SCHOOLS THAT LEAD

Summative Evaluation

October 1, 2025

Prepared by the Education Policy Initiative at Carolina (EPIC)
Julie T. Marks & Andrew Otte

EPIC EDUCATION POLICY INITIATIVE *at* CAROLINA

SCHOOLS THAT LEAD EVALUATION - EXECUTIVE SUMMARY

Background

The North Carolina General Assembly authorized the Department of Public Instruction to contract with Schools That Lead (STL) to deliver a professional development program structured around networked improvement communities, and grounded in principles of improvement science. The legislation directed STL to work with schools across grade spans—high schools to improve on-time graduation, middle schools to reduce ninth-grade retention, and elementary schools to reduce the incidence of early warning indicators related to course failures, absences, and discipline. Unlike traditional top-down school improvement models, STL emphasizes capacity-building by equipping educators with improvement science tools and embedding them in a networked improvement community.

Sample: The evaluation sample for the current cohort of schools partnering with STL is comprised of three elementary schools, one middle school, an early college high school, and a traditional high school. Schools vary by size, urbanicity, and sociodemographics.

Data Sources: This evaluation of Schools That Lead (STL) employed a concurrent mixed-methods design, integrating qualitative and quantitative evidence to assess implementation, effectiveness, and impact.

Four primary categories of data were analyzed:

- Improvement Review Interviews: Semi-structured interviews where schools reported on their improvement strategies, successes, challenges, and impacts on students.
- Professional Development and Planning Artifacts: Close to 200 professional development (PD) presentations, agendas, planning documents, and implementation resources were categorized and analyzed to provide a detailed and holistic understanding of the STL PD.
- School-Level Watch List Data: Schools provided longitudinal "Watch List" tracking the number of students flagged for EWIs (attendance, behavior, and course performance).
- Administrative Data from NCDPI: North Carolina Department of Public Instruction (NCDPI) administrative records were used to track legislated outcomes of chronic absenteeism, suspensions, and graduation.

Program Implementation**:** Across participating schools, STL professional development offered a coherent framework for identifying problems of practice, testing interventions, and learning from results. Training in tools such as Plan-Do-Study-Act (PDSA) cycles, fishbone diagrams, and early warning indicator (EWI) watch lists gave educators structured methods to address pressing

challenges. These processes were reinforced through sequenced learning sessions, ongoing coaching, and the exchange of ideas within a networked improvement communi**ty.**

Impacts: Substantial improvements in student outcomes were seen across each of the early warning indicator (EWI) targets:

- Attendance: Reductions ranging from 45% to 82% in chronic absences with strategies like reframed letters, well-check calls, and same-day parent contact.

- Behavior: Meaningful drops in office referrals, suspensions, and disruptive incidents (e.g., one student's suspensions fell from 19 days to zero).

- Course Performance: Consistent grade improvements across settings, with failing students moving to passing or achieving gains between 10 and 30 percentage points on assessments.

Challenges: The evaluation also surfaced important challenges that could shape STL's impact trajectory. Schools struggled with limited staff capacity and time, uneven buy-in across faculty, and delays in accessing timely data. External disruptions such as leadership turnover threatened to diminish progress, but ultimately did not due to the structures and processes that schools developed and routinized through their work with STL.

Sustainability: Additional time will be needed to track true sustainability, but there is evidence of its potential. Many schools have embedded STL practices into daily operations, such as using watch lists, proactively reaching out to families, and holding regular data meetings. Principals and teacher-leaders have developed the capacity to carry this work forward, and in some cases, districts have aligned STL methods with their accountability systems.

Key Takeaway: STL represents a sustainable, transformative model that leverages educator expertise, strengthens school capacity, and improves student outcomes. Additional research is needed to see the extent these positive findings persist when scaled.

SCHOOLS THAT LEAD EVALUATION REPORT

I. BACKGROUND

In July 2018, the North Carolina General Assembly passed legislation requiring the Department of Public Instruction to contract with Schools That Lead (STL) to pilot a professional development program structured around the networked improvement communities model, and grounded in principles of improvement science. The three-year pilot concluded at the end of the 2020-21 school year, and showed strong evidence of implementation uptake, sustainable capacity building, and measurable impacts on key indicators that predict student success. As such, STL was granted another round of funding to expand their improvement network and continue to test and scale evidence-based school improvement efforts, continuing through 2026-27 school year.

The legislation directed STL to support K-12 schools on critical problems of practice, specifically: improving on-time graduation in high schools; reducing 9th grade retention out of middle school, and reducing the incidence of early warning indicators in elementary school related to course failures, absences, and discipline. The Education Policy Initiative at Carolina (EPIC) serves as the independent external evaluator of STL program. The scope of this report spans 2022-23 school year through summer of 2025, and will examine legislated outcomes as well the mechanisms of change that contribute to sustainable and scalable improvement.

II. EVALUATION DESIGN

The North Carolina General Assembly required an independent evaluation of Schools That Lead (STL) to determine its effectiveness in improving student outcomes in participating schools. The Education Policy Initiative at Carolina (EPIC) was awarded this contract for both phases of STL legislative funding. This evaluation addresses both implementation and impact, focusing on the extent to which schools adopted STL's improvement science methods and the degree to which there were measurable improvements in targeted student outcomes.

Sample: The evaluation sample for the current cohort of NIC schools spans all grade bands, and is comprised of three elementary schools, one middle school, an early college high school, and a traditional high school. Together, these schools serve a diverse student population that spans elementary, middle, and high school levels, with enrollments ranging from under 200 students in small, specialized schools to more than 1,000 students in larger suburban schools. This variation in school context is notable as it provides the opportunity to investigate whether there are enabling or inhibiting conditions for optimal program implementation. School characteristics and sociodemographics can be found in Appendix C.

Data Sources: This evaluation of Schools That Lead (STL) employed a concurrent mixed-methods design, integrating qualitative and quantitative evidence to assess implementation, effectiveness, and impact. Four primary categories of data were analyzed:

Improvement Review Interviews: Semi-structured interviews were conducted with four of the six participating schools. The remaining two schools did not complete interviews due to the cadence of their intervention cycles, which precluded reviews prior to this analysis. Interviews were transcribed and coded thematically, with analytic codes linked to: (a) early warning indicator (EWI) focal area, (b) type of intervention, (c) specific strategies employed, and (d) empirical evidence of student impact. This process enabled triangulation across schools and generated insight into the mechanisms underlying reported changes.

