

RECOMMENDATIONS FOR

North Carolina's Opportunity Scholarship Program

PURSUANT TO S.L. 2024-57

Prepared for the NC Collaboratory and the Joint Legislative
Education Oversight Committee, NC General Assembly

Dr. Matthew Pepper
Basis Policy Research

Dr. Anna Egalite
North Carolina State University
Visiting Fellow, Institute for Governance and Civics, Florida State University

Dr. Matthew G. Springer
Basis Policy Research

DECEMBER 2025



Basis Policy Research

Basis Policy Research is an independent research firm advancing evidence-based solutions in education, workforce, and child and family welfare. Through rigorous applied research, evaluation, and analytics, we partner with over 125 mission-driven organizations nationwide, including government agencies, foundations, and nonprofits, to inform decisions, strengthen programs, and demonstrate measurable impact.



North Carolina Collaboratory

The North Carolina Collaboratory is a research funding agency that partners with academic institutions and government entities to transform research into practical information and technologies for use by the State and local governments and the communities they serve. Since its authorization in 2016 by the General Assembly (see 31A N.C.G.S. Sec. 116-255), the Collaboratory has stewarded approximately \$250 million in appropriations from the legislature, investing in over 700 research projects that have the potential to develop innovative, evidence-based solutions that benefit the State and its residents.



North Carolina State University

As a research-extensive land-grant university, North Carolina State University is dedicated to excellent teaching, the creation and application of knowledge, and engagement with public and private partners. By uniting our strength in science and technology with a commitment to excellence in a comprehensive range of disciplines, NC State promotes an integrated approach to problem solving that transforms lives and provides leadership for social, economic and technological development across North Carolina and around the world.

This project is supported by the North Carolina Collaboratory at The University of North Carolina at Chapel Hill with funding appropriated by the North Carolina General Assembly.

TABLE OF CONTENTS

List of Tables, Figures, and Appendices	2
Acknowledgements	3
Executive Summary	4
1. Introduction	7
2. State Approaches for Assessing Opportunity Scholarship Students	11
3. Recommendations for a Nationally-Normed Assessment	18
4. Recommendation for Alignment of Nationally Norm-Referenced Assessments to the North Carolina Standard Course of Study	25
5. Recommendations for Developing a Through-Grade Assessment	29
6. Conclusions and Next Steps	35
References	37
Appendix A: List of Interviewees	40
Appendix B: Overview of the Study Methodology	42
Appendix C: State-by-State Comparison of K-12 Choice Programs	47
Appendix D: Data and Information Sources for Table 3, Characteristics of Nationally-Normed Assessments	49
Endnotes	54

LIST OF TABLES, FIGURES, AND APPENDICES

Table 1: 2025-26 Opportunity Scholarship Eligibility Tiers

Table 2: Comparison of K-12 Scholarship Programs in North Carolina, Indiana, Florida, and Texas

Table 3: Characteristics of Nationally Norm-Referenced Assessments

Table 4: Current Assessment Context in North Carolina Private and Public Schools

Table 5: Policies Considered but Not Recommended

Table 6: Recommendations Considered but Outside the Scope of the Legislative Study

Table 7: Comparison of Assessment System Types Referenced in Policy Discussions

Table 8: Challenges Associated with Developing a North Carolina Through-Grade Assessment System

Figure 1: Opportunity Scholarship Participation and Funding, 2014-15 through 2025-26 School Years

Appendix A: List of Interviewees

Appendix B: Overview of the Study Methodology

Appendix C: State-by-State Comparison of K-12 Choice Programs

Appendix D: Data and Information Sources for Table 3, Characteristics of Nationally-Normed Assessments

ACKNOWLEDGEMENTS

We sincerely thank the numerous educators, leaders, advocates, and policymakers who generously shared their valuable perspectives on the Opportunity Scholarship (OS) Program and the methods used to measure student achievement for both OS recipients and public school students. We are grateful to Dr. Jeni Corn and Ms. Liza Rodler from the NC Collaboratory for their insightful guidance and steadfast support throughout this work. Finally, we deeply appreciate the thoughtful comments and constructive feedback provided by anonymous reviewers, which greatly strengthened the early drafts of this report.

EXECUTIVE SUMMARY

As North Carolina's Opportunity Scholarship Program (OS Program) has expanded, so has the desire for clear information about student learning in private schools that accept Opportunity Scholarships (OS). In response, the North Carolina General Assembly (NCGA) enacted S.L. 2024-57, directing the North Carolina Collaboratory's Office of Learning Research (Collaboratory) to study assessment options for third- and eighth-grade students receiving OS. The legislation charged the Collaboratory with partnering with researchers to recommend nationally norm-referenced assessments for use across the public and private sectors, evaluate the alignment of those assessments with North Carolina's Standard Course of Study (SCOS), and assess the feasibility of developing a through-grade assessment system.

In response to this charge, the Collaboratory commissioned a research team from North Carolina State University and Basis Policy Research to conduct a multi-pronged study examining assessment practices in public and private schools in North Carolina and across private school choice programs nationwide. The study included: (a) a review of state and national policies governing assessments within voucher and Education Savings Accounts (ESA) programs; (b) an analysis of nationally norm-referenced assessments for potential adoption; (c) a review of the technical, instructional, and regulatory considerations associated with developing a through-grade assessment system; and (d) a review of relevant research literature. The research team also engaged more than 75 stakeholders to better understand practical perspectives, operational constraints, and reporting needs in the context of S.L. 2024-57. These stakeholders included state and local education leaders, school and district administrators from public and private schools, legislators from both chambers, assessment specialists, and representatives from faith-based, policy, and advocacy organizations.

The OS Program is now one of the most extensive universal private school choice programs in the nation, serving more than 104,000 students annually and distributing over \$575 million in public funds. While students in traditional public schools participate in a uniform, standards-aligned assessment system, students receiving OS are assessed using a wide array of nationally norm-referenced assessments selected by individual private schools. This divergence creates fundamental challenges for policymakers seeking comparability: students across sectors experience notable differences in the assessments, curricula, and instructional pacing to which they are exposed. These challenges are compounded by psychometric limitations when comparing results across different assessments or when equating scores from nationally norm-referenced assessments with North Carolina's standards-based End-of-Grade (EOGs) and End-of-Course (EOCs) assessments.

The authors conclude that no single assessment solution can simultaneously maximize comparability, minimize testing burden, and preserve instructional autonomy. There are, however, alternative approaches that maximize the validity and value of administered assessments and provide parents/guardians with meaningful information about their student's performance as they move between sectors.

FINDINGS AND RECOMMENDATIONS

Legislative Charge #1 - Nationally Norm-Referenced Assessments. Meaningful student-level performance comparisons across sectors should only be made when student achievement is assessed using the same test. However, requiring public schools to abandon EOGs and EOCs or mandating that private schools adopt the state assessment would entail high costs, instructional disruption, and significant policy consequences. The research team instead recommends a more practical approach in which the NCGA designates a limited number of high-quality, nationally norm-referenced assessments in reading and mathematics for third- and eighth-grade students receiving Opportunity Scholarships. The research team offers criteria and a list of assessments that meet them, along with considerations for selection. Participating private schools would choose and administer one or more assessments from this list.

This recommendation preserves private school autonomy, aligns with current private school practices, and ensures that OS students are assessed using nationally norm-referenced assessments with through-grade capabilities. However, this approach does not allow for direct performance comparisons between OS and public school students or support equating results across different nationally norm-referenced assessments. Consistent with the current psychometric literature, the research team strongly cautions against such comparisons (e.g., Kolen and Brennan, 2014; von Davier, 2011).

Legislative Charge #2 - Alignment with the Standard Course of Study. Nationally norm-referenced assessments are not designed to measure state-specific standards, and vendor claims of alignment vary widely in their rigor, purpose, and independence. If a nationally norm-referenced assessment were adopted through legislative action, the research team recommends that the North Carolina Department of Public Instruction (NCDPI) commission an independent alignment study to evaluate its alignment with the SCOS in third- and eighth-grades. An external party should conduct such a study through a competitive procurement process and employ established best-practice methodologies (preferably the GAAT framework) to produce detailed crosswalks and documentation of content coverage and cognitive demand. This type of independent study would evaluate the extent to which a nationally norm-referenced assessment aligns with SCOS. Given that private schools may adopt different content standards, curricula, and instructional practices, it is important to acknowledge that aligning an assessment with SCOS in third- and eighth-grades does not necessarily mean that the courses of study used in private schools will align in practice with the SCOS.

Legislative Charge #3 - Through-Grade Assessment Feasibility. Federal law permits states to administer multiple interim assessments that aggregate into a single summative score, and the development and implementation of a through-grade assessment system in North Carolina is technically feasible. However, the research team does not recommend developing a new statewide through-grade assessment to replace EOGs and EOCs. It would represent a substantial departure from the state's current assessment framework, resulting in more student testing events, eliminating the formative instructional benefits of existing tools such as NC Check-Ins 2.0, requiring significant financial investment, and introducing unresolved challenges related to opportunity-to-learn and fairness. As such, the research team recommends leveraging commercially available, nationally norm-referenced assessments with formative capabilities. They also suggest exploring the expansion and strategic use of classroom-level instructional tools that complement the state's current assessment system, rather than mandating a broad-scale overhaul.

SUMMARY AND NEXT STEPS

The Collaboratory-commissioned study was directed to address a narrowly defined legislative mandate that highlights broader policy issues as the OS Program continues to grow. Enhancing transparency around OS students' performance requires careful consideration of test comparability, burden, instructional relevance, cost, and alignment with the state's existing assessment system. As such, this study concludes that:

- No single assessment approach can fully satisfy all legislative goals defined in S.L. 2024-57, as each policy option explored has non-negligible tradeoffs.
- Designating a limited menu of high-quality, nationally norm-referenced assessments is the best path forward for the OS Program. Allowing private schools to select from an approved list preserves autonomy, aligns with current practice, and provides parents/guardians with nationally norm-referenced information.
- Valid, student-level comparisons require a representative sample of students to complete the same assessment from the groups to be compared. The research literature is not shy about warning against attempts to link or equate scores across different assessments and different populations due to established, substantive psychometric limitations.
- Aligning a new assessment with the SCOS requires an independent, state-specific study to ensure that curriculum, instruction, and assessments target the same objectives.
- Even though developing a through-grade assessment system is technically feasible, transforming formative tools into a summative, accountability-driven model would likely compromise the instructional benefits of those formative assessments and incur substantial costs.

As the NCGA considers next steps, clarity of purpose will be critical. That purpose should guide final decisions on the assessments to deploy, the target audience for assessment results, and the policy questions those results might confidently answer. This report offers guidance for that decision-making and outlines options for implementing proposed solutions that would work within the state's existing assessment system beginning in the 2026–27 school year.

1. INTRODUCTION

Established by the NCGA in 2013, the North Carolina Opportunity Scholarship Program (OS Program) has become one of the most extensive universal school choice programs in the United States, with a budget of over \$575 million and more than 104,000 scholarships awarded during the 2025-26 school year (North Carolina State Education Assistance Authority, 2025).¹ As the OS Program has expanded, so has the demand for clear information about student learning in schools that accept Opportunity Scholarships (OS). In response, the North Carolina General Assembly (NCGA) enacted S.L. 2024-57, directing the North Carolina Collaboratory's Office of Learning Research (Collaboratory) to study assessment options for third- and eighth-grade students receiving OS. The legislation charged the Collaboratory with recommending nationally norm-referenced assessments for use across public and private sectors, evaluating the alignment of those assessments with North Carolina's Standard Course of Study (SCOS), and assessing the feasibility of developing a through-grade assessment system. This report synthesizes those findings and outlines options for implementing proposed solutions that would work within the state's existing assessment system beginning in the 2026-27 school year.

1.1 THE LEGISLATION

The NCGA enacted S.L. 2024-57, which directs the Collaboratory to study and report the following to the General Assembly's Joint Legislative Education Oversight Committee:

1. For the purpose of comparing student performance, recommendations for nationally standardized tests for use in third grade and eighth grade that would be appropriate for administering to (i) students in nonpublic schools who are receiving Opportunity Scholarships beginning with the 2026-27 school year and (ii) students attending schools in public school units.² To the extent practicable, the Office of Learning Research shall recommend only one test for use in third grade and one test for use in eighth grade.
2. Alignment between the nationally standardized tests selected pursuant to subdivision (1) of this subsection and the standard course of study for third grade and eighth grade, respectively, including a crosswalk between the standards assessed by the nationally standardized tests and the standard course of study.
3. Feasibility of developing a through-grade assessment for third and eighth grade that would meet the following criteria:
 - a. Assess mastery of the standard course of study.
 - b. Consists of multiple testing events throughout the year that are aggregated into a summary score.
 - c. Replace the current end-of-grade assessments for third and eighth grade.
 - d. Yield data that can be used with the Education Value-Added Assessment System (EVAAS).³
 - e. Comply with federal law.

