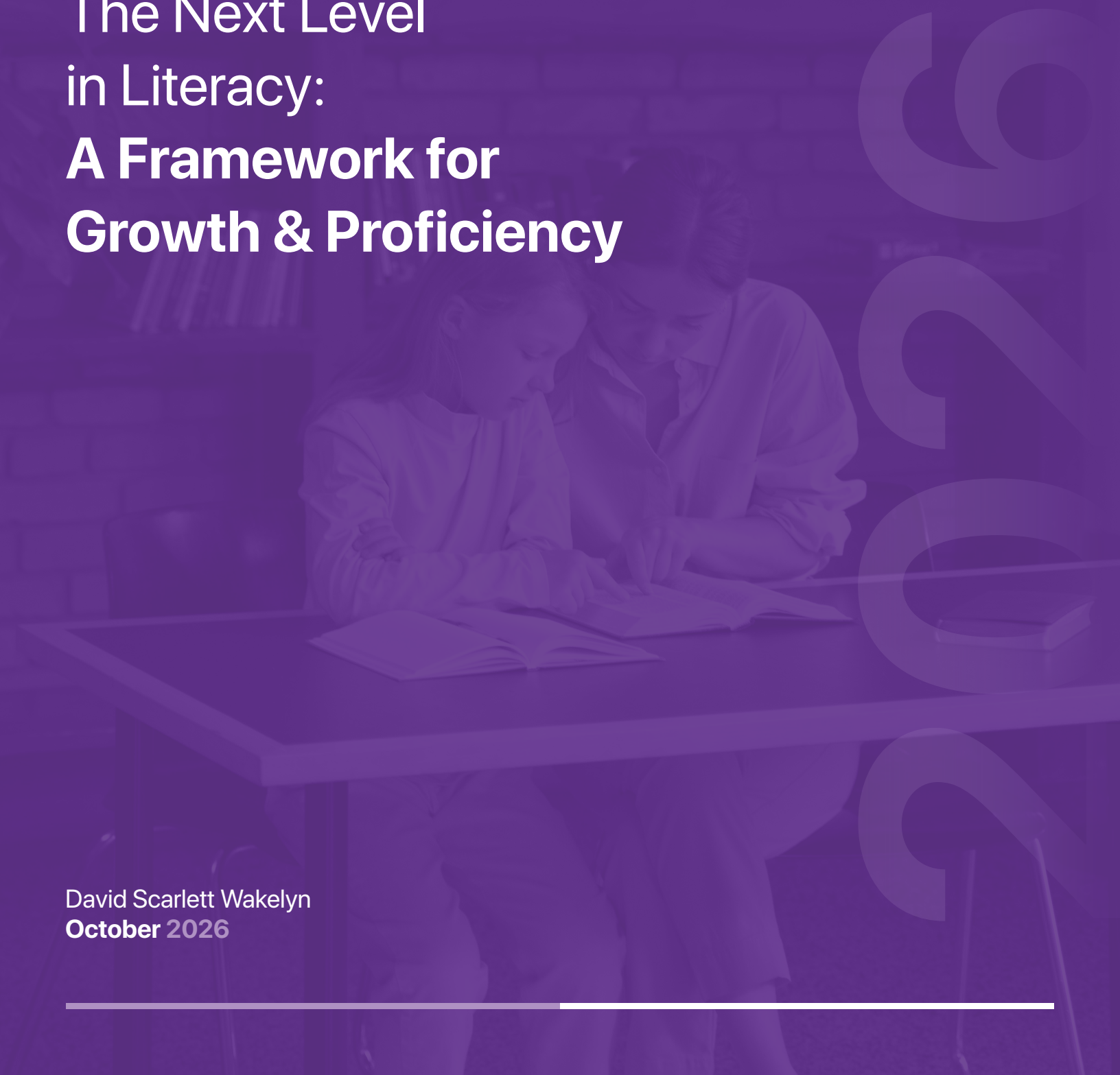
UpswingLabs

The Next Level in Literacy: **A Framework for Growth & Proficiency**



David Scarlett Wakelyn
October 2026

Executive Summary

This report is written for chief academic officers, assistant superintendents for teaching and learning, and district leaders responsible for ELA curriculum and instruction, assessment, data and coaching—especially those whose literacy achievement hasn't grown as much as you hoped in the last four years and don't want to stay stuck.

This report will help you:

- **Get clear on the destination** where your district should be headed with literacy achievement. Recent research from Georgetown University offers one evidence-informed marker.
- **Set annual growth targets that are ambitious but realistic**
- **Use a five-level framework** to understand where your current performance sits and how far you must travel to reach the next level. If your district's performance is *Good*, you can see how far it must go to be *Great*.
- Learn from steadily improving districts, especially those starting from a similar place.

The ultimate goal is to ensure that students, even in elementary school, are on track to do the reading required for success in college, careers, and adult life.

Introduction

School districts regularly celebrate being above average or that they are improving literacy achievement faster than others. But neither tells us whether enough students are getting where they need to be. Are they on track for success as adults?

For instance, Cape Henlopen, a rural district of 10 schools in southern Delaware touted being "well above" the state average on 2025's Smarter Balanced test. Fifty-eight percent of their elementary and middle school students read and write at grade level proficiency, 17 percentage points higher than the state average.