Professional Development and Planning Artifacts: A total of 191 artifacts produced between 2022–2025 were reviewed, including professional development (PD) presentations, agendas, planning documents, and implementation resources. Analysis of these artifacts emphasized PD session materials, allowing the evaluation team to trace the progression of learning objectives, focal areas, and implementation activities. Artifacts were then synthesized thematically to capture STL's overarching professional learning goals across the lifespan of the intervention.

School-Level Watch List Data

Six schools provided longitudinal "watch lists" documenting the number of students flagged for EWIs (attendance, behavior, course performance, and course performance). These data were examined for changes over time, with particular attention to the trajectory of identification (e.g., increases reflecting systematization of tracking) and subsequent reductions indicating targeted improvements.

Administrative Data from NCDPI: North Carolina Department of Public Instruction (NCDPI) administrative records were used to track schools' progress on legislated outcomes. These included graduation rates, school performance grades, student proficiency rates, and low-performing status. Due to delays in state-level reporting, attendance, discipline, and demographic data were not available at the time of this report. Where appropriate, school-level outcomes were benchmarked against statewide averages.

Analytic Approach

All qualitative data were open-coded to identify recurring themes and synthesized into categories aligned with STL's improvement aims. Quantitative analyses focused on descriptive trends in both watch list and administrative datasets. Together, these methods provided a robust

basis for assessing program implementation, progress on early warning indicators, and movement toward legislated outcomes.

III. PROGRAM DESCRIPTION

STL Model

Schools That Lead represents a distinctive approach to school improvement in North Carolina. Rather than prescribing a uniform set of practices, STL equips schools with tools and training grounded in the principles of improvement science and the structure of networked improvement communities (NICs). Improvement science emphasizes disciplined cycles of testing, learning, and adaptation, enabling educators to identify problems of practice and refine interventions through Plan-Do-Study-Act (PDSA) cycles (Bryk, 2018). NICs extend this capacity by linking schools together in structured networks, where educators share strategies, adapt promising practices, and accelerate collective learning (LeMahieu et al., 2017).

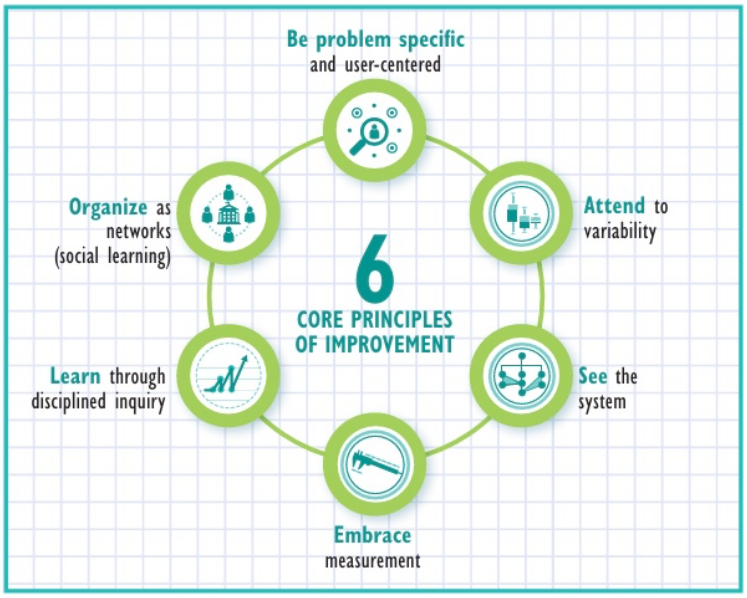
This approach contrasts with more traditional top-down school improvement models, which are often tied to the duration of external funding. STL places teachers and school leaders at the center of the change process, positioning them as the most proximate agents for improving student outcomes. In doing so, STL aligns with a growing body of research that emphasizes the importance of practitioner-led innovation and continuous improvement in education (Dolle et al., 2013).

Networked Improvement Communities (NIC) & Improvement Science

The STL approach is grounded in a Networked Improvement Communities (NIC) framework, a blend of improvement science and networked science, developed by the Carnegie Foundation. The hallmark of STL's continuous improvement initiative is a focus on incremental changes to address identified student learning issues.

The six principles of improvement science underlying the NIC model are as follows:

1) Make the work problem-specific and user-centered
2) Focus on variation in performance
3) See the system that produces outcomes
4) Improve at scale what you can measure
5) Use disciplined inquiry to drive improvement
6) Accelerate learning through networked communities



Bryk, A. S. (2018, April 3). Advancing quality in continuous improvement. Carnegie Foundation Summit on Improvement in Education, San Francisco, CA.

Early Warning Indicators

A central feature of STL's model is its focus on "early warning indicators" (EWIs) — measurable, research-based signals that identify students at risk for academic failure or dropping out. These indicators are most predictive when considered in combination, forming what is often referred to as the "ABCs" of student success: Attendance, Behavior, and Course Performance.

Montgomery County Public Schools (MCPS) conducted one of the landmark studies applying the EWI approach to its student population (West, 2013). Using longitudinal data from more than 11,000 students, MCPS identified specific cut-points in early grades that strongly predicted high school dropout (see Table 1). Importantly, these predictors are observable as early as Grade 1, and by Grade 6 nearly 9 out of 10 eventual dropouts could be identified using EWIs. The study reinforced earlier Johns Hopkins research (Balfanz & Byrnes, 2010; Allensworth & Easton, 2007) showing that the trajectory toward dropping out is detectable long before high school.

Table 1. Early Warning Indicator Risks (West, 2013)

Grade	Early Warning Indicator	Correlation with Dropping Out of School
1st grade, 3rd Marking Period	Absent 9 or more times	2x more likely to drop out
	Suspended	5x more likely to drop out
	Below grade-level in Math/ELA	2x more likely to drop out
	GPA below 1.2	2x more likely to drop out
3rd Grade, 1st Marking Period	Absent 3 or more times	2x more likely to drop out
	Suspended	9x more likely to drop out
	Below grade-level in Math/ELA	2x more likely to drop out
	GPA below 3.0	2x more likely to drop out

Each of these signals is independently predictive of later academic difficulty, but their predictive power is strongest when combined. For example, students who miss significant school time and also fail a core course are at markedly elevated risk of retention, disengagement, and eventual dropout.

