

15. IMPROVING LITERACY

Investing in Reading Coaching and Evaluation in Tennessee

Nate Schwartz is the chief research and strategy officer for the Tennessee Department of Education.

Laura Booker is the executive director of research for the Tennessee Department of Education.

Adam Rollins is a senior research analyst for the Tennessee Department of Education.



Between 2010 and 2015, Tennessee celebrated large gains across a wide variety of K-12 student outcomes. High school graduation rates increased steadily. More students enrolled in advanced coursework, and state test scores climbed. On the National Assessment of Education Progress (NAEP), Tennessee grew faster than any other state.¹

Elementary reading was the exception. In 2011, the state deemed 46 percent of third-graders proficient on the state exam that measured reading comprehension, vocabulary, and language skills and knowledge.² Five years later, the percentage meeting the bar was the exact same. In contrast, over the same time period, for students in third through fifth grades, proficiency rates in mathematics soared from 39 percent to 59 percent and average ACT scores climbed by half a point.³

The lack of growth in early reading proficiency raised doubts about the sustainability of the state's academic gains. Research from Tennessee and across the nation has repeatedly demonstrated the tight connection between early reading ability and later outcomes. By the middle grades, these differences start to harden. In Tennessee, fewer than 10 percent of eighth-graders who are reading below grade level meet the college-readiness benchmark on the 11th grade ACT exam. Without improvements in reading proficiency, Tennessee's long-term academic progress seemed likely to grind to a halt.

ISSUE BACKGROUND

In 2007, the U.S. Chamber of Commerce awarded Tennessee a failing grade in the category of "truth-in-advertising" to highlight the misleading information coming from the high proficiency rates on Tennessee's statewide assessment (around 91 percent proficiency in English language arts) as compared with the more objective standard measured by the NAEP (28 percent proficiency).⁴

Supported by a favorable political climate and two governors who made education their primary focus, the state responded by doubling down on more rigorous standards and graduation requirements over the next several years. Between 2007 and 2012, Tennessee raised standards in core academic subjects and made substantial investments in teacher training and performance evaluation, funded by a \$500 million Race to the Top grant.⁵ Improvements in most student outcomes across the time period seemed to vindicate this approach, making the lack of growth in elementary reading proficiency all the more problematic.

Tennessee's static reading results in the early grades were directly at odds with the level of effort districts were devoting to the issue. Each year, the vast majority of Tennessee districts placed elementary reading improvement as one of their highest priorities in the required annual plans they submitted to the state, and they supported these priorities through investments in trainings and resources.

When a new commissioner of education took office in January 2015, she announced that early grades literacy would be a major area of action for the Tennessee Department of Education. Her commitment to the issue was driven by evidence that being a proficient reader by third grade is foundational to future success and by the data showing Tennessee's lagging scores in this area.

EVIDENCE AVAILABILITY

Making research and data use a priority at the outset, the department launched its early reading initiative with a study of the landscape of teaching and learning in Tennessee K-3 classrooms. This needs assessment served to ground improvement efforts in a deeper understanding of the quality and variation in classroom instruction across the state.

The department outlined its dive into classroom practice in *Setting the Foundation*, the first in a series of reading reports that highlighted critical classroom needs.⁶ Classroom observations suggested that kindergarten through third-grade reading teachers across the state were overly focused on skills-based competencies like alphabet knowledge, fluency, spelling, and print concepts, with not enough time spent on the deeper comprehension and vocabulary—teaching practices that decades of literacy research suggest matter most for long-term student gains.⁷ Moreover, teachers rarely exposed students to high-quality texts that built their knowledge as well as their decoding and comprehension skills.