To meet these legislative directives, the research team designed a multi-pronged study examining assessment practices in public and private schools in North Carolina and across private school choice programs nationwide. The study included: (a) a review of state and national policies governing assessments within voucher and Education Savings Accounts (ESA) programs; (b) an analysis of nationally norm-referenced assessments for potential adoption; (c) a review of the technical, instructional, and regulatory considerations associated with developing a through-grade assessment system; and (d) a review of relevant research literature. The research team also engaged more than 75 stakeholders to better understand practical perspectives, operational constraints, and reporting needs in the context of S.L. 2024-57. These stakeholders included state and local education leaders, school and district administrators from public and private schools, legislators from both chambers, assessment specialists, and representatives from faith-based, policy, and advocacy organizations.

1.2 DESCRIPTION OF THE OPPORTUNITY SCHOLARSHIP PROGRAM

Established in 2013 by the NCGA, the OS Program is a universal private school choice program, meaning it is open to all North Carolina families with students entering kindergarten through 12th grade. Voucher amounts range from approximately \$3,000 to \$7,000, depending on household income, with priority given to the lowest income families (per Table 1). OS funds can be used towards tuition and fees at registered private schools that participate in the OS Program. As of fall 2025, more than 700 private schools were registered with the North Carolina State Education Assistance Authority (NCSEAA)⁴ as “Direct Payment Schools,” meaning these schools accept OS funds. Home schools are not eligible to participate in the OS Program.

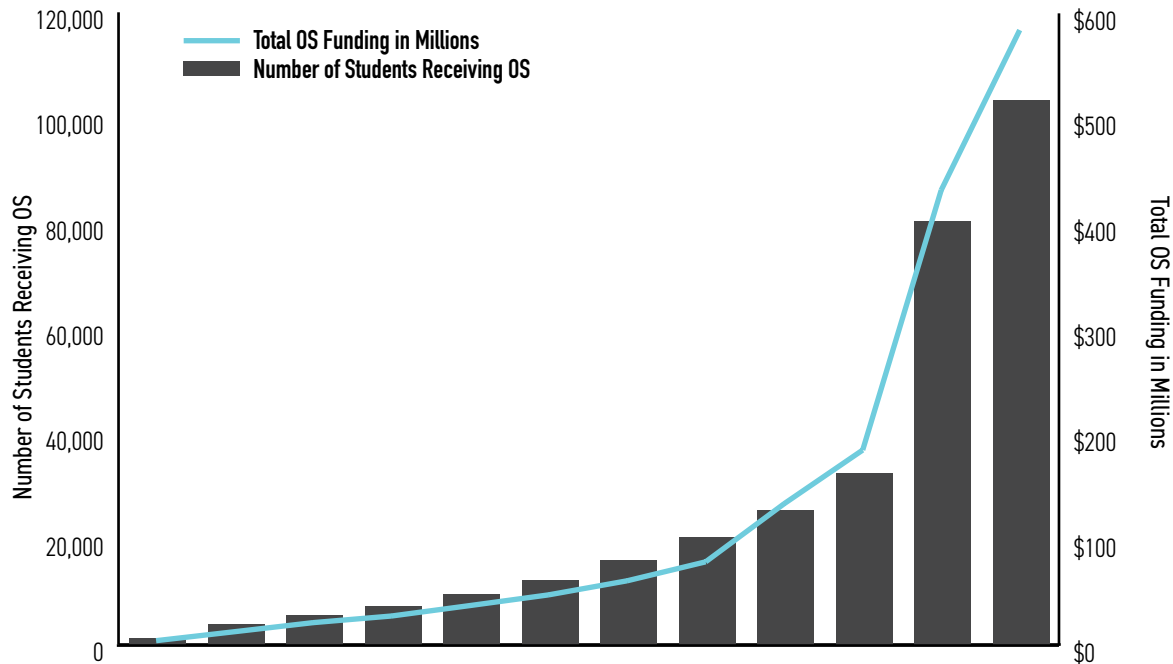
Table 1: 2025–26 Opportunity Scholarship Eligibility Tiers

Tier	Award Limit	Household Income (Family of Four)
1	\$7,686	\$59,478
2	\$6,918	\$118,956
3	\$4,612	\$267,651
4	\$3,458	No income limit

Source: <https://k12.ncseaa.edu/media/faqan5tq/incometiersnps25.pdf>

As shown in Figure 1, program participation and funding have grown exponentially, with the number of participating students nearly tripling over the last two school years. In the 2025-26 school year, approximately 104,000 students were participating in the OS Program, with program expenditures expected to exceed \$584 million.⁵ Two-thirds (66 percent) of these students are renewal students; that is, they participated in the OS Program in the prior school year.

Figure 1: Opportunity Scholarship Participation and Funding, 2014-15 through 2025-26 School Years



Source: North Carolina State Education Assistance Authority, Opportunity Scholarships Summary Data, Downloaded December 15, 2025 from <https://www.ncseaa.edu/opportunity-scholarship-summary-of-data/>.

With universal eligibility, rapid uptake, and hundreds of millions in annual public funding, the OS Program plays a pivotal role in determining where North Carolina’s students learn and how public resources are allocated. Future research is needed to better understand the following questions: Who participates in the program? How have parent/guardian perspectives of the OS Program evolved? What impact does the program have on these students’ academic achievement and durable skills? Do OS students have better secondary, post-secondary, and labor-market outcomes? How does the OS Program affect the broader public school system?

1.3 STUDY OVERVIEW

The Collaboratory commissioned researchers from North Carolina State University (NC State) and Basis Policy Research (Basis) to carry out the legislatively mandated study. In partnership with the Collaboratory, the research team developed a comprehensive study comprising three strands of work aligned with the three sections of S.L. 2024-57.

The first strand involved a comprehensive review of state and national policies governing assessments in school choice programs, with particular emphasis on recently adopted voucher and ESA initiatives. The research team also identified and analyzed nationally

norm-referenced assessments, examining their structure, scope, and key characteristics, alongside a thorough review of relevant literature. Discussions with more than 75 key stakeholders and technical experts over five months further informed the findings and recommendations in this study (see Appendix A for a complete list).

For the second strand of work, the research team reviewed the content, subject, and grade coverage of nationally norm-referenced assessments, including alignment studies published by independent sources and assessment companies. The team also reviewed the various methods used by content experts to measure alignment and then cataloged the processes required to conduct an independent alignment study between identified assessments and the North Carolina SCOS. Once again, the discussions with stakeholders and technical experts informed this strand of work.

The third strand of work involved reviewing the current assessment requirements and processes within North Carolina public schools to understand the implications of developing and implementing a through-grade assessment system, including those related to federal regulations, modifications to educator policies (e.g., EVAAS calculations and educator compensation), and instructional relevancy.

See Appendix B for additional details on the study's research design, data sources, analytic methods, and stakeholder engagement procedures.

2. STATE APPROACHES FOR ASSESSING OPPORTUNITY SCHOLARSHIP STUDENTS

Across states, policymakers must carefully navigate a complex balance between academic accountability, administrative feasibility, and institutional autonomy. With roughly 40 percent of students nationwide now eligible to participate in private school choice programs (Cohen and DiMarco, 2024), these programs – and their approaches to assessing student performance – vary markedly.⁶ For example, Indiana requires participating private schools to administer the same state assessments as public schools, thereby emphasizing cross-sector accountability and transparency. In contrast, states like Arkansas and Virginia offer schools a selection of nationally norm-referenced assessments, prioritizing flexibility and reduced regulatory burden. Meanwhile, states such as Arizona impose no formal testing requirements, particularly for disability-focused Education Savings Accounts (ESAs).

Below, we provide more details on three states with comparable private school choice programs: Indiana and Florida, each with close to 80,000 students and facing similar testing questions as North Carolina and Texas, which is considering new through-grade assessments directly related to S.L. 2024-57.

2.1 INDIANA: PRIORITIZING CROSS-SECTOR TRANSPARENCY, LIMITING FLEXIBILITY

Indiana represents one end of this policy spectrum by prioritizing cross-sector transparency and limiting flexibility. The nearly 80,000 students enrolled in the state's Choice Scholarship Program in the 2025–26 school year are required to participate in the same statewide assessments administered to public school students, including the Indiana Learning Evaluation Assessment Readiness Network (ILEARN) summative assessment. Indiana also administers a through-grade assessment in grades 3–8 mathematics and English language arts consisting of three checkpoints that supplement rather than replace the end-of-year summative testing. Private schools are not responsible for the cost of administering these assessments. Joining Iowa and Tennessee, Indiana is one of only three states that require students participating in private school choice programs to take the same assessment as public school students, reflecting a strong emphasis on comparability and cross-sector transparency (Roy, Schwartz, & Gable, 2024). While facilitating direct comparisons, this approach also imposes constrictive alignment and compliance requirements on participating private schools.

2.2 FLORIDA: MAXIMIZING FEASIBILITY AND PARENTAL CHOICE, LIMITING COMPARABILITY

Florida, in contrast, offers a particularly instructive example of a minimally complex assessment model within a large-scale private school choice system. Florida operates one of the nation's largest private school choice programs, serving more than 80,000 students in the 2025–26 school year through the Tax Credit Scholarship Program alone (EdChoice, 2026), and spends more on private school choice relative to total K–12 expenditures than any other state. Under Florida law, participating private schools may administer any nationally norm-referenced assessment from an approved list of 28 vendors (Florida Department of Education, N.D.). As an alternative, private schools or individual families may opt to administer Florida's statewide standardized assessment to fulfill this requirement, though participation is limited to end-of-year tests and does not include interim progress-monitoring assessments. Assessment results are reported to parents/guardians and submitted to an independent research organization at Florida State University, which aggregates student performance using Normal Curve Equivalent (NCE) scores. Because students take different assessments (e.g., Stanford Achievement, Iowa Test, TerraNova), the state does not attempt to equate results to Florida's state accountability assessment or compare performance between sectors. Instead, the state produces an annual program-level report summarizing student performance relative to national norms and the distribution of gains and losses across the scholarship population.

This approach is technically straightforward, low-cost, scalable, and imposes minimal regulatory burden given Florida's large and growing scholarship population. By avoiding the complexity associated with test equating, growth modeling, and school-level accountability determinations across heterogeneous assessment systems, Florida has established a transparent reporting framework that provides a descriptive snapshot of how scholarship students perform relative to national norms. Yet that framework does not support inferences about private school students' learning relative to Florida standards, nor enable precise comparisons of instructional effectiveness and student- and school-level performance between the public and private school sectors. As the program's evaluators themselves note, there is “no correspondence” between Florida's statewide assessment and the array of norm-referenced assessments taken by scholarship students (Zuilkowski et al., 2024). And any cross-sector comparisons are to be avoided entirely. Florida's model prioritizes feasibility and parental choice over comparability and accountability.

2.3 TEXAS: DIVERGING ASSESSMENT RULES ACROSS SECTORS — MORE TESTING FOR PUBLIC SCHOOLS, FLEXIBILITY FOR ESAs

Recent policy developments in Texas illustrate the different strategies being used to monitor student learning across sectors. For public schools, House Bill 8 will introduce significant changes to the Texas Assessment Program, replacing the State of Texas Assessments of Academic Readiness (STAAR) with three shorter tests administered at the beginning, middle, and end of the school year (i.e., a through-grade approach). Governor Greg Abbott signed the bill on September 17, 2025, and it will affect students starting in the 2027–28 school year. “House Bill 8 ends the high stakes and high stress nature of one test, one day,” said Rep. Brad Buckley, the bill’s author (Dey, 2025). “This is unprecedented oversight of the assessment and accountability system by this body” (Dey, 2025).

Lawmakers positioned this bill as a response to the frustration of families and teachers who say the existing testing regime puts undue pressure on students and that preparation for year-end testing consumes too much instructional time during the latter part of each school year. This proposal has spurred mixed reactions from educational commentators, with the President of the Thomas B. Fordham Institute, Michael Petrilli (2025), writing, “I sincerely don’t understand why we think parents, teachers, and kids are going to like three high-stakes standardized tests a year instead of one.” This critique highlights a central design challenge of through-grade assessments. While they may reduce the salience of any single testing event, they also risk increasing the cumulative testing footprint and burden if each administration carries accountability consequences.

At the same time, Texas has adopted a markedly different approach for students participating in its new Education Savings Account program, the Texas Education Freedom Accounts (TEFA), enacted in 2025 and set to launch in fall 2026 with an initial \$1 billion appropriation. Under TEFA, participating private school students in third through twelfth grades are required to take an annual, nationally norm-referenced assessment.

In effect, Texas is moving toward a state-designed through-year testing strategy for public-school accountability, while relying on norm-referenced testing for ESA oversight, reflecting different priorities for comparability and burden across sectors. This approach illustrates the practical appeal and the internal tension of applying different measurement philosophies across sectors. On the ESA side, the state’s norm-referenced testing requirement is straightforward and low-burden, offering private schools flexibility like Florida’s approach while still preserving a basic accountability signal for policymakers and families. On the public-school side, however, the shift to a three-event, through-year testing system represents a significant increase in design complexity and operational demands, and it introduces new questions about comparability, administrative burden, and the role each testing event will play in accountability. In short, Texas is pairing a relatively feasible, flexible assessment requirement for ESA oversight with a far more ambitious redesign of the statewide public-school assessment system. This approach may satisfy different constituencies but also risks heightening public concern about the overall testing burden and complicating the state’s accountability architecture.