Watch Lists

A central tenet of the STL model is that by systematically monitoring EWI thresholds (see Table 2) through school-level data systems and teacher input, educators can identify struggling students early and intervene before it negatively impacts their academic trajectory. In STL's model, the consistent use of EWIs allows schools to maintain dynamic "watch lists" that track students across time (see Figure X for exemplar) watch lists provide the foundation for testing interventions, studying their effects, and iteratively refining approaches in ways that are responsive to student needs.

Table 2. Early Warning Indicator (EWI) Thresholds

Early Warning Indicator	Threshold
Attendance	Missing 10 percent or more of instructional days (approximately 18 days in a 180-day school year).
Behavior	Receiving one or more out-of-school suspensions or multiple office discipline referrals in a given year.
Course Performance	Failing a core course such as English language arts, mathematics, or science during the school year.

IV. IMPLEMENTATION FINDINGS

Program Components

Schools That Lead (STL) provides a structured yet adaptive professional development model that positions schools to engage in continuous improvement. Rather than prescribing a curriculum or set of standardized practices, STL builds local capacity by equipping educators with the tools, routines, and dispositions necessary to identify problems, test solutions, and sustain what works. The model integrates principles of improvement science with school-based leadership structures and networked learning, organized around the following topical trainings. (See Appendix A for PD Session sequence and scope)

Early Warning Indicators and Watch Lists: A central feature of STL is centered around improving the "A,B,Cs" of Early Warning Indicators (EWIs) - attendance, behavior, and course performance - which are evidence-based predictors of student disengagement and dropout. Schools are trained to create and maintain a dynamic Watch List of students who meet the threshold for each of these indicators (see Figure 1), updated at regular intervals across the school year. The professional development provided by STL emphasizes the technical aspects of data management (frequency, accuracy, subgroup analysis) and the interpretive practices that help educators identify specific targets for intervention. Furthermore, piloting these interventions among a small group of students over a short period of time provides important information about barriers and contextual factors that get masked in large-scale change initiatives.

Figure 1. Early Warning Indicator (EWI) Watch List

Year 2 Data Revised

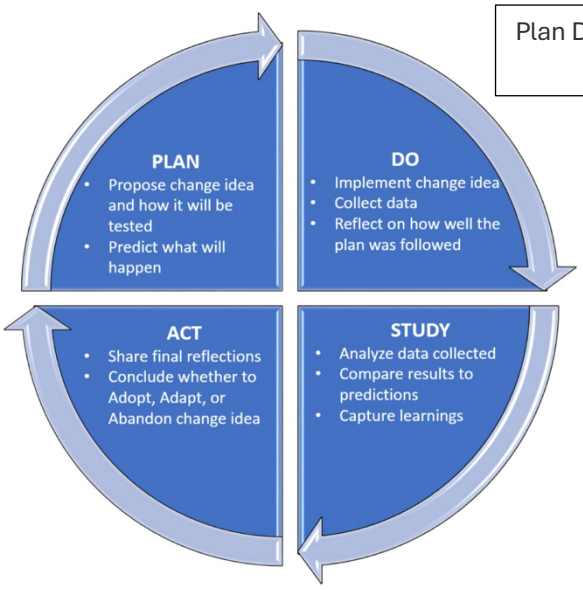
				Attendance				Behavior				Math				ELA			
Data Point	Date	Total	Watch List	6th	7th	8th	Total	6th	7th	8th	Total	6th	7th	8th	Total	6th	7th	8th	Total
BOY	9/19/23	196	15	2	7	4	13	0	1	1	2	0	0	0	0	0	0	0	0
Progress Report Q1	10/23/23	196	15	1	4	4	9	2	4	0	6	0	0	0	0	0	0	0	0
Report Card Q1	11/27/23	195	20	1	3	4	8	2	3	7	12	0	0	0	0	0	0	0	0
Progress Report Q2		193	17	1	1	3	5	2	4	6	12	0	0	0	0	0	0	0	0
Report Card Q2	1/30/24	193	17	1	0	2	3	3	5	6	14	0	0	0	0	0	0	0	0
Progress Report Q3		193	25	3	3	5	11	2	5	7	14	0	0	0	0	0	0	0	0
Report Card Q3	04/07/24	193	29	4	6	6	16	1	4	8	13	0	0	0	0	0	0	0	0
Progress Report Q4		193	18	3	4	6	13	2	3	0	5	0	0	0	0	0	0	0	0
Report Card Q4	06/06/24	193	15	3	3	6	12	0	3	0	3	0	0	0	0	0	0	0	0
EOY		193	39	3	3	6	12	4	10	13	27	0	0	0	0	0	0	0	0

