WORKING PAPER · NO. 2025-103

What Drives Educational Technology Adoption in Classrooms Serving Young Children? Evidence from Two Experiments

Daniela Bresciani, Ariel Kalil, Michelle Michelini, and Rohen Shah

AUGUST 2025

THE UNIVERSITY OF CHICAGO
BECKER FRIEDMAN INSTITUTE FOR ECONOMICS

5757 S. University Ave.
Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

What Drives Educational Technology Adoption in Classrooms Serving Young Children? Evidence from Two Experiments.*

Daniela Bresciani [a]
Ariel Kalil [a]
Michelle Michelini [a]
Rohen Shah [b]

[a] *Harris School of Public Policy, University of Chicago, 1307 E. 60th St., Chicago IL 60637, USA*
[b] *Department of Economics, Yale University, 87 Trumbull St., New Haven, CT 06511, USA*

August 8 2025

Abstract:

As new education technology applications emerge, teachers and school leaders must decide which ones to endorse. What influences these decisions? This study examines how information about two aspects of ed tech apps—efficacy and popularity—affects educators' choices to adopt them. We present causal estimates from two experiments: a within-subject survey of 1,104 teachers (pre-K through grade 4) and a between-subjects experiment with 154 school leaders. In the first experiment, providing teachers with a cue about strong research evidence increased their likelihood of recommending a math learning app by 0.24 SD (compared to a no-cue control), while a peer-popularity cue raised it by 0.21 SD; using both cues together led to a 0.30 SD increase. In the second experiment, showing school leaders an informational video about a math app's popularity and research evidence did not significantly increase their likelihood of recommending the app or their willingness to pay for it.

JEL Codes: C91, D91, I28.
Keywords: Educational Technology, Bandwagon Effect, Within-Subjects Experiment.

*This study was pre-registered on the AEA RCT Registry (AEARCTR-0013267 and AEARCTR-0013482) and overseen by the Social and Behavioral Sciences IRB from the University of Chicago (IRB24-0151). We thank Carina Initiatives for funding. Online Appendix: https://sites.google.com/view/rohen-shah/projects/ed-tech-adoption
Email Addresses: dbresciani@uchicago.edu (D. Bresciani), akalil@uchicago.edu (A. Kalil), meesh@uchicago.edu (M. Michelini), rohen.shah@yale.edu (R. Shah).
Corresponding author: Rohen Shah

I. Introduction

The number of educational technology applications available to students is constantly increasing. Teachers must decide whether to endorse them, and principals choose whether to support or fund them. Several factors can influence these decisions, including costs, perceived effectiveness, and district policies. In this paper, we use two survey experiments to examine whether an app's adoption is affected by two factors: its popularity and its efficacy.

In the first experiment, we use a within-subjects design where teachers are shown short descriptions of fictional math learning apps and asked to rate how likely they are to recommend the app to students for home use. The four app descriptions are followed by either 1) a statement about research evidence showing the app's effectiveness, 2) a statement that the app is popular among teachers, 3) both statements on efficacy and popularity, or 4) no additional information. To minimize any effects of sequence or the order of app descriptions, we randomize both the order in which treatments are shown to teachers and which app descriptions are paired with each treatment. We find that the likelihood of a teacher recommending a math app increases by 0.24 SD when they are informed that there is research evidence that the app improves math skills and by 0.21 SD when teachers are told that the app is popular among educators. Telling teachers about both the app's popularity and its efficacy results in a 0.30 SD increase.

In the second experiment, Pre-K school leaders were randomly assigned to either watch an informational video on research evidence supporting a digital game or view a control video describing analog learning materials without discussing research evidence. We find no significant effect on the likelihood of recommending digital apps and a marginally significant (at the $p < .10$ level) increase in their willingness to spend more on digital apps from their school budgets.

