

The Effectiveness of Educational Technology Applications for Enhancing Reading Achievement in K-12 Classrooms: A Meta-Analysis

Educator's Summary Updated April 2012

This review examines research on the effects of technology use on reading achievement in K-12 classrooms. It applies consistent inclusion standards to focus on studies that met high methodological standards. A total of 84 qualified studies based on over 60,000 K-12 participants were included in the final analysis. Four major categories of education technology are reviewed:

- (1) **Computer-managed learning**, which included only *Accelerated Reader*. This program uses computers to assess students' reading levels, assigning reading materials at students' levels, scoring tests on those readings, and charting students' progress. Students do not work directly on the computer with *Accelerated Reader*.
- (2) **Innovative technology applications**, such as *Fast ForWord*, *Reading Reels*, and *Lightspan*.
- (3) **Comprehensive models**, such as *READ 180*, *Writing to Read*, and *Voyager Passport*. These programs use computer-assisted instruction along with non-computer activities as students' core reading approach.
- (4) **Supplemental technology**, such as *Destination Reading*, *Plato Focus*, *Waterford*, and *WICAT*. These programs provide additional instruction at students' assessed levels of need to supplement traditional classroom instruction.

Review Methods

A literature search of articles written between 1980 and 2010 was carried out to find studies that met the following inclusion criteria:

- Students taught in classes using a given technology-assisted reading program had to be compared to randomly assigned or well-matched control groups.

- Pretest data had to be provided, unless studies used random assignment of at least 30 units (individuals, classes, or schools) and there were no indications of initial inequality. Studies with pretest differences of more than 50% of a standard deviation were excluded.
- Dependent measures needed to be quantitative measures of reading performance, such as standardized reading measures and informal reading assessments.
- Study duration had to be at least 12 weeks.
- Studies had to have at least two teachers in each condition to avoid possible teacher effect.
- Studied programs needed to be replicable in realistic school settings. Studies providing experimental classes with extraordinary amounts of assistance that could not be provided in ordinary applications were excluded.

Key Findings

Researchers examined the relationship between education technology effectiveness and five key study variables: grade levels, types of intervention, program intensity, level of implementation, and socio-economic status. Key findings were as follows:

Grade Levels. Studies were organized in three grade levels: Kindergarten (N=8), Elementary (N=59), and Secondary (N=18). The effect sizes for kindergarten, elementary, and secondary levels were +0.15, +0.10, and +0.31, respectively. The between-group difference ($Q_B = 9.52$, $df=2$, $p<0.01$) was significant. The post hoc test suggests that the effect size at the secondary level was significantly higher than that at the elementary levels.

Types of intervention. In an analysis of the studies by program type, a marginally significant between-group effect ($Q_B = 7.15$, $df=3$, $p<0.07$) was found, indicating some variations among the four types of programs. The 18 comprehensive model studies produced the largest effect size, +0.28, and the four computer-managed learning and the six innovative technology applications produced similar moderate effect sizes of +0.19 and +0.18, respectively. The average effect size for the 56 supplemental technology programs (traditional CAI) was only +0.11. The results of the analyses of comprehensive and innovative programs have to be considered carefully, however, due to the small number of studies in these categories.

Program intensity. Program intensity may help explain some of the variation in the model. Program intensity was divided into two categories: low intensity (the use of technology less than

15 minutes a day or less than 75 minutes a week) and high intensity (over 15 minutes a day or 75 minutes a week). Analyzing the use of technology as a moderator variable, only a marginally significant difference was found between the two intensity categories ($Q_B=3.04$, $df=1$, $p=0.08$). The effect sizes for low and high intensity are +0.11 and +0.19, respectively.

Level of implementation. Significant differences were found among low, medium, and high levels of implementation, as reported by the researchers. The mean effect sizes for low, medium, and high implementation were +0.01, +0.18, and, +0.22, respectively. Over half of the studies (53%) did not provide sufficient information about implementation. It is clear from the findings that no effect was found when implementation was described as low. A significant and positive effect was detected for groups that had a medium or high level of implementation rating. The implementation ratings must be considered cautiously, however, because authors who knew that there were no experimental-control differences may have described poor implementation as the reason, while those with positive effects might be less likely to describe implementation as poor.

Socio-economic status (SES). Studies were divided into three categories: low, mixed, and high SES. Low SES refers to studies that had 40% or more students receiving free and reduced-price lunch, and high SES refers to studies in which less than 40% of students received free lunches. Four studies that involved a diverse population, including both low and high SES students, were excluded in these analyses. The p-value (0.31) of the test of heterogeneity in effect sizes suggests that the variance in the sample of effect sizes were within the range that could be expected based on sampling error alone. The effect sizes for low and high SES were +0.17 and +0.12, respectively, indicating a minimal effect of SES. In addition to the between-study comparison, we also looked at the differential impact of instructional technology on students with different SES background within studies. There were a total of ten studies identified. Educational technology had a slightly higher positive impact on low SES students with an average effect of +0.31, whereas the effect for high SES students was +0.20. Due to the small number of studies, however, no significant difference was found between low SES and high SES groups.

Within-Study Subgroup Analyses

Subgroup analyses of comparisons within studies were also conducted to compute differential mean effect sizes based on student demographic characteristics such as student ability, gender, race, and language. Key findings were as follows:

Ability. Out of the 84 qualifying studies, there were a total of 13 studies that examined the impact of instructional technology on students with different academic abilities, yielding 29 effect sizes. The mean effect sizes for low, middle, and high ability students were +0.37, +0.27, and +0.08, respectively. The post hoc tests suggest that instructional technology had a more positive impact on low and middle ability students than it did on high ability students.

Gender. Instructional technology generated a more positive impact among males than females. The effect sizes for males and females were +0.28 and +0.12, respectively. No significant difference according to gender was found, however, due to the small number of studies reporting effects by gender.

Race. A total of nine studies examined the interaction effect of race with the use of educational technology. The mean effect sizes for students who were African American, Hispanic, and White were +0.12, +0.42, and +0.11. The numbers of studies with each group was small, however, and there was only one study on a Hispanic population.

English Language Learners. Only three studies examined the effect of instructional technology on English language learners. The effect size was +0.29 ($p < 0.05$).

Conclusions

The findings of this review support those of earlier reviews by other researchers. The classroom use of educational technology will undoubtedly continue to expand and play an increasingly significant role in public education in the years to come as technology becomes more sophisticated and more cost effective. This review highlights the need for more randomized studies. In addition, schools and districts should make concerted efforts to identify and adopt research-proven educational technology programs to improve student academic achievement as well as to close the ability and language gaps in their schools.

The technology approaches most widely used in schools, especially supplemental computer-assisted instruction, have the least evidence of effectiveness. Alternative uses of technology applications have greater promise. For example, the integration of non-technology classroom instruction and computer-assisted instruction and the utility of video, computer content, and embedded multimedia as a component of beginning reading instruction have shown particular promise. Government and foundation funders should continue to invest in evaluation of innovative programs and in creation of new technology applications. For example, interactive whiteboards have become increasingly popular. Yet there is little experimental research on their outcomes or on effective ways of using these and other whole-class technologies.

Full Report

Cheung, A., Slavin, R.E. (2012, April). *The Effectiveness of Educational Technology Applications for Enhancing Reading Achievement in K-12 Classrooms: A Meta-Analysis*. Baltimore, MD: Johns Hopkins University, Center for Research and Reform in Education.

The full report can be downloaded at www.bestevidence.org/reading/tech/tech.html