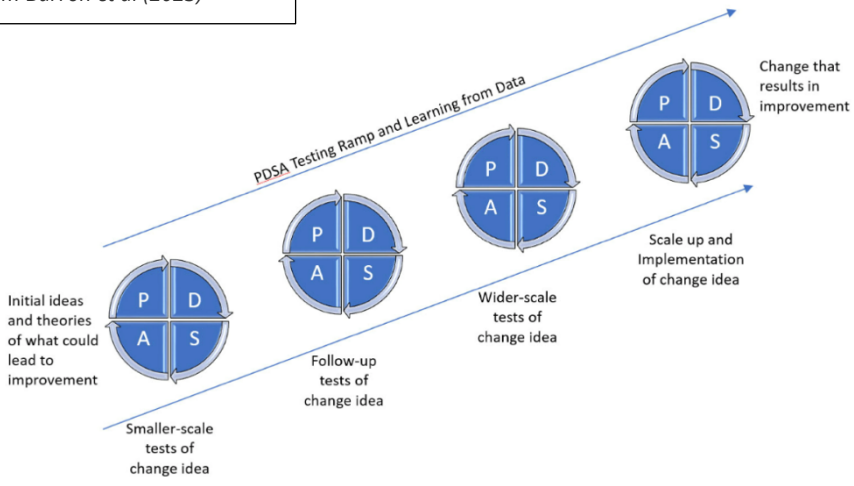
Year 1 Data Revised

				Attendance				Behavior				Math Course Performance				ELA Course Performance			
		Total	Watch List	6th Grade	7th Grade	8th Grade	Totals	6th Grade	7th Grade	8th Grade	Totals	6th Grade	7th Grade	8th Grade	Totals	6th Grade	7th Grade	8th Grade	Totals
Progress Report Q1	10/14/22	188	114	33	30	31	**94**	3	5	6	**14**	3	2	0	**5**	0	0	1	**1**
Report Card Q1	11/21/22	190	127	38	31	29	**98**	6	10	8	**24**	3	2	0	**5**	0	0	0	**0**
Progress Report Q2	6/19/23	190	105	29	29	26	**84**	5	6	3	**14**	2	2	1	**5**	1	0	1	**2**
Report Card Q2	6/19/23	190	85	23	23	21	**67**	4	6	3	**13**	1	1	1	**3**	1	0	1	**2**
Progress Report Q3	3/3/23	193	81	16	16	14	**46**	6	10	9	**25**	4	2	1	**7**	1	0	2	**3**
Report Card Q3	6/19/23	188	54	13	14	14	**41**	3	2	2	**7**	1	1	1	**3**	1	1	1	**3**
Progress Report Q4	6/19/23	188	48	12	10	10	**32**	4	3	3	**10**	1	1	1	**3**	1	1	1	**3**
Report Card Q4	6/19/23	188	40	9	8	8	**25**	3	3	3	**9**	1	1	1	**3**	1	1	1	**3**

Continuous Improvement through PDSA Cycles: STL situates its professional learning within the improvement science framework of Plan-Do-Study-Act (PDSA) cycles. Schools are encouraged to conduct short (3–8 week) cycles in which they test "change ideas" with small groups of students, collect data on outcomes, and then decide whether to adopt, adapt, or abandon the intervention. This structured process reduces the risk of prematurely scaling ineffective practices, and builds an evidence-base about what works, for whom, and under what conditions.

Plan Do Study Act (PDSA) Cycles
From Barron et al (2023)





Huddles and Snowflakes: Sustained improvement requires organizational structures to monitor and support program implementation. STL provides guidance to schools on developing and maintaining two complementary structures. 1) *Huddles* are short, biweekly meetings of the principal, improvement facilitator, and teacher-leaders. The primary focus of these meetings is to maintain momentum for program activities, review Watch List data, and plan next steps; and 2) *Snowflakes* serve to extend improvement testing to teachers outside the core improvement team, enabling broader participation and more opportunities to test promising improvement strategies in multiple settings.

Sustainability and Scale: Later sessions in the professional development sequence address sustainability. Schools are trained to monitor momentum using the heuristic "mass × velocity," where mass represents the number of testers engaged and velocity represents the frequency of PDSA cycles. They are also prepared to synthesize their learning into Improvement Reviews, where data, documented tests, and qualitative insights are assembled into a cohesive account of progress. These reviews serve both as an accountability mechanism and as a means of building internal capacity for reflective practice and communication of results.

Table 3. STL Core Improvement Processes

Process	What	How Often	Why
Watch List	Data collected and shared with all teachers	Quarterly	Early warning indicators identify which students are off-track; provides focus for improvement ideas and demonstrates progress
PDSA	Improvement ideas being tested by huddle members	3-8 weeks	Learning what works, for whom, in what context; gain confidence and know-how; be able to share with others for further testing
Huddle	Principal, Improvement Facilitator and 2 teacher-leaders	At least every other week; 15 minutes	Share and make sense of data; discuss theory of improvement; share PDSAs and outcomes; determine next steps; maintain momentum
Snowflake	Teachers outside the huddle	Short sprints – 3-5 weeks	Learning more about what works, for whom, in what context; determines warrant if the

			improvement idea reliably leads to improvement in multiple contexts

V. IMPACT FINDINGS

Within School Impacts

A unique aspect of the STL model is the prioritization of real-time data collection while an improvement strategy is being implemented. The following table highlights multiple improvement approaches within each school, and their observed impacts on early warning indicator targets.

Within-school Improvement Strategy	Early Warning Indicators	Observed Impacts
- Morning check-ins, Breakfast Club, and positive parent phone calls - Lunch Bunch fluency groups - Sentence stems for participation in ELA - Immediate parent contact for absences - Aligned Watch List with MTSS	-Attendance -Behavior -Course Performance	- Watch List reduced from 78% to 57%. - Fluency Lunch Bunch students gained >40 words per minute; some exited MTSS. - Sentence stems boosted participation and academic averages - Rapid attendance improvements; systemic barriers noted - Improved empathy and relationships with families
- Positive phone calls - Attendance letters - Mentoring/"adopting" students - Empathy interviews and perspective-taking PD - PLCs shifted toward strategy-focused	-Attendance -Behavior -Course Performance	- Improved tracking of student attendance, substantially increased number of students identified as "at risk" - 25 staff conducted empathy interviews, reporting improved relationships with students and colleagues - Students formed new mentoring relationships

discussions		- Despite staff turnover, core team sustained STL through PD and mentoring
- Attendance letters for mid-level absences - Retest request forms for assessments - Bathroom escort tracking via Ring Central - Charting missing assignments - Adopt-a-Kid mentoring	-Attendance -Behavior -Course Performance	- Attendance letters produced short-term increases in attendance, additional strategies needed for sustained improvements - Retest forms strengthened student persistence; some student scores improved consistently over time - Bathroom escort tracking reduced time out of class and improved communication - Charting missing assignments was dependent on buy-in, which was mixed early on - Adopt-a-Kid slowed disciplinary referrals, improved academics outcomes for some students.
- Reframed attendance letters (Spanish translation) - Immediate parent phone calls, transportation offers - Student-parent communication via biweekly student emails - Positive behavior contracts, proactive staff support - Extended recess for work completion - Interactive reading logs - Family engagement via local Episcopal Church and Latino resource fair	-Attendance -Behavior -Course Performance	- Watch List reduced from 45% to 23% - Spanish letters improved communication - phone calls + transportation boosted attendance same-day - Students with missing work decreased by over 50% by end of year - Reading logs boosted interest and ownership in reading - Increased empathy toward families; staff recognized structural barriers rather than lack of caring

Cross-network Improvement Strategies

This matrix highlights the value of a networked improvement community (NIC), where multiple schools are investigating common problems of practice through an improvement science lens. Schools benefit from a rapidly expanding evidence base of effective practices, with strategies often being tested concurrently across multiple contexts and student populations. Network participation also enables opportunity for cross-site knowledge exchanges, which educators frequently described as an uncommon undertaking in an often-siloed profession. See Appendix B for further description of improvement strategies.