This work adds to the growing literature on what influences decision-making in various contexts, especially among public leaders (Carattini et al., 2024; Garcia-Hombrados et al., 2024; Mayer et al., 2021; Vivalt et al., 2025; Vivalt & Coville, 2023). Several studies specifically explore the role of research evidence in shaping decisions. Hjort et al. (2021) find that mayors in Brazil are willing to pay for study results, update their beliefs, and value large sample sizes in studies. Conversely, Briscese & List (2024) discover that policymakers initially overly optimistic about a program's effectiveness tend to temper their views based on evidence, yet show decreased demand for experimentation, indicating experiment aversion when results defy expectations. Evidence also exists on whether the strength of research evidence can predict policy adoption. How policy evaluation results are communicated is also important — for example, using numbers versus words (Thaler et al., 2024). Additionally, tools that enable "side by side" comparisons and combine multiple impact features into a single metric increase the likelihood that research findings influence decision-making (Toma & Bell, 2024).

At the same time, social influences impact adoption choices. The bandwagon effect (or herding bias) explains the tendency to adopt an idea or product simply because others have done so (Henshel & Johnston, 1987). Evidence indicates that social proof cues can influence behavior in both consumer and public settings; informing people about peer choices has been shown to "nudge" the adoption of new behaviors across various fields, from energy conservation to health practices (Bergman, 2016). However, research providing causal evidence about bandwagon cues in professional decision-making environments remains limited. This gap is especially significant in education, where decisions to adopt curricula or technologies might be driven as much by peer popularity signals as by formal evidence of their effectiveness.

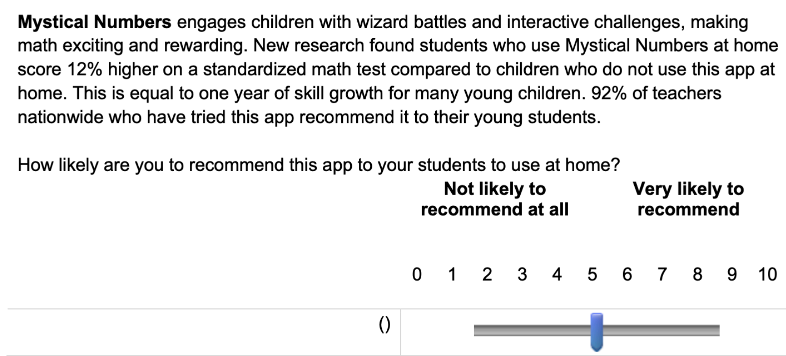
II. RCT 1: Within-Subjects Experiment with Teachers

The first survey experiment was conducted online from April to June 2024 with teachers from 11 school districts serving about 161,000 preschool and elementary students. To build a sample for collecting structured survey and experimental data, we directly contacted superintendents via email listed on the district websites, using the Institute of Education Sciences (IES) National Center for Education Statistics (NCES) list of public-school districts with enrollments over 15,000 students as of Fall 2021. We emailed 493 superintendents and received 89 responses. This led to 24 districts having Zoom conversations with us, and 11 of those districts agreed to participate. Teachers were invited by their school or district to complete the online survey and received a $40 digital Amazon gift card upon completion. Our final sample included 1,104 teachers, including 289 Pre-K teachers, 206 Kindergarten teachers, and about 150 teachers from each of grades 1-4. The ages of teachers in our sample (96% of whom were female) ranged roughly uniformly from 25 to 60 years old.

The experiment was part of a 10-minute survey completed by teachers, which also asked about their familiarity with and barriers to recommending digital math games. The full survey results and replication files are available in the online appendix. Conducted over a 10-week period, the survey included four experimental questions designed as a 2×2 factorial design. This design randomized the presence of two endorsement signals when describing a digital math app: (i) a reference to research evidence supporting the app's effectiveness, and (ii) the popularity of the app among other teachers. This resulted in four experimental conditions: Efficacy; Popularity; Efficacy & Popularity; and Control (no additional information).

The four experimental questions first described a math app, then asked teachers to rate how likely they were to recommend the app. We gave these four fictional math apps realistic names and descriptions. The questions included the following informational cues: 1) “Control”: only a description of the app; 2) “Efficacy”: the app description plus information about positive research results showing the app’s effectiveness for children’s learning; 3) “Popularity”: the app description plus information on the app’s endorsement by most teachers; and 4) “Efficacy & Popularity”: the app description plus both positive research results and endorsement by most teachers. After each description, teachers rated how likely they were to recommend the app for student use at home on a scale of 1 to 10. An example question is shown in Figure 1.

