



Annenberg Institute  
BROWN UNIVERSITY

# Effective Teacher Learning Through A's and B's

Partnerships Between Researcher and Educational Technology Firms that Generate Evidence about Professional Learning Design Features

May 2026

**Heather C. Hill**  
Harvard Graduate School of Education

**Olympia Francis Taylor**  
Harvard Graduate School of Education

**Jing Liu**  
University of Maryland, College Park

**Dora Demszky**  
Stanford Graduate School of Education

**Hannah Carter**  
Harvard Graduate School of Education

# Introduction

Every year, U.S. school systems allocate billions of dollars toward deepening teachers' knowledge and skill. This substantial investment reflects the fact that teachers represent the most crucial school-based influence on both academic achievement and students' social-emotional development. Like any talent-driven field, enhancing educational outcomes necessitates significant investments in the teaching workforce's capabilities, knowledge, and skills. Moreover, despite common misconceptions, research demonstrates that teacher professional learning (PL) can enhance instructional effectiveness and student performance.<sup>1</sup>

Given the size of district investment in teacher PL, most school leaders wish to spend their professional development dollars efficiently, making decisions about PL design that maximize teacher and student learning. Yet information about best design can be hard to come by. Although researchers have conducted hundreds of experimental trials of programs that contain at least some PL element, these trials typically provide only a thumbs up or down on a given program. And, while research syntheses can indicate effective *design features* – the who, what, where, when and how of PL is delivered – these syntheses cannot make causal claims. This leaves district and school leaders without strong information about design features that positively affect teacher and student learning. Because most teacher PL is locally created, information about effective design features can help accelerate teacher and student learning.

To achieve this goal, researchers have begun to conduct A/B studies of teacher professional learning. In these experimental studies, researchers systematically vary PL design features, assigning some teachers to PL with feature A and others to the nearly identical PL except with feature B, answering theoretical and practical questions. A/B studies are particularly beneficial because they can make causal claims about particular features. A study under way, for instance, examines the impact of more directive versus more reflective coaching on teacher engagement and skill development by holding constant coaching topics and most routines, but varying the role of the coach in setting the agenda and determining next steps. Another study, recently completed and the subject of a [companion RPPL brief](#), varies whether offering instructors some choice over the kind of feedback they receive enhances PL engagement and the development of dialogic teaching practices.<sup>2</sup> Both studies address persistent questions faced by district and school leaders, but that cannot be answered in an evaluation of a single program.

---

<sup>1</sup> Lisa Didion, Jessica R. Toste, and Marissa J. Filderman, "Teacher Professional Development and Student Reading Achievement: A Meta-Analytic Review of the Effects," *Journal of Research on Educational Effectiveness* 13, no. 1 (2020): 29-66; Kathleen Lynch, Heather C. Hill, Kathryn Gonzalez, and Cynthia Pollard, "Strengthening STEM Instruction in Schools: Learning from Research," *Policy Insights from the Behavioral and Brain Sciences* 6, no. 2 (2019): 236-242, <https://doi.org/10.1177/2372732219864385>; Michael A. Kraft, David Blazar, and Dylan Hogan, "The Effect of Teaching Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence," *Review of Educational Research* 88, no. 4 (2018): 547-588.; Hill, H. C., Lynch, K., Gonzalez, K. E., & Pollard, C. (2020). Professional development that improves STEM outcomes. *Phi Delta Kappan*, 101(5), 50-56.

<sup>2</sup> Dorottya Demszky, Heather C. Hill, Eric S. Taylor, Ashlee Kupor, Deepak Varuvel Dennison, and Chris Piech, "Does Increased Agency Improve the Effectiveness of Self-Directed Professional Learning for Educators?" (EdWorkingPaper: 25-1162, Annenberg Institute at Brown University, 2025), <https://doi.org/10.26300/04kc-7085>.

This brief describes an innovative way of getting A/B studies completed: by leveraging educational technology firms (hereafter ‘edtech firms’) whose mission and product includes improving teacher learning. As we show below, these platforms present tremendous advantages to researchers interested in testing PL design features, and can bring benefits to the edtech firms as well.

## Working with Edtech Firms: Benefits to PL Researchers

Edtech firms can provide direct benefits to researchers at multiple stages of project development.

**First, companies can provide access to hundreds if not thousands of teachers and students, meaning researchers can conduct studies using samples far larger than would be feasible if recruitment occurred through usual channels.** Stanford’s Dora Demszky, for example, partnered with Schoolhouse.world, a non-profit founded by Sal Khan that connects 125K+ students around the world with free, peer-to-peer tutoring. In this partnership, they conducted experimental studies involving between 500-700 tutors to test the impact of feedback to novice tutors and learners.<sup>3</sup> Samples this large are particularly useful in the context of A/B studies. Unlike most randomized trials, which compare participating and non-participating teachers, all teachers in A/B studies receive some form of the intervention. Often, the difference between the program with feature A and feature B is relatively small and thus requires a larger sample to estimate with precision. Further, the A/B format appeals to company partners who often want all participants to receive the professional development provided by the partner or research team.