Cross-school Improvement Strategy	Schools Implementing	Observed Impacts
Attendance letters (reframed/customized + Spanish translation)	Foothills, Franklin, South Newton	Short-term improvements in attendance for students receiving letters. Sustained change likely requires repeated follow-ups over an extended period. Translating materials into Spanish improved two-way communication with Spanish-speaking families.
Rapid response to student absences (immediate parent contact, staff-provided transportation in some cases)	Foothills, South Newton, West Marion	Directly increased attendance, often resulting in same-day student return.
Positive parent phone calls (behavior-focused)	Foothills, West Marion, East Garner	Reduced frequency of minor behavioral infractions; improved parent receptivity to teacher communication; shifted expectations toward valuing positive feedback.
Incentives (house system, recess, prize wheel)	Foothills, South Newton, West Marion	Increased number of students with perfect attendance. Decreased number of students with missing work. Reported increases in student motivation to complete high-quality work on time.
Academic supports (fluency groups,	West Marion, Franklin, South Newton, Foothills	Fluency gains >40 WPM. Greater student interest in reading and

sentence stems, retest forms, reading logs)		engagement. Strengthened academic persistence.
Family engagement (empathy interviews, student-parent communication, resource fairs)	East Garner, South Newton	Increased teacher empathy; stronger parent–school connections; empowered parents despite structural barriers.

Improvement Science Snapshot:
Foothills Community School

School Context

Foothills Community School is a small STEM-focused public school located in McDowell County. The school serves a diverse population and joined Schools That Lead to strengthen its capacity to use improvement science to address persistent challenges with attendance and student engagement. The leadership team set an ambitious aim to reduce the percentage of students on the early warning indicator Watch List, with particular attention to chronic absenteeism and behavior.

Improvement Strategies

Foothills leveraged their improvement science tools and processes to develop and implement multiple Plan-Do-Study-Act (PDSA) cycles for the purpose of testing and iterating improvement ideas – small, targeted, and measurable changes - with a potential to impact multiple early warning indicators.

Promising Practices

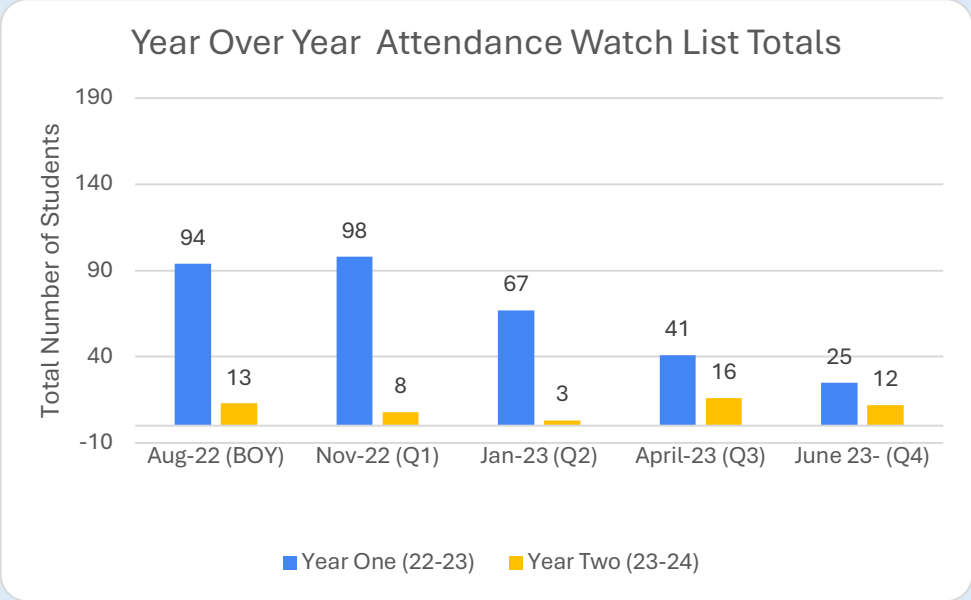
- ✓ Launching positive phone calls home for students with repeated minor behavior incidents.
- ✓ Introducing a house system with rewards for perfect attendance, including prize wheel spins.
- ✓ Using a Mobile STEM Lab to engage disengaged students by involving them in outreach to elementary schools.
- ✓ Establishing daily huddles and systematic data checks to build staff-wide responsibility for improvement.

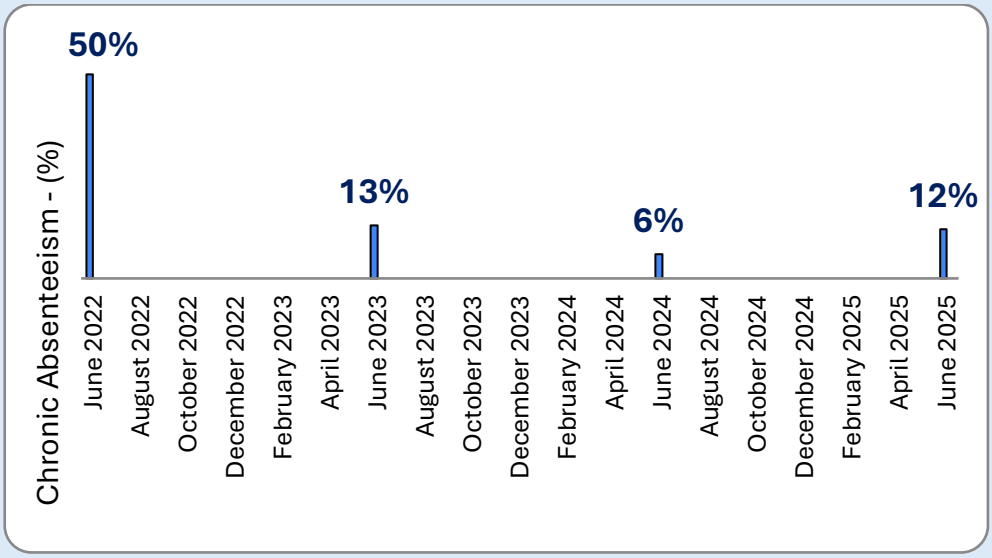
Observed Impacts

Foothills reported significant progress on its aims. Attendance improved dramatically, with the percentage of students on the Watch List dropping from nearly 60 percent to just 8.2 percent. Behavior referrals also decreased, with parents expressing satisfaction with positive communications about their child. As one staff member reflected, “The parent was starting to expect that good thing,” showing a cultural shift in home-school relationships.