Figure 1: Example of a Survey Question (“Efficacy & Popularity” Treatment)


Mystical Numbers engages children with wizard battles and interactive challenges, making math exciting and rewarding. New research found students who use Mystical Numbers at home score 12% higher on a standardized math test compared to children who do not use this app at home. This is equal to one year of skill growth for many young children. 92% of teachers nationwide who have tried this app recommend it to their young students.

How likely are you to recommend this app to your students to use at home?

	Not likely to recommend at all	Very likely to recommend
	0 1 2 3 4 5 6 7 8 9 10	
()		

In this “within-subjects” design, each participant experiences all treatment conditions (List, 2025a, 2025b). To minimize threats to valid inference, we randomize the pairing of the “app” description and treatment cue, as well as the order of treatments. Table 1 shows the treatment effect of each cue on the likelihood of recommending each app. Column 1 presents the primary, pre-registered regression results, including app- and teacher-fixed effects. As a robustness check, column 2 shows the results when we only consider the first question for each participant. Since the treatment order was randomized for each participant, focusing only on the first question effectively creates a “between-subjects” experimental design.

Table 1: Treatment Effect of Efficacy & Popularity Cues on Likelihood of Recommending Apps

	(1) Likelihood of Recommending	(2) Likelihood of Recommending
Efficacy	0.63***	0.53**
	(0.07)	(0.23)
Popularity	0.55***	0.31
	(0.07)	(0.23)
Efficacy & Popularity	0.79***	0.76***
	(0.07)	(0.22)
Constant	6.05***	5.98***
	(0.06)	(0.23)
App Fixed Effect	Yes	Yes
Teacher Fixed Effect	Yes	No
Observations	4,416	1,104

Note: The outcome, "Likelihood of Recommending", is measured on a scale from 1 to 10. In column 1, each teacher appears 4 times in the dataset, resulting in 4,416 observations. Robust standard errors are in parentheses. *** p<.01, ** p<.05, * p<.1.

We find that the "Efficacy" treatment increased the likelihood of recommending the math app by 0.63 points on a 1-10 scale, which corresponds to a 0.24 standard deviation treatment effect. The "Popularity" treatment resulted in a 0.21 SD increase, and the combined "Efficacy & Popularity" treatment led to a 0.30 SD increase. The effect of the combined treatment is significantly higher than either Efficacy alone or Popularity alone. Column 2 shows that the "first question only" results are consistent with the overall findings from column 1, although the "Popularity" treatment effect is not statistically significant.

Lastly, one benefit of a "within-subjects" design is that it allows us to analyze treatment effects for each individual. Across all treatments, about 51% of teachers show no effect, 34% experience a positive effect, and 15% experience a negative effect.

III. RCT 2: Information Experiment

The school leader survey experiment took place over a 7-week period from April to June 2024. As part of the survey, participants were asked to watch a short, two-minute video. Respondents were randomly assigned to view either a treatment or a control video. The treatment video explained different types of analog and digital math learning materials and highlighted positive research findings on the effectiveness of digital math games compared to analog math materials for pre-K children's learning (Kalil et al., 2023, 2024, 2025; Mayer et al., 2023). The control video described various types of analog and digital math learning materials but did not include any research results.

After the video, pre-K leaders were asked about 1) their likelihood of endorsing a digital math game among other learning resources for pre-K students to use at home, rated on a Likert scale from "Very Unlikely" to "Very Likely,' and 2) their willingness to pay for a digital math game for each pre-K student using school funds, on a scale of \$0-\$10.

Tables 2 and 3 show how watching the treatment video affects school leaders' likelihood to recommend each type of material and their willingness to pay, respectively. Overall, we find no statistically significant effects at the $p < .05$ level for either outcome. At the $p < .10$ level, the treatment video seems to decrease the likelihood of recommending physical worksheets and increase the willingness to pay for digital games, as shown in Figure 2.

