

# NORTH CAROLINA K-12 ACHIEVEMENT

## EXECUTIVE SUMMARY

When considering a postsecondary attainment goal for North Carolina, it is critical to examine key elementary and secondary academic benchmarks. Students who perform well on standardized tests will have more opportunities for postsecondary education and in the workforce than those who perform less well.

Most economically disadvantaged students and some minority groups do not have ACT scores high enough to gain access to many of the more selective UNC campuses. Because test scores are highly correlated across time, it is possible to predict with some accuracy how a student will perform on the ACT based on prior test score performance.

There is some evidence of improvements in test scores in the early grades across cohorts, but this improvement generally does not hold in middle school. Our state made great progress in raising test scores between 1990 and 2000, but there has been much less progress since 2000. We now perform at about the national mean in 4th and 8th grade reading and math on NAEP.

As in all states, we have a particularly large proficiency gap between black and white students, which only grew in math with the introduction of newer, more difficult math standards in the 2012-2013 school year.

Finally, early literacy book levels are highly predictive of meeting Read to Achieve standards in 3rd grade, but because of disparities in this likelihood between ED and non-ED students, the 2nd grade book level target may need to be increased.

*my*FUTURENC

# NORTH CAROLINA K-12 ACHIEVEMENT

## POLICY BRIEF

### Introduction

To aid in the establishment of a postsecondary attainment goal for North Carolina, as well as strategies for meeting that goal, this policy brief summarizes key data points about elementary and secondary academic achievement. The brief focuses on academic achievement as measured by test scores, but not because we think that is the only student outcome measure of interest. There are many student capacities that educators foster, such as socio-emotional, leadership, creativity, and teamwork skills, to name just a few. Unfortunately many of these are not measured statewide in a valid and reliable way. Academic achievement is measured statewide, however, and teachers and principals agree that raising academic achievement is a key aim of schooling.

Academic achievement is an important end in and of itself. It measures what students know and can do in a particular grade and subject. As education policy experts, policymakers, and parents we should know how our students are doing and what skills are required for postsecondary education and the world of work. Achievement is also important because those who perform well on standardized tests will have more opportunities for postsecondary education and in the workforce than those who perform less well.

This brief works backward through the age span of students from postsecondary, to high school, middle school and elementary school. We organize the brief in this way to emphasize that educational achievement is strongly linked across time. Learning begets learning, which means that raising and maintaining achievement takes concerted action across many different levels of schooling.

### ACT and Admission to a UNC System Campus

Take, for example, performance on the ACT. As part of North Carolina's shift to college and career content standards, all 11th-grade students have been required to take the ACT college admissions exam since Spring 2012. This makes it possible to examine the relationship between this very important high school test score and postsecondary outcomes such as acceptance to a UNC system campus. In addition, and unlike all but about a dozen other states, we can do this for a complete cohort of students rather than for just a self-selected sample of test-takers. The ACT also serves as an important benchmark for secondary education, as students who earn a composite score<sup>1</sup> of 17 or better are considered to meet the state's definition of college-ready.

Figure 1 estimates the probability that a North Carolina high school student with a given ACT score was accepted to a university in the UNC system for the 2017-18 school year. We plot three separate probability curves, one for each of three levels of selectivity as defined by *U.S. News and World Report*: more/most selective, selective, and less/least selective.

The horizontal line on the chart indicates a 50% probability of acceptance to a university. The vertical line indicates an applicant with an ACT composite score of 17. We can see that these two lines intersect at about the point at which the probability of acceptance at a selective university reaches about 50%. In other words, in order to have a coin-flip chance to gain acceptance at one of



#### AUTHOR

**Douglas Lee Lauen**  
Associate Professor  
of Public Policy;  
UNC Chapel Hill



#### AUTHOR

**Thomas R. Tomberlin**  
Director, School  
Research, Data, and  
Reporting; NCDPI



P-12



POST-SECONDARY



WORKFORCE

NORTH CAROLINA K-12  
ACHIEVEMENT

myFUTURENC

UNC's selective universities, students need an ACT of about 17, on average. An ACT of 17 would garner a much higher acceptance probability from a less selective university (about 74%) and a much lower acceptance chance from a more selective university (about 12%).

Given the strong relationship between ACT performance and acceptance to a UNC system university, one might want to know how many 11th graders were able to meet the minimum ACT score of 17. Of white 11th graders in North Carolina, 73% met or exceeded the benchmark score of 17 in the 2016-17 school year, but only 35% and 44% of black and Hispanic students, respectively, met or exceeded this threshold.