The house system also paid dividends, with perfect attendance climbing from the teens in the prior year to more than 80 students by May. A particularly striking example came from an eighth grader who was failing and chronically disengaged. After being given a role in the Mobile STEM Lab, he completed all assignments, improving from 10 missing assignments to zero. The school

recognized him as the most improved student, underscoring the whole child impact of these interventions.





Conclusion

Foothills Community School demonstrates how improvement science can translate into measurable gains through small but powerful interventions. By correcting data systems, testing targeted strategies, and fostering a culture of trust and collaboration, the school moved substantially closer to its goals. The experience highlights how building staff ownership and creatively engaging students—particularly through STEM opportunities—can reduce early warning indicators and improve outcomes. The case illustrates STL's broader promise: that schools, empowered to test and refine their own solutions, can meaningfully improve student success.

V. CHALLENGES

While Schools That Lead (STL) has demonstrated promising impacts on student outcomes, the evaluation also surfaced challenges that schools faced in implementing and sustaining improvement initiatives. These challenges highlight the contextual barriers that influence program success and provide insight into where additional supports may be needed for broader scaling.

Capacity and Time Constraints: A cross-cutting challenge described by schools was the need to balance competing priorities for staff time, many of whom are already at capacity. Teachers and principals appreciated that the STL model focused on building evidence at a smaller scale, compared with past initiatives that overhauled their entire system and hoped for change. However, they noted that the process of continuous improvement requires a time commitment to test change ideas, collect data, and reflect on results. Understandably, some educators expressed it would be a challenge to layer anything on top of their already demanding workloads.

Data Quality and Access: Schools reflected on start-up challenges surrounding the availability and usability of data. Difficulties included delays in data availability, and needing to adapt data from their student information system to fit the Early Warning Indicator tracking for their Watch Lists.

Variability in Staff Buy-In: Although STL's professional development model emphasizes collective responsibility, schools reported varying levels of staff engagement. In some cases, teacher buy-in was strong within smaller teams but uneven across the broader faculty. This variability limited the spread of certain initiatives beyond early adopters and created difficulties in scaling and sustaining improvement work beyond a core group of educators.

External Shocks and Contextual Factors: Several schools experienced external disruptions that hindered implementation, such as leadership turnover or competing district-level initiatives. While described as setbacks, educators relayed that they persevered and were able to successfully continue with improvement work due to the established structures and processes through their work with STL.

Taken together, this suggests that the impact of the STL model may be moderated by organizational conditions within participating schools. Addressing barriers around time, data access, and staff buy-in will be critical for ensuring that promising practices diffuse beyond the early adopters and ultimately contribute to systemic change.

VI. SUSTAINABILITY

Evidence from participating schools suggests that several aspects of the Schools That Lead (STL) model show strong potential for sustainability beyond the funded period.

Leadership and teacher-leader development emerged as a key facilitator of sustainability. Teacher leaders facilitate use of improvement processes within their schools, such as guiding colleagues through Plan-Do-Study-Act (PDSA) cycles or leading data-focused conversations. Shared "ownership" of STL implementation enabled practices to persist even amid administrator turnover.

District-level alignment can provide structural support for sustaining the STL work. Some districts began incorporating STL tools into their broader accountability and planning frameworks, reinforcing schools' ability to sustain STL-related work. This institutionalization at the district level increases the likelihood that STL practices will continue across multiple years.

Incorporating key practices into daily school operations signals a commitment to lasting change. Practices such as the systematic use of early warning indicator (EWI) watch lists, proactive family outreach, and structured data meetings were reported to be increasingly embedded into schools' daily operations. Teachers and principals described these practices as part of their regular routines rather than temporary initiatives, a critical element for organizational transformation.

At the same time, sustainability is not guaranteed. Schools identified ongoing risks, including staff turnover, leadership changes, and limited resources to maintain the intensity of professional development and coaching support. Without continued investment, there is a danger that STL practices could erode over time, particularly in schools already facing capacity challenges.

Overall, the evaluation finds that STL has laid the groundwork for durable change by embedding improvement science practices into school culture. The extent to which this progress is sustained will depend on maintaining leadership capacity, district alignment, and adequate resources to support ongoing professional learning.

VII. CONCLUSION & TAKEAWAYS

Key takeaways

Taken together, the evidence indicates that STL represents a distinctive and promising approach to school improvement in North Carolina. Unlike one-size-fits-all interventions, STL builds capacity within schools to identify problems, test strategies, and adapt based on evidence. This model not only addresses immediate challenges in attendance, behavior, and course performance but also creates the conditions for long-term, systemic improvement.

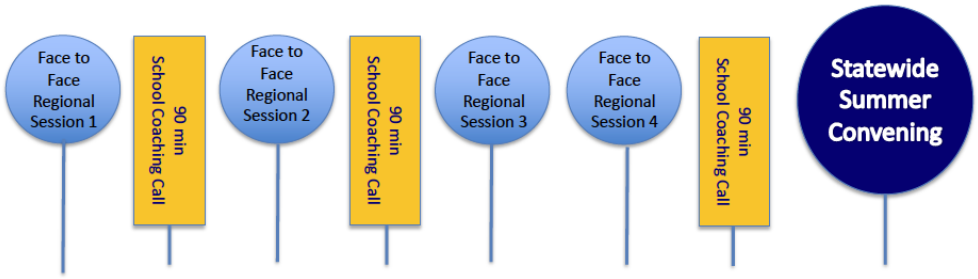
While this evaluation did not include a formal cost-analysis, it is clear that STL offers a strong return on investment due to the very nature of the program. Namely, STL provides tools and structures that empower schools to solve their own problems, thereby reducing reliance on costly external mandates.