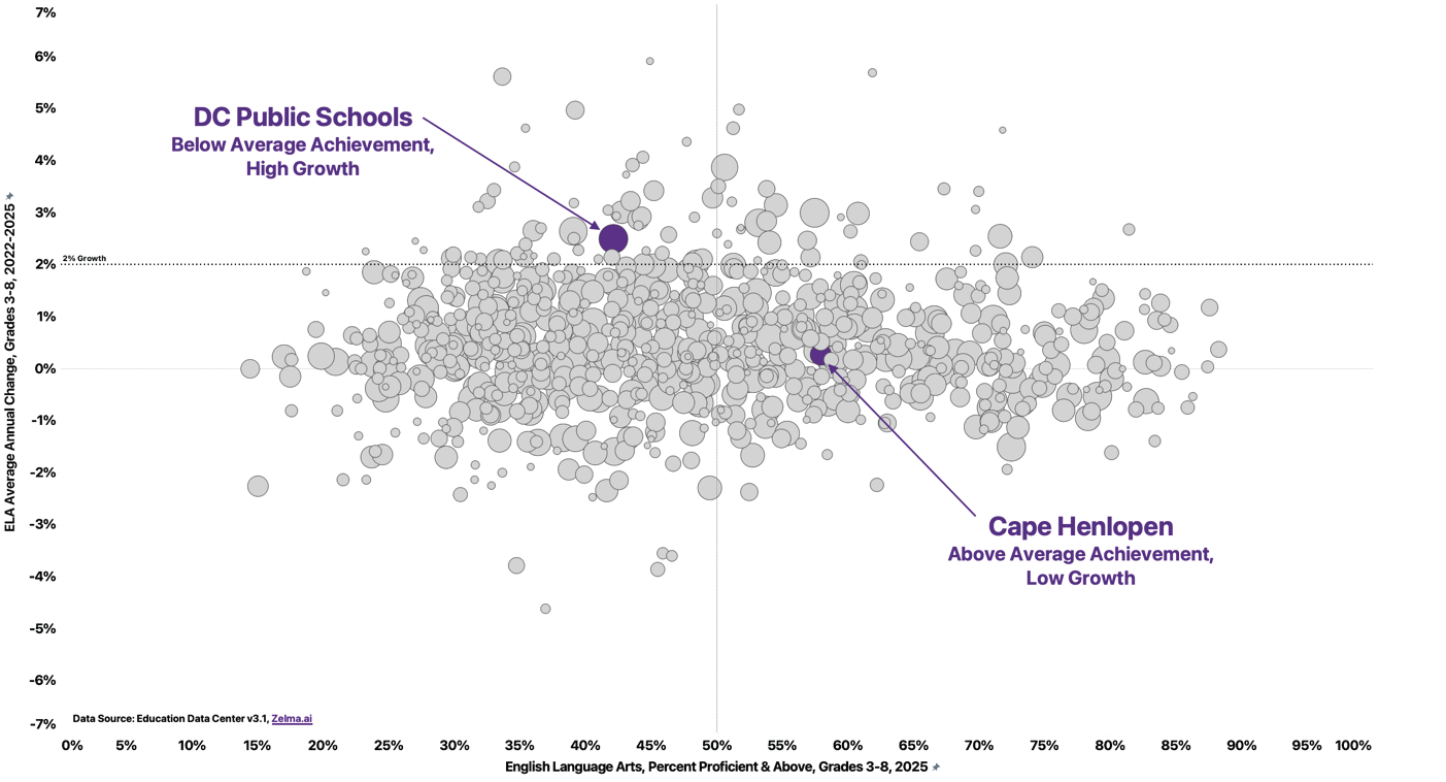
Similarly, Washington, DC celebrated growth in both literacy and math achievement and was recognized as a "district on the rise." But Steven Wilson, a senior fellow at the Pioneer Institute, said about DCPS, "Gains of any magnitude are a good thing, but when most students are not functioning at grade level, this is nothing to applaud."

DCPS shows above-average growth but remains below the national average in literacy achievement. Cape Henlopen has above-average achievement but relatively no growth in the last four years.

FIGURE 1

Where Should We Be Headed with Literacy Achievement & How Much Growth Do We Need?

Average district grew <1% annually from 2022 to 2025; 35% of districts moved backwards



These examples illustrate two fundamental roadblocks to improving student achievement. First, we have lacked a shared understanding of what our destination should be, our end goal. District leaders often think "above average" is good enough. But is it?

Second, we have lacked a clear understanding of how much progress we need along the way. As Figure 1 above shows, most districts have been stuck, growing less than one percentage point annually.

Recent results point in the same direction. Harvard and Stanford's Education Scorecard shows that recovery from COVID remains uneven across districts. PISA 2025 also offers a mixed picture: half of American 15-year-olds are among the world's best readers, but overall achievement has fallen by the equivalent of three-fourths of a school year.

In light of these trends, is 2-3 percentage points of annual growth ambitious or not ambitious enough? And how can we help the districts in the lower-left quadrant, with below average achievement and losing ground, begin moving to the next level?

This report answers those questions by proposing a clear, evidence-informed destination for literacy achievement, a five-level framework for interpreting current performance, and annual growth targets that help states and districts see how far they must grow to reach the next level. Each district of at least 600 students will be able to see what level of achievement it has reached, how far it must go to reach the next level, and whether it is improving fast enough to get there.

Our end goals and annual targets ought to be based on whether students are on track for success in careers and college.

Finally, the report identifies the systems worth studying, those performing at consistently high levels and those annually improving literacy at a rate most schools and districts need to match. Over the next few years, the field needs much deeper understanding of what districts at each level are doing, and how individual students are, or are not, making progress.

Getting Clear on the Destination

Part of our problem in K-12 education is that we have lacked a clear, credible destination for student achievement. Goals 2000 urged the nation to lead the world in math and science, and our ranking slipped relative to other countries. No Child Left Behind (NCLB) demanded 100% proficiency, but many states lowered their standards and not even the best-resourced school districts met these goals.

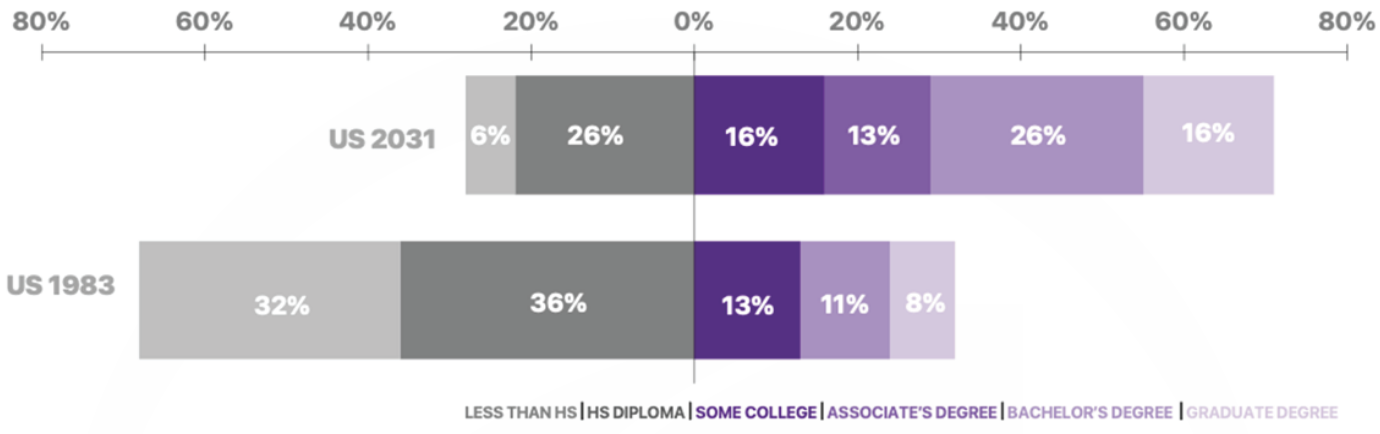
Research from Georgetown University's Center on Education and the Workforce provides one possible destination. The center projects that nationally, 72% of jobs in the next decade will require at least some college, with 55% requiring an associate degree or higher. As Figure 2 shows, that is a sharp reversal from 40 years ago, when 70% of jobs required a high school diploma or less.

