The Effectiveness of Educational Technology Applications for Enhancing Reading Achievement: A Meta-Analysis

Educator's Summary

Robert E Slavin **Johns Hopkins University and University of York**
Alan Cheung **Johns Hopkins University**

Last updated April 2012

Best Evidence Encyclopaedia
Empowering educators with evidence

This review summarises research on the effects of technology use on reading achievement in primary and secondary school classrooms.

We conducted a literature search of all articles written between 1980 and 2010, and applied consistent inclusion standards to focus on studies that met high methodological standards. A total of 84 studies were included in the final analysis, based on over 60,000 primary and secondary school participants

Four major categories of education technology are reviewed:

1. **Computer-managed learning,** which included only *Accelerated Reader*. This programme uses computers to assess children's reading levels, assigning reading materials at their levels, scoring tests on those readings, and charting pupil progress. Children 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 programmes use computer-assisted teaching along with non-computer activities as pupils' core reading approach.

4. **Supplemental technology,** such as *Destination Reading*, *Plato Focus*, *Waterford*, and *WICAT*. These programmes provide additional instruction at pupils' assessed levels of need to supplement traditional classroom teaching.

The full report (which this review summarises) is available at www.bestevidence.org

Key Findings

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

Year levels Studies were organised in three year levels: Key Stage 1 (N=8), primary (N=59), and secondary (N=18). The effect sizes for Key Stage 1, primary, and secondary 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 Key Stage 1 and primary levels.

Types of intervention In an analysis of the studies by programme 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 intervention (computer-managed learning, innovative technology applications, comprehensive models, and supplemental technology). The 18 studies of comprehensive models produced the largest effect size, +0.28, and four computer-managed learning and six innovative technology application studies produced similar moderate effect sizes of +0.19 and +0.18, respectively. The average effect size for 56 studies of supplemental technology programmes was only +0.11. The results of the analyses of computer-managed learning and innovative technology

application data have to be considered carefully, however, due to the small number of studies in these categories.

Programme intensity Programme intensity was divided into two categories: low intensity (the use of technology for 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). Analysing the use of technology as a moderator variable, no significant difference was found between the two intensity categories (Q_B =3.04, df=1, p=0.08). This result suggests that more technology use does not necessarily result in better outcomes. The effect sizes for low and high intensity were +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 pupils receiving free and reduced-price lunch and high SES less than 40%. Four studies that involved a diverse population, including both low and high SES pupils, 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 pupils with different SES backgrounds within studies. A total of 10 studies were identified. Findings showed that education technology had a slightly higher positive impact on low SES pupils with an average effect of +0.31, whereas the effect for high SES pupils was +0.20. Due to small numbers of studies, 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 pupil demographic characteristics such as pupil 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 pupils with different academic abilities, yielding 29 effect sizes. The mean effect sizes for low, middle, and high ability pupils 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 pupils than it did on high ability pupils.

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 low power.

Race A total of seven studies examined the interaction effect of race with the use of education technology. The mean effect sizes for pupils 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.

Pupils with English as an Additional Language Only three studies examined the effect of instructional technology on pupils with English as an Additional Language. The effect size was +0.29 ($p<0.05$).

Conclusions

Consistent with previous reviews of similar focus, the findings of this review suggest that education technology generally produces a positive, though small, effect (ES=+0.16) in comparison to traditional methods. However, the effects may vary by education technology type. In particular, the types of supplementary computer-assisted teaching programmes that have dominated the classroom use of education technology in the past few decades are not producing educationally meaningful effects in reading for pupils. In contrast, innovative technology applications and integrated literacy interventions with the support of extensive professional development showed somewhat promising evidence. However, too few randomised studies for these promising approaches are available at this point for firm conclusions.

Review Methods

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

- Pupils taught in classes using a given technology-assisted reading programme 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 standardised 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 programmes 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.

The Full Report

The full report, which this review summarises, is Cheung A, and Slavin RE (2011), *The Effectiveness of Education Technology for Enhancing Reading Achievement: 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.uk