2.4 NORTH CAROLINA: MAINTAINING PRIVATE-SECTOR FLEXIBILITY, LIMITING CROSS-SECTOR COMPARABILITY

North Carolina currently requires private schools receiving OS funds to administer a nationally norm-referenced assessment each year. However, the policy does not require participating private schools to use the same assessment instrument statewide, nor does it require alignment with the state’s public-school assessment system (i.e., EOGs and EOCs assessments). This approach reflects a policy balance that prioritizes institutional autonomy and feasibility for private schools, while limiting the extent to which results can be used for comparing student performance across schools and sectors.

In practice, assessment practices vary substantively both within and between private and public sectors. Public schools operate under a uniform statewide accountability system built around standardized end-of-grade and end-of-course assessments aligned to the Standard Course of Study, designed explicitly for comparability and growth modeling. Private schools, in contrast, may draw from a range of norm-referenced instruments that vary in content coverage, scaling, administration conditions, and reporting metrics. These differences mean that OS assessment results can provide descriptive information about student performance relative to national norms, but do not support inferences about performance relative to North Carolina’s academic standards or differences in achievement across schools and sectors.

Table 2: Comparison of K–12 Scholarship Programs in North Carolina, Indiana, Florida, and Texas

Program Type(s)	TESTING REQUIREMENTS FOR...		Eligibility Criteria	# Choice Students (school year)
	Private Schools	Public Schools		
North Carolina				
School vouchers (Opportunity Scholarship); Education Savings Account (ESA)	Voucher students must take a nationally norm-referenced standardized test each year	• Annual, summative, EOG in ELA/Reading & Math (grades 3–8) and science (grades 5 & 8) • Annual, summative, EOCs in Math 1, English II, and Biology (high school)	• Open to all K–12 students statewide without an income limit • Priority goes to renewals and lower-income applicants if applications exceed the budget • ESA program is limited to students with disabilities who have an IEP	104,073 (2025–26)
Indiana				
School vouchers (Choice Scholarship); Tax-credit scholarships (SGO); Individual tax deduction	Voucher students must take Indiana’s statewide assessments	• Annual statewide ILEARN assessments in ELA & Math (grades 3–8) administered through-grade via 3 checkpoints plus an end-of-year (EOY) summative component • Science: Grades 4 & 6 (EOY) • Social Studies: Grade 5 (EOY) • High school: Biology EOC	• Universal eligibility for all K–12 students (as of 2024–25); prior income limits removed • SGO scholarships remain income-capped at 300 percent of FRL	76,067 (2024–25)

Program Type(s)	TESTING REQUIREMENTS FOR...		Eligibility Criteria	# Choice Students (school year)
	Private Schools	Public Schools		
Florida				
ESA (Family Empowerment Scholarship for Unique Abilities); Tax-credit ESAs (Florida Tax Credit Scholarship)	Scholarship students must take either Florida's state assessments or a nationally norm-referenced test	- Annual statewide, norm-referenced assessments in ELA (grades 3–10) and mathematics (grades 3–8) End-of-course (EOC) assessments in Algebra 1 and Geometry - Statewide science/social studies testing includes comprehensive science assessments and EOCs such as Biology 1, Civics, and U.S. History	Universal eligibility for all K–12 students (as of 2023), with priority to low-income and foster students	502,705 (2025–26)
Texas				
ESA (Texas Education Freedom Accounts, begins 2026)	Participating students in grades 3–12 must take a nationally norm-referenced standardized test each year. (No state exam requirement)	- Annual, statewide STAAR summative assessments in grades 3–8 (Math & Reading/Language Arts; Science in grades 5 & 8; Social Studies in grade 8) - STAAR end-of-course exams in Algebra I, Biology, English I, English II, and U.S. History - Beginning in 2027–28, HB 8 replaces STAAR with three shorter statewide testing events for grades 3–8 (a through-grade approach)	- Universal ESA program effective 2026–27 - All K–12 students eligible, up to \$10,474 per student	Does not currently operate a program

Sources: EdChoice. (2026). *The ABCs of school choice: The comprehensive guide to every private school choice program in America (2026 ed.)*. <https://www.edchoice.org/wp-content/uploads/2025/12/The-ABCs-of-School-Choice-2026-WEB.pdf>; United States Department of Education, Office of Elementary and Secondary Education (n.d.). Key documents: School support and accountability. U.S. Department of Education. Retrieved December 12, 2025, from <https://www.ed.gov/about/ed-offices/oese/key-documents-school-support-and-accountability>. Author.

2.5 THE ABSENCE OF A STRAIGHTFORWARD SOLUTION

The diverse approaches to assessing student performance across states underscore the challenges the NCGA faces in obtaining comparable, valid information on student learning while minimizing testing burden, preserving instructional relevance, and controlling costs. For example, states that prioritize cross-sector comparability tend to adopt more prescriptive and complex assessment requirements. However, imposing strict assessment mandates on students receiving OS funds can introduce additional regulatory burdens and raise concerns about alignment with the mission and institutional goals of private schools. Conversely, states that emphasize practicality, sustainability, and institutional diversity tend to rely on nationally norm-referenced assessments, which limit the depth and scope of insights that can be drawn from the data.

This state review, combined with insights from North Carolina practitioners and policymakers and expert guidance from assessment and measurement experts, demonstrates that no single approach can meet all legislative goals outlined in S.L. 2024-57. Potential options such as using a single nationally norm-referenced assessment per grade, commissioning alignment studies, or developing a through-grade assessment system each have trade-offs in terms of cost, timeline, acceptability, and the usefulness of the results.

Policy recommendations and legislative decisions must carefully account for system dependencies, capacities, costs, and measurement challenges when determining the path forward, as summarized further below and referenced throughout the recommendation sections of this report.

2.5.1 SYSTEM DEPENDENCIES AND CAPACITIES

- **Testing Burden:** Proposals to change assessments in third- and eighth-grade would either replace current tests or add new ones, increasing the overall testing burden on students. Replacing existing assessment practices poses challenges because these tests are deeply integrated into schools' core teaching and learning systems (e.g., teachers and administrators typically participate in regular training to interpret results and adjust instruction accordingly). Expanding testing requirements raises serious concerns for families, as 78 percent of public school parents/guardians feel their children already take too many standardized assessments, and 81 percent believe testing consumes too much time that could otherwise be spent on teaching and learning (North Carolina Department of Public Instruction, 2019).
- **Transition Challenges:** Most North Carolina private schools have well-established assessment systems. Some educators with whom the research team spoke described years of investment in staff training and communication supporting these deeply embedded assessment practices.
- **State Capacity:** Current staffing levels and infrastructure at the NCSEAA are already operating at full capacity to fulfill existing legislative reporting requirements. Introducing new reporting requirements and/or maintaining a designated assessment menu would require commensurate increases in operational support.

2.5.2 COSTS

- **Transition Costs:** Implementing a new assessment requires substantial financial and transition-related investments, such as training, communications, and data security. While recent advances in generative AI have improved efficiencies in item creation, high-quality test development remains a fundamentally political process. Decisions about which knowledge, skills, and values to assess – reflected in test content – require negotiation among state leaders, content experts, advocacy groups, and the public. Even with AI tools accelerating technical tasks, the complex, resource-intensive work of building consensus on test purpose, content priorities, performance expectations, and accountability stakes remains essential. Thus, adopting a new assessment system requires not only financial and operational readiness but also broad stakeholder buy-in to ensure legitimacy and alignment with state educational goals.
- **Additional Costs, Additional Value:** Current per-student costs for North Carolina's EOGs and EOCs assessments are generally lower than those for widely used nationally norm-referenced assessments.^{7,8} However, the additional costs associated with nationally norm-referenced assessments may reflect additional supports, such as enhanced psychometrics, user-friendly dashboards, communication tools for parents, guardians, and educators, training, and professional development, and offset both local costs and demands.

2.5.3 MEASUREMENT CHALLENGES

- **Non-Comparability Across Assessments:** Achieving statistical concordance between different assessments, even nationally norm-referenced ones, is complex, and the research literature on the topic strongly cautions against it.⁹ Leading measurement experts generally advise against comparing or attempting to equate scores across different nationally norm-referenced assessments due to technical limitations such as differences in test content and constructs, scaling and score metrics, norming samples, test administration conditions, and the absence of common anchor items or shared equating designs needed to place scores on a truly comparable scale.
- **Statistical Validity of Cross-Sector Comparisons:** Even when using the same assessment, accounting for differences in the students being tested across contexts is challenging. For example, families' decision to participate in the OS Program is likely strongly correlated with factors such as parental education, income, motivation, educational values, prior achievement, and access to supplemental learning resources – all of which influence student outcomes (Woolridge, 2010). Consequently, observed differences in student performance often reflect preexisting differences in the students themselves rather than actual differences in school effectiveness. Without rigorous controls for selection bias through experimental or quasi-experimental methods, estimates of student and school performance risk being biased (Neal, 2002). Other areas of concern include opportunity-to-learn differences, administration variability, and accommodations that can bias score equating and, in turn, growth estimates derived from those scores.
- **Innate Differences by Sector:** Differences in curriculum scope and sequencing, instructional pacing, and pedagogical emphasis complicate comparisons across diverse groups, languages, and cultures, even when the same assessment is used (Braun, Jenkins, & Grigg, 2006). As a result, straightforward performance comparisons between public and private school students may overstate sector differences and obscure the underlying factors driving those differences.

3. RECOMMENDATIONS FOR A NATIONALLY-NORMED ASSESSMENT

3.1 LEGISLATIVE CHARGE

The first of the three legislative charges in S.L. 2024-57 was to identify an assessment (or assessments) with the following characteristics:

- The assessment(s) must be nationally norm-referenced, which is defined as performance benchmarks derived from a nationally representative sample of students that serve as the reference group against which scores from a state, district, or school are compared.
- The assessment(s) must be recommended for use in third- and eighth-grade.
- The assessment(s) must be appropriate for use in private schools and public schools and capable of judging a student's continued academic growth if the student moves from one sector to the other.

3.2 RECOMMENDATION

The research team recommends that the NCGA designate a list of high-quality, nationally norm-referenced assessments for reading and math in third- and eighth-grade. Private schools accepting OS dollars would select and administer one or more of those assessments to all OS students, with the state paying those costs.

This recommendation aligns with current assessment practices in the state, preserves the autonomy and flexibility valued by private schools, and narrows the list of recommended assessments to those that have through-grade assessment capabilities. This recommendation also enables the state to monitor the ongoing performance of OS students and benchmark that performance relative to national norms.

While the research team is confident in this recommendation, there are, however, some downsides. For example, private schools that are not presently administering one of the approved assessments would need to add an entirely new assessment, replace one or more of their extant assessments, and/or withdraw from the OS Program owing to switching costs associated with altering their testing regimes. In addition, there is no psychometrically valid approach for evaluating student-level performance between private and public schools unless a representative sample of public school students also takes each test administered in private schools accepting OS students.

3.3 SUMMARY OF SUPPORTING EVIDENCE

In response to the first legislative charge, the research team conducted a comprehensive review of state and national policies governing assessment in school choice programs, identified and analyzed nationally norm-referenced assessments, examining their structure, scope, and key characteristics, alongside a thorough review of relevant literature. Discussions with more than 75 key stakeholders and technical experts over five months were also instrumental in shaping the recommendations and findings in this study.

3.3.1 REVIEW OF NATIONALLY NORM-REFERENCED ASSESSMENTS

The research team conducted a review of nationally norm-referenced assessments using a structured review matrix. This matrix evaluated each candidate assessment against the requirements of S.L. 2024-57, focusing on whether it: (a) applies to grades 3 and 8, (b) provides through-grade assessment and summative scoring, and (c) is supported by valid evidence of alignment with SCOS. Assessments were excluded if they were teacher-developed; not nationally normed; primarily used as screeners rather than comprehensive measures of student achievement; intended solely for instructional support within a limited scope or curricular sequence rather than comprehensively measuring student mastery; designed exclusively for instructional diagnostics without clear concordance with state summative assessments; or were diagnostic/adaptive tests intended to predict performance on assessments that were otherwise precluded from consideration (e.g., not nationally-normed).¹⁰

As shown in Table 3, our review indicates that i-Ready (Standards Mastery), NWEA MAP, and Renaissance Star meet all evaluated criteria. Each of these assessments is available in both third- and eighth-grade, includes through-grade and summative testing options, and enables student results to be compared to national norms. Their cost per pupil ranges from approximately \$13.50 to \$17.00, which is about 40 to 62 percent higher than the EOGs and EOCs currently administered in North Carolina.¹¹

Table 3: Characteristics of Nationally Norm-Referenced Assessments

Assessment	Available in Third- and Eighth-grade	Through-Grade and Summative	Costs	Accommodations	National Norms Study (Year)	Used by North Carolina Educators	Evidence of Alignment with NC SCOS
Panel A: Meets All Criteria							
i-Ready (Standards Mastery)	✓	Yes, typically 3 times per year	\$14.50 per student	✓	✓ 2021	✓	✓
NWEA MAP	✓	Yes, typically 3 times per year	\$13.50 to \$14.50 per student	✓	✓ 2025	✓	✓
Renaissance Star	✓	Yes, 2 to 5 times per year	\$14.00 to \$17.00 per student	✓	✓ 2025 (English) 2023 (Reading, Literacy, Math)	✓	✓
Panel B: Meets Some Criteria							
California Achievement Test (CAT 6)	✓	No, typically as a one-time annual assessment	Unavailable	✓	✗ 2005	✓	✗
CTP (ERB)	✓	Yes, when used with the Interim Milestones assessments	Unavailable	✓	✓ 2024	✓	✗
Iowa Test of Basic Skills (ITBS)	✓	No, typically a one-time annual assessment	\$14.50 per student	✓	✓ 2025	✓	✗
Stanford Achievement Test	✓	No, typically a one-time annual assessment	\$18.00 per student	✓	✗ 2018	✓	✗
TerraNova NEXT	✓	Yes, when used with the DRC Beacon	\$24.90 per student	✓	✗ 2017	✓	✗
Panel C: North Carolina End-of-Grade Assessment							
North Carolina End-of-Course	✓	Yes, when used with NC Check-Ins	\$9 per student	✓	Not nationally norm-referenced	✓	✓

Sources: Data on assessments used in North Carolina public schools obtained from 2-24-25 Local Testing Calendars submitted to the North Carolina Department of Public Instruction under §115C-174.12(d) and a data extract from the Private School Annual Report obtained from the Division of Non-Public Education. For all other sources, see Appendix D for a comprehensive listing. The per student cost estimate for TerraNova NEXT incorporates the DRC Beacon assessment. The i-Ready alignment study in North Carolina is based on approximately 40,000 student observations. For NWEA MAP and Renaissance Star, the number of student observations was approximately 80,000 and 10,000, respectively. The per student cost estimate for EOGs and EOCs does not include development and administration costs for NC Check-Ins.

