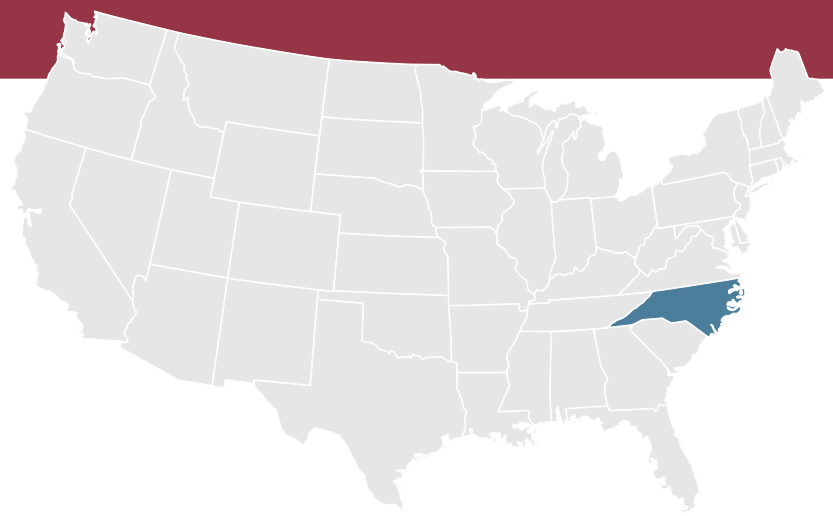


# *An Independent Review of ESSA State Plans*



## North Carolina

### *Project Overview*

**B**ellwether Education Partners, in partnership with the Collaborative for Student Success, convened an objective, independent panel of accountability experts to review ESSA state plans. We sought out a diverse group of peer reviewers with a range of political viewpoints and backgrounds, and we asked them to review each state's accountability plan with an eye toward capturing strengths and weaknesses.

We aimed to provide constructive feedback to the states, and to serve as a source of straightforward information to the public so that they are better able to engage policymakers if and how they see fit. Inherently, this independent process could not take into account the numerous political and situational challenges that occur in every state. We are in no way attempting to diminish those challenges, but the scope of this review was to compare the rigor and comprehensive nature of state accountability plans.

Peers worked in small teams to review the plans that states formally submitted to the U.S. Department of Education. After reviewing independently, the peers met for two days to discuss their individual reviews and work together on the collaborative draft you'll see below. The teams were asked to use their discretion and expertise to respond to and score each rubric item, and those scores were normed across states and peers.

Each state was given the opportunity to review the draft peer analysis and to provide substantive additions and corrections. Still, the reviews should be considered a snapshot of state plans as of September–November 2017, and we anticipate that states will continue to update their plans going forward.

To read more about the project, as well as a list of the expert peer reviewers, visit the Bellwether website [here](#).

## Overall Strengths and Weaknesses

**Strengths:** What are the most promising aspects of the state’s plan? What parts are worth emulating by other states?

North Carolina grounds its plan in robust data, which supports goals and openly points out significant areas for improvement. Given the relatively low baseline of achievement data for many student subgroups documented in the plan, North Carolina rightly includes strong emphasis on academic achievement indicators. The state has selected a simple list of straightforward indicators that focus on key academic outcomes for students.

North Carolina also describes a robust intervention and support system for schools. The state documents its statewide tiered support system and a comprehensive set of resources and supports available to schools and districts.

The state commits to continuous improvement to monitor implementation and make adjustments as necessary and indicates it will consider additional indicators that may provide a more holistic view of school performance beyond test scores. North Carolina documents strong efforts at stakeholder engagement in the development of its plan, and assuming this commitment continues, this approach will benefit future efforts to revisit and revise.

**Weaknesses:** What are the most pressing areas for the state to improve in its plan? What aspects should other states avoid?

North Carolina’s plan suffers from a lack of coherence between state goals, accountability indicators, and the public accountability structure. The state’s goals focus on improving the proportion of students achieving college and career readiness, but its chosen academic achievement indicator measures proficiency at a level below true college and career readiness.