**Second, edtech firms have an existing content delivery platform for researchers to build from, significantly reducing the initial cost associated with intervention design.** TeachFX, for instance, has a fully developed pipeline for teachers to upload lesson recordings, receive feedback on specific instructional practices, and to support learning about those practices. Making adjustments to an existing content delivery platform has proven far easier and cost-efficient, say researchers, than building one from scratch.

**Third, in many cases, the edtech firm’s platform doubles as both an intervention and as a data collection instrument.** A teacher or tutor might, for instance, participate in a study for six or eight sessions; instructional recordings collected pre-intervention can be used as a baseline measure and those collected post-intervention can be analyzed for changes in practice. Critically, these platforms not only collect the data, they also create the measures used in the experiment. For instance, a partnership between Dora Demszky, Sean Geraghty and Cuemath, an online 1:1 tutoring company, provided feedback to both students and tutors about their talk time in hopes of increasing the amount of student speech. This measure served as not only the impetus for tutor and student reflection but also the main

---

<sup>3</sup> Preregistration: <https://www.socialscienceregistry.org/trials/14050>; Work in progress paper: Yun, J., Hicke, Y., Olson, M., & Demszky, D. (2024, July). Enhancing Tutoring Effectiveness Through Automated Feedback: Preliminary Findings from a Pilot Randomized Controlled Trial on SAT Tutoring. In *Proceedings of the Eleventh ACM Conference on Learning@ Scale* (pp. 422-426).

outcome of the experiment.<sup>4</sup> Because collecting and analyzing information about instruction can be costly – study budgets for these tasks typically run into the tens if not hundreds of thousands of dollars – a study design that can use existing data as an outcome measure is a significant advantage for researchers.

**Fourth, edtech firms can be helpful thought partners in study design.** Firms have extensive knowledge about users’ challenges and priorities that can meaningfully inform intervention design. Similarly, edtech firm staff can also be partners in iterative study design –delivering not only the large-scale intervention and measures but informing the design of the study through proof-of-concept piloting and small randomized trials. Firms can also put researchers in touch with instructors and students who can participate in qualitative interviews, further helping shape the design of the intervention and study procedures. For example, in the CueMath partnership, Dora Demszky piloted her study with a small number of tutors whom her team interviewed and asked for feedback. Their pilot studies were critical to designing the randomized control trial.

**Fifth, edtech firms can be great scaling partners beyond the study.** For example, if an intervention is effective, there is already a “ready-made” implementation of it within the edtech platform, which can continue using and scaling it for research impact.

**Finally, some companies have developed software that can seamlessly support A/B testing, making it easy for firms and researchers to run A/B testing with minimal disruptions to coding.** Carnegie Learning, for instance, supports A/B testing with its Upgrade software, which sits atop its existing learning management system and allows researchers to randomly assign teachers and students to new versions of content seamlessly. Source code for Upgrade is [available via Carnegie Learning](#).

## Benefits for Edtech Companies

Conducting studies with academic researchers can also benefit edtech firms.

Although many firms have their own in-house research staff, many of those staff noted in interviews that receiving outside expertise and perspective can be invaluable. Academics’ rigorous research methodology can also provide a level of detail and objectivity that rapid internal research may not be able to. Researchers can also help obtain grant funding that can further edtech firms’ goals by building research infrastructure within the company, supporting innovations in the platform, and more.

In working with researchers, tech companies are also exposed to a pool of talent from which they frequently hire. Dora Demszky, for instance, reported having multiple students who were hired by TeachFX after participating in research partnerships.

Finally, researchers have strong incentives to publish and disseminate the work, meaning the research done by the partnership will contribute to the more general knowledge base about teacher

---

<sup>4</sup> Dorottya Demszky, Rachel Wang, Sean Geraghty, and Connie Yu, "Does Feedback on Talk Time Increase Student Engagement? Evidence from a Randomized Controlled Trial on a Math Tutoring Platform," in Proceedings of the 14th Learning Analytics and Knowledge Conference (March 2024): 632-644.

learning. This can greatly increase an edtech firm’s visibility and reputation; firms benefit from their visibility among academics and their reputation can be enhanced through university research partnerships.

## Making Firm–Research Partnerships Work

Interviews with both researchers and representatives of edtech firms suggest the following best practices catalyze the most effective partnerships. Many of these best practices echo that in the literature on research-practice partnerships (e.g., Penuel et al., 2025).