Most states have upgraded their tests in the past decade to measure whether students are on a path to college and career readiness. We recommend they take the next step and align their K-12 achievement goals with their projected economic needs. The Georgetown research offers a pragmatic benchmark.

FIGURE 2

72% of Future Jobs Will Require "Some College"

At least 55% will require an AA degree or more



Data Source: *After Everything: Projections of Jobs, Education and Training Requirements through 2031.* Georgetown University Center on Education & Work (https://cew.georgetown.edu/cew-reports/projections2031/)

UpswingLabs

In the NCLB era, testing experts advised against focusing exclusively on state proficiency scores. They rightfully worried about unintended consequences, such as narrowing of curriculum, and schools trying to game scores by targeting kids on the bubble. The focus on achievement levels rather than growth measures also led to disproportionately penalizing schools with large numbers of low-income students.

State tests are not designed to tell teachers what to do tomorrow, but they are designed for monitoring achievement and growth across students, schools and districts. While they don't capture everything about a school's quality, recent evidence offers new support for focusing on whether students are meeting state proficiency standards in reading and writing.

Students' performance on the Common Core-aligned state test in California predicts later college success. Research by Kurlaender & Dykeman at UC Davis finds that across the state university systems, the Smarter Balanced test is essentially equivalent to the SAT in predicting first-year college GPA.

We've also learned that increasing middle school literacy achievement improves a student's chances of earning a college (AA+) degree. A study by Darrin DeChane and colleagues at the University of Missouri followed three groups of students over time as they moved through school. The odds of earning a college degree roughly double for any student who moves from a score of Basic to Proficient on any of the exams.

They estimate that if all 8th grade students who scored at Basic on the Missouri state test raised their scores to Proficient, the number earning college degrees (AA+) would increase roughly 50 percent. Black and Hispanic students earning a college degree would increase by roughly 150 and 75 percent.

Students born in states with large increases in student achievement also see long-term economic and social benefits. Recent research by Elena Doty and colleagues finds that these students experience larger increases in earnings and declines in teen motherhood, incarceration and arrest rates compared to those in states with smaller increases.

States and districts should align their literacy achievement goals with the share of students who need to be prepared for postsecondary education, training, and work. The Georgetown averages nationally are at 55 and then 70%, but the state projections vary considerably. In Arkansas, 35% of future jobs will require at least an AA, while in DC, 80% of jobs will require at least some college.

For the purposes of setting a common destination, we apply the median in the state projections, framing them as floors to reach, not as ceilings. Since 50% of future jobs will require an associate degree or more, then at least 50% of all students should be reading and writing proficiently. Once that first floor is reached, districts can then aspire to reach 68-70% proficiency, since that many students need at least some postsecondary education beyond high school.

The proficiency destinations suggested here have not been empirically proven. Fifty and 68 percent are pragmatic benchmarks, informed by evidence from the Georgetown research. Goals of any sort need to be meaningful. Here, we suggest that one, but not the only possible source of meaning, is the level of education and training students are likely to need as adults. District leaders could choose a different destination, but the goal should connect to an outcome that matters for students' lives.

The next section discusses what it looks like to apply these performance levels in districts across most of the nation.

Five Levels for Classifying Student Achievement

McKinsey's 2010 report, *How the world's most improved systems keep getting better*, offered a staged approach to understanding growth and performance in 20 school systems around the world. Included in their analysis were the improvement journeys of two American school districts, Boston and Long Beach, CA, as they moved from "Fair" to "Great" over a decade.[1]

We adapt McKinsey's framework, classifying districts into one of five performance levels: Disappointing, Fair, Good, Great, and Elite.[2]

The levels established here start with 50% as *Good* and 68% as *Great* because they correspond to meaningful postsecondary attainment benchmarks informed by Georgetown. A lower level, *Fair* is set 18 percentage points below the 50% benchmark and ranges from 32-49%. This creates a parallel distance between the three levels: *"Fair," "Good"* and *"Great"*. Anything below 32% proficient is *"Disappointing"*. The final level, 80% marks *"Elite"* performance. These are districts whose achievement is two standard deviations above the current national average.

Returning to the examples: 58% proficiency puts Cape Henlopen above the initial floor, and it can be considered a *Good* district. But to be *Great*, it should aim to grow another 10 percentage points. As for Washington, DC, if it keeps growing at a rate of 2-3 percentage points a year, it will reach the floor for being *Good* in three years.

When we organize literacy proficiency levels this way, we see this pattern of ratings across the 5,407 school districts with at least 600 students in grades 3-8:

FIGURE 3

Districts at Each Level of Performance in ELA, 2025

Level	Number of Districts	Percent of All Districts	Number of Students (K-8)
Elite	126	2%	0.2m
Great	477	9%	1.3m
Good	1,759	33%	6.8m
Fair	2,237	41%	6.9m
Disappointing	808	15%	2.2m
Total	**5,407**	**100%**	**17.4m**

Data Source: Education Data Center v3.1, Zelma.ai

[1]Both Boston and Long Beach were also winners of the Broad Prize in Urban Education for making great improvements in student achievement.

[2]McKinsey uses the labels "Poor" and "Excellent" at the tail ends of its distribution. We substitute Disappointing and Elite in their place.

We can then chart each district and its performance level against the number of students eligible for free and reduced-price meals. What stands out in Figure 4 below is that while there's a correlation between achievement and family income (-0.67), there's considerable variation within the *Good* and *Fair* levels.

Two-thirds of the *Fair* districts range from 47 to 89% of their students receiving free meals. Almost all the *"Fair"* districts have a counterpart with similar demographics also scoring at the *Good* level. This suggests that for districts at these levels, movement is possible. In-school factors such as the quality of instruction can provide leverage.

FIGURE 4

Districts at Each Level of Performance on State Tests

Data Source: Education Data Center v3.1, Zelma.ai

Thinking in terms of what it takes to reach the next level may be especially useful for districts in certain states. For example, Minnesota has 48% of its districts at the *Good* level, but they haven't been growing over the last four years. Similarly, North Carolina has 47% of districts stuck in the *Fair* level, also growing less than 1 percentage point annually since 2022. Perhaps this framework can spark a new round of discussions about what it takes to move from *Fair* to *Good* and then *Good* to *Great?*

Adjusting for Easier State Standards

In setting these levels, we should account for differences in states' definitions of grade-level proficiency. The movement to common standards, and in some states, common tests, has meant that most states raised their expectations and eliminated what Fordham two decades ago called "the proficiency illusion."

Nonetheless, we know that some states continue to have ELA standards that are much easier than average. Harvard and Stanford University's Education Opportunity Project equates state tests to NAEP,

whereby they convert each state's proficiency rates to common "grade levels" of achievement. This enables us to see what might happen if we moved students from one district to another state whose ELA performance standards are more in line with the national average.