Further, the proposed system over-weights achievement today relative to progress over time, which both encourages schools to focus narrowly on student performance around the proficiency cut score and fails to recognize low-performing schools that are making rapid, significant progress. Such high-growth schools may in turn be identified for state support, when they may not be the schools most in need of support.

North Carolina will assign A to F grades to schools, but the plan does not always articulate a clear relationship between those letter grades and the state accountability system. For example, the state will identify schools for comprehensive support based on their relative ranking according to the accountability

score measuring performance against the indicators. The plan does not provide data on the grades of schools likely to be identified under this methodology. As a result, it is unclear whether the ranking method would capture all or only some of the schools graded F. Additionally, this disconnect between the A to F system and the method for identifying schools for support lacks transparency and may be confusing for schools, districts, and parents.

While the state’s methodology for identifying schools with low-performing subgroups is better aligned with the A to F system, the timeline for identification lacks urgency. It is also unclear how many schools are likely to be identified.

North Carolina also proposes weak exit criteria for support. Schools can exit comprehensive support based on purely normative criteria, making it possible for schools to exit because other schools performed worse, not because of any actual improvement.

## Plan Components

Each state’s plan has been rated on a scale of 1 (“This practice should be avoided by other states”) to 5 (“This could be a potential model for other states”).

**Goals:** Are the state’s vision, goals, and interim targets aligned, ambitious, and attainable? Why or why not?



North Carolina’s vision focuses on preparing all students for their chosen path after graduation to college, career, or the military. Its theory of action for attaining this mission focuses primarily on innovation, emphasizing the state’s model for piloting innovative practices and then expanding them based on proven success. It envisions transforming the education system to a more personalized model in which students engage with “digital-age” learning.

North Carolina has taken a strong, data-based approach to setting those long-term goals and interim targets for academic achievement. The state has used data in a very thoughtful way to ensure that the goals and interim targets are attainable based on past performance, while also providing enough ambition to push schools forward. This use of data is a model for other states to consider. The state aims to increase the statewide percentage of students meeting college and career benchmarks on reading and math assessments from 43.5 to 51 percent to 65 to 70 percent, depending on grade level and subject.

The state sets interim goals for all students and for each subgroup, and they expect lower-performing subgroups to make faster progress than higher-performing groups. However, due to substantial initial gaps, even if subgroups hit their targets for each subject and grade range, gaps in test scores will remain substantial. Even with the remaining proficiency gap challenges, the goals are ambitious, with some subgroups improving proficiency rates by nearly 40 percentage points or more over 10 years. Historic performance data suggests that gains of the magnitude established in the goals may be extremely challenging to attain absent dramatic changes to instruction, supports, and resources.

The long-term goal for high school graduation rates is based solely on the four-year cohort rate and seeks 95 percent on-time graduation for all students and subgroups within 10 years, with no remaining achievement gap. This very ambitious goal is aspirational, but may be unrealistic based on the baseline graduation rates for some subgroups. For example, the most recent graduation rate available for English learners is 56.2 percent and for students with disabilities is 68.9 percent. Coupled with a long-term goal of 50 percent of English learners progressing toward proficiency (up from less than 18 percent currently), a 95 percent graduation rate for all students and subgroups seems extremely ambitious within the state’s declared timeframe.

**Standards and Assessments:** Is the state’s accountability system built on high-quality standards and assessments aligned to college and career readiness? Why or why not?

- 1
- 2
- ✓ 3
- 4
- 5

North Carolina’s standards and assessments are rigorous and aligned with college and career readiness across grades and subjects. In addition, the system sets high standards for the achievement of students with disabilities. However, it makes inadequate accommodation for English learners, creating the potential for inaccurately assessing students’ achievement by confounding language and subject matter proficiency. The system could be stronger if the state offered assessments in Spanish.

The North Carolina Standard Course of Study is based on the Common Core State Standards for math and English language arts and on state-developed learning standards for science, all developed through a rigorous process involving expert writing teams, engagement with the higher education community, and extensive engagement with educators, districts, and other stakeholders.