Of course, many students and parents might want better than a 12% chance of getting into one of UNC's more selective schools. To increase their chances to above 50% to a more selective institution, students should aim for an ACT composite of 24, or better, on average. Only 27% of white, 4% of black, and 7% of Hispanic 11th graders met or exceeded this ACT benchmark.

## ACT and 8th Grade Test Scores and UNC Enrollment

So, this first bit of evidence should be no surprise: college entrance exams are strong predictors of gaining acceptance to college. What may be more surprising is that 8th grade test scores (reading or math) are about as strong a predictor of postsecondary enrollment in a UNC system school as the ACT score itself.

Figure 2 shows the predicted change in probability of higher test scores<sup>2</sup> for three different tests (8th grade reading, 8th grade math, 11th grade ACT) on two different outcomes: enrolling in a North Carolina community college or enrolling in a UNC campus. Higher test scores in 8th grade math or reading increase the probability of enrolling in a UNC system school by 19 percentage points. The same increase in ACT test scores is correlated with a 20 percentage point increase.<sup>3</sup> There is very little relationship between test scores and enrollment in the community college system, but campuses in the North Carolina Community College System have open admissions and do not select students based on test scores. As a result, they serve students with all test score achievement backgrounds. This is an important reminder that academic test score benchmarks may be useful for increasing the pipeline to some postsecondary destinations, but not for all attainment goals. On the other hand, students with very low levels of academic literacy could face more limited postsecondary and employment options than students with higher levels of literacy.

FIG. 1: AVERAGE PROBABILITY OF UNC-SYSTEM ACCEPTANCE BY ACT COMPOSITE SCORE FOR NC RESIDENTS (2017-18)

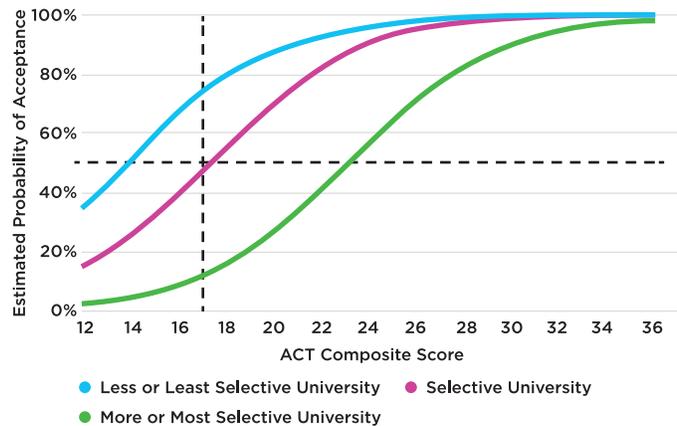
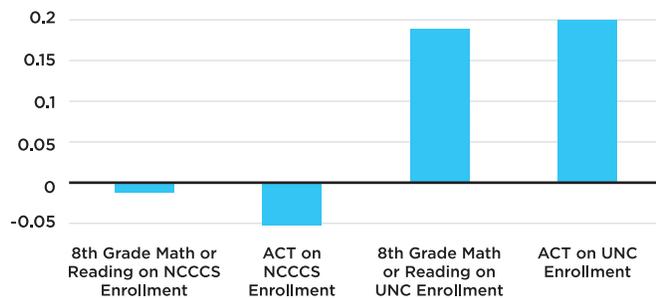


FIG. 2: CHANGE IN PROBABILITY IN POSTSECONDARY ENROLLMENT FROM HIGHER TEST SCORES



Note: A statewide sample of students who entered 9th grade in 2010 and 2011. Outcome is enrollment (not acceptance, as in Figure 1). Postsecondary enrollment tracked as of five years of being a high school freshman.



P-12



POST-SECONDARY



WORKFORCE

NORTH CAROLINA K-12 ACHIEVEMENT

myFUTURENC

The importance of this analysis is not rooted in the belief that all students should attend a highly selective four-year university, or even that all students should go to college immediately upon finishing high school. These data represent the relative access of students to opportunities beyond K-12 education. Certainly, there are many students who will elect to forego college for the workforce. Likewise, there are students who may choose to attend a less selective university for a variety of reasons. What is of critical importance is that NC public schools prepare students academically so that students can make choices from a variety of options, rather than be forced down a path due to a lack of opportunity.