Table 2: Treatment Effect on School Leader Recommendation of Digital & Analog Materials

	(1) Math Master (digital game)	(2) Number Rumble (physical game)	(3) i-Math (digital worksheet)	(4) Notebook Math (physical worksheet)
Treatment	0.51	−0.24	−0.06	−0.81*
	(0.39)	(0.39)	(0.52)	(0.49)
Constant	6.62***	6.98***	4.93***	4.62***
	(0.29)	(0.30)	(0.36)	(0.34)
Observations	169	169	169	169

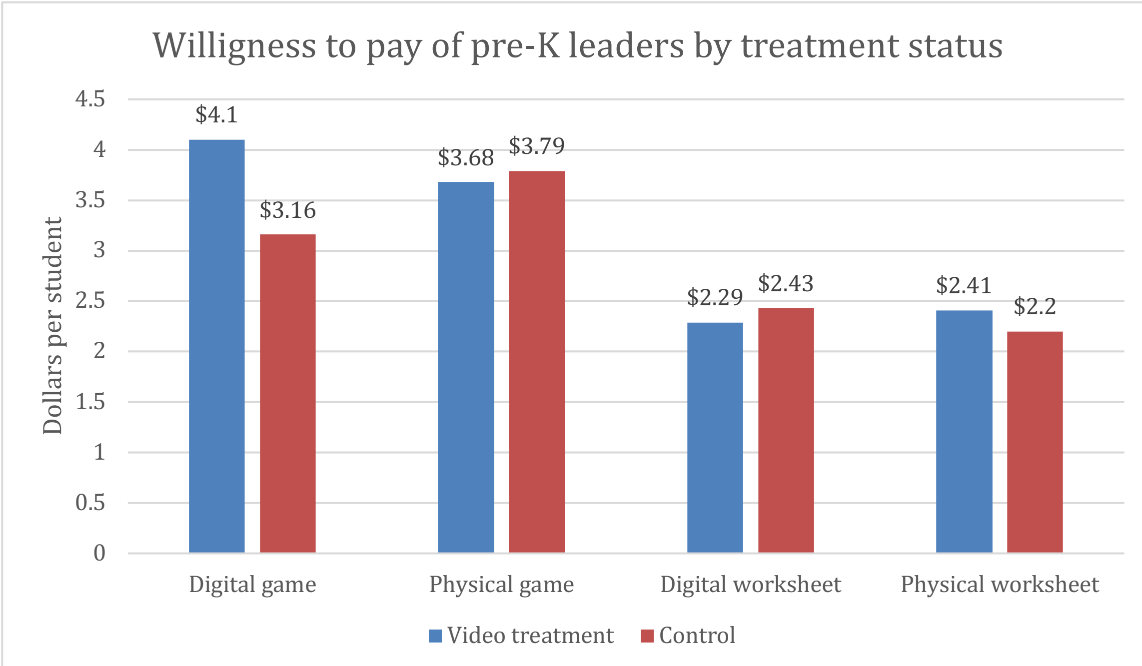
Note: The outcome is the likelihood of recommending (scale of 1-10) the type of digital and non-digital material described in each column. Robust standard errors are in parentheses. *** p<.01, ** p<.05, * p<.1.

Table 3: Treatment Effect on School Leader Willingness to Pay for Digital & Analog Materials

	(1) Math Master (digital game)	(2) Number Rumble (physical game)	(3) i-Math (digital worksheet)	(4) Notebook Math (physical worksheet)
Treatment	0.82*	−0.17	−0.20	0.02
	(0.48)	(0.45)	(0.45)	(0.46)
Constant	3.26***	3.78***	2.48***	2.38***
	(0.31)	(0.31)	(0.31)	(0.32)
Observations	169	169	169	169

Note: The outcome is the Willingness to Pay per student (measured in dollars) for the type of digital and non-digital material described in each column. Robust standard errors are in parentheses. *** p<.01, ** p<.05, * p<.1.

Figure 2: School Leader Willingness to Pay for Digital & Analog Materials



IV. Discussion

The statistically nonsignificant results in the second experiment might suggest that school leaders are not easily influenceable. However, the power calculations for this experiment were based on observed treatment effects of information experiments in other settings. It is possible that this experiment is simply underpowered if the treatment effect is positive but smaller in size than that seen in other information experiments.