The state developed its own system of end-of-grade tests in grades 3 through 8 for reading and math and in grades 5 and 8 for science. Required high school assessments include end-of-course exams in Math I, English II, and Biology. The state will develop a Math III end-of-course examination to be administered beginning in 2018-19 to serve as the high school assessment for students who complete the Math I assessment prior to high school.

As a matter of state law, all North Carolina assessments are conducted in English only. The state’s English Learner Advisory Council, which includes English learner coordinators, educators, parents, and representatives

from the higher education community, engaged in the development of the state’s ESSA plan and concluded that focusing state resources on improving instruction and services for English learners would be a higher priority than developing state assessment instruments in languages other than English. However, it is worth noting that legislative action would be required for the North Carolina Department of Public Instruction (NCDPI) even to have the authority to develop and administer state assessments in languages other than English.

The vast majority of students with disabilities participate in the general assessment program with or without accommodation, which helps ensure that schools hold the same expectations and are held accountable in the same way for the achievement of students with disabilities. However, North Carolina could strengthen its plan by providing more information about its alternate achievement standards and aligned assessments for students with the most severe cognitive disabilities. Additionally, the state could strengthen its plan by providing the steps it will take to ensure that it does not exceed the 1 percent cap on participation in alternate assessments for students with the most significant cognitive disabilities.

**Indicators:** Are the state’s chosen accountability indicators aligned to ensure targets and goals are met and likely to lead to improved educational outcomes for students? Why or why not?

- 1
- 2
- ✓ 3
- 4
- 5

North Carolina identifies a short set of indicators focused on important student outcomes, including student proficiency, growth, graduation rates, progress toward English language proficiency, and a broad set of markers of college and career readiness and a measure of academic rigor at the high school level. However, the state’s academic indicators are weakened by focusing narrowly on proficiency rates, rather than crediting multiple levels of achievement, and by measuring proficiency at a lower level than the college- and career-ready standard referenced in the state’s long-term goals.

The focus on the “Level 3” standard, which reflects “grade-level proficiency” but not college and career readiness, will unnecessarily inflate school grades, create misalignment with state goals, and confuse parents who should rightfully expect schools to be held to a standard that prepares students for postsecondary opportunities.

Additionally, all the indicators applied to grades 3 through 8 focus on test scores—a missed opportunity to incorporate a more holistic assessment of school quality.

North Carolina’s high school indicators include scores on the ACT and WorkKeys aligned with admissions standards for state postsecondary programs or specific career and technical education benchmarks, which are notable for acknowledging multiple pathways to postsecondary success.

North Carolina also includes a measure of high school mathematics rigor that assesses the percentage of students passing an advanced math course. While including a measure of high school mathematics rigor is commendable, both as a marker of postsecondary preparedness and as an example of the state’s focus on STEM and advanced career and technical education, the measure would be stronger if were based on some externally validated metric rather than merely completing a course with a passing grade.

**Academic Progress:** Has the state created sufficient incentives for schools to care about both student proficiency and student growth over time? Why or why not?



North Carolina undervalues growth relative to proficiency in its system, a condition exacerbated by weaknesses in its proficiency indicators.

The state’s academic achievement indicator measures the percentage of students achieving at Level 3, which signifies grade-level proficiency, but not college and career readiness (indicated by Level 4 and 5 performance). This challenge is compounded by the proficiency indicators’ narrow focus on a single cut score, which excludes the higher college- and career-readiness-aligned standard from the scoring methodology entirely and may encourage schools to focus narrowly on the performance of students scoring near that benchmark, rather than incentivizing a broader focus on achievement at all levels.

Further, the weighting system for summative accountability scores includes the state’s growth measure at only 20 percent of the total. All other indicators combined account for 80 percent. This challenge applies at all levels, but is particularly acute for grades 3 through 8, in which proficiency rates and progress toward English language proficiency make up the entirety of this 80 percent. The 20 percent weight gives little incentive for improving the performance of students who are fairly far behind, particularly English learners and students with disabilities.