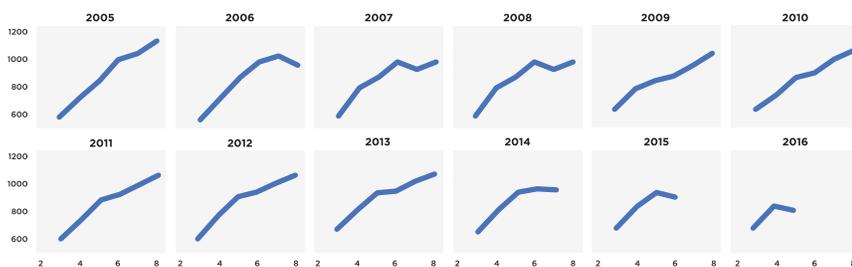
## Student Growth across Cohorts

If 8th grade scores are strong predictors of ACT and, thus, of college enrollment, one might wonder if test scores before this point are also strongly related to each other. It turns out they are. The correlation from one year to the next is about .80 (very strong). Even five years apart there is a strong relationship: A student's 8th grade score is correlated with a student's 3rd grade score at about .72, so scores are fairly stable across time within the same student.

To say that test scores are stable over time is not the same thing as saying that educators cannot raise student test scores. In many states measuring student growth is a difficult thing to assess when standards and the tests change. In our state, because of the desire to assess growth over time for schools, teachers, and students, our tests have been *vertically equated* across grades and also *equated across test editions* over time. What this means is that we can compare test scores across long periods of time and across many different cohorts. In addition, our state tests are equated to two absolute scales: the Lexile scale, measuring the text complexity of a reading passage, and the Quantile scale, measuring mathematical difficulty.<sup>4</sup>

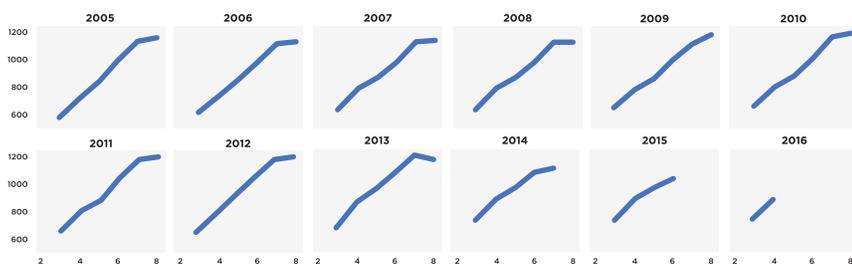
Figure 3 and Figure 4 show test score trajectories of students in twelve different cohorts spanning from 2005 8th graders to 2016 8th graders. They plot the average scale scores in reading and math across grade levels. For the most part these trajectories are upward sloping because as students move from grade to grade they learn more. Each cohort also starts at a slightly higher

FIG. 3: MATH QUANTILE



*Growth trajectories in math (top) & reading (below) achievement across cohorts defined by the spring of the year in which the cohort finished 8th grade. This data series ends in 2013, so some cohorts do not have middle school data points.*

FIG. 4: READING LEXILE



point with each successive cohort. By the final two cohorts, 3rd grade students are starting out about one grade level higher than where students started in the first cohort. That is the good news. The bad news is that in both subjects, the average scores of 8th graders from later cohorts are not much higher in reading and are in fact lower in math than were the scores for earlier cohorts. In other words, students are now better prepared than they used to be as they enter upper elementary, but they are not growing as much between 4th and 8th grade as their older peers did.

Ideally each line would be upward sloping and linear, which would indicate equal growth across grade levels. When the trajectory flattens out, we should pay attention. If this flattening happens cohort after cohort, we should pay even closer attention. In reading, growth slows in 8th grade. In math, it slows in 6th grade. In summary, it appears that progress made across cohorts in the early grades in raising test scores in reading and math is slowing considerably in middle school. One possibility to investigate is whether state and district grade-level academic standards and actual teaching practices align well with what is being tested at these particular grade levels.

## Trends in Average 4th and 8th Grade NAEP Scores

While it is important to know about variation across cohorts in achievement on the state tests, it is also important to know how our state's students are doing over time and relative to nationwide averages. All students in North Carolina and across the country take the National Assessment of Educational Progress (NAEP) test. This national test provides an important, and consistent, benchmark because states vary in their standards, making state testing results hard to compare. It is important to keep in mind, however, that average test scores also reflect the family backgrounds (e.g., poverty, race, linguistic background) of the students taking the tests, and North Carolina's population is somewhat more disadvantaged than the national average.

