, ethnicity and whether or not they received free or reduced lunch.

The students who used BrainPOP showed significant gains from pre- to posttest in Science, Language, Reading comprehension, and Vocabulary. From January 2009 to June 2009, students in BrainPOP classrooms increased their SAT 10 scale-scores by 11-24 points. Students in BrainPOP classrooms did significantly better than their peers in classrooms that did not use BrainPOP, scoring 5-16 points higher than their peers on the SAT 10, when adjusted for initial pretest differences. This represents an effect size of +.29 for Science, +.34 for Reading Comprehension, +.17 for Language and +.09 for Vocbulary. This is a substantial effect, indicating that the students who used BrainPOP performed better than those the students who did not use BrainPOP, particularly in Reading Comprehension, Science and Language.

These findings are particularly significant for two reasons: First, the study was conducted for only a single semester, from January to June. The level of growth seen is particularly significant in light of the fact that it reflects the effects of only 16-20 weeks of instruction. Second, the SAT 10 (the pre- and posttest measure used) is a standardized, national, norm-referenced test; it measures achievement in broad strokes to allow comparison across diverse curricular and instructional settings. Therefore, the SAT 10 may not be as instructionally sensitive to the specific areas addressed in the BrainPOP instructional program. So, the level of growth seen for BrainPOP users is very significant in light of the nature of the SAT 10.

The findings of this study demonstrate that students using BrainPOP can make significant gains in Reading, Language, Vocabulary and Science during one school year's time and make significantly greater gains than students who do not use BrainPOP.

References

Baddeley, A. (1992) "Working memory." *Science*, 255, pp.556-559.

Baddeley, A. (1999) *Human memory*. Boston: Allyn and Bacon.

Betrancourt, M. (2005) "The animation and interactivity principles in multimedia learning." In R. E. Mayer (Ed.). *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.

Chandler, Panda Sweller, J. (1991). "Cognitive load theory and the format of instruction." *Cognition and Instruction*, 8, pp.293-332.

Chi, M., Glaser R., and Rees, E. (1982) "Expertise in problem solving." In R. Sternberg (Ed.) *Advances in the psychology of human intelligence* (pp.7-75) Hillsdale, NJ. Erlbaum

Clark, R.E. and Feldon, D.F. (2005) "Five common but questionable principles of multimedia learning." In R. E. Mayer (Ed.). *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.

Craig, S.D., Gholson B., and Driscoll, D.M. (2002). "Animated pedagogical agent in multimedia education environments: Effects of agent properties, picture features, and redundancy." *Journal of Educational Psychology*, 94, pp.428-434.

Gardner, H. (1993) *Frames of Mind: the theory of multiple intelligences*. New York: Basic Books.

Gardner, H. (1999) *Intelligence reframed: Multiple intelligences for the 21st Century*. New York: Basic Books.

Gee, J.P. (2005): "Learning by design: Good video games as learning machines." *E-Learning*. Vol.2 pp. 5-16.

Harcourt Assessment (2002) Stanford *Achievement Test Series™, Tenth Edition, Technical manual*: San Antonio, Texas: Harcourt Assessment.

Institute for Learning Styles (2008) "Overview of the seven perceptual styles." Retrieved August 28, 2008 from http://www.learning styles .org

Kalyuga, S. (2005) "Prior knowledge principle in multimedia learning." In R. E. Mayer (Ed.). *The Cambridge Handbook of Multimedia Learning.* New York: Cambridge University Press.

Kalyuga, S.,Chandler, P. and Sweller, J. (1999). "Managing spilt attention and redundancy in multimedia instruction." *Applied Cognitive Psychology*, 13, pp.351-371.

Kolb, D. A., Boyatzis, R., & Mainemelis, C. (2001). "Experiential learning theory: Previous research and new directions." In R. Sternberg and L. Zhang (Eds.) *Perspectives on cognitive learning, and thinking styles.* Mahwah, NJ: Erlbaum

Lee, J., Grigg, W., and Dion, G. (2007). *The Nation's Report Card: Mathematics 2007(NCES 2007–494).*National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

Lowe, K. (2002) *What's the story: making meaning in primary classrooms.* ERIC Document No. ED468691

Miller, (1956) "The magic number of seven plus or minus two. Some limits on our capacity for processing information." *Psychological Review*; Vole, 63, 31-97.

Mayer, R.E. (2003*) Learning and Instruction.* Upper Saddle River, NJ: Prentice Hall.

Mayer, R.E. (2005) "Introduction to multimedia learning." in R. E. Mayer (Ed.). *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.

Mayer, R.E. (2005a). "Principles of multimedia learning based on social cues: personalization, voice, and image principles." In R. E. Mayer, (Ed.) *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.

Mayer, R.E. and Chandler, P. (2001) "When learning is just a click away: Does simple interaction foster deeper understanding of multimedia messages?" *Journal of Educational Psychology*, Vol. 93, pp. 390-397.

Mayer, R.E. and Sims, V.K. (1994) "For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning." *Journal of Educational Psychology*, 86, pp.389-401

Mayer, R.E., Sobko, K., and Mautone, P.D. (2003) "Social cues in multimedia learning: role of speaker's voice." *Journal of Educational Psychology*, 95, pp. 419-425.

Park, O. (1994). "Dynamic visual displays in media-based instruction." *Educational Technology*, pp. 21–25.

Perkins, D. (1992) *Smart Schools: Better thinking and learning for every child*. New York: The Free Press.

Pollock, E., Chandler, P., and Sweller J. (2002) "Assimilating complex information." *Learning and Instruction*, 12, pp. 61-86.

Sweller, J. (2003) "Evolution of human cognitive architecture." In B. Ross (Ed.), *The Psychology of learning and Motivation* (Vol. 43, pp.215-266). San Diego, CA: Academic Press.

Sweller, J. (2005) "Implications of cognitive load theory for multimedia learning." in R. E. Mayer (Ed.). *The Cambridge Handbook of Multimedia Learning.* New York: Cambridge University Press.

Tversky, B., Bauer-Morrison, J., & Betrancourt, M. (2002). "Animation: can it facilitate?" *International Journal of Human-Computer Studies*, Vole 57, 247-262.