1. **Pursue research questions of mutual interest.** Research questions must both help edtech firms improve their product and be of general interest to the broader academic community. Jing Liu notes that “if we can structure the research in a way that benefits both parties, it makes collaboration much easier.”
2. **Avoid projects that create significant additional work for internal firm teams.** Sean Geraghty notes that “if what you're trying to test creates undue operational headaches for other verticals within the organization, that can become a third rail – not because of controversy, but because the firm simply does not have the capacity to implement it.”
3. **Provide research funding to someone internal to the company** to do data harvesting and build any applicable features into the platform. Researchers repeatedly noted that having someone “on the inside” makes projects far more successful. Jing Liu noted that foundations can play a role in making this happen.
4. **Design the study within the boundaries of edtech firms' existing practice.** For example, for a study with TeachFX, Liu had to randomize on a rolling basis as teachers registered on their app. This created other variables Liu needed to adjust given differential starting times for the intervention.
5. **Cultivate founder support and commitment.** Researchers must often work with several teams across a company, and having a founder’s buy-in to a project is critical to maintaining internal support. Across all projects interviewed, in fact, researchers and firm representatives noted the strong commitment of the founder – and thus the edtech firm generally – to conducting research.

Many also noted a few more concrete steps researchers and tech firms can take to support a smooth collaboration.

6. **Ensure that data pipelines are robust.** It's critically important and nontrivial to make sure that the data is properly **stored** and **linked up** (e.g. that the metadata for a particular recording is linked to that recording). Depending on the needs of the project, it may also be necessary to ensure that firms’ technology provides accurate speaker identification of classroom recordings.

7. **Provide product demos to researchers.** An initial demonstration of the platform can help ensure that researchers and firms are working from a place of concrete shared understanding, not abstract concepts.
8. **Run a small pilot version** of a study before a full study to ensure that study-related modifications to the platform, data infrastructure and randomization processes are functioning accurately.

## Limitations to Firm-Research Partnerships

1. **Tech firms' data privacy concerns.** Companies are understandably concerned about data sharing and any corresponding risk to their reputation. In particular, data breaches that involve student information could be particularly damaging, both in terms of managing the fallout and securing future funding. Researchers noted that managing this risk may involve having a member of the research team “on the inside” at the company, so that the processing of any identifiable data can take place within company firewalls.
2. **Baseline data can be difficult to acquire.** To make experiments more efficient, researchers often collect baseline data – for example, what a teacher’s or tutor’s practice looked like prior to any new learning experience. In order to get accurate baseline data from teachers, TeachFX estimates that researchers would need a minimum of three and ideally five to six classroom recordings. Following that, experiments can sometimes last for five to six weeks. But teachers may not participate in data collection for that long, meaning researchers must forgo or shorten baseline data collection to ensure teachers’ full participation in the learning experience.
3. **Researchers’ need for background data about students and teachers.** Researchers must often collect background data on students and teachers, both for reporting sample characteristics and also for improving the efficiency of experiments. However, it is often not feasible to collect that type of data in a business environment.
4. **Researchers’ pace and inflexibility.** It is critical that researchers remain flexible; research projects are about tradeoffs. Researchers and firms must negotiate to set a pace that works with the demands of both academia and the private technology sector.
5. **Limited generalizability.** Because participants are existing users of a particular tool, samples are typically not representative or generalizable. When participants are not K12 teachers (e.g., tutors or online course instructors), generalizability may be similarly limited.
6. **Administering opt-in consent can be challenging.** Edtech platforms might be hesitant to administer research consent forms to users, as it may erode trust (users might wonder, “Who are these researchers, and what do they have to do with this edtech platform?”). Studies that actively consent students or teachers may also exacerbate selection bias. Some researchers have solved this problem by analyzing de-identified data after the fact, but doing so makes it especially critical to have a liaison within the tech platform who can validate the experimental setup and

process the data. And, for user interviews in which researchers are involved, opt-in consent is a must, and researchers have to do their best to identify effective incentives to mitigate low response rates and selection bias.

7. **Bureaucratic hoops of three-way partnerships.** Some studies may involve partnerships with districts, which require their own human subjects review and data use agreements for research to be conducted at their site. For example, Susanna Loeb (Stanford GSE) and Demszky partnered with districts that used the FEV tutoring platform to conduct randomized studies. Establishing these district partnerships took several months, and some fell through at the end. Critically, in some situations, industry partners can set up agreements with districts they work with that allow them to share their data with researchers. These agreements make it much easier for researchers to work with tech firms and to access their data across districts. Researchers should start thinking about these requirements well in advance, remain flexible and not count too much on a single partnership.
8. **Interventions are constrained by the bounds and experiences of an edtech platform.** It's difficult to do neatly controlled experiments. "Business as usual" still means that it's business as usual within that particular tech experience.

## Conclusion

Partnerships between researchers and education technology firms to conduct A/B studies represents a powerful opportunity to advance our understanding of effective teacher professional learning. These collaborations leverage the unique strengths of both academia and industry—combining researchers' methodological rigor with edtech platforms' scale, infrastructure, and real-world applications. When thoughtfully structured around mutual interests and clear agreements, these partnerships can generate causal evidence about specific design features that impact teacher learning, addressing a critical gap in educational knowledge.