Figure 5 and Figure 6 show math and reading scale scores for 4th and 8th graders. The figures show considerable growth over 11 years between 1992 and 2003 in 4th and 8th grade math, as well as for 4th graders in reading. In math over this time period, this growth brought the state from below the national average to above the national average. But over the next 14 years, there has been very little, if any, growth in NAEP math scores. Fourth grade reading scores continued to grow, but at a much slower pace than they did during the previous period. Eighth grade reading scores have been flat in our state between 1998 and 2017 (the 8th grade reading test was not

FIG. 5: AVERAGE NAEP MATH SCALE SCORES

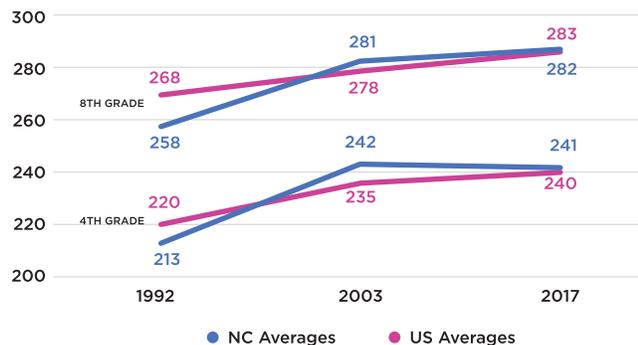
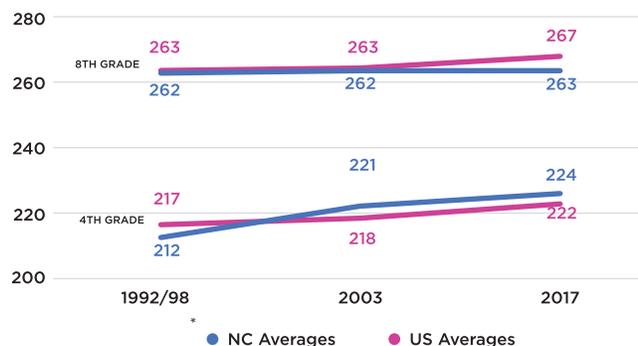


FIG. 6: AVERAGE NAEP READING SCALE SCORES



\*There were no 8th grade reading scores in NC in 1992, so the graph plots the 1998 instead.



P-12



POST-SECONDARY



WORKFORCE

NORTH CAROLINA K-12  
ACHIEVEMENT

myFUTURENC

administered in our state in 1992). With the exception of 4th grade reading, our state ranks among the bottom among all states in scale score gain across the country between 2000 and 2017.

In summary, North Carolina punches above its weight in the sense that our students achieve at national averages despite our state having somewhat higher poverty rates and lower median incomes than the national average. On the other hand, the lack of improvement over the past decade and a half deserves some attention. Moreover, while the state tests show some improvement in 3rd grade scores across time during the same time period covered by our NAEP figures, we get a more mixed view from NAEP scores: slight improvement in reading and a small decline in math.

## Changes in Test Score Gaps for One Cohort

Tracking trends in scale scores can tell us about progress over time, but averages can hide important differences in test score gaps. Tracking gaps in scores has been an important point of emphasis in accountability models since the No Child Left Behind Act of 2001 (NCLB). Scrutiny of test score gaps is important for many reasons. One is that as we have seen above, there are large gaps in ACT achievement, and thus likelihood of acceptance in a UNC system campus, among student subgroups. Another reason is that the 2015 reauthorization of the Elementary and Secondary Education Act (ESEA), known as the Every Student Succeeds Act (ESSA), returned authority over accountability models back to the states. While the new federal law allows states to develop their own accountability models, the federal government requires that states report on, and address, gaps in student achievement for historically low-performing student populations (e.g., economically disadvantaged, students with disabilities, and minority students). As North Carolina prepares to implement its accountability model under the new federal law, it is important to understand the extent of the achievement gap for North Carolina students.

Here we highlight the gaps in student proficiency for a cohort of white, African-American, and Hispanic students in 3rd through 8th grades in both reading and mathematics. These analyses follow these students from the 3rd grade in the 2011-12 school year to the 8th grade in the 2016-17 school year. Results from state assessments in the 3rd grade (2011-12) shown in Figure 7 reveal a large disparity in grade-level proficiency between white and black students: 28 percentage points in reading and 21 percentage points in math. In these students' 4th-grade year (2012-13), the achievement gap in mathematics increased to 34 percentage points. The gap in reading increased as well, but far less dramatically.