The *Setting the Foundation* report also suggested, as a statewide goal, that 75 percent of Tennessee third-graders be proficient in reading by 2025. Yet it wasn't clear how to shift teachers' practice in reaction to these needs in order to meet this goal. While professional-development programs tend to be popular among teachers and administrators, numerous experimental evaluations of teacher training programs have suggested that the programs often do not meet expectations for impacts on student learning or on teacher practice.⁸

Studies of teacher coaching are something of an outlier in this literature. As one meta-analysis including only randomized control trials and quasi-experimental evaluations concluded: “Coaching works. With coaching, the quality of teachers’ instruction improves by as much as—or more than—the difference in effectiveness between a novice and teacher with five to 10 years of experience, a more positive estimated effect than traditional [professional development] and most other school-based interventions.”⁹

Based on reviews of this research and the increasing prevalence of instructional coaches in Tennessee schools, the Tennessee Department of Education started to envision a strategy based around supporting the work of literacy instructional coaches across the state.¹⁰ Many districts and schools in the state had invested in instructional coaches as a lever for improving instruction, but coaches often reported little training on how to actually be a coach.

Nevertheless, this initial direction only served as the starting point for a series of far more difficult, practical decisions about the program itself. The challenge here was practice, not policy. While the evidence broadly suggested that coaching could work, these impacts tended to emerge in studies of smaller, more-intensive programs, and large-scale programs tended to be far less reliable.

The department did not have the option of replicating and scaling up a small-scale model. Funding literacy coaches in a large contingent of schools across the state was financially out of reach, even with a committed legislature. However, legislators were looking for a program that could reach the majority of districts in the state. As is the case with many such programs, state leaders faced a tension between the political need to create a program with a large enough reach to justify its expense and the reality that there was not a blueprint for success on a large scale.

In order to meaningfully shift classroom practice toward higher-quality literacy instruction across around 15,000 K-3 classrooms and 146 school districts, the department needed to use existing structures with innovative, evidence-based approaches. The department would also need to use ongoing evidence collection to inform continuous improvement in implementation.

EVIDENCE USE

The coaching strategy proposed by the Tennessee Department of Education to the legislature relied on a small group of centralized leaders—15 reading coach consultants across eight regions—who would build relationships with a far larger group of district-employed instructional coaches, who would then work with individual teachers across the state to build a shared understanding of early grades literacy practice. Each of the regional coaches would work with about 13 district coaches who in turn focused on supporting 15 teachers in their district. The district coaches would convene twice per year for focused training on reading instruction and coaching best practices.

This immediate setup was hurried along by legislative and budget timelines that freed up money in the late spring for a program directed to launch at the beginning of the following school year. It led to a program with the broad goal of transforming teaching across hundreds of classrooms but with only a rough outline of how to get there.

At the same time, program leaders across the department shared a long-term and genuine commitment to continuous improvement and openness to research and evaluation that is relatively rare in the political world of state agencies. Even as program leaders first began to piece together the outline of the coaching initiative, they worked with members of the department's embedded research and strategy team to lay out a framework for ongoing monitoring of short-term and long-term outcomes. This included thinking through a logic model for the initiative tied to concrete ways of assessing the stated outcomes. A member of the research team also accompanied program leaders to information sessions around the state for district leaders to share the data driving the program and to communicate about the evaluation plan. This partnership meant that ongoing evaluation became a primary element of program design, with discussions about data and evidence built directly into the planning and implementation process.

Embedded evaluation

As part of an embedded evaluation, the department committed to classroom observations that would allow monitoring of actual classroom practices over the three years of the coaching network, rather than relying on after-the-fact analyses of test scores to measure program effects. Through a stratified sampling plan to ensure representation on a variety of crucial school and district characteristics, the department identified 18 schools that experts could visit over time to track changes in teaching practices.

The department also built a system of continuous feedback from teachers and coaches, drawing on processes that were already in place across the state, specifically annual surveys of educators and district leaders, but adapting questions to meet evaluation needs. The team also developed new pre- and post-surveys for each of the six semesters of initiative. Making use of the statewide annual educator survey allowed for comparisons between teachers receiving coaching and a control group of uncoached teachers, and the supplemental surveys for participating coaches and teachers meant deeper information related to the program content. Tracking intermediate measures like changes to teacher and coach knowledge and practice was critical to knowing whether the program was on track to achieve its ultimate goal of moving the needle on student reading proficiency. These intermediate measures were especially important because statewide testing of students in Tennessee begins in third grade, meaning there was limited achievement data for assessing the ultimate outcome.