3.3.2 REVIEW OF ASSESSMENTS CURRENTLY ADMINISTERED IN NORTH CAROLINA

The research team conducted a comprehensive review of the assessments administered by public and private schools in North Carolina during the 2024–25 school year. As shown in Table 4 and described earlier, public school students are required to take the EOGs and EOCs assessments administered by the NCDPI. Private schools receiving OS dollars, on the other hand, must administer a nationally norm-referenced assessment selected by the senior executive responsible for overseeing the school’s educational programs.

Table 4 highlights the most common assessments in private schools, as well as state-mandated and supplemental assessments in public schools. The limited overlap between these assessments creates significant challenges for comparing student performance. Two factors further complicate this reality: (a) EOGs and EOCs assessments are not nationally normed, leaving no psychometrically valid way to make comparisons with nationally norm-referenced assessments, and (b) private schools have invested substantial resources in staff training for specific assessment systems, making any transition to alternative assessments costly and complex.

In addition to reviewing relevant research literature, the research team also discussed these challenges and potential solutions with state and national psychometric experts to explore whether there is a methodologically sound approach for comparing: (a) private school student performance on nationally norm-referenced assessments with public school student performance on the EOGs and EOCs, and (b) private school student performance between and across different nationally norm-referenced assessments. The review and conversations with experts confirmed that no state currently uses concordance between assessments on a statewide or large-scale level, and that developing, testing, and validating solutions to overcome these limitations remains a leading – and highly debated – area of psychometric research.¹² Consequently, until such a solution is established, the research team strongly advises against making such comparisons.

Table 4: Current Assessment Context in North Carolina Private and Public Schools

	Private Schools (2024-25: 930 schools)	Public Districts (2024-25: 115 districts; 2,758 schools)
Required Grades Assessed	ALL PRIVATE SCHOOLS: Grades 3, 6, 9, and 11; PRIVATE SCHOOLS RECEIVING OS FUNDS: Grades 3-12	Grades 3 to 9, 11 (Math) Grades 3 to 8, 10 (Reading/English) Grades 5, 8, 10 (Science/Biology)
Assessments Administered	Nationally norm-referenced assessment selected by the school's chief administrative officer	EOGs and EOCs
Reporting of Results	ALL PRIVATE SCHOOLS: Not publicly reported; PRIVATE SCHOOLS RECEIVING OS FUNDS: Results submitted to NCSEAA, but not publicly reported	By school and district, through North Carolina School Report Cards
Common Assessments Used (2024-25 School Year)	# of schools Utilize listed assessment	# of districts Utilize listed assessment
	218 Iowa (ITBS)	115 EOGs and EOCs
	119 NWEA MAP	45 iReady
	63 ACT	19 Renaissance Star
	51 Stanford Achievement Test	16 mClass
	46 CTP by ERB	8 CogAT
	34 Terra Nova	7 Schoolnet
	32 California Achievement Test	6 Exact Path

Source: Information about testing requirements in private schools obtained from NCGS § 115C-557, NCGS § 115C-558, and NCGS § 115C-562.5.

3.3.3 POLICY OPTIONS CONSIDERED, BUT NOT RECOMMENDED

The research team considered several alternative policy options but did not recommend them due to various challenges or failure to meet the criteria defined in the first legislative charge. Details, including benefits and drawbacks, are outlined in Table 5, which also presents two alternative policies to improve private school transparency for parents/guardians. The research team ultimately concluded that changing public school assessments now is not immediately feasible, as it would require overhauling the state's accountability system and impact major policy and program areas, such as school report cards, growth metrics, and staff evaluations. However, as shown in Table 4, many public school districts already administer nationally norm-referenced assessments as formative tools. Nevertheless, current law mandates that all public schools administer End-of-Grade (EOG) and End-of-Course (EOC) assessments, which creates a disincentive for schools to fully transition to nationally norm-referenced assessments.

Table 5: Policies Considered but Not Recommended

Policy Option	Benefits	Drawbacks
1. Require all private school students using OS to participate in math and reading EOGs and EOCs assessments in third- and eighth-grade.	- Represents a cost-effective policy option as EOGs and EOCs are significantly cheaper compared to high-quality nationally norm-referenced assessments. - Could allow comparisons between students accepting OS and public school students if a representative sample of OS students took it.	- Private schools may use curricula and/or instructional pacing that do not align with the NC Standard Course of Study. - Does not allow private schools enrolling OS students to independently select a high-quality nationally norm-referenced assessment. - Does not provide insight into how students accepting OS perform relative to national benchmarks.
2. Contract with a nationally recognized assessment firm like MetaMetrics to develop comparative analyses between student performance on 1 to 3 nationally norm-referenced math and reading assessments for students accepting OS.	In theory, this approach allows for comparisons between students accepting OS even with different nationally norm-referenced assessments.	- Research studies and leading assessment experts caution that this remains frontier work that has yet to be implemented at scale. - Significant concerns were raised about using concordance analyses to compare student performance across different assessments. - Cost estimates are approximately \$160,000 the first year, and \$120,000 annually thereafter.
3. Contract with a company like the SAS Institute to estimate growth on nationally norm-referenced assessments for students accepting OS.	Using a common metric enables valid comparisons within the private sector.	- Research studies and leading assessment experts caution that this remains frontier work that has yet to be implemented at scale. - Significant concerns were raised about comparing growth measures across different assessments. - EVAAS could only be calculated when minimum sample sizes of students at each subject and grade level were met, which excludes many schools serving students receiving OS that did not meet these thresholds. - Cost estimates are approximately \$190,000 for within-test progress measures each year.

Table 6: Recommendations Considered but Outside the Study Scope of the Legislative Study

Policy Options	Benefits	Drawbacks
4. Require private schools accepting OS to be accredited in lieu of new assessment requirements.	• Create an accreditation board for schools accepting OS without placing additional burden on students. • Does not increase costs for North Carolina.	• Does not strengthen data infrastructure across schools accepting OS. • Does not allow direct comparisons between students in public and private schools.
5. Create public-facing data dashboards summarizing grade-level performance for students at schools receiving OS.	• Increase data transparency for policymakers and families/guardians to make educated decisions about where they spend OS dollars.	• Does not allow direct comparisons between students in public and private schools, though it does provide more public information about private school performance. • Requires staffing commensurate with additional responsibilities or contracting with an external party.
6. Conduct a one-time, rigorous evaluation of the OS Program. Administer a single assessment to a statistically representative sample of students receiving OS and a matched group of public school students.	• Provide a single, point-in-time estimate of the impact of the OS Program. Measure the impact of Scholarships on student performance. • Imposes testing burden on a much smaller number of students in public and private schools.	• Does not provide ongoing oversight of students receiving OS, or comparisons with public school students. • Does not strengthen data infrastructure or accountability systems across schools accepting OS.

4. RECOMMENDATION FOR ALIGNMENT OF NATIONALLY NORM-REFERENCED ASSESSMENTS TO THE NORTH CAROLINA STANDARD COURSE OF STUDY

4.1 LEGISLATIVE CHARGE

The second legislative charge was to assess how well the nationally norm-referenced assessments proposed in response to the first charge align with the North Carolina Standard Course of Study (SCOS) for third- and eighth-grade.

4.2 RECOMMENDATION

The research team recommends that NCDPI issue a Request for Proposals (RFP) to procure an independent vendor to conduct a formal alignment study upon the identification and selection of high-quality, nationally norm-referenced assessments for use in schools accepting OS funds.

4.3 SUMMARY OF SUPPORTING EVIDENCE

In response to the second legislative charge, the research team reviewed the content, subject, and grade coverage of nationally norm-referenced assessments, including alignment studies published by independent sources and assessment companies. The team also reviewed the various methods used by content experts to measure alignment and cataloged the processes required to conduct an independent alignment study between recommended assessments and the North Carolina SCOS.

4.3.1 WHY STANDARDS-ASSESSMENT ALIGNMENT MATTERS AND HOW IT IS ESTABLISHED

Evidence-based alignment between content standards and student assessments is foundational to valid, standards-based testing. Strong alignment enables conclusions on student achievement and growth to be drawn confidently from large-scale assessment data. In the absence of this evidence, use of assessment results to evaluate student performance is both less defensible and less informative.

In North Carolina, the State Board of Education (SBE) is responsible for adopting the content standards that guide instruction in all public schools. State-mandated EOGs and EOCs were developed to explicitly align with these content standards, ensuring that statewide assessments validly measure the knowledge and skills expected to be taught in public school classrooms across the state.

Commercial test vendors employ a mix of approaches to demonstrate alignment between their assessments and state standards, though the rigor and purpose of these approaches vary markedly. Examples include: (a) vendor-led mapping of state standards to assessment items; (b) creation and periodic updating (often annual) of state-specific learning progressions that mirror state standards; (c) linking or concordance studies that predict student performance on an assessment so that results can be interpreted on a common scale; (d) correlational analyses to understand the extent to which their assessments' results relate to states' summative assessments; and (e) independent alignment studies using established methods (preferably the GAAT framework) and/or state-led blueprint reviews.

These approaches and their underlying documentation vary in credibility, technical depth, and intended use. Some commercial test vendors document evidence of linking and alignment. In contrast, others caution that assessments not built to a given state's standards will have measurement, diagnostic, and accountability limitations, even when they are mapped to those standards after the fact. Accordingly, the research team treats these vendor claims as preliminary indicators of potential fit, rather than as substitutes for a more formal state-commissioned alignment process.

When North Carolina adopts new content standards, the state must develop and/or identify assessments to measure student mastery of revised expectations. Subsequent to the initial administration of each new assessment, the state commissions an external alignment study to evaluate the relationship between assessment items and state standards.¹³ North Carolina's present reading and math assessments are in their fifth edition, and science assessments are in their third edition. All extant NC assessments have undergone rigorous independent review by external experts to verify their alignment with state-adopted content standards.

S.L. 2024-57 requires alignment between the SCOS and one or more nationally norm-referenced assessments. However, nationally norm-referenced assessments do not inherently measure state-specific content. In the absence of an independent conducted alignment study, the state needs to commission an independent study to determine: (a) the degree to which assessment blueprints cover the breadth and depth of SCOS content; (b) whether constructs (e.g., reading comprehension subdomains or mathematical strands) overlap sufficiently to support valid interpretations; (c) the presence of gaps (e.g., unassessed content areas and/or cognitive processes); and (d) the utility and/or appropriateness of the assessment for high-stakes, comparative, and/or accountability purposes.

North Carolina’s most recent EOGs and EOCs alignment studies in reading and mathematics adhere to Webb’s criteria, which remain the national standard for evaluating assessment alignment. Webb’s four core alignment criteria include:

- *Categorical Concurrence*: The degree to which the domains or categories of knowledge assessed correspond to those defined in the standards.
- *Depth of Knowledge Consistency*: The extent to which the cognitive complexity of assessment items reflects the intended rigor of the standards.
- *Range of Knowledge Correspondence*: The breadth and diversity of content covered by the assessment in relation to the standards.
- *Balance of Representation*: The equitable representation of content areas and cognitive demands across the assessment to ensure that all facets of the standards are measured proportionately and in accordance with their intended emphasis.

To ensure credibility and technical rigor, and to document compliance with state and federal regulations, North Carolina contracts with external experts through a competitive procurement process to conduct independent, rigorous alignment studies. NCDPI oversees the contractual and operational aspects of these studies but does not participate in the evaluative judgments of alignment outcomes. Commercial assessment vendors must engage national content experts and North Carolina teachers to independently review assessment items and determine the degree of alignment between those items and the SCOS.