The 2012-13 school year was significant because the state shifted from the ABCs of Public Education accountability model to the READY accountability model, which aligned state assessments to college- and career-ready content standards. Due to these more difficult standards, proficiency rates in math and reading declined for all students in the state, but African-American and Hispanic students experienced a much greater decline in their proficiency rates than did white students. It is not clear to us why this happened. Black and Hispanic students are more likely to attend high-poverty schools. It could be that educators in high-poverty schools struggled more than educators in low-poverty schools to teach to the more difficult standards.

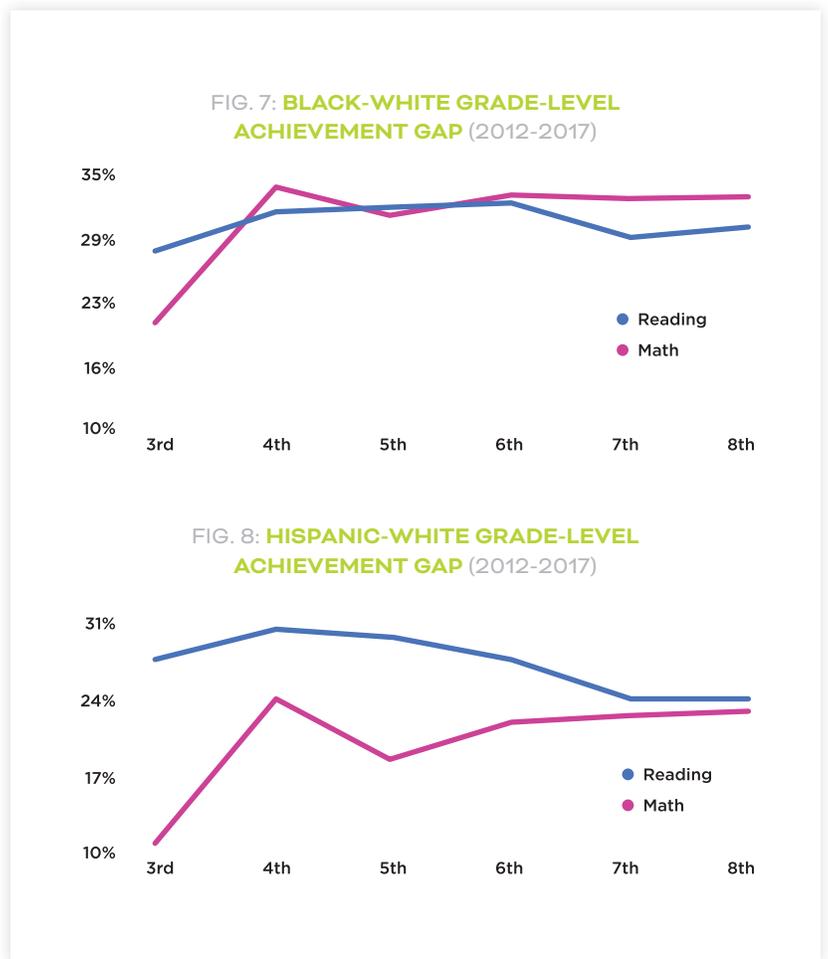
From Figure 7, we can see that the achievement gap between white and black students in this cohort has leveled off, but in neither math nor reading has the gap returned to lower rates observed in the 2011-12 school year. When these students left 8th grade at the end of the 2016-17 school year, the proficiency rate gap for black students in mathematics was 12 percentage points greater than it was in the 3rd grade; for reading the gap was 2 percentage points larger than the baseline year.



In Figure 8, the 3rd grade gap in grade-level proficiency between white and Hispanic students is virtually the same as that of white and black students (27.6%). After an initial increase in the achievement gap in the 4th and 5th grades, the achievement gap begins to narrow. By the 8th grade, the achievement gap in reading between white and Hispanic students is lower than in 3rd grade in the baseline year.

The achievement gap in mathematics between white and Hispanic students, however, widens over time. It is relatively small (compared to the white vs. black gap in math) in the 3rd grade, but widens dramatically in the 4th grade. After a dip in the 5th grade, the gap in mathematics achievement between white and Hispanic students rises steadily until the end of 8th grade. By the time these students entered high school, the achievement gap in math had more than doubled from its initial size.