Cedar Rapids, Iowa, has 63 percent of its students in grades 3-8 reading proficiently on the ISASP. In the Stanford database, this translates to a score of 4.3 grade levels. If we moved the students from Cedar Rapids to Mississippi or California, we estimate achievement would drop from 63 percent to 40 percent proficient.

Please see the Appendix for more details on how we adjust for easier state standards in five states: Alabama, Iowa, Ohio, Maine, and for the time being, Virginia.

If you want to see which level your district has reached in literacy, visit this Tableau Public site which is updated as states release their 2026 achievement results.

How to Set Ambitious but Realistic Growth Targets

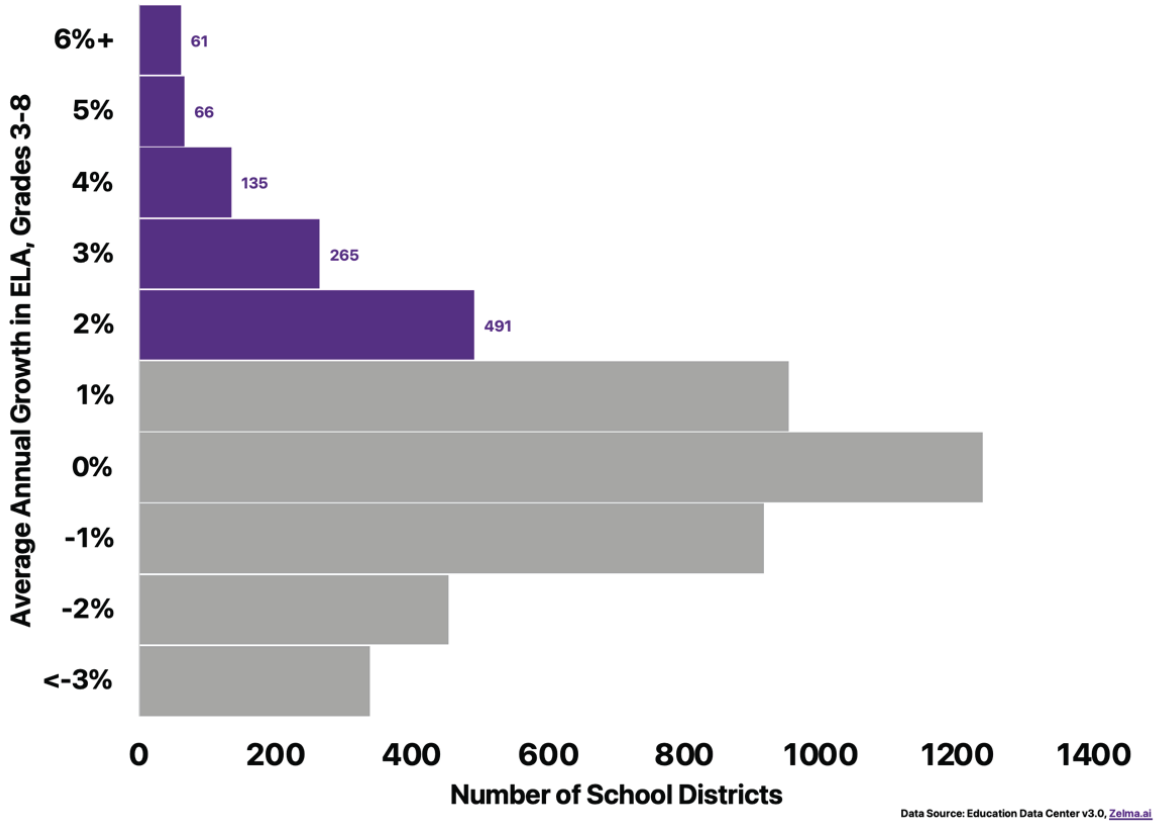
In March 2026, Cape Henlopen School District in Delaware set a very ambitious goal: by 2027-2028, 100% of students will reach reading proficiency by the end of third grade. In 2025, the district had 56.5% of third graders reading proficiently, but growth had been flat (+0.2%) for the previous four years. To reach their goal, the district will need to grow 15 percentage points a year for the next three years. In August 2026, newly posted scores showed no change in third grade reading.

Setting aside the fact that none of the ~5,300 districts in our study are at 100 percent proficiency, how likely is any district to grow by 15 percentage points? What's typical growth? What's within reach but requires a real stretch?

Figure 5 below shows the distribution of average growth rates in districts where we have consistent data over the last three to four years. As mentioned earlier, the typical district grew less than one percentage point annually. Approximately 20 percent of districts grew literacy achievement by two percentage points or more and 11 percent grew three or more.

FIGURE 5

Twenty Percent of Districts (~1,000)Are Growing at Least 2 Percentage Points Annually



Data Source: Education Data Center v3.0, Zelma.ai

Four districts had growth rates of 15 percentage points, but they're all in states—Alabama, Arkansas, New York and Texas—that have recently revised their tests.[3] We can't be sure if the gains represent real learning or if it's about understanding the new format of the test.

There is a legitimate case for bold, seemingly impossible goals: they can focus attention and inspire greater effort. But research on stretch goals finds that extreme targets work best when organizations already have strong performance and enough capacity to experiment; organizations struggling to grow are less likely to benefit from them.

Measurement expert Robert Linn of the University of Colorado recommended an approach to goal setting in the early days of NCLB that's worth replicating now. Look at schools that make large gains as "existence proofs of ambitious, but realistic targets," he suggested, focusing on the top 10% of schools with the largest increases in proficiency.

Linn said, "it would require major improvement for most schools. But the challenge would not be so far out of touch with reality as the targets that were plucked out of air and dropped into legislation."[4]

For a district with 32% of its students currently at proficiency, sitting at the *Fair* level, a growth target of 2-3 percentage points annually puts it on the path to reaching the next level *Good* in six years, by the time the current class of 3rd graders finishes middle school.[5] This may not seem impressive to most school leaders, but it meets Linn's criteria of "ambitious but realistic." But helping massive numbers of districts reach the next level will require us to learn more than we currently know about districts' improvement journeys.

What We Know About Getting to the Next Level

Once we know where districts stand relative to a meaningful destination, and how fast they need to grow to get there, the next question is what districts at each level are doing to move upward?

The examples that follow serve different purposes: Steubenville shows what sustained high performance can look like; Marietta, Allegany, and DC Prep provide a glimpse of what districts do to move from one level to the next; and SRI identifies the most pressing implementation obstacle in front of growing districts. McKinsey, whose 2010 report inspires this framework, suggests that the strategies needed differ depending on a district's starting point.

There has been surprisingly little inquiry about what *Great* districts do differently. A notable exception has been Steubenville, Ohio. The small city district in northeastern Ohio has been featured in Karin Chenoweth's 2021 book, Districts That Succeed and in episodes 12 & 13 of Emily Hanford's Sold a Story podcast for APM Reports.