4.3.2 METHODS FOR CONDUCTING AN ALIGNMENT STUDY

Alignment evidence is a core source of validity support in standards-based testing because it links the inferences we want to draw from scores to the content and cognitive demands students are expected to learn. This is particularly important in large-scale K–12 testing contexts, which is why the federal education law (ESSA) requires states to administer high-quality annual academic assessments in reading, mathematics, and science, and specifies that these assessments must be aligned with the state’s academic content and achievement standards.

There are many established approaches for documenting alignment, and the research literature reflects progression towards methods that are more systemic, transparent, and dependable. Historically, three of the most common include the Surveys of Enacted Curriculum (SEC; Porter & Smithson, 2001), the Achieve method (Rothman et al., 2002), and Webb’s alignment method (Webb, 1997, 1999), with Webb’s method often cited as the most widely used for mandated statewide achievement tests. Building on the strengths of these prior approaches, the research team supports the findings of Cizek and colleagues (2018) and recommends that North Carolina use the Generalized Assessment Alignment Tool (GAAT) as the primary framework for documenting alignment. GAAT is explicitly designed to support alignment claims in large-scale assessment contexts by producing structured, comparable, and purpose-sensitive evidence.

Alternative approaches are very feasible and bring their own strengths to an alignment study. The SEC approach is distinctive in its emphasis on the relationship among standards, instruction, and assessment, using a common content “language” (often a two-dimensional

classification of topics by cognitive demand) to compare what is taught and what is tested. The Achieve method focuses on whether an assessment reflects the central expectations of the standards, typically using structured review processes that consider content match, challenge, and balance, to determine whether the test supports valid claims about student performance relative to standards. Webb's method, in contrast, is widely used in statewide contexts and relies on clear, criterion-based evidence across multiple dimensions, such as categorical concurrence, depth-of-knowledge consistency, range-of-knowledge correspondence, and balance of representation. These claims map directly onto questions policymakers ask about whether a test adequately covers the breadth and rigor of standards. Although the methods differ in emphasis (e.g., links to instruction, centrality of content, cognitive demand, or breadth of coverage), they share a key feature: they rely on structured judgments by subject matter experts who know the standards and the intended test population.

GAAT extends and systematizes these traditions by organizing alignment evidence into four distinct facets: curriculum coverage, construct comprehensiveness, cognitive complexity, and content concentration. Each of these facets is summarized using quantitative indices intended to be transparent, interpretable, and comparable across contexts, which allows reviewers and policymakers to distinguish among different dimensions of alignment strength and weakness. An additional advantage of GAAT is that it can reduce the judgment burden on expert reviewers and uses more nuanced rating scales, which can strengthen the dependability of alignment conclusions. By producing clear, comparable, and purpose-sensitive alignment evidence, GAAT strengthens the state's ability to make defensible claims about assessment quality and alignment to adopted academic standards.

5. RECOMMENDATIONS FOR DEVELOPING A THROUGH-GRADE ASSESSMENT

5.1 LEGISLATIVE CHARGE

The final of the three legislative charges was to determine the feasibility of developing a through-grade assessment system for use in third- and eighth-grades that (a) assesses mastery of the standard courses of study, (b) consists of multiple testing events that aggregate to a summative score, (c) replaces current EOGs and EOCs assessments, (d) yields data compatible with the EVAAS, and (e) complies with federal law.

5.2 RECOMMENDATIONS

Rather than developing a new through-grade assessment, the research team recommends leveraging commercially available, nationally norm-referenced assessments with formative features and scaling the strategic use of classroom-level instructional tools to offer a more practical and cost-effective pathway for meeting the NCGA's objectives. This approach preserves existing state investments, minimizes operational risk, and maintains a clear distinction between assessments designed for accountability and tools designed to support teaching and learning. Ultimately, it allows the state to advance the legislative intent of S.L. 2024-57 while avoiding the financial, technical, and operational burdens associated with developing and validating a new statewide through-grade model.

5.3 SUMMARY OF SUPPORTING EVIDENCE

In response to the third legislative charge, the research team examined the current assessment requirements and processes within North Carolina public schools to understand the implications of developing and implementing a through-grade assessment system, including matters related to federal regulations, instructional relevance, and consequences on current policies and practices such as EVAAS calculations and educator compensation.

5.3.1 FEDERAL CONTEXT AND LEGAL FEASIBILITY

Under Section 1111 of the Every Student Succeeds Act, states may administer a series of statewide interim assessments that together produce a single summative score, provided the system yields valid, reliable, and transparent information on achievement or growth. In principle, this provision allows states to replace a single end-of-year assessment with a through-grade assessment comprising multiple testing events.

While ESSA permits such systems, and some states are piloting through-grade assessment systems that combine interim checkpoints with end-of-year summative tests, there is currently no wholly implemented statewide assessment that has completed federal peer review and received official approval from the U.S. Department of Education to replace existing statewide summative assessments for accountability purposes under ESSA. As a result, the legal pathway exists, but practical implementation remains complex and uncertain.

5.3.2 INNOVATIVE ASSESSMENT WORK IN NORTH CAROLINA

North Carolina has been a leader in pioneering innovative assessment strategies. In 2018, the U.S. Department of Education established the Innovative Assessment Demonstration Authority (IADA) to give states greater flexibility in developing novel assessment approaches, such as through-grade systems. As one of the first five states to receive this authority, North Carolina has since made remarkable strides in creating a more balanced and effective assessment system.

A key component of this work is the North Carolina Personalized Assessment Tool (NCPAT). The NCPAT framework includes the NC Check-Ins 2.0 and a multistage adaptive EOGs and EOCs summative assessment. NC Check-Ins 2.0 are administered periodically throughout the school year and are explicitly designed as classroom-level formative instruments. They provide teachers with detailed, timely information to support instructional planning and within-year adjustments, such as item-level results, performance by standard and domain, and comparative information at the class, school, and state levels. Teachers have access to assessment items and student responses, and schools retain flexibility in scheduling administrations based on local pacing.

Although the NC Check-Ins 2.0 system demonstrates North Carolina's technical and operational capacity for frequent online assessment and rapid reporting, its formative design differs fundamentally from what would be required of a true through-grade accountability system. Any through-grade assessment system intended to replace EOGs and EOCs would require standardized administration conditions, secure items, and psychometric linking across testing events to support valid summative scores, growth modeling, and accountability determinations.

Repurposing NC Check-Ins 2.0 for this function would require substantial redesign. Changes would remove teachers' access to items, restrict classroom flexibility, and shift the assessments away from their formative purpose. In effect, converting a formative system into a summative accountability instrument would alter the features that currently make that assessment instructionally useful.

A complementary strategy worth consideration is to expand the use of classroom-level and AI-enabled instructional tools that support personalized learning and rapid feedback, supplementing existing statewide summative assessments. Under this approach, statewide EOGs and EOCs would continue to serve their accountability and growth-measurement functions, while instructional tools would provide information on student performance at multiple points throughout the academic year. This would enable schools to better monitor student progress and support instructional decisions by leveraging a nationally norm-referenced assessment with formative capabilities. Some commercially available assessments support regular testing throughout the year, provide timely, instructionally helpful feedback, and produce summative results.

5.3.3 DISTINGUISHING THROUGH-GRADE ASSESSMENTS FROM AI-ENABLED LEARNING TOOLS

Some lawmakers have expressed strong interest in assessment approaches that provide continuous, technology-enabled insight into student learning throughout the school year. This interest aligns with the growing availability of instructional tools that rely on low-stakes, classroom-embedded assessment data to support personalized learning.

One example is Khanmigo, an AI-powered tutoring and instructional support tool. Khanmigo uses student interactions with instructional content to generate real-time feedback for teachers and students, supporting lesson planning, targeted assistance, and immediate instructional adjustments. These systems exemplify a broader category of classroom-level, rapid-feedback assessment tools designed to support learning rather than measurement.

Another example is Teach to One, a middle school mathematics program implemented in New York City. Teach to One relies on frequent, low-stakes assessments embedded in daily instruction to inform a machine-learning-based scheduling system that produces individualized daily learning plans, or “playlists,” for students. The model uses assessment data to support personalized pacing, grouping, and instructional delivery within a defined curriculum, rapidly adjusting instruction based on student performance.

While these tools offer significant instructional value, they are distinct in purpose and design from through-grade assessments as defined in statute and federal accountability law. Through-grade assessments are statewide measurement systems intentionally designed to aggregate results across multiple testing events into a single summative score that can replace an EOGs and EOCs, support comparability across schools, and yield data suitable for growth modeling systems such as EVAAS.

In contrast, instructional models and AI-enabled learning tools such as Khanmigo and Teach to One rely on adaptive, context-specific formative assessment data used locally. These assessments are neither standardized across students and schools, nor appropriate for aggregation into summative scores, and they do not satisfy federal accountability or value-added mandates. Scaling such tools statewide would not constitute a through-grade assessment system capable of replacing EOGs and EOCs under ESSA.

Statewide summative assessments, innovative pilot assessments, and classroom-level instructional tools are designed to serve distinct functions within an assessment ecosystem. Challenges arise when tools built to support learning and instruction are expected to fulfill accountability functions for which they were not designed.

5.3.4 THREE TYPES OF ASSESSMENT SYSTEMS

Table 7 identifies and summarizes key differences among the three referenced types of assessment systems: (a) statewide summative assessments authorized under ESSA; (b) innovative pilot formative assessments, and (c) classroom-level assessments used in personalized instructional models. The table demonstrates that each assessment category fulfills a distinct role within a comprehensive assessment system. While all three provide valuable insights into student learning, they are not interchangeable. Designing a through-grade assessment system to replace EOGs and EOCs necessitates prioritizing standardization, comparability, and accountability, whereas instructional tools prioritize immediacy, adaptability, and local decision-making. Clear alignment between purpose and design is therefore essential when considering future assessment policy decisions.

5.3.5 CHALLENGES ASSOCIATED WITH A NEW NORTH CAROLINA THROUGH-GRADE SYSTEM

Table 8 summarizes the challenges associated with designing and implementing a new through-grade assessment system, which would represent a significant shift from North Carolina's current assessment framework rather than an incremental modification. Although federal law permits states to develop and administer multiple interim assessments that aggregate into a single summative score, creating a new through-grade assessment system requires navigating a broad array of policy, operational, financial, and psychometric challenges that extend well beyond test design alone. This change would likely increase the number of tests students must take annually, eliminate the formative instructional benefits of existing tools such as NC Check-Ins 2.0, and require a substantial financial investment. Furthermore, key features that make current tools instructionally valuable (e.g., teacher access to test items, scheduling flexibility, and rapid formative feedback) would be constrained in a summative accountability context. Additionally, producing EVAAS-compatible results from the new system would require extensive psychometric development and close coordination with SAS, while unresolved challenges remain regarding opportunity-to-learn and fairness.

Table 7: Comparison of Assessment System Types Referenced in Policy Discussions

	Existing Statewide Summative Assessments (ESSA-authorized)	Innovative Pilot Formative Assessments (e.g., NCPAT)	Classroom-Level Assessments for Personalized Instruction (e.g., Khanmigo)
Primary Purpose	Accountability, proficiency determination, and growth measurement	Instructional support and system learning	Real-time instructional adjustment and personalization
Assessment Stakes	High stakes	Low- to moderate-stakes	Low stakes
Administration	Standardized statewide conditions	State-designed with controlled flexibility	Embedded in daily instruction
Frequency	Once annually or multiple times per year that can be aggregated	Multiple times per year	Continuous or near continuous
Item Security	Fully secure	Partially secure	Not secure
Teacher Access to Items and Responses	✗	✓	✓
Standardized across Schools	Required	Limited / pilot-based	Not required
Produce Summative Score	✓	No (formative use only)	✗
Compatibility with Growth Models (e.g., EVAAS)	✓	✗	✗
Federal Accountability Compliance	✓	Pilot only	✗
Instructional Flexibility	Low	Moderate	High
Users of Results	State, districts, schools	Teachers, schools, state (diagnostic)	Teachers and students

Table 8: Challenges Associated with Developing a North Carolina Through-Grade Assessment System

Feasibility Consideration	Description
1. Increased testing burden	A through-grade assessment system increases the number of standardized testing events, all of which require secure administration comparable to EOG and EOC protocols (e.g., scripts, proctors, covered materials, controlled environments).
2. Decreased instructional feedback	Because summative items must remain secure, teachers would no longer have access to items or student responses, eliminating parts of the formative value central to NC Check-Ins 2.0.
3. Opportunity-to-Learn and fairness	Students assessed early in the year may not yet have been taught evaluated content, raising fairness and validity concerns consistent with the Standards for Educational and Psychological Testing.
4. Pacing guide requirements	A uniform statewide pacing guide would be required to ensure a consistent opportunity to learn. Local pacing flexibility currently allowed under NC Check-Ins 2.0 would no longer be permissible.
5. Measuring learning across time	Composite scores based on early-year assessments may fail to capture learning that occurs later. Students who master content later in the school year could be disadvantaged.
6. State law governing assessment timing	Current law requires EOGs and EOCs to be administered in the final 5–10 days of the year to maximize instructional time. A through-grade assessment system operates on a fundamentally different principle.
7. Implications for EVAAS	Growth models rely on full-year achievement data. Interim-based composites require substantial redesign and coordination with SAS to ensure accurate growth calculations.
8. Potential for moral hazard	Assessments in a state's accountability system can create a moral hazard, as educators may encourage initial underperformance to inflate gains in later administrations.
9. Technical and operational feasibility	Although combining interim scores is psychometrically feasible, strong statewide rules for missing data, mobility, and score comparability is required to ensure reliability and interpretability.
10. Financial	Developing unbiased, validated assessment items is costly; for example, the average cost of creating a new NAEP test question was estimated at \$3,700 per item.