ESSA presents tremendous opportunities and challenges for the North Carolina public schools. NCLB focused attention on test score gaps, but obviously the gaps have not yet closed. ESSA provides another bite at the apple, forcing states to focus attention and resources on students who may need more support to reach the more rigorous college and career content standards. Additionally, analyses like the one presented here will allow the state and local districts to find examples of schools or classrooms where the gaps in proficiency among student subgroups has widened, narrowed, or has even been eliminated. The challenge is that closing test score gaps will require effort and political will on the part of educators and policymakers at the state and local level.



## Benchmarks for Early Literacy and 3rd Grade Reading Proficiency

To close gaps and make better progress on raising test scores and other measures of student success, we must view student development as linked across many levels—elementary, middle, secondary, and postsecondary. Perhaps the most important building block of all is early literacy. If students cannot read well, it is very hard for them to learn any subject well. For this reason and others, in 2012, the North Carolina General Assembly passed the Excellent Public Schools Act (Senate Bill 795) which established the *North Carolina Read to Achieve* (RtA) program. One of the articulated goals of the program is to ensure that all NC students are reading at grade-level proficiency by the end of the 3rd grade. NC students have a number of options to meet the RtA proficiency standard, but the most common method is by demonstrating proficiency on the 3rd grade reading beginning of grade (BOG) or end of grade (EOG) assessment. In order to measure



progress toward meeting this goal, North Carolina instituted a K-2 diagnostic reading assessment, mCLASS Reading 3D. The mCLASS assessment sets statewide benchmarks for on-grade proficiency in kindergarten through 2nd grade.

This section focuses on the extent to which the mCLASS diagnostic reading assessment predicts success on the NC 3rd-grade BOG or EOG<sup>5</sup> assessment for economically disadvantaged (ED) students and non-economically disadvantaged (non-ED) students. In the 2016-17 school year, only 58% of the state's 122,825 3rd-grade students met the RtA proficiency requirement by passing either the 3rd-grade BOG or EOG. If the mCLASS assessment is meant to prepare K-2 students for success on the statewide reading assessment, then it is important to understand how well the state's benchmarks predict success on the BOG/EOG. For this report, we use mCLASS benchmark data at the end of the 2nd grade to estimate the probability that a student will meet the proficiency requirement on the BOG or EOG.

North Carolina uses a guided reading level system where students proceed upward through reading complexity through books graded by letters, with kindergarteners reading book levels A through D, 1st graders reading book levels D through J, and so on. Figure 9 shows the estimated probability that students will achieve proficiency in 3rd-grade reading by mCLASS book level at the end of the 2nd grade.

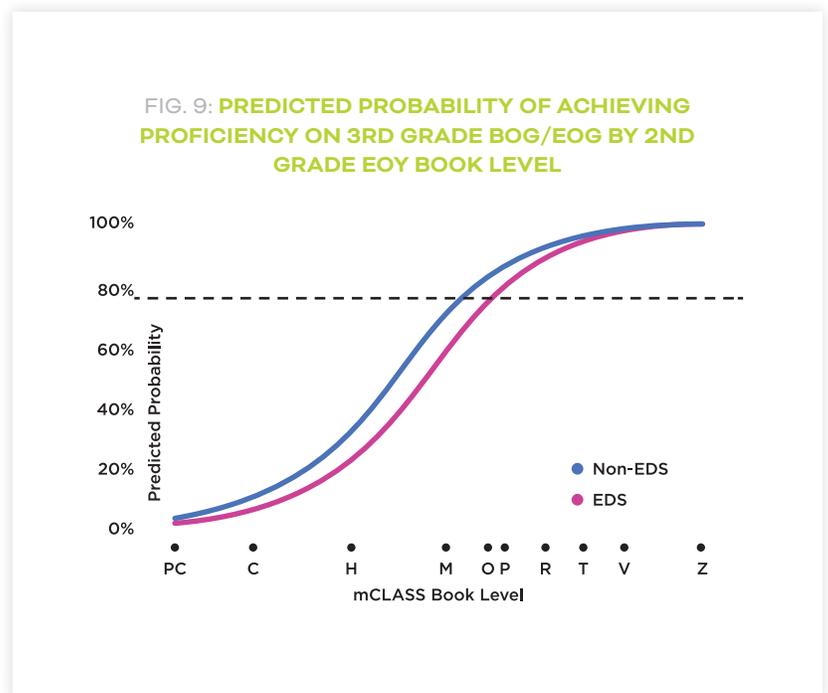
The graph shows two curves: one for ED students and one for non-ED students. In the book range of H to R there is a higher probability that non-ED students will meet the proficiency requirement than will ED students. Book level M is the grade-level proficiency indicator for 2nd grade. While non-ED students at book level M at the end of 2nd grade have a 70% probability of meeting proficiency on the 3rd-grade BOG/EOG, ED students at the same book level have only a 59% probability of meeting the goal. Furthermore, only 43% of ED students are reaching the M book level by the end of the 2nd grade compared to 67% of non-ED students.