At the close of the first year, classroom observations suggested little to no positive movement across a variety of metrics of teacher practice. As program leaders studied these results, it became increasingly clear that the initial program aimed to cover so many elements of strong instructional practice that it left coaches and coach consultants with little focus on concrete classroom outcomes.

Meanwhile, the feedback from surveys and observations zeroed in on reading texts and curriculum as a crucial input to the success of the coaching program. On the early literacy branch of the state's annual educator survey, K-3 English language-arts teachers reported that misalignment between instructional standards and district curricula meant that they were spending an average of four and a half hours per week just sourcing the right texts and materials to use in their classrooms, leaving little to none of their crucial prep time for building a deeper understanding of their craft. Also teachers used coaches mostly as purveyors of instructional materials rather than as teaching experts. Moreover, surveys of coach time showed that many coaches struggled to actually spend time with teachers or to focus on observing and providing feedback, often because they were pulled into other duties.

Programmatic changes

Through ongoing conversations that were originally meant to ensure that classroom observers collected the right evidence, the program team ended up redefining the program itself.

Using the observational data-collection tool as a guide, program leaders narrowed in on the changes they'd like to see teachers make across the state, and modified the program to focus on (a) the quality of classroom texts; (b) question sequences and tasks selected to build conceptual knowledge; and (c) systematic foundational skills instruction that incorporated opportunities to practice through reading and writing. Interactions between regional coach consultants and district instructional coaches in year two aimed to solidify teachers' understanding of these key outcomes. District coaches also began to focus their interactions with teachers more solidly on these outcomes. The research team supported this effort by developing one-pagers that unpacked each of these three priority areas.¹¹

In response to findings from observations and surveys that the quality of instructional materials was a barrier to increasing the quality of instructional practice, in year two department leaders expanded the scope of the initiative to include a strategy focused on the materials and resources available to K-3 teachers. This included a short-term strategy of developing short units with tasks and question sequences built around a series of highly recommended text sets. It also included a long-term strategy to modify the textbook adoption process to ensure stronger alignment between the teaching techniques built into the coaching model and the curricula that the state recommended for adoption. Finally, in an attempt to bring school leaders directly into the work and thereby create more time and space for coaches to carry out their duties, the department held a series of regional principal meetings to invest principals in the vision.

Program impact

By the start of year three, observation results suggested that these program modifications had begun to yield meaningful changes in classroom practice. Observed teachers were increasingly introducing students to foundational skills within the context of genuine reading and comprehension exercises rather than creating an artificial divide between skills practice and the act of reading. Text quality in classrooms shifted, with more teachers using demanding texts that would introduce students to meaningful ideas and vocabulary. In

classrooms using the newly developed short units, almost all teachers were now using high-quality texts and one-third were using high-quality question sequences compared with only 10 percent before implementation of the units.

All the same, several crucial metrics of classroom practice, including the quality of student tasks had not moved. More problematic still, there were no changes yet in student test scores that evaluators could tie directly to the program. Student growth was difficult to measure since tests only began in grade three and the program focused primarily on grades K-2. Still, this remained the primary outcome that department leaders hoped to move.

The third and final year that the program had guaranteed legislative funding brought a new governor to the statehouse and a new commissioner to the department. While the program had achieved substantial popularity among district superintendents, it did not surface as a policy priority during the governor's election campaign. Instead, the governor, a fiscal conservative, had trained his sights on funding to improve career and technical education in high school and to increase school choice. As of this writing, it remains unclear whether the initiative will receive continued legislative funding. Yet despite the uncertainty, schools and districts across the state remain committed to the instructional improvements that the statewide coaching initiative spurred, and the work to increase the quality and rigor of instructional materials continues to be a priority of the new commissioner.