Notes: The Standards for Educational and Psychological Testing were developed by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education. The standards are open access at <https://www.testingstandards.net/open-access-files.html>. For more information on the cost of test items, see National Academies of Sciences, Engineering, and Medicine. (2022). *A Pragmatic Future for NAEP: Containing Costs and Updating Technologies*. The National Academies Press.

6. CONCLUSIONS AND NEXT STEPS

This study was initiated and designed in response to S.L. 2024-57, enacted by the NCGA, which directed the Collaboratory to examine assessment options for third- and eighth-grade students receiving OS. Despite the relatively narrow focus of these requirements, the findings highlight broader policy considerations as the OS Program expands in scale, scope, and fiscal significance. To illustrate, S.L. 2024-57 requires the design of an assessment system for OS, underscoring the NCGA's commitment to transparency and responsible stewardship of public funds. Yet, the study also reveals that selecting an assessment strategy aligning to S.L. 2024-57 is a complex policy decision with substantial instructional, financial, and policy consequences. Meeting all requirements would necessitate an overhaul of the state's current assessment framework.

Identification alone of a nationally norm-referenced assessment for use in both private and public schools would entail prohibitive constraints. Either public schools would need to abandon EOGs and EOCs or private schools would need to adopt the state's existing testing regime, with both options involving financial costs, instructional disruption, and significant policy consequences.

Accordingly, the research team instead recommends that the NCGA designate a limited number of high-quality, nationally norm-referenced assessments with formative assessment properties in reading and mathematics for third- and eighth-grade students, for use by private schools enrolling OS students. This approach would ensure that OS students take high-quality assessments with through-grade capabilities and would provide parents/guardians and legislators with clearer information about the performance of students who use public funds to attend private schools. It also preserves private school autonomy and aligns with current private school practices.

With respect to the second legislative charge, which requires aligning assessments with the SCOS, the study finds opportunity to enhance assessment validity through a thoughtful sequencing and review process. Nationally norm-referenced assessments typically do not reflect state- or private school-specific standards. Should the NCGA choose to identify one or more nationally norm-referenced assessments for third- and eighth-grade students for schools accepting OS funds to select from, the research team recommends that NCDPI commission an independent alignment study. This study should be conducted through a competitive procurement process and utilize established best-practice methodologies (preferably the GAAT framework) to develop detailed crosswalks and thorough documentation of content coverage and cognitive demand.

While conducting an alignment study with the SCOS does not guarantee that private schools' courses of study will align in practice, given that private schools may follow different content standards, curricula, and instructional practices, it still provides valuable insights that can benefit public schools currently administering these assessments. Moreover, this approach allows private schools the flexibility to continue innovating with their curricula and instructional practices, while fostering transparency and comparability in student achievement.

The third and final legislative charge related to assessing the feasibility of a through-grade assessment system highlights an important distinction between technical feasibility and policy desirability. North Carolina has demonstrated its interest in and commitment to through-grade and interim assessments through its participation in the Innovative Assessment Demonstration Authority and development of NC Check-Ins 2.0. However, the design, implementation, and refinement of a summative, accountability-bearing through-grade assessment system would require a fundamental redesign of the state's existing testing regime and instructional systems, negate key instructional benefits associated with these systems, and introduce challenges related to opportunity-to-learn, fairness, and growth modeling. While such a system could be built, doing so would represent a significant departure from the state's current assessment philosophy and entail substantial costs and risks.

The research team instead recommends leveraging commercially available, nationally norm-referenced assessments with formative capabilities. Exploring the expansion and strategic use of classroom-level instructional tools would complement and elevate the state's current assessment system, rather than mandating a broad-scale overhaul. This approach balances the need for rigorous, comparable data with practical considerations of cost, implementation time, and minimizing disruption to instruction. By integrating these tools thoughtfully, private and public schools can enhance ongoing student learning and provide educators with actionable insights.

To conclude, as the NCGA considers next steps, clarity of purpose will be essential. Decisions about assessments should be guided not only by technical feasibility but also by a shared understanding of how the results will be used, who the primary audiences are, and which policy questions the assessments are intended to address. Whether the state prioritizes descriptive comparisons, accountability enforcement, instructional improvement, or long-term program evaluation will fundamentally shape the tradeoffs it is willing to accept.

REFERENCES

- Bolt, D. (2021). Commentary on 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(2), 168-172.
- Braun, H., Jenkins, F., & Grigg, W. (2006). *Comparing private schools and public schools using hierarchical linear modeling*. U.S. Department of Education, National Center for Education Statistics.
- Carolina Demography at the Carolina Population Center, UNC-CH, & Demographic Analytics Advisors. (2024–2025). *Enrollment forecast: Five-year projection*. <https://www.ednc.org/wp-content/uploads/2025/05/Dollar-report.pdf>.
- Chingos, M. M. (2012). *Strength in numbers: State spending on K–12 assessment systems*. Brookings Institution.
- Cizek, G. J., Kosh, A. E., & Toutkoushian, E. (2018). Generalized assessment alignment tool (GAAT): A framework and indices for evaluating alignment between academic standards and assessments. *Journal of Educational Measurement*, 55(4), 477–512.
- Cohen, J., & DiMarco, N. (2024). *Early returns: First results from the new wave of public funding of private schooling*. FutureEd. <https://www.future-ed.org/early-returns-first-results-from-the-new-wave-of-public-funding-of-private-sc>.
- Coleman, J. S., Hoffer, T., & Kilgore, S. (1982). *High school achievement: Public, Catholic, and private schools compared*. Basic Books.
- Cronbach, L. J. (1970). *Essentials of psychological testing* (3rd ed.). Harper & Row.
- Davison, M. L. (2021). Commentary on 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(2), 173-186.
- Dey, S. (2025). Texas is poised to replace STAAR. Here is what schools' new standardized tests would look like. *The Texas Tribune*, September 30. Accessed from <https://www.texastribune.org/2025/09/30/texas-staar-standardized-test-accountability/>.
- EdChoice. (2026). *The ABCs of school choice: The comprehensive guide to every private school choice program in America (2026 ed.)*. <https://www.edchoice.org/wp-content/uploads/2025/12/The-ABCs-of-School-Choice-2026-WEB.pdf>.
- Egalite, A., Kisida, B., & Bowen, D. (2025). *Factors Driving School Choice in North Carolina During the COVID-19 Pandemic*. <https://ced.ncsu.edu/elphd/wp-content/uploads/sites/16/2025/06/OS-Eval-Report-9.pdf>.
- Egan, K., Davidson, A., & Rabinowitz, S. (2023). *Alignment Study for North Carolina End-of-Grade and End-of-Course Assessments in Grades 3-8 and High School*. Edmetric, LLC.
- Florida Department of Education (N.D.). Annual Assessment Requirements. Tallahassee. Accessed from <https://www.fldoe.org/schools/school-choice/k-12-scholarship-programs/ftc/annual-assessment-requirement.stml>.

- Forte, E. (2017). *Evaluating alignment in large-scale standards-based assessment systems*. Council of Chief State School Officers.
- Ho, A. D., Reardon, S.F., & Kalogrides, D. (2021). Rejoinder to commentaries on 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(2), 209-218.
- Kolen, M. J., & Brennan, R. L. (2014). *Test Equating, Scaling, and Linking*. Springer.
<https://doi.org/10.1007/978-1-4939-0317-7>.
- Levin, H. M. (2002). A comprehensive framework for evaluating educational vouchers. *Educational Evaluation and Policy Analysis*, 24(3), 159-174.
- Lueken, M. F., & Bedrick, J. (2024). *Early returns: First results from the new wave of public funding of private schooling*. EdChoice.
- Moses, T., & Doran, N. J. (2021). Commentary on 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(2), 187-202.
- National Academies of Sciences, Engineering, and Medicine. (2022). *A Pragmatic Future for NAEP: Containing Costs and Updating Technologies*. The National Academies Press.
- National Research Council. 1999. *Uncommon Measures: Equivalence and Linkage Among Educational Tests*. Washington, DC: The National Academies Press.
<https://doi.org/10.17226/6332>.
- Neal, D. (2002). How vouchers could change the market for education. *Journal of Economic Perspectives*, 16(4), 25-44. <https://doi.org/10.1257/089533002320950966>.
- North Carolina Department of Administration. (2025). *Statistical summary for private schools, 2024-2025*. <https://www.doa.nc.gov/media/18150/open>.
- North Carolina Department of Public Instruction. (n.d.). *Overview of EducationNC*.
<https://www.dpi.nc.gov/documents/overview-education-nc/open>.
- North Carolina Department of Public Instruction. (n.d.). *Test specifications*.
<https://www.dpi.nc.gov/document-terms/test-specifications>.
- North Carolina Department of Public Instruction. (2019). State Superintendent Mark Johnson Announces Steps to Reduce Testing for Students in North Carolina. January 15. Accessed from <https://www.dpi.nc.gov/news/press-releases/2019/01/15/state-superintendent-mark-johnson-announces-steps-reduce-testing-students-north-carolina>.
- North Carolina State Education Assistance Authority. (2025). *Summary data*.
<https://www.ncseaa.edu/opportunity-scholarship-summary-of-data/>.
- Petrilli, M. J. (2025). *SCHOOLED: Opening Bell*. September 9. Accessed from <https://schooledbymikepetrilli.substack.com/p/schooled>.

- Porter, A. C., & Smithson, J. L. (2001). *Defining, developing, and using curriculum indicators* (CPRE Research Report Series No. RR-048). Philadelphia: University of Pennsylvania: Consortium for Policy Research in Education.
- Reardon, S. F., Kalogrides, D., & Ho, A. D. (2021). 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(2), 138–167.
- Rothman, R., Slattery, J. B., Vranek, J. L., & Resnick, L. B. (2002). *Benchmarking and alignment of standards and testing* (CSE Technical Report 566). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.
- Roy, S., Schwartz, H. L., & Gable, A. (2024). *The path toward evaluating the impacts of education savings accounts on academic achievement outcomes* (RR-A3431-1). RAND Corporation.
- United States Department of Education, Office of Elementary and Secondary Education (n.d.). Key documents: School support and accountability. U.S. Department of Education. Retrieved December 12, 2025, from <https://www.ed.gov/about/ed-offices/oese/key-documents-school-support-and-accountability>. Author.
- Von Davier, A. A. (Ed.). (2011). *Statistical Models for Test Equating, Scaling, and Linking*. Springer New York. <https://doi.org/10.1007/978-0-387-98138-3>.
- Von Davier, A. A. (2021). Commentary on 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(2), 203-208.
- Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education* (Research Monograph No. 8). Washington, DC: Council of Chief State School Officers.
- Webb, N. L. (1999). *Alignment of science and mathematics standards and assessments in four states* (Research Monograph No. 18). Madison: Wisconsin Center for Education Research.
- Woolridge, J. M. (2010). *Econometric Analysis of Cross Section Panel Data* (2nd ed.). MIT Press.
- Zuilkowski, S., Kim, Y., & Akinrinmade, B. (2024). *Evaluation of the Florida Tax Credit Scholarship Program*. Learning Systems Institute, Florida State University.

APPENDIX A: LIST OF INTERVIEWEES

The individuals listed below met with the research team to share their perspectives on assessments, the OS Program, and S.L. 2024-57.