The state’s growth measure, the proprietary Educational Value-Added Assessment System (EVAAS), is a value-added growth methodology that compares current achievement with all prior achievement and indicates whether a student’s current growth meets, exceeds, or does not meet expectations based on past historical results. The EVAAS applies to both end-of-grade and end-of-course assessments, allowing growth to be measured for elementary, middle, and high schools. While its weight in the system is small relative to the focus on proficiency, the measure itself is commendable both for its rigor and its ability to measure growth at the high school level.

**All Students:** Does the state system mask the performance of some subgroups of students, or does it have adequate checks in place to ensure all students (including all subgroups of students) receive a high-quality education? Why or why not?

1

2

✓ 3

4

5

North Carolina’s proposed system translates accountability scores to summative A to F grades. Additionally, the state calculates and reports A to F grades at the school level for each student subgroup; however, the summative school grade does not include a specific weighting devoted to the performance of individual subgroups.

Schools’ subgroup grades will be used to identify schools for targeted support, and reporting this disaggregated performance information promotes public transparency around subgroup performance. However, excluding subgroup performance as a factor in summative scoring and grading is likely to mask underperformance at the school level.

The state will continue to rely on a minimum subgroup size of 30 for accountability purposes and 10 for reporting. While the smaller group size for reporting purposes supports public transparency, the large subgroup size for accountability seems likely to mask underperformance among subgroups, particularly in the state’s many small, rural schools.

The state provides statistical documentation of the impact of selecting the 30-member group size, which indicates that a significant percentage of schools will be excluded from accountability for students with disabilities and English learners, in particular. The plan’s documentation suggests significant gains at various lower group sizes in the percentage of schools and students included across groups, which calls into question the value of the tradeoffs in statistical validity and reliability cited as the rationale for the state’s choice.

Further, the state is requesting a waiver for identification of schools for targeted support so that identification for the first cohort can be based on two years of data beginning at the end of the 2018-19 school year. The delay in identification of schools with critically underperforming subgroups will also delay the pressure to improve that consequential oversight could bring to bear.

North Carolina’s plan indicates that for schools with less than 95 percent participation in state assessments among all students or at the subgroup level, the denominator used to calculate accountability scores will be equivalent to 95 percent of students. By automatically lowering accountability scores based on low participation rates, North Carolina provides an incentive to schools to ensure strong participation.

**Identifying Schools:** Is the state’s plan to identify schools for comprehensive and targeted support likely to identify the schools and student groups most in need?

1

2

✓ 3

4

5

North Carolina’s identification methodologies are basic and compliant with the law. However, in some cases, the method of identifying schools may be out of sync with the state’s A to F grading system. In addition, the relative weighting of indicators may result in the state failing to identify schools most in need of intervention and support.

North Carolina grades schools on an A to F scale, but those grades do not factor into the identification of schools for comprehensive support. Rather, in addition to identifying high schools with low graduation rates, the state will identify the bottom 5 percent of schools based on composite accountability scores. No data is provided regarding how this identification methodology correlates with the state’s grading system, so it is unclear whether a school could receive a school grade of F and not be identified for improvement. The potential disconnect between the A to F system and the identification methodology may be confusing for schools and lacks transparency. Although the state adopted an A to F scale to better define schools for families and the public, the scale as defined in this plan is confusing. For example, how can a school with letter grade F not receive comprehensive support to help students and educators?

Another challenge to this methodology is the potential impact of the relative weight of various indicators. For one, since the state greatly emphasizes academic achievement over all other indicators, it is possible that the identified schools could be helping students make significant growth while still having low proficiency rates. These schools may not be the neediest candidates for intervention, assuming they continue to produce significant growth without it, and resources may be better dedicated to schools with more stagnant outcomes.