LESSONS

- ***Partnerships can benefit program quality.*** The partnership between the early literacy team and the research team at the department demonstrated the value of long-term, embedded evaluation. When the research team first formed in the Tennessee Department of Education, officials often called on team members to evaluate programs after the fact. The research team would do whatever was possible with existing data, which often led to evaluations of outcomes not tightly linked to programmatic goals. By involving the research team from the very beginning, we were able to create a joint partnership that simultaneously improved the quality of both the research and the program.
- ***Tie ongoing evaluation to the initiative.*** Program leaders within the department had reason to continually review their theory of action, since they were working to align evaluation metrics to that theory. By developing pre- and post-knowledge surveys for each of the in-person trainings offered to the district coaches, the program team had to could solidify their expectations. The research team also gave presentations of the ongoing findings to the regional coaches so that those implementing the program had the opportunity to grapple with the data.

- **Learn to be okay with not getting positive results.** By forcing difficult conversations, the continuous flow of sometimes negative evidence helped to give direction to year-by-year adjustments that resulted in a stronger program over time.
- **Evidence use doesn't guarantee funding.** Tennessee continues to grapple with the ways that the political realities of funding make it difficult to sustain a truly iterative process of continuous improvement over the long term.
- **Internal and external evaluations are both valuable and serve different purposes.** It is also important to note that there are benefits to having an internal research team like the one described in this case study, but there are also real reasons to have many aspects of the state's long-term evaluation work carried out by an independent and external organization. In Tennessee's case, the state greatly benefited from the Tennessee Education Research Alliance at Vanderbilt University, which works in partnership with the department to build knowledge around central priorities.¹² ■

- 1 State Collaborative on Reforming Education. *Priorities for Progress: 2018-2019 State of Education in Tennessee*. Available at: http://tnscore.org/wp-content/uploads/2019/02/2018-19-Annual-Report_18P1959_digital-FINAL.pdf.
- 2 Tennessee Department of Education. *Setting the Foundation: A Report on Elementary Grades Reading in Tennessee*. March 2016. Available at: https://www.tn.gov/content/dam/tn/readready/documents/coaching_network_docs/Setting_the_Foundation_Reading_Report.pdf.
- 3 Ibid.
- 4 Sheryll Poe. *Does Your State Make the Grade?* July 2014. Available at: <https://www.uschamber.com/does-your-state-make-grade>.
- 5 U.S. Department of Education. *Race to the Top: Tennessee Report Year 1: School Year 2010-2011*. January 2012. Available at: <https://www2.ed.gov/programs/racetothetop/performance/tennessee-year-1.pdf>.
- 6 Tennessee Department of Education. *Setting the Foundation: A Report on Elementary Grades Reading in Tennessee*. March 2016. Available at: https://www.tn.gov/content/dam/tn/readready/documents/coaching_network_docs/Setting_the_Foundation_Reading_Report.pdf.
- 7 Harvard University Lead for Literacy Initiative. "Literacy Unpacked: What Do We Mean By Literacy." Memo No. 4, 2012, 1. Available at: https://langlit.gse.harvard.edu/files/lesaux/files/04literacy_unpacked_langlit.pdf.
- 8 See, for example: M. Garet, et al. *The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement*. Institute of Education Sciences, September 2008. Available at: <https://ies.ed.gov/ncee/pdf/20084030.pdf>.
- 9 Matthew A. Kraft, D. Blazar, and D. Hogan. "The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence." *Review of Educational Research*, 88(4): 547-588, 2018.
- 10 Among Tennessee districts, 80 percent reported having instructional coaches, according to a 2015 survey.
- 11 Tennessee Department of Education. *First Steps: A Report on Elementary Grades Reading in Tennessee*. 2018.
- 12 Vanderbilt University. "Tennessee Education Research Alliance." Available at: <https://peabody.vanderbilt.edu/TERA/index.php>.