A.1 INDIVIDUALS

- Mr. Dalton Bailey, Research and Data Manager, North Carolina Rural Center
- Mr. Chuck Baldecchi, Head of School, Charlotte Latin School
- Ms. Brenda Berg, President & CEO, BEST NC
- Representative Brian Biggs, North Carolina House of Representatives
- Representative Hugh Blackwell, North Carolina House of Representatives
- Mr. Eric Bradley, Head of School, GRACE Christian School
- Ms. Margaret Bradsher, President, North Carolina School Boards Association
- Dr. Anna Bragg, Director of Marketing and Enrollment, Diocese of Raleigh
- Ms. Dominique Burgess, Founder and Head of School, Burbrella Learning Academy
- Senator Jay Chaudhuri, North Carolina State Senate
- Mr. Eric Davis, Chairman, State Board of Education
- Mr. Bobby Dixon, Director, NC Department of Administration's Non-Public Education Division
- Mr. Alan Duncan, Vice Chair, State Board of Education
- Mr. Don D'Ambrosi, Research Assistant to State Representative Dean Arp
- Representative Jeffrey Elmore, former member, North Carolina House of Representatives
- Dr. Stephen Fisher, Superintendent, Cleveland County Schools
- Ms. Cameron Florio, Founder, Grace Community University-Model
- Mr. Rupen Fofaria, Director, State Board of Education
- Dr. James Ford, Founding Executive Director, Center for Racial Equity in Education
- Ms. Beverly Fowler, Head of School, Salisbury Academy
- Dr. Lauren Fox, Senior Director of Policy and Research, Public School Forum of North Carolina
- Mr. Bruce Friend, Co-Founder of Coast to Mountains Preparatory Academy and Chair, North Carolina Charter Schools Review Board
- Mr. Mo Green, Superintendent, North Carolina Department of Public Instruction
- Ms. Allison Guenther, Vice President of External Affairs, Parents for Educational Freedom in North Carolina
- Dr. Anthony D. Jackson, Superintendent, Chatham County Schools
- Ms. Katherine Joyce, Executive Director, North Carolina Association of School Administrators

- Mr. David Kaiser, Senior Director of Policy, Advocacy, and Innovation, North Carolina Rural Center
- Ms. Lindalyn Kakadelis, Executive Director, North Carolina Coalition for Charter Schools
- Ms. Stephanie Keaney, Executive Director, North Carolina Association of Independent Schools
- Mr. John Lassiter, Principal, Hertford Grammar School
- Mr. Mike Long, President, Parents for Educational Freedom in North Carolina
- Dr. Michael Maher, Chief Accountability Officer, North Carolina Department of Public Instruction
- Dr. Katheryn Marker, Director of K-12 Scholarship Programs, North Carolina State Education Assistance Authority
- Dr. Debbie Marsh, President-Elect, North Carolina School Boards Association
- Mr. Michael Mastrocinque, Director of Instruction, Diocese of Raleigh
- Senator Graig Meyer, North Carolina State Senate
- Dr. Melanie Mikusa, Head of School, Morganton Day School
- Ms. Sara Miller, Director of Student Programs, Hill Learning Center
- Dr. Gregory Monroe, Superintendent, Diocese of Charlotte
- Mr. David Moody, Head of School, Arendell Parrott Academy
- Ms. Marcia Edge Navarro, Assistant Superintendent of Leadership, Diocese of Raleigh
- Mr. Kris Nordstrom, Senior Policy Analyst, North Carolina Justice Center
- Representative Erin Paré, North Carolina House of Representatives
- Dr. Anthony Sgro, Head of School, Asheville School
- Ms. Mary Shuping, Executive Director and Director of Governmental and External Affairs, North Carolina State Education Assistance Authority
- Mr. Gregg Sinderson, Co-Founder of Coast to Mountains Preparatory Academy
- Ms. Maureen Stover, Vice President of Policy & Engagement, BEST NC
- Dr. Tara Terry, Head of School, The Fletcher School
- Mr. Derrick Willard, Head of School, Ravenscroft School
- Ms. Leanne Winner, Executive Director for the North Carolina School Boards Association

A.2 GROUPS

- Members of the North Carolina Department of Public Instruction Technical Advisory Group
- Superintendents from the North Carolina Large District Consortium
- Attendees at the October 2025 Meeting of the North Carolina Center for Effective Education at the John Locke Foundation

APPENDIX B: OVERVIEW OF THE STUDY METHODOLOGY

B.1 LEGISLATIVE CHARGE AND CLARIFICATION OF SCOPE

The NCGA tasked the NC Collaboratory with studying and developing recommendations aligned with S.L. 2024-57, as set forth below.

1. For the purpose of comparing student performance, recommendations for nationally standardized tests for use in third grade and eighth grade that would be appropriate for administering to (i) students in nonpublic schools who are receiving Opportunity Scholarships beginning with the 2026-27 school year and (ii) students attending schools in public school units.¹⁴ To the extent practicable, the Office of Learning Research shall recommend only one test for use in third grade and one test for use in eighth grade.
2. Alignment between the nationally standardized tests selected pursuant to subdivision (1) of this subsection and the standard course of study for third grade and eighth grade, respectively, including a crosswalk between the standards assessed by the nationally standardized tests and the standard course of study.
3. Feasibility of developing a through-grade assessment for third and eighth grade that would meet the following criteria:
 - a. Assess mastery of the standard course of study.
 - b. Consists of multiple testing events throughout the year that are aggregated into a summary score.
 - c. Replace the current end-of-grade assessments for third and eighth grade.
 - d. Yield data that can be used with the Education Value-Added Assessment System (EVAAS).
 - e. Comply with federal law.

In partnership with the Collaboratory, the research team designed a study to understand existing assessment practices, collate and review existing alignment studies, evaluate the technical viability of different comparability approaches with the goal of making clear policy recommendations, and surface the implementation conditions and constraints that would shape any future policy decision. The goal was to triangulate evidence rather than rely on a single method, recognizing that questions of comparability, alignment, and assessment feasibility require both technical and practical perspectives. The study did not involve new student testing, conducting alignment studies for one or more nationally norm-referenced assessments, or developing a through-grade assessment system.

B.2 STUDY DESIGN

The study employed a multi-method approach organized that examined assessment practices in both public and private schools in North Carolina and across private school choice programs nationwide. It included:

- A review of state and national policies governing assessments within voucher and Education Savings Accounts (ESA) programs;
- An analysis of nationally norm-referenced assessments for potential adoption; and
- A review of the technical, instructional, and regulatory considerations associated with developing a through-grade assessment system.

The research team also engaged more than 75 stakeholders to understand practical perspectives, operational constraints, and reporting needs in the context of S.L. 2024-57, and reviewed relevant research on the topic. These stakeholders included state and local education leaders, school and district administrators from public and private schools, legislators from both chambers, assessment specialists, and representatives from faith-based, policy, and advocacy organizations.

B.3 DATA SOURCES

B.3.1 STATE AND AGENCY DOCUMENTS

- Annual reports on private school assessment practices obtain from the Division of Non-Public Education;
- Opportunity Scholarship Program data from the North Carolina State Education Assistance Authority
- Publicly available information on the Local School Testing Reports from the North Carolina Department of Public Instruction;
- Policies governing assessment, alignment, and peer review from the North Carolina State Board of Education;
- Technical manuals, alignment studies, and psychometric documentation from test vendors.

B.3.2 NATIONAL POLICY AND RESEARCH REVIEW

- A 50-state review of choice and ESA assessment requirements, including statutory language, reporting expectations, and transparency practices;
- A review of assessments administered in both universal and disability-focused programs;
- A review of published technical evidence and guides from vendors and state departments of education;
- A review of published alignment studies conducted for state departments of education by external vendors and research teams;
- A review of assessment manuals documenting reliability, validity, scaling, and score interpretation.

B.3.4 EXPERT AND PRACTITIONER CONSULTATION

Between June and November 2025, the research team engaged more than 75 stakeholders statewide and nationally. Consultations included:

- Public school district leaders
- Private school principals and association representatives (including faith-based networks)
- Legislators from both chambers and parties
- NCDPI and NCSEAA officials
- National psychometricians and content experts
- Advocates and policy organizations representing a wide range of perspectives

Most of these engagements were conducted through semi-structured interviews, technical briefings, expert advisory discussions, and targeted follow-ups as findings emerged.

Stakeholder input was essential for understanding:

- Instructional practices and pacing
- Existing assessment system reliance (including teacher training investments)
- Transition cost and administrative feasibility
- Perceptions of comparability and reporting needs
- Requirements for maintaining EVAAS compatibility

This input shaped not only technical judgments but also the feasibility of various recommendations and the identification of trade-offs.

B.4. OVERVIEW OF METHODS BY RESEARCH STRAND

B.4.1 STRAND 1: COMPARATIVE ANALYSIS OF ASSESSMENTS IN SCHOOL CHOICE PROGRAMS

The first strand examined assessment requirements in school choice programs nationally and analyzed current assessment practices in North Carolina's public and private sectors. For each assessment, this included:

- Documenting assessments used by states as well as public and private schools in North Carolina
- Cataloging the reporting mandates
- Understanding sector-specific priorities such as public transparency vs. instructional autonomy, test burden, etc.

This descriptive work established the scale and variation in North Carolina's current testing landscape and clarified the policy conditions under which comparability could be achieved.

The research team conducted an extensive review of major, nationally norm-referenced assessments using a structured review matrix. The review matrix was designed to capture whether an assessment met the requirements of S.L. 2024-57, including if it (a) applies to grades 3 and 8, (b) provides through-grade assessment and summative scoring, and (c) offers valid evidence of alignment with North Carolina State Standards. In addition to these criteria, an assessment is identified as not meeting all the requirements if it: (d) is teacher-developed, lacks national norms, or is primarily a screener rather than a comprehensive achievement measure; (e) is intended to support instruction within a specific scope or curriculum sequence instead of serving as a stand-alone measure of student mastery; (f) is designed mainly for instructional diagnostics or as a predictor of another summative assessment; or (g) is a diagnostic or adaptive test aimed at predicting performance on a different, non-norm-referenced assessment.

The review excluded information from technical manuals, training resources, or user guides published by assessment firms. While these resources may provide details on administration and technical properties, their quality and completeness vary widely, making an objective, comprehensive review unfeasible within this study's scope. Instead, the study relied solely on publicly available and verifiable information about each assessment's design, purpose, and technical features.

B.4.2 STRAND 2: ALIGNMENT EVALUATION

To address statutory directives regarding alignment, the team reviewed:

- Existing alignment studies for nationally norm-referenced assessments
- Methodologies used by external vendors (including Webb alignment criteria)
- Experts' expectations for independent alignment evidence
- Timelines and process for procuring a formal alignment study, if needed

The psychometricians we engaged confirmed that alignment is essential to interpretability and validity if an assessment is meant to reflect the SCOS. Where no external alignment evidence exists, an independent alignment study would be necessary before large-scale implementation.

B.4.3 STRAND 3: THROUGH-GRADE FEASIBILITY ASSESSMENT

The third strand evaluated the technical and instructional feasibility of replacing EOGs and EOCs with a through-grade assessment system, drawing on:

- North Carolina's current assessment landscape
- Federal ESSA requirements
- Field standards for educational and psychological testing
- Interviews with psychometric experts and SAS (regarding EVAAS implications)

B.5 LIMITATIONS

Findings should be interpreted with the following boundaries:

- The authors did not attempt to equate across nationally norm-referenced assessments. Leading experts confirmed that such approaches remain on the research frontier and are not yet ready to serve as a foundation for statewide policy.
- Private school historical data are incomplete due to a lack of standardized identifiers and incomplete participation within the North Carolina Division of Non-Public Education Private School Annual Report, limiting any retrospective trend analyses.
- Stakeholder perspectives were diverse and informed the identification of both policy opportunities and operational constraints.
- Assessment characteristics (Table 3) are based on publicly available data, some of which were limited. While the research team sent requests for additional information to test vendors, not all requests were answered.
- Public school district local testing information relies on self-reported data from local school districts and includes some inconsistencies between districts due to how test types are operationalized within the database.

APPENDIX C: STATE-BY-STATE COMPARISON OF K-12 CHOICE PROGRAMS

School Choice Program Type(s)	Testing Requirement(s)	Eligibility Criteria	Choice Students (latest year)
North Carolina			
School vouchers (Opportunity Scholarship); Education Savings Account (ESA)	Voucher students must take a nationally norm-referenced standardized test each year. (No state exam requirement)	Opportunity Scholarships are available to all K-12 students statewide; if applications exceed the budget cap, renewals and lower-income applicants receive priority. (ESA program is limited to students with disabilities, requiring an IEP).	104,073 (2025-26)
Alabama			
ESA (CHOOSE Act 2024); Tax-credit scholarship; Refundable tax credit (failing-school transfer)	Participants must take either the state assessment or a nationally norm-referenced test (students with significant disabilities may be exempt).	CHOOSE ESA: Alabama residents ages 5-19 with household income $\leq$ 300 percent FPL (phasing to universal by 2027). Tax-credit Scholarships: Income $\leq$ 250 percent FPL initially; priority to students zoned to low-performing public schools.	22,466 (2025-26)
Florida			
ESA (Family Empowerment Scholarship for Unique Abilities); Tax-credit ESAs (Florida Tax Credit Scholarship)	Scholarship students must take either Florida's state assessments or a nationally norm-referenced test.	Universal eligibility for all K-12 students (as of 2023); priority to low-income and foster students.	502,705 (2025-26)
Georgia			
School voucher (Special Needs Scholarship); Tax-credit scholarship (Qualified Education Expense Tax Credit)	Voucher schools must administer annual pre- and post-assessments; no state test required.	Special Needs Voucher: For students with IEP/504 plans; one year in public school required. Tax-credit Scholarship: For students from families meeting income limits set by SGOs.	36,255 (2023-24)
Mississippi			
School vouchers (Dyslexia Therapy; Nate Rogers); ESA (Equal Opportunity for Students with Special Needs)	No state testing mandate for these programs.	Dyslexia & Nate Rogers Scholarships: For students diagnosed with eligible disabilities; public-school attendance typically required. ESA: For students with special needs and active IEPs.	728 total (2025-26)
Ohio			
School vouchers (EdChoice, EdChoice Expansion, Cleveland, Autism, Jon Peterson)	Students must take all state standardized assessments for their grade.	Universal eligibility (as of 2023) for EdChoice Expansion; specialized programs for students with IEPs or autism.	166,563 (2024-25)