In order to give ED students a probability of meeting the RtA requirement that is on par with their non-ED counterparts, ED students would need to reach a book level of O by the end of the 2nd grade. Currently, only about 17% of ED 2nd-graders are meeting this benchmark.

The state of North Carolina would benefit from a more robust study of the relationship between mCLASS benchmarking of K-2 students and meeting the RtA proficiency requirement. Aligning the benchmark goals with increased probability of meeting the RtA requirement can help schools and teachers set high expectations for student performance and may lead to increased success in student outcomes in 3rd grade and perhaps beyond.

## Conclusion

In summary, currently most economically disadvantaged students and some minority groups do not have ACT scores high enough to gain access to many of the more selective UNC campuses. Because test scores are highly correlated across time, it is possible to predict with some accuracy how a student will perform



on the ACT based on prior test score performance. There is some evidence of improvements in test scores in the early grades across cohorts, but this improvement generally does not hold in middle school. Our state made great progress in raising test scores between 1990 and 2000, but there has been much less progress since 2000. We now perform at about the national mean in 4th and 8th grade reading and math on NAEP. As in all states, we have a particularly large proficiency gap between black and white students, which only grew in math with the introduction of newer, more difficult math standards in the 2012-2013 school year. Finally, early literacy book levels are highly predictive of meeting Read to Achieve standards in 3rd grade, but because of disparities in this likelihood between ED and non-ED students, the 2nd grade book level target may need to be increased.

## Discussion Points

### Standards and accountability

- Should the ACT benchmark for college readiness be increased from a 17 to something higher?
- If the ACT is an important benchmark for high school accountability, should districts identify students with low 8th grade test scores as being at risk of scoring low on the ACT for intervention?
- Should the state or districts consider raising the 2nd grade book standard from M to O?
- What accounts for the dip in reading test score growth in 8th grade?
- What accounts for the dip in math test score growth in 6th grade?
- Why did black-white test score gaps increase so much with the 4th edition math test?

### Supports to increase achievement

- What supports are particularly effective in elementary, middle, and high school?
- Are the same, or different, supports needed in high-poverty schools?
- How can professional learning communities be structured to help teachers collaborate on important questions of practice?
- What role should principals play in evaluating the quality of teaching?
- Are financial resources adequate and flexible enough to focus on areas of instructional need?
- What role should instructional coaches play and what is an appropriate caseload for each coach?
- What other interventions have strong research support and large effects on student achievement overall and closing gaps in particular?

### Contextual factors and teacher's role

- Do students have adequate pre-kindergarten experience for school readiness?
- Are high-quality teachers adequately distributed across schools and classrooms?
- Are teachers adequately supported and developed over time?



## Endnotes

1. The composite contains tests of English, mathematics, reading, and science.
2. We chose a one standard deviation increase for this figure. A standard deviation is a measure of variation in data. A one standard deviation increase from the mean of normally distributed data is equivalent to an increase from the 50th to the 84th percentile of all test takers' scores.
3. This is partially explained by the fact that 8th grade test scores and ACT scores are positively correlated at a high level ( $r = .74$  on a scale of -1 to +1).
4. The Lexile and Quantile scales were created by Metametrics (Durham, NC). See, for example, Smith, M. and Gary Williamson (2016). "Aligning the K-12 Academic Journey with a Postsecondary Destination." The State Education Standard: January, pp. 14-18.
5. Retrieved from <https://simbli.eboardsolutions.com/Meetings/ViewMeetingOrder.aspx?S=10399&MID=3712>



P-12



POST-SECONDARY



WORKFORCE

NORTH CAROLINA K-12  
ACHIEVEMENT

myFUTURENC