Both accounts describe Steubenville's use of the heavily researched Success for All curriculum in K-8, developed by the late Robert Slavin and colleagues at Johns Hopkins University. Success for All includes a strong emphasis on building foundational skills and frequently assesses students to group and re-group them for reading instruction every eight weeks. That is, early elementary students are grouped not by age but by current achievement levels. A 3rd grader can be in a "mid-2nd grade" group for Tier 1 instruction and moves to a new group when the data says they are ready.

All the professionals push in to teach reading at the same time in Steubenville. The most highly skilled teachers are assigned to the students in the lower groups to help accelerate them as fast as possible.

As a result, every student gets a full 90 minutes of reading instruction.

Steubenville is also an outlier in math achievement and Chenoweth summarizes their approach as "programs plus culture." Steady leadership has a vision for problem solving and engages in careful management of scarce resources. Steubenville is a community with a high level of trust where everyone takes great pride in their accomplishments.

In our data, there are 21 other districts like Steubenville, with high literacy achievement and above-average free and reduced meal rates. We lack a full understanding of the various pathways they take to 68-70 percent proficiency. We're presently writing a case study on one of these districts. Worcester County is a rural district that is the highest performer in Maryland.

[3]A recent analysis by Emily Oster suggests New York's 2025 growth was an error in scoring by the vendor.
[4]Linn guessed that the top 10% of schools were growing 4 percentage points annually. In our data, the top decile of districts averages 3 percentage points or more.
[5]McKinsey also found that systems could make substantial moves from one level to the next in as few as six years.

FIGURE 6

The Other Steubenvilles: Great Districts with High Poverty

State	School District	Locale	Percent Proficient, ELA, Grades 3-8 (2025)	SEDA Score ELA (2024)**	Number of Students Tested, Grades 3-8	Number of Schools K-12	Percent Free & Reduced Lunch (2024)
IL	North Palos SD 117	Suburbs	86%	--	2,025	5	62%
AL	Saraland City	Suburbs	86%	8.0	1,535	4	52%
OH	Poland Local	Suburbs	82%	6.9	773	4	53%
OH	Steubenville City	Urban	81%	7.1	1,140	6	73%
NJ	College Achieve Paterson Charter School	Suburbs	77%	8.6	734	1	87%
SC	Anderson 04	Rural	75%	--	1,414	6	69%
TX	Harmony Public Schools-Houston West	Urban	75%	7.2	3,105	6	60%
SC	Anderson 01	Rural	74%	--	4,856	14	53%
NJ	Lodi School District	Suburbs	74%	7.9	1,363	7	57%
NY	Valley Stream 30 UFSD	Suburbs	71%	--	642	3	54%
IL	Carterville CUSD 5	Rural	71%	--	1,017	4	52%
WA	Cashmere School District	Suburbs	71%	7.0	699	3	58%
KY	Corbin Independent	Rural	71%	7.6	1,325	9	59%
WA	University Place School District	Suburbs	70%	7.1	2,602	8	51%
IL	Hillsboro CUSD 3	Rural	70%	--	627	4	55%
IL	Massac UD 1	Rural	69%	--	795	7	69%
NJ	BelovED Community Charter	Urban	69%	7.1	719	1	76%
MD	Worcester County Public Schools	Rural	69%	7.3	2,866	14	52%
SC	Anderson 03	Rural	68%	--	1,179	5	70%
NJ	North Star Academy Charter School	Urban	68%	7.1	3,302	1	85%
MA*	Brooke Charter School	Urban	65%	8.4	1,137	1	69%
MA*	Excel Academy Charter	Urban	55%	6.8	694	1	70%

* MA districts are included because of their SEDA scores. The state has the highest proficiency standards in the nation.

** The Stanford Education Data Archive sets the national average at 5.5. If a district is missing a SEDA score, it is because the state recently revised its ELA tests.

Improving Districts: What We Know and Still Need to Learn

In research published last summer, we identified three improving districts that made the jump to the next level. We looked closely at three very different systems. Marietta City (a midsize district in metro Atlanta) and Allegany County, Maryland (an Appalachian small town) moved from *Fair* to *Good*. DC Prep, a K–8 public charter network in Washington, D.C. moved from Disappointing to *Fair*. Briefly, here's what we learned about their improvement journeys:

Start with the problem, not a vision. Leaders in these systems did not begin by writing a vision statement. They asked a harder question: *"What is our most important problem, and how do we solve it?"* Their answer: through no fault of their own, many teachers had not been prepared in the full range of evidence-based reading instruction.

Build teacher knowledge deeply and over time. Each system then launched multiyear professional learning for all elementary teachers.

Marietta used Top Ten Tools (short courses on core elements of literacy) and then added a year of The Writing Revolution to strengthen expository writing. Allegany began a two-year, districtwide engagement with LETRS, covering phonological awareness, phonics, fluency, vocabulary and comprehension.

Make more time for literacy — and protect it. DC Prep extended its literacy block to 1 hour, 45 minutes, and added a second adult in the room to give targeted help at students' desks rather than pulling them out. Allegany set two hours for whole-class literacy plus 30 minutes for small-group intervention and enrichment. Marietta redesigned its day, so students now receive roughly two to three hours of literacy daily: about 90 minutes of whole-class reading that weaves in grade-level science and social studies texts; 20-30 minutes of explicit phonics and word study; and 30-60 minutes of small-group work.

Staff to the challenge. Marietta hired reading specialists across eight schools to work with small groups of about 10 students. Allegany redefined the literacy coach role to provide more direct, classroom-level feedback. DC Prep converted an assistant principal position into a full-time instructional coach.

Use grade-level books in a great curriculum. Both Marietta and Allegany upgraded their K–5 ELA programs. Marietta moved away from leveled books that students could already read and adopted Wit & Wisdom alongside a dedicated foundational-skills block. Allegany replaced Treasure Island with Core Knowledge Language Arts (CKLA). Independent reviewers rate both highly, and they are designed to build broad background knowledge with grade-level texts in history, science, literature, and the arts. DC Prep used home-grown materials but has realized it needs to adopt a high-quality curriculum.

Tighten assessments — and the response. These systems protected daily phonics time, but they also trained teachers to diagnose problems precisely, because not every reading difficulty is a phonics problem. Short, regular checks inform who needs which types of support, and when.

More recent research by SRI reveals a deeper implementation challenge improving districts face as they implement a new, high-quality curriculum. In a study of four districts, including three districts trying to move from *Disappointing* to *Fair*-- Baltimore City, Maryland, Aldine, Texas, and Richmond, Virginia-- researchers found that most ELA lessons stayed on the surface. That is, teachers asked many literal questions and led perfunctory discussions about what happened and when in a book. "Deep comprehension" discussions occurred in only a quarter of the lessons. Students were not spending enough time on the bigger ideas and themes needed to help them make meaning as they read.