The reliance on data to test improvement strategies has enabled tracking of measurable improvements in critical elements of children's schooling, and the emphasis on improvement science and networked learning provides the infrastructure for those improvements to deepen and spread. In short, Schools That Lead stands apart from most state-level initiatives as a transformative model of school improvement that positions classroom educators and building leaders as the architects and drivers of change.

Additional research is needed to establish whether these positive outcomes hold at increased scale, however current evidence clearly demonstrates the capacity of STL schools to design and test context-specific strategies that address the root causes of underperformance. These findings provide compelling evidence that the impacts of STL are consistent with its legislative goals, fostering measurable improvements in indicators most predictive of student success.

APPENDIX A. STL PD SCOPE & SEQUENCE

Over the course of a school year, participating schools attend four full-day professional development sessions, participate in three 90-minute coaching calls, and end their year with a summer convening. Each session builds on the last, with the following scope and sequence.



Session 1: Foundations and Watch Lists - Schools learn to create and maintain a Watch List using Early Warning Indicators (attendance, behavior, and course performance). They set year-end aims, establish data submission routines, and begin testing small improvement ideas using Plan-Do-Study-Act (PDSA) cycles.

Session 2: Using Data and Strengthening Huddles - Schools deepen their analysis of Watch List data, learning to subgroup students to identify patterns and target strategies. They clarify the purpose and rhythm of Huddles, and begin identifying additional teachers to spread improvement efforts.

Session 3: Making Learning Public and Engaging Faculty - Schools prepare to introduce Watch Lists to the full faculty, develop communication plans, and anticipate resistance. They share PDSA results more widely, practice declaring whether ideas should be adopted, adapted, or abandoned, and build momentum by expanding the circle of testers.

Session 4: Sustaining Momentum and Scaling Practices - Schools assess their progress against aims, update Watch List data, and plan how to sustain or regain momentum. Using the metaphor "momentum = mass × velocity," they focus on engaging more testers (mass) and running more frequent cycles (velocity). They also plan for upcoming Improvement Reviews to demonstrate their progress.

Summer Convening: Reflection, Storytelling, and Planning for Year Two: Each year culminates in a summer convening where schools consolidate their learning, tell their "data stories," and set aims for the coming year. Schools hear from peers further along in the network, reflect on lessons learned, and design plans for a strong kickoff in the fall.

APPENDIX B. TESTED IMPROVEMENT STRATEGIES & MEASURED IMPACTS

EWI: Attendance

Attendance Letters: Customized letters sent to families of students with mid-level or chronic absences, often translated into home languages such as Spanish. Letters were tailored to highlight the importance of attendance and to notify parents of their child's specific number of missed days.
Impacts: In one school, 15 of 18 targeted students improved attendance in the following quarter. However, gains often faded by the next reporting period if letters were not paired with follow-up strategies.

Immediate Parent Phone Calls: Teachers or staff contacted parents as soon as a student was marked absent, sometimes offering direct support such as transportation. Calls emphasized problem-solving and encouragement rather than punishment.
Impacts: Increased same-day attendance in multiple cases. For example, at one school, students who missed the bus were brought to school later that day following a parent call and transportation support.

Attendance Incentives (House System, Prize Wheel, Extra Recess): Schools implemented reward systems such as house competitions, spins on a prize wheel, or extended recess for perfect attendance. These incentives were designed to make consistent attendance a visible and celebrated achievement.
Impacts: At Foothills, perfect attendance rose from the 'teens' the prior year to more than 80 students per month by May. Other schools reported steady increases in the number of students qualifying for recognition.

EWI: Behavior

Positive Phone Calls Home: Weekly calls to parents highlighting positive student behaviors and achievements. Calls were deliberately designed to counterbalance typical negative-only contact from schools.
Impacts: In multiple schools, minor behavior incidents and referrals decreased. One principal reported that 'the parent was starting to expect that good thing,' illustrating a cultural shift in school-home communication.

Behavior Contracts: Agreements between students and staff outlining specific behavior expectations and rewards for meeting them. Contracts were often paired with regular check-ins and reinforcement.

Impacts: Helped students develop self-regulation strategies. At South Newton, students with contracts successfully exited the behavior Watch List after sustained progress.

Adopt-a-Student / Mentoring: Staff members 'claimed' specific students on the Watch List and built consistent mentoring relationships through check-ins and encouragement. This strategy personalized adult-student connections.
Impacts: Reduced discipline referrals and suspensions in schools such as Franklin. In East Garner, students sought out their mentors independently, with one student visiting a teacher before leaving campus for the summer.

EWI: Course Performance

Lunch Bunch Fluency Groups: Small-group reading fluency practice sessions held during lunch, combining skill practice with relationship-building. Groups targeted struggling readers flagged by MTSS.
Impacts: Students gained more than 40 words per minute within a semester. Several students exited MTSS following participation, and teachers reported increased confidence.

Sentence Stems for Participation: Structured prompts such as 'I think...,' 'I wonder...,' and 'I agree...' were provided to encourage hesitant students to contribute in class discussions.
Impacts: Voluntary participation jumped from 2–3 times per week to 10–11 times. Course averages also improved, with some students moving off the academic Watch List.

Retest Request Form: Students completed a reflection form before retaking an assessment, identifying mistakes, and outlining new preparation strategies.
Impacts: Students who used the form consistently improved scores. In one case, a student retook nearly every assessment and raised overall grades from C/Ds to higher levels. Teachers also noted stronger persistence and self-advocacy among participating students.

Missing Work Interventions (Charts, Incentives): Teachers tracked missing assignments for small groups of students and paired accountability charts with incentives such as extended recess.
Impacts: At one school, the number of students with chronic missing assignments declined from 17–18 to 6 by year's end—a reduction of over 60 percent.

Interactive Reading Logs: Reading logs were redesigned to include comprehension questions and reflective prompts, making them more engaging and meaningful for students.
Impacts: Increased student ownership and interest in reading. One school expanded the strategy across an entire grade level after observing steady student gains in comprehension and engagement.