In the teacher survey experiment, the treatment effect on Pre-K teachers is half as strong as K-4 teachers, which could indicate that preschool teachers' opinions are less easily swayed than those of elementary school teachers. There may also be evidence of the "bandwagon" effect, as teachers are equally influenced by peer popularity and research evidence. However, this is not necessarily a "bias," since peer popularity might legitimately reflect that the app is helpful for learning. Overall, this study shows us that teachers are influenced by research evidence, which highlights the importance of rigorously evaluating ed tech products.

REFERENCES

Bergman, P. L. S. (2016). *Technology adoption in education: Usage, spillovers and student achievement*. CESifo Working Paper. https://www.econstor.eu/handle/10419/147355

Briscese, G., & List, J. A. (2024). *Toward an Understanding of the Political Economy of Using Field Experiments in Policymaking*. National Bureau of Economic Research.

Carattini, S., Dur, R., & List, J. (2024). Policy evaluation and the causal analysis of public support. *Science*, *386*(6721), 490–492. https://doi.org/10.1126/science.adp7497

DellaVigna, S., Kim, W., & Linos, E. (2024). Bottlenecks for Evidence Adoption. *Journal of Political Economy*, *132*(8), 2748–2789. https://doi.org/10.1086/729447

Garcia-Hombrados, J., Jansen, M., Martínez, Á., Özcan, B., Rey-Biel, P., & Roldán-Monés, A. (2024). *Ideological Alignment and Evidence-Based Policy Adoption*.

Henshel, R. L., & Johnston, W. (1987). The Emergence of Bandwagon Effects: A Theory. *The Sociological Quarterly*, *28*(4), 493–511. https://doi.org/10.1111/j.1533-8525.1987.tb00308.x

Hjort, J., Moreira, D., Rao, G., & Santini, J. F. (2021). How research affects policy: Experimental evidence from 2,150 brazilian municipalities. *American Economic Review*, *111*(5), 1442–1480.

Kalil, A., Liu, H., Mayer, S., Rury, D., & Shah, R. (2023). Nudging or Nagging? Conflicting Effects of Behavioral Tools. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, *2023–02*.

Kalil, A., Liu, H., Mayer, S., Rury, D., & Shah, R. (2025). A Digital Library for Parent-Child Shared Reading Improves Literacy Skills for Young Disadvantaged Children. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, *2025–27*.

Kalil, A., Mayer, S., Oreopoulos, P., & Shah, R. (2024). *Making a song and dance about it: The effectiveness of teaching children vocabulary with animated music videos*. National Bureau of Economic Research. https://www.nber.org/papers/w32132

List, J. A. (2025a). *A course in experimental economics*. University of Chicago Press.

List, J. A. (2025b). *The Experimentalist Looks Within: Toward an Understanding of Within-Subject Experimental Designs*. National Bureau of Economic Research.

Mayer, S., Kalil, A., Delgado, W., Liu, H., Rury, D., & Shah, R. (2023). Boosting Parent-Child Math Engagement and Preschool Children's Math Skills: Evidence from an RCT with Low-Income Families. *Economics of Education Review*, *95*.

Mayer, S., Shah, R., & Kalil, A. (2021). How cognitive biases can undermine program scale-up decisions. In *The Scale-Up Effect in Early Childhood and Public Policy* (pp. 41–57). Routledge.

Thaler, M., Toma, M., & Wang, V. Y. (2024). *Numbers Tell, Words Sell*.

Toma, M., & Bell, E. (2024). Understanding and increasing policymakers' sensitivity to program impact. *Journal of Public Economics*, *234*, 105096.

Vivalt, E., & Coville, A. (2023). How do policymakers update their beliefs? *Journal of Development Economics*, *165*, 103121.

Vivalt, E., Coville, A., & Sampada, K. C. (2025). Local knowledge, formal evidence, and policy decisions. *Journal of Development Economics*, *174*, 103425.

Zelizer, A. (2018). How Responsive Are Legislators to Policy Information? Evidence from a Field Experiment in a State Legislature. *Legislative Studies Quarterly*, *43*(4), 595–618.