McKinsey, in its 2010 report, writes: "School system leaders, when looking for direction, are too often told what to do from a starting point that's different than their own. Educators in a moderately performing system would be better off seeking inspiration from similar systems that are managing to improve, rather than those that are configured and positioned differently, even if they are the world's best performing ones."

McKinsey found in its study of improving systems that they enact a set of strategies appropriate to a particular performance stage. Systems trying to move from the bottom level to *Fair* are focused on trying to reduce variation among classrooms and providing scaffolding for low-skill teachers. Systems trying to move from *Fair* to *Good* create new teaching models, use data differently, and develop new organizational structures. As systems like Singapore and Ontario, Canada advance to higher levels, they're trying to open up the ceiling. McKinsey finds that they provide more autonomy and loosen centralized guidance to schools.

How districts grow from one level to the next in the U.S. should not be a black box. Moving forward, we need more cases of districts on each kind of improvement journey, in a variety of geographical contexts.

A Framework for Improvement but Not Accountability

In recent articles about how America's schools have declined in early reading since the highs of 2015 and 2017, a group of researchers and policy analysts have expressed a preference for accountability systems to emphasize growth over proficiency levels. This report lays out a framework where both measures are equally important and straightforward for everyone to act upon.

Proficiency and average annual growth tell leaders whether the district is moving toward the next performance level. At the student level, districts should also set growth goals that ask whether each student—especially those below proficiency—is making enough progress to get on track. Louisiana has a model of how to do this with an approach known as "growth to proficiency". Mississippi does something similar, while also creating an extra incentive for progress of students at the lowest levels. Delaware also shows how it can be done in a state using a test with a vertically aligned scale, which makes it easier to follow growth over time. Christy Hovanetz of ExcelinEd makes a strong case for growth to mastery approaches in a recent article.

The report also articulates one standard for all groups of students. It does not lay out separate standards for different groups. It does not emphasize who is "beating the odds" because exceeding a predicted regression line is not the standard. Students who start furthest behind need more support, but that should not mean expecting less of them.

The projections from Georgetown don't suggest that only 33% of students from economically disadvantaged families need an AA degree or more. It says that 50% of future jobs will require this level of knowledge and skills.

All schools should aim first to reach the 50 percent proficiency floor for all students, then advance toward 68–70 percent as the next benchmark on the path to college and career success.

This framework is not intended to replace current accountability policy but instead is a way to anchor improvement discussions at a time when the majority of districts aren't growing. For instance:

- Where are we now? What would it take to reach the next level, to go from *Good* to *Great?*
- What improvement did we expect over the year, and what actually happened?
- What explains our success? Or what explains the gap between the results we expected and achieved?
- What should we sustain, change, or stop doing as a result?

Rankings of states within regions, or districts within states, are useful but incomplete. They tell us who's in the lead, or who outran whom. But they don't answer these questions.

This framework also reveals what is lacking in one of the more prominent accountability policies: assigning A-F letter grades. For example, consider Arizona, which modeled its letter grade accountability system after Florida's. The five *Elite* and *Great* districts in the state all earn As. But for the 39 Disappointing districts, where elementary and middle school literacy achievement is less than 32%, 30 of them receive Bs. Only one of them receives a D. Conventional letter grades are concealing low literacy achievement. Grade inflation is as much a problem for state policymakers as it is for teachers and students.

Conclusion

Literacy achievement has been stuck for the last five years. How do we jumpstart progress? Districts and states have pursued several responses such as passing science of reading laws, investing in deep training of how to teach foundational skills, and adopting high-quality curricula.

This report suggests we also need to be clear about what schools and districts are trying to accomplish and proposes one evidence-informed destination. At a minimum, 50 percent of all students should be reading proficiently. Once that threshold is reached, 68-70% provides a second benchmark, informed by the share of future jobs projected to require education beyond high school.

With these destinations in mind, we then lay out the annual growth targets needed to get there. The field has been plagued by efforts that aim high and fall flat. Over time, missing goals saps motivation and effort throughout schools. Here, we suggest grounded annual targets that are ambitious but realistic.

About 12% of districts are at 68-70 percent proficient or higher. Another group is demonstrating sustained annual gains of three percentage points. If gains at this pace became more widespread, they would meaningfully change national outcomes. However, these school systems remain largely unexamined. The next frontier is to study them in depth and gain the implementation knowledge needed to replicate success at scale.

Sources

Bellwether. *Proficiency Confusion: Making Sense of Student Math and Reading Scores.* 2026.

Carnevale, Anthony P., Nicole Smith, Martin Van Der Werf, and Michael C. Quinn. *After Everything: Projections of Jobs, Education, and Training Requirements through 2031.* Washington, DC: Georgetown University Center on Education and the Workforce, 2023.

Chenoweth, Karin. *Districts That Succeed: Breaking the Correlation Between Race, Poverty, and Achievement.* Cambridge, MA: Harvard Education Press, 2021.

DeChane, Darrin, Takako Nomi, and Michael Podgursky. *College and Career Ready: How Well Does 8th Grade MAP Performance Predict Post-Secondary Educational Attainment?* CALDER Working Paper No. 300-0524. Arlington, VA: Center for Analysis of Longitudinal Data in Education Research, 2024.

Dewey, Daniel C., Erin M. Fahle, Thomas J. Kane, Sean F. Reardon, and Douglas O. Staiger. *From Learning Recession to Learning Recovery: Understanding the Sources of U.S. K–12 Improvement.* Education Scorecard, 2026.

Doty, Elena, Thomas J. Kane, Tyler Patterson, and Douglas O. Staiger. *What Do Changes in State Test Scores Imply for Later Life Outcomes?* NBER Working Paper No. 30701. Cambridge, MA: National Bureau of Economic Research, 2022.

Hanford, Emily. "The Evidence" and "The List." *Sold a Story,* episodes 12 & 13. APM Reports, March 6, 2025. Podcast audio.

Hovanetz, Christy. "Measuring Academic Growth for School Accountability." *State Education Standard,* 26, no. 3 (September 2026).

Kurlaender, Michal, Kramer A. Dykeman, and Erika Jackson. *Predicting College Success: How Do Different High School Assessments Measure Up?* Stanford, CA: Policy Analysis for California Education, 2018.

Linn, Robert L. "Accountability: Responsibility and Reasonable Expectations." *Educational Researcher* 32, no. 7 (2003): 3–13

Locke, Edwin A., and Gary P. Latham, eds. New Developments in Goal Setting and Task Performance. New York: Routledge, 2013.

Mourshed, Mona, Chinezi Chijioke, and Michael Barber. How the World's Most Improved School Systems Keep Getting Better. McKinsey & Company, 2010.