Further, the method of combining individual indicators into a summative accountability score could result in different indicators generating different levels of weight across schools. Each indicator included in the composite is based on the percentage of students achieving specified benchmarks, and the plan alludes to the relative size of the population to which each measure applies, providing for a self-correcting weighting mechanism. For example, for an elementary school that serves a relatively small population of English learners, the percentage of students making progress toward English language proficiency would carry less weight in combination with student proficiency rates compared to a school serving relatively more English learners. The plan provides no concrete examples of how this methodology would work in practice and no analysis of how variations in student populations or the distribution of students across grades and programs may affect the relative weight of various indicators. This lack of clarity leaves open questions as to whether this system will uniformly, accurately, and consistently identify the lowest-performing schools.



The state is seeking a waiver to delay identifying schools with low-performing subgroups for targeted support until the end of the 2018-19 school year, when it plans to identify schools receiving a grade of “F” in the state’s School Performance Grades system for any subgroup, based on two years of data. After the first cohort, identification for targeted support will be based on three years of failure. During the waiver year, the state proposes putting schools on a “watch list,” but it is unclear whether watch list status serves any purpose other than putting schools on notice for potential future identification. The plan does not indicate how many schools are likely to be identified for the watch list or for future targeted support status. The delayed identification and the reliance on three years of failure both delay targeted interventions in schools struggling to serve subgroups of students.

**Supporting Schools:** Are the state’s planned interventions in comprehensive and targeted support schools evidence-based and sufficiently rigorous to match the challenges those schools face? Why or why not?

- 1
- 2
- 3
- ✓ 4
- 5

North Carolina identifies a robust set of intervention and support resources and emphasizes data-driven and evidence-based strategies. However, the state plan is less specific about which strategies will be employed when, indicating that the sampling of support strategies “may” be enacted. Additionally, the plan does not stipulate what action the state may take for perpetual underperformance.

North Carolina’s Statewide System of Support defines three levels of support, ranging from optional general support available to all schools to much more intensive supports that will be mandatory for schools identified for comprehensive and targeted support. The plan documents a menu of strategies that could apply under each level of support, but the selection of specific interventions is driven by local needs assessments and improvement plans.

The plan also references North Carolina’s newly created Innovative School District, under which low-performing schools may either be taken over by a charter school operator under state oversight, or groups of low-performing schools can form an innovation zone under local management separate from the district. The plan indicates that action through the Innovative School District structure may be an option for intervention, but does not establish a timeline or commitment for takeover or closure of chronically underperforming schools.

North Carolina could strengthen its plan by including discussion of how the state will use the 7 percent of funds set aside from Title I to support schools. Additionally, the state should indicate if and how it intends to provide direct student services using the optional 3 percent set-aside, which provides an additional opportunity for the state to align school improvement activities with its statewide goals.

**Exiting Improvement Status:** Are the state’s criteria for schools to exit comprehensive and targeted support status sufficient to demonstrate sustained improvements? Why or why not?

- 1
- ✓ 2
- 3
- 4
- 5

Schools will exit comprehensive support status once their ranking exceeds the bottom 5 percent threshold. This relative criterion without any accompanying demonstration of improvement means that a low-performing school could exit comprehensive support based on worse performance by other schools rather than actual improvement.

Schools will exit targeted support status once the identified subgroup(s) meet or exceed their growth status on their three-year growth score, or meet their interim targets. Requiring these schools to demonstrate some level of improvement is a positive practice, and the focus in subgroup-level growth is positive.

**Continuous Improvement:** Has the state outlined a clear plan to learn from its implementation efforts and modify its actions accordingly, including through continued consultation and engagement of key stakeholders? If not, what steps could the state take to do so?

- 1
- 2
- 3
- ✓ 4
- 5

The plan indicates that the state will study and consider additional indicators identified through stakeholder engagement, such as chronic absenteeism, early childhood education, physical education, school climate, and college and career readiness. The state also plans to review other states’ systems to understand how other indicators are incorporated. However, the plan for continuous improvement is not particularly detailed and does not cover other areas of the system. Additionally, the plan does not define a specific role for engagement of the state’s higher education community in achieving ambitious college and career readiness goals, a missed opportunity for a state with numerous top-tier public and private institutions.

The plan documents an extensive stakeholder engagement process in the creation of the plan, and the state should continue this robust approach as it considers revisions and improvements down the line.