APPENDIX C. LEGISLATED IMPACTS

At the time of this report, NCDPI had not completed their full annual transfer of data to EPIC. The delayed transfer includes data for two of the legislated outcomes, chronic absenteeism and short-term suspensions. That data is expected to be available to EPIC in October, at which time an amended report will be resubmitted.

I. CHRONIC ABSENTEEISM *

	% Chronic Absenteeism		
	2021-22	**2022-23**	**2023-24**
West Marion Elementary School	53.1	32.5	28.8
East Garner Middle	23.5	18.1	28.0
Foothills Community School	29.1	21.6	20.9
South Newton Elementary	29.6	17.2	14.2
McDowell Early College	2.3	2.2	3.5
McDowell Academy of Innovation	10.4	3.6	11.0
State Average	31.7	26.5	25.0

**[To be updated upon receipt of 2024-25 data]*

- Administrative data reflected the substantial improvements in absenteeism reported by case study schools who focused on reducing absences
- Between 2021-2022 and 2023-2024:
    - Absenteeism rates for one elementary school was reduced by half, from 29% to 14%
    - Another elementary school saw absenteeism decrease almost in half from 53% to 28%.
    - Absenteeism rates for the middle school were reduced by a third from 29% to 20%.

- For the three remaining schools:
    - one maintained absenteeism rates below 5% across all three years
    - two had trends where absenteeism increased from 2022-23 to 2023-24. ‘

An important consideration when interpreting these latter results is that, by design, the improvement science model requires schools to select a single area where they would like to pilot improvement ideas. The two schools where absenteeism did not decrease had the two lowest levels of absenteeism to start and thus may have opted to focus on different early warning indicators

II. SHORT-TERM SUSPENSIONS*

	Short-term Suspension Rates		
	2021-22	**2022-23**	**2023-24**
West Marion	2.84	0.00	84.47
East Garner Middle	229.93	287.05	242.42
Foothills Community School	11.17	42.33	62.18
South Newton Elementary	211.77	86.05	33.11
McDowell Early College	0.00	0.00	26.43
McDowell Academy of Innovation	77.67	9.01	24.00
State Average	146.57	164.57	162.34

Short-term suspension rates are per 1000 students

**[To be updated upon receipt of 2024-25 data]*

- Short-term suspension data lacked any clear patterns
- This is consistent with findings from prior phases of the evaluation, likely reflecting that the STS measure is not sensitive to the types of changes in discipline or behavior being targeted by improvement efforts
- The evaluation team proposes collecting primary data through surveys, focus groups, and interviews before drawing any generalizable conclusions around impacts on discipline.

III. HIGH SCHOOL GRADUATION RATES

	2021-22	**2022-23**	**2023-24**	**2024-25**
McDowell Early College	98.5%	>95*%	>95%	>95%
McDowell Academy of Innovation	89.4%	>95%	>95%	>95%
Statewide Average Graduate Rate	84.3%	86.5%	87.0%	87.8

NCDPI is no longer reporting specific graduation rates above 95% or below 5%

- STL high schools have sustained graduation rates above 95% since 2022–23, consistently outperforming the state average.
- This stability is notable given pandemic-related disruptions and ongoing statewide variation.

IV. STUDENT PROFICIENCY

	% Proficient Math 2023-24	% Proficient Math 2024-25	% Proficient ELA 2023-24	% Proficient ELA 2024-25	% Proficient Science 2023-24	% Proficient Science* 2024-25
West Marion Elementary	D	49.7	45.1	46.6	76.4	60.3
East Garner Middle	35.9	43.1	35.5	43.1	55.2	54.6
Foothills Community School	63.4	67.9	62.3	74.5	88.0	75.0
South Newton Elementary	56.9	62.3	45.1	50.0	54.3	36.0
McDowell Early College	90.0	90.2	>95	>95	84.4	77.8
McDowell Academy of Innovation	>95	92.0	>95	>95	>95	76.5
Statewide	53.1	55.7	51.5	53.5	63.0	56.7

*New science standards implemented in 2024-25

- ELA: The overall percentage of students GLP in ELA varied considerably across STL schools, ranging from 43% up to >95%
    - All STL schools below 95% proficiency in ELA increased the percentage of students proficient from 2023-24 and 2024-25, however half the schools were still below the state average of 53%.
- Percentage of students GLP in Math increased in 4 of the 6 schools, and 4 of 6 STL schools had a greater proportion of students proficient in math than the state average rates compared with the state.
- All science scores dropped between years, which is reflective of trends seen statewide due to the change in science standard.

Tables below are not STL legislated outcomes

IV. SCHOOL PERFORMANCE GRADES

	2021-22	2022-23	2023-24	2024-25
West Marion Elementary	D	C	C	D
East Garner Middle	F	F	D	D
Foothills Community School (MS)	B	B	C	B
South Newton Elementary	D	D	C	C
McDowell Early College (HS)	A	A	A	A
McDowell Academy of Innovation (HS)	B	A	A	A
State Average	C	C	C	C

IV. NC NIC SCHOOL GROWTH STATUS

	2022-23	2023-24	2024-25
West Marion Elementary School	Met	Not Met	Not Met
East Garner Middle	Not Met	Met	Met
Foothills Community School	Met	Not Met	Met
South Newton Elementary	Not Met	Met	Not Met
McDowell Early College	Exceeded	Exceeded	Exceeded
McDowell Academy of Innovation	Exceeded	Exceeded	Exceeded
State Average	Met	Met	Met

V. 2023-24 SCHOOL-LEVEL DEMOGRAPHICS*

	School Size	% Caucasian	% African American	% Latino	Rural/Urban	% Low Income	% Teacher Turnover
West Marion Elementary	367	77	4	13	Rural, Fringe	70.8	15.6
East Garner Middle	1089	14	49	31	City, Large	72.4	37.6
Foothills Community School	193	87	2	12	Rural, Fringe	49.0	8.3
South Newton Elementary	302	36	17	28	Suburb, Midsize	75.8	29.2
McDowell Early College	227	75	2	19	Rural, Fringe	42.5	16.7
McDowell Academy of Innovation	125	68	6	23	Rural, Fringe	53.2	12.5
NC	1,508,194	43.7	25.4	29.5	--	58.7	23.5

**[To be updated upon receipt of 2024-25 data]*