

Evidence Brief: Examining the Effects of Preschool and Early Elementary Educational Apps on Student Literacy and Math Skills

The Model of Reading Engagement (MORE) is an approach to content literacy¹ developed by the [Harvard READS Lab](#) in collaboration with teachers in our partner districts. A key component of the MORE program is a **personalized literacy app** that uses words from the MORE lesson books to give students exposure to and opportunities to play with topic-related words that complement the MORE lessons. Using apps to promote student learning has rapidly expanded over the last decade. There are currently more than 200,000 educational and reference apps available on Apple's App Store, and schools spend billions on educational software each year. These expenditures have only grown since the spread of COVID-19. However, little is known about the effectiveness of educational apps for preschool and lower elementary children or which app or study characteristics predict their effectiveness.



Our question: What is the effect of educational apps, how much do apps vary in their effectiveness, and what app features predict bigger or smaller effects?

While software companies often conduct their own research on an app's effectiveness, many of these studies are flawed and lead to conclusions that benefit the companies. However, there is an emerging set of studies that employ rigorous designs to test app effectiveness. We conducted a comprehensive search of such app studies and identified 228 studies of preschool and early elementary apps. Thirty-six of these met our criteria on rigor and tested literacy or math outcomes. We reviewed these studies to assess the overall effectiveness of the apps for this age group. To analyze the effects of the studies, we used a statistical technique called meta-analysis. Meta-analysis uses statistics to review the combined effects of multiple studies. This improves the likelihood of finding a statistically significant effect even when individual studies did not. It also helps to understand differences in effects across the studies. Our study answers three key questions:

1. What is the overall effect of educational apps?
2. How much do apps vary in their effectiveness?
3. What app or study features predict bigger or smaller effects on student test performance?

What we found: We find positive effects of preschool and early elementary apps. However, features of the studies or the student population explain large differences across studies

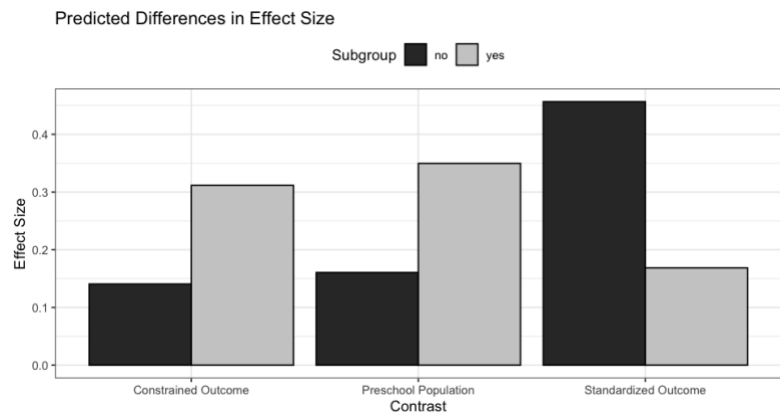
While the outcome used to assess each app's effectiveness would ideally be the same across all the studies, this is not the case. One way to overcome this and combine effects across studies is to translate them into a statistical "effect size" that is comparable across studies.² Using this method, we find a positive weighted

¹ **Content literacy approaches** like MORE help students build "rich and connected ideas" about science, social studies, and other content areas, while also supporting students' development of the reading, writing, and speaking skills necessary to engage deeply with these ideas (Catts, H. W. (2022). Rethinking How to Promote Reading Comprehension. *American Educator*, 45(4), 26).

² Effect size is a measure of how much the app improved performance. An effect size of 0 means that the treatment students receiving the app intervention performed no differently than the comparison students who did not use the app. A positive effect size means that app students scored higher than comparison students, and a negative effect size means that app students scored lower than the comparison students.

average effect across all the studies, meaning that on average, apps provided a boost to student performance. However, focusing on the overall effect of educational apps paints an overly simplistic and optimistic portrait of the findings. While the average effect is positive, the effect of individual apps in our analysis ranged from *negative* to hugely positive. Meta-analysis allows us to unpack the range of effects by app features.

Our analysis identified three features that predict differences in app effects. The graph shows the effect for particular elements of the studies or student populations. The vertical y-axis of the graph represents the average effect size. First, app effects were larger for studies that measured “constrained” skills. These are skills that grow fast for a while until the student masters them. Some examples include letter naming or counting. “Unconstrained” skills continue to grow throughout life, like vocabulary. Second, app effects were larger for studies involving preschool rather than K-3 students, which is not all that surprising as skills tend to be easier to learn when students are younger. Finally, app effects were larger for studies using tests developed by the researchers rather than standardized tests. Standardized tests measure a student’s ability to transfer knowledge into novel contexts, and these are the most relevant tests for policymakers and families.



Our findings are just as intriguing in showing what *doesn’t* predict app effectiveness. For example, “dosage”, or the amount of time on the app, was not systematically related to app effectiveness, suggesting that apps are best thought of as a supplement to, rather than a replacement for, traditional classroom instruction.

Conclusion: Educational Apps are Potentially Beneficial, but Policymakers and Parents Should Approach them with Caution

Educational apps have great potential. They are often free or low-cost, easily scalable, and provide targeted opportunities for students to practice concrete skills as a supplement to classroom or home learning. However, many app developers never commission formal studies, and even among those we reviewed that did, less than 40 percent measured literacy and math outcomes or could draw conclusions that the app caused the effects they reported. Furthermore, even within these rigorous studies, we found evidence that suggests that apps focused on short-term and targeted skills, such as constrained skills, might yield the largest benefits for children between Preschool and third grade. This supplemental type of practice appears to benefit younger children the most, likely because teaching skills such as hearing separate sounds based upon letters or combinations of letters could be better taught with an app than more complex skills like language comprehension. These results are both promising and suggest some caution in the adoption of new educational apps.

This brief describes work done for READS Lab at the Harvard Graduate School of Education based upon Kim, Gilbert, Yu, and Gale (2021), *Measures Matter: A Meta-Analysis of the Effects of Educational Apps on Preschool to Grade 3 Children’s Literacy and Math Skills*. The research reported here was supported by the Chan Zuckerberg Initiative.