National Center for Education Statistics. Mapping State Proficiency Standards Onto the NAEP Scales: Results From the 2022 NAEP Reading and Mathematics Assessments. NCES 2026-014. Washington, DC: U.S. Department of Education, Institute of Education Sciences, 2026.

National Center for the Improvement of Educational Assessment. The Case for State Testing. March 2025.
OECD. PISA 2025 Results (Volume I): Future-Ready Students. Paris: OECD Publishing, 2026.

Reardon, Sean F., Demetra Kalogrides, and Andrew D. Ho. "Validation Methods for Aggregate-Level Test Scale Linking: A Case Study Mapping School District Test Score Distributions to a Common Scale." Journal of Educational and Behavioral Statistics 46, no. 2 (2021): 138–167.

Reynolds, Dan, Sara Rutherford-Quach, Lauren J. Cassidy, Anna Jennerjohn, and Katrina Woodworth. Beyond the Surface: Leveraging High-Quality Instructional Materials for Robust Reading Comprehension. Learning brief. SRI, 2025.

Sitkin, Sim B., Kelly E. See, C. Chet Miller, Michael Lawless, and Andrew Carton. "The Paradox of Stretch Goals: Organizations in Pursuit of the Seemingly Impossible." Academy of Management Review, Vol. 36, no. 3 (2011): 544–566.

Appendix

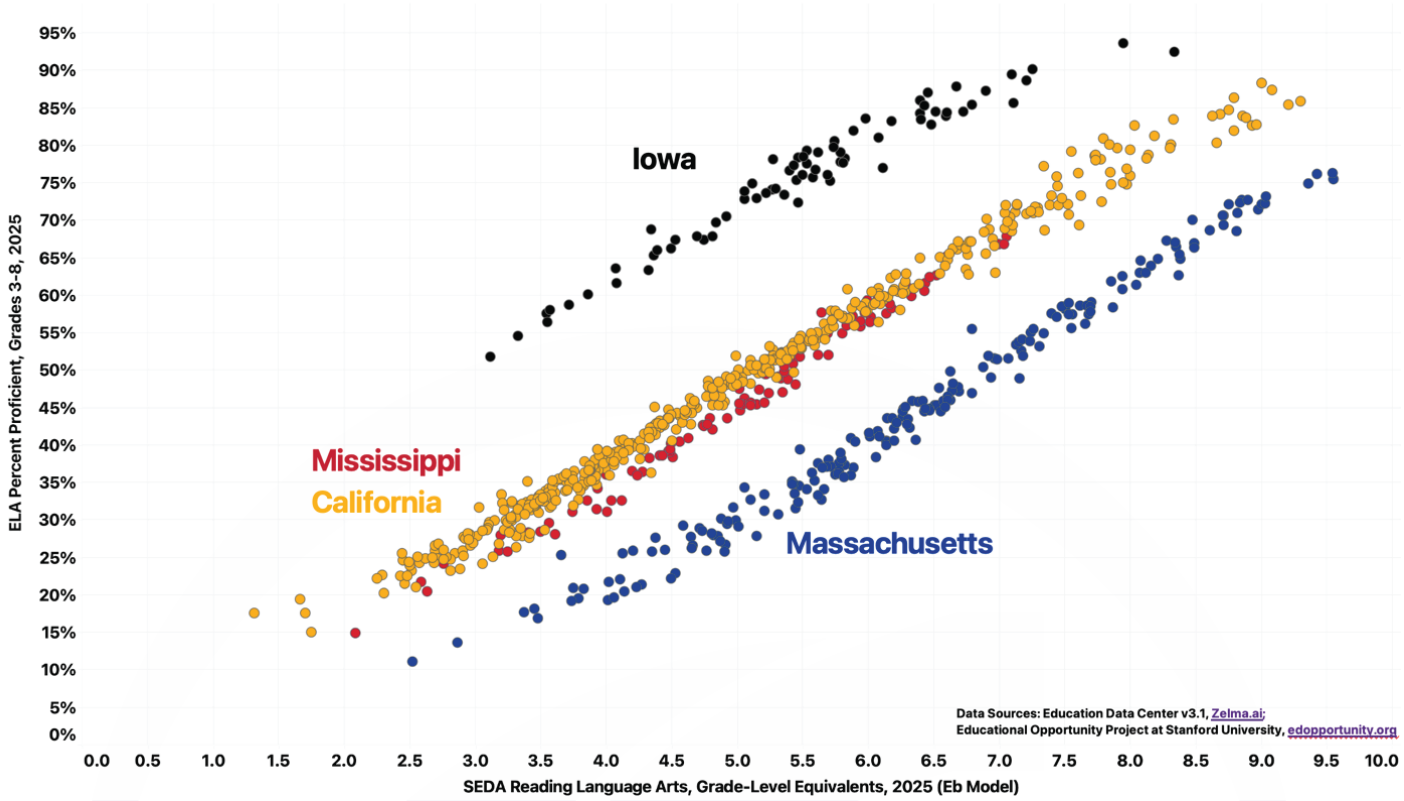
Using 2022 data[6] the National Center for Education Statistics placed state tests onto the NAEP (National Assessment of Educational Progress) scale, a common metric for all states, and produced a NAEP equivalent score.[7] Looking at that data online, we can see that Iowa's NAEP equivalent score for 4th grade reading is 197, which maps to a level below NAEP Basic. On its state test that year, Iowa reported 73 percent of fourth grade students reading proficiently. Most state tests map nearly 30 scale score points higher, or approximately three grade levels, to mid-*NAEP Basic*.

Five states presently have early literacy standards that map to the *NAEP Proficient* level. NCES has been very clear that NAEP Proficiency means "competency over challenging subject matter" and is *not the same as being "on grade level."* In a 2017 Christian Science Monitor article, former NCES Commissioner Peggy Carr said that to understand how many students are performing on grade level, it is far more accurate to look at "basic" scores than at "proficient".

For several years, Sean Reardon and colleagues at Stanford University's Educational Opportunity Project run the analysis one step further. They use the 4th and 8th grade results from the 2019 NAEP to place the state proficiency thresholds on the NAEP scale. With this linkage between state scores and NAEP, they generate grade-level equivalents for student achievement at the state and district levels. (For more, see Reardon, Kalogrides, and Ho, 2021)[8]

FIGURE A1

State Proficiency & Grade-Level Equivalents in Four States



[6]NCES's analysis of 2024 data was not public at the time of writing this report, but it promises to provide a more current view of where state standards sit on a common scale.

[7]NCES stresses that "The mapping of the state standards for proficient performance does not imply that the NAEP achievement levels are more valid than the state standards or that states should emulate NAEP standards."

[8]Because they assume that the NAEP tests and state tests apply to the same population, they exclude states where fewer than 95% of students took state tests (NY & CO), as well as states that do not post sufficiently detailed data in all performance categories. There are "SEDA Scores" for 38 states and Washington, D.C.

