A Meta-analysis of Pre-K to Grade 3 Educational Apps: Are There Meaningful Effects on Student Achievement Outcomes?

James Kim, Joshua Gilbert, Vicky Yu, Charles Gale

HARVARD GRADUATE SCHOOL OF EDUCATION | READS LAB

Abstract

This meta-analytic review of prekindergarten to Grade 3 educational apps in the domains of literacy and mathematics synthesizes the results from 32 intervention studies that examined causal effects off app-based interventions on student outcomes. The synthesis examined overall effects of educational apps on student outcomes in literacy and math, methodological and intervention moderators of effects, and associations between overall effects and the extent to which apps fostered active, engaged, meaningful, and socially interactive learning and had clear learning goals (Hirsch-Pasek et al., 2015). In contrast to correlational studies, this meta-analytic review revealed that educational apps subjected to experimental or quasi-experimental evaluations produced positive and moderately large effect sizes on standardized achievement outcomes (Hedge's g = .33, 95% CI [.21, .46]), with similar effects when results were broken down by reading outcomes (g = .35) and math (g = .33). Results from a meta-regression revealed a positive association between effect sizes and the degree to which apps foster learning that was active, meaningful, engaging and socially interactive, though the effect was not statistically significant. Findings are situated in both theoretical work on the science of learning and empirical research on the relations between app usage and student learning.

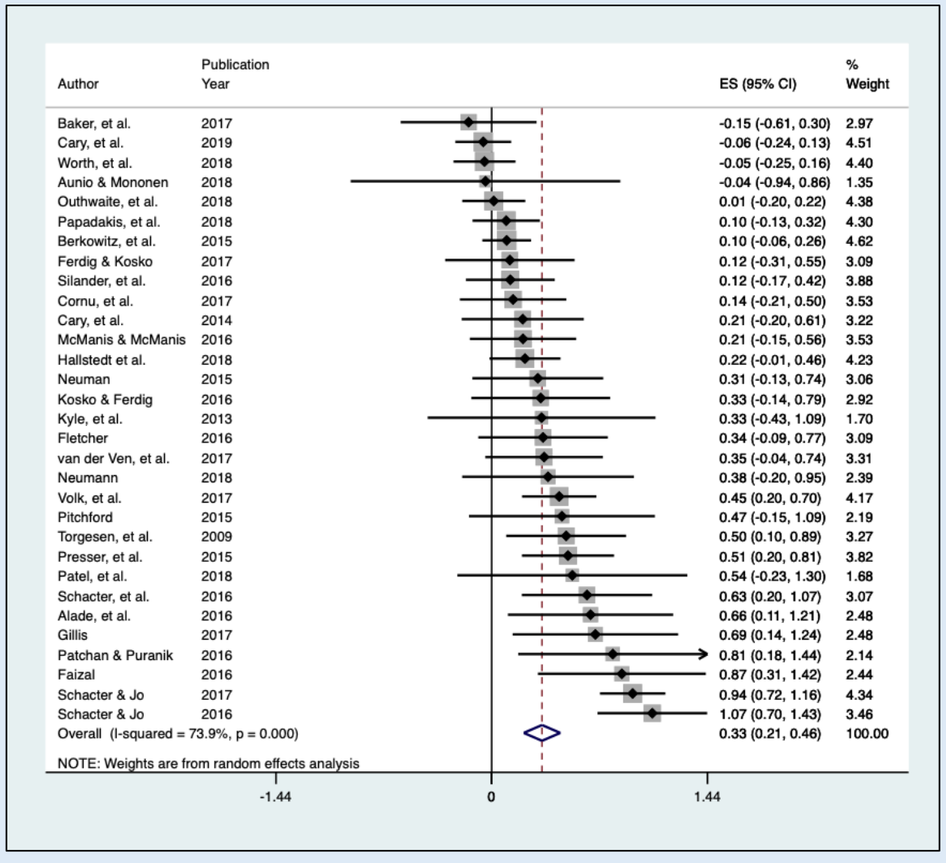
Inclusion Criteria

1. The study employed an experimental or quasi-experimental research design
2. The intervention was centered on an educational app for touchscreen device (e.g., a smartphone or tablet) or computer

Overall Findings

The weighted effect size (Hedge's *g*) from a random effects meta-analysis of the 32 studies included in our review was .33 (95% CI = [.21, .46]). The results are stable when literacy and math outcomes are considered separately. The distribution of effect sizes across studies varied widely, ranging from a low of -0.15 to a high of +1.07. The distribution of effect sizes is displayed in the forest plot in Figure 1 below.

Figure 1. Forest Plot of a Random Effects Meta-Analysis of Combined Literacy and Math Outcomes



Given the wide distribution of effects around the average, we conducted moderator analyses to determine to what extent this variation could be explained by systematic app- or study-level variables.

Exploring Treatment Effect Heterogeneity through Moderator Effects

We explored the following moderator variables to explain variability in effect size: alignment with Hirsch-Pasek et al.'s (2015) criteria for educational app quality, app "dosage" (operationalized as duration of intervention, frequency of intervention sessions, and intensity of each session), school or home context, peer-review status, age of the students in the sample, and various interactions between moderators. Most moderators were non-significant predictors of effect size variation, with the exception of studies of PreK students (PreK studies had on average larger effect sizes), and whether the app was judged to foster "Meaningful" learning. Results of additional moderator analyses are presented below.

Figure 2. Meta-Regression of Effect Size on Continuous Measure of Alignment with Hirsch-Pasek et al.'s Criteria for App Quality (Active, Engaging, Meaningful, Social, and Learning Goal)

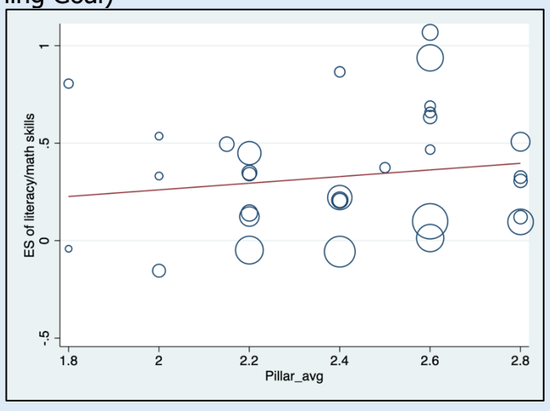
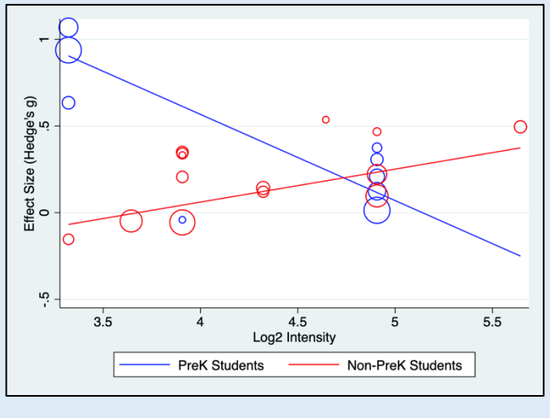


Figure 3. Meta-Regression of Effect Size on Intensity of Intervention Sessions (in Log_2 Minutes) Interacted with Studies of PreK children (Blue=PreK, Red=Non-PreK).



3. The study measured a literacy or math outcome
4. The participants were aged three to nine years old (i.e., Pre-K to 3rd grade)
5. The study was published between January 2008 and June 2019

Contact Information
James Kim: james_kim@gse.harvard.edu
Josh Gilbert: joshua_gilbert@gse.harvard.edu