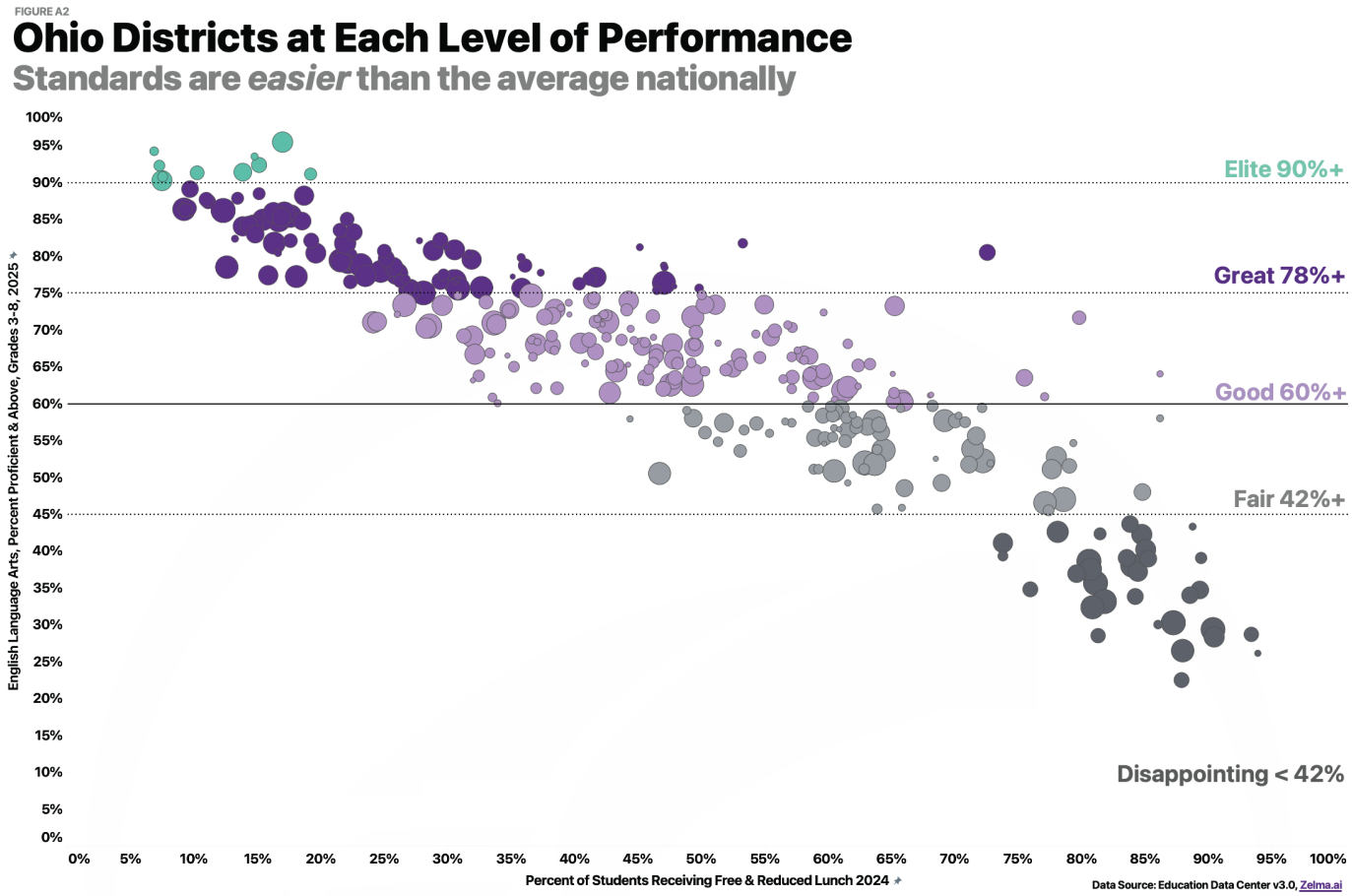
We plot these grade level equivalents from SEDA here on the x-axis against the 2025 state proficiency scores for grades 3-8 on the y-axis. The chart above shows these correlations in both Mississippi and California compared to Iowa and Massachusetts.

For districts in both California and Mississippi, the correlations are tight: a grade-level equivalent of 5.5, (the midpoint between grades 3 and 8), lines up with approximately 50 percent of elementary and middle school students reading and writing proficiently on the state tests. Having kids from one state take the other state's tests wouldn't change performance on average. This fits, as the NCES mapping study finds their ELA state tests in line with a score that's at the midpoint between NAEP Basic and Proficient.

The chart shows the correlations between state proficiency results and the SEDA grade level equivalents in Iowa too. There, the divergence between districts with the same grade-level equivalents is sharp. In Iowa, a grade-level equivalent of 5.5 corresponds to 73-74% on the state test. Move Iowa kids to California or Mississippi for testing and their score would drop by more than 20 percentage points.

At this time, we use the regressions generated by the correlations to adjust the levels upward for the states with easier standards. As a result, to be considered *Good* in Iowa, a district needs to be at 73% proficiency (an adjustment of 23 percentage points), rather than 50% in most states.
All told, there are five states with "easier" literacy standards in grades 3-8: Virginia, Iowa, Alabama, Maine, and Ohio.

FIGURE A2



We leave alone the states with harder standards. For instance, Rhode Island and Massachusetts state tests have a NAEP equivalent of 242, four points higher than *NAEP Proficient*. In the most recent NAEP Reading results, Massachusetts' 4th grade scores were 10 points higher than Iowa's, almost a full grade level's difference.

On their state tests, 38 percent of Massachusetts 4th graders were said to be meeting standards. But the high bars in Massachusetts, Rhode Island, Alaska, Arkansas, Colorado, Indiana, Oklahoma, and Tennessee are policy choices by state lawmakers, and so we do not adjust the expected levels.

While the SEDA grade-level equivalents are a common ruler, state proficiency rates are the language district leaders and parents communicate with and understand best. These estimates are practical translations intended to make the framework usable; they are not formal psychometric equating.

The estimated adjustment for the levels in Ohio is 10 percentage points for each level. To be Good in Ohio, a district needs to be at 60% proficiency, not 50%.

About the Author

David Scarlett Wakelyn *is founder of Upswing Labs and a former middle school ELA teacher in South Central Los Angeles. At the National Governors Association, he led the nation's largest effort to expand AP classes for low-income and minority students, and he worked on the team that developed common state standards and assessments. David also served as Deputy Secretary of Education in New York. He can be reached at david@upswinglabs.org*

About Upswing Labs

Upswing Labs works with school districts, public charter schools, and states to improve literacy achievement.

Usage

The non-commercial use, reproduction, and distribution of this report is permitted. ©2026 Upswing Labs