School Choice Program Type(s)	Testing Requirement(s)	Eligibility Criteria	Choice Students (latest year)
South Carolina			
ESA (Education Scholarship Trust Fund, 2024); Tax-credit scholarship (Exceptional Needs Children)	ESA students must take state or nationally norm-referenced assessments; tax-credit scholarship schools must also test and report aggregate results.	ESA phased rollout by income (≤ 200 percent FPL in 2024–25, up to 400 percent by 2026–27). Exceptional Needs program for students with disabilities.	11,322 (2025–26)
Tennessee			
ESAs (Education Savings Account pilot; Education Freedom Scholarship 2025); IEA for special needs	ESA students must take the Tennessee state achievement assessments; IEA students take state or national assessments (grades 3–8).	ESA pilot limited to Memphis/Nashville/Chattanooga; income ≤ 200 percent FRL. IEA for students with IEPs; 2025 ESA expansion will be universal.	25,795 (2025–26)
Virginia			
Tax-credit scholarships (Education Improvement Scholarships Tax Credits)	Private schools must administer a nationally norm-referenced achievement test yearly.	Students from families ≤ 300 percent FPL (400 percent for special needs); must be entering kindergarten/1st or transferring from public school.	5,820 (2023–24)
Indiana			
School vouchers (Choice Scholarship); Tax-credit scholarships (SGO); Individual tax deduction	All voucher students must take Indiana's statewide assessments.	Universal eligibility for all K–12 students (as of 2024–25); prior income limits removed. SGO scholarships remain income-capped at 300 percent FRL.	93,131 (2025–26)
Arkansas			
ESA (Educational Freedom Account); School voucher (Succeed Scholarship – closed/absorbed)	ESA students must take Arkansas state assessments or approved nationally norm-referenced alternatives.	ESA phasing to universal by 2025–26; initial priority for students with disabilities, foster, military, or in F-rated schools.	46,987 (2025–26)
Oklahoma			
School voucher (Lindsey Nicole Henry); Tax-credit scholarships; Refundable tax credit (2024)	No state testing requirements for private participants.	LNH: Students with disabilities and IEPs. Tax-credit and refundable credits available to families (refundable credit up to \$5,000 per student).	4,478 (2022–23)
Texas			
Education Freedom Accounts (ESA, begins 2026)	Participating students in grades 3–12 must take a nationally norm-referenced standardized test each year. (No state exam requirement)	Universal ESA program effective 2026–27; all K–12 students eligible, up to \$10,474 per student.	No current program

Sources: North Carolina Department of Public Instruction's 2024–25 Opportunity Scholarship Report accessed from <https://www.nc.gov/dpi/reports/opportunity-scholarship-2025>; Alabama Department of Revenue's CHOOSE Act February 2025 press release accessed from <https://www.revenue.alabama.gov/choose-act>; Florida Department of Education's school choice announcements accessed from <https://www.fldoe.org/schools/school-choice>; Step Up for Students private school application information accessed from <https://www.stepupforstudents.org>; the Georgia Governor's Office of Student Achievement's summary of the Georgia Special Needs Scholarship Report accessed from <https://gosa.georgia.gov/georgia-special-needs-scholarship-program>; EdChoice's data on school choice programs accessed from <https://www.edchoice.org/school-choice/programs>; Ohio Department of Education's 2024–25 voucher program statistics accessed from <https://education.ohio.gov>; South Carolina Department of Education ESA Program description accessed from <https://ed.sc.gov/esa>; Exceptional SC's summary accessed from <https://www.exceptionalsc.org/>; Tennessee Department of Education's EDA and IEA program overview accessed from <https://www.tn.gov/education/lea-operations/esa.html>; Virginia Department of Education's overview of scholarship tax credits accessed from <https://www.doe.virginia.gov/parents-students/private-schools/scholarship-tax-credits>; Indiana Department of Education's 2024–25 choice report accessed from <https://www.in.gov/doe/choice/> and Indiana Capital Chronicle: <https://indianacapitalchronicle.com>; Arkansas Department of Education's Education Freedom Account report accessed from <https://dese.ade.arkansas.gov>; Oklahoma Council of Public Affairs data accessed from <https://www.ocpathink.org>; and Reuters coverage of Texas' ESA legislation dated May 3, 2025 accessed from <https://www.reuters.com/business/finance/texas-governor-sign-largest-us-school-voucher-law-marking-conservative-shift-2025-05-03/>.

APPENDIX D: DATA AND INFORMATION SOURCES FOR TABLE 3, CHARACTERISTICS OF NATIONALLY-NORMED ASSESSMENTS

D.1 i-READY (STANDARDS OF MASTERY)

Information on the i-Ready assessment was obtained from the following sources:

- Curriculum Associates i-Ready Assessment Components Diagnostics Flyer accessed from <https://cdn.bfldr.com/LS6J0F7/at/sbfx3qwq6wgkp5t5h6j5tp9/iready-assessment-components-diagnostic-flyer.pdf>
- Curriculum Associates overview of the i-Ready accessed from <https://www.curriculumassociates.com/reviews/ireadyaccessibility/i-ready-assessment> and <https://www.curriculumassociates.com/research-and-efficacy/nc-eog-linking-study>
- Curriculum Associates i-Ready Diagnostics National Norms Tables for Reading and Mathematics for Grades K-8, 2021-22 accessed from <https://www.fldoe.org/core/fileparse.php/7539/urlt/iready-norms-tables-K-8-2020.pdf>
- National Center on Intensive Intervention at the American Institutes for Research accessed from <https://charts.intensiveintervention.org/progressmonitoring/tool/?id=bf6c3db4458b6bbb>, <https://charts.intensiveintervention.org/progressmonitoring/tool/?id=4c9f5945c8c5dcc7>

D.2 NWEA MAP

Information on the NWEA MAP assessment was obtained from the following sources:

- NWEA's MAP Growth accessed from <https://www.nwea.org/map-growth/>
- NWEA's MAP primer for parents/guardians accessed from <https://www.nwea.org/blog/2024/map-growth-101-everything-families-need-to-know/>
- NWEA's Frequently Asked Questions on Accessibility and Accommodations in MAP Growth accessed from <https://www.nwea.org/uploads/2019/12/NWEA-Accessibility-and-Accommodations-FAQ-JAN2020.pdf>

- NWEA's 2025 Norms Quick Reference Fact Sheet accessed from https://www.nwea.org/resource-center/fact-sheet/87992/MAP-Growth-2025-norms-quick-reference_NWEA_onesheet.pdf/
- NWEA Psychometrics and Analytics 2025 report on Predicting Performance on the North Carolina End-of-Grade (NC EOG) Assessments Based on NWEA MAP Growth Scores accessed from https://www.nwea.org/uploads/NC-MAP-Growth-Linking-Study-Report_EOG_2025.pdf
- National Center on Intensive Intervention at the American Institutes for Research summary of MAP accessed from <https://charts.intensiveintervention.org/screening/tool/?id=576cd73956493b98>
- ACSI Member Pricing document accessed from <https://acsipdp.s3.amazonaws.com/Assessment/MAP+ACSI+Member+Pricing+2021-2022.pdf>

D.3 RENAISSANCE STAR

Information on Renaissance Star was obtained from the following sources:

- Renaissance's Consolidated Summary Report accessed from https://support.renaissance.com/s/article/Consolidated-Summary-Report-1752674158578?language=en_US
- Renaissance's Star Annual Progress Report accessed from https://support.renaissance.com/s/article/Star-Annual-Progress-Report-1752674095144?language=en_US
- Renaissance's 2022 product update summary accessed from https://www.renaissance.com/product_update/an-easier-way-to-enable-accommodations-in-star-assessments/
- Renaissance's 2024-25 summary update on Star assessment national norms accessed from https://support.renaissance.com/s/article/Updated-Star-Assessment-National-Norms-1752673776142?language=en_US
- Renaissance's technical working paper linking Star reading and math to the North Carolina EOGs (2025-26) accessed from <https://renaissance.widen.net/view/pdf/vvkx0vtlmg/R45814.pdf?t.download=true&u=zceria>
- Pricing quote from Gridley Unified School District (CA) accessed from <https://www.gusd.org/documents/Agendas%20and%20Minutes/2021-22/August%2018%202021%20Regular/Renaissance-Quote-2021-2627597.pdf>
- Pricing quote accessed from the Santa Barbara Unified School District (CA) accessed from [https://go.boarddocs.com/ca/sbunified/Board.nsf/files/D6EPFA645010/\\$file/Renaissance%20Learning%2C%20Inc.pdf](https://go.boarddocs.com/ca/sbunified/Board.nsf/files/D6EPFA645010/$file/Renaissance%20Learning%2C%20Inc.pdf)
- Pricing quote accessed from the Wayne County Public Schools (NC) accessed from [https://go.boarddocs.com/nc/wcpsnc/Board.nsf/files/D6JQRQ6A97A6/\\$file/Renaissance.pdf](https://go.boarddocs.com/nc/wcpsnc/Board.nsf/files/D6JQRQ6A97A6/$file/Renaissance.pdf)

D.4 CALIFORNIA ACHIEVEMENT TEST (CAT 6)

Information on the California Achievement Test (CAT 6) was obtained from the following sources:

- Seton Testing Services overview of the CAT 6 accessed from <https://www.setontesting.com/product/terranova/>
- Test costs for the CAT 6, Stanford, and TerraNova NEXT can be obtained from third-party brokers who provide these assessments to homeschool families. See, for example, information reported by the Family Learning Organization, <https://www.familylearning.org/testing/terranova-2>.
- A TerraNova NEXT salesperson provided test costs reported in Table 3.

D.5 CTP (ERB)

Information on the CTB (ERB) was obtained from the following sources:

- ERB's overview of the CTB accessed from <https://www.erblearn.org/educators/comprehensive-assessments/ctps/>
- ERB's overview of the testing season accessed from <https://www.erblearn.org/educators/comprehensive-assessments/testing-seasons/>
- ERB's milestone assessment FAQ document accessed from https://cdn.erblearn.org/pdf/www/20200611_ERB_Milestones_FAQ_Extended.pdf
- ERB's blog update on the CTPs norms update accessed from <https://www.erblearn.org/blog/spring-2024-ctp-norms-updates/>
- The research team received the requested cost information from the publisher, but this query was not returned.

D.6 IOWA TEST OF BASIC SKILLS (ITBS)

Information on the ITBS was obtained from the following sources:

- Seton Testing's product information form accessed from <https://www.setontesting.com/product/iowa-assessments-form-e/>
- National Center on Intensive Intervention at the American Institutes for Research summary of the Iowa assessments accessed from <https://charts.intensiveintervention.org/screening/tool/?id=5864bd42688bcf45>
- Iowa Assessment Accommodations document accessed from <https://itp.education.uiowa.edu/documents/Accommodations-for-the-Iowa-Assessments.pdf>
- Riverside Insights summary of ITBS' new norms accessed from <https://blog.riversideinsights.com/iowa-assessments-new-norms-2025>

D.7 STANFORD ACHIEVEMENT TEST

Information on the Stanford Achievement Test was obtained from the following sources:

- Pearson Assessment's scope and sequence summary accessed from <https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/sat10/sat10-scope-sequence.pdf>
- Pearson Assessment's assessment report on accommodations on Stanford 10 for students with disabilities accessed from <https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/sat10/sat10-accommodations-for-student-disabilities.pdf>
- Seaton Testing Services overview of the Stanford 10 Online accessed from <https://www.setontesting.com/product/stanford-10-online/>

D.8 TERRANOVA NEXT

Information on the TerraNova NEXT was obtained from the following sources:

- Terra Nova NEXT scope accessed from <https://terranovanext.com/terranova-next-2/>
- Information about the formative nature of the DRC Beacon and its alignment with TerraNova Next accessed from https://terranovanext.com/PDFs/Assessment_Accommodations_Supplement.pdf
- Norming information of the TerraNova NEXT assessment was accessed from two sources. This first describes that TerraNova NEXT uses the same norms as TerraNova3. https://terranovanext.com/wp-content/uploads/2024/02/TNN_Broch_020124.pdf. The second source describes the last TerraNova3 norming, which occurred in 2017: <https://terranovanext.com/terranova-complete-battery/>.

D.9 NORTH CAROLINA EOGs AND EOCs

Information on the EOGs and EOCs was obtained from the following sources:

- North Carolina Department of Public Instruction's summary of the NC Check-Ins 2.0 accessed from <https://www.dpi.nc.gov/districts-schools/accountability-and-testing/state-tests/nc-check-ins-20>
- North Carolina Department of Public Instruction's summary of testing accommodations accessed from <https://www.dpi.nc.gov/districts-schools/accountability-and-testing/testing-policy-and-operations/testing-accommodations>

- EdMetrics’s 2023 alignment study for the EOGs and EOCs assessments accessed from <https://www.dpi.nc.gov/documents/accountability/testing/technotes/technical-report-alignment-study-nc-eog-and-eoc-assessments/open>
- Costs are based on authors’ calculations using annualized EOGs and EOCs costs of \$9,432,117 per year, which are administered to approximately 1.05 million students in third through eleventh grades in North Carolina.

ENDNOTES

- 1 The state's Division of Non-Public Education reported 135,738 students enrolled in North Carolina private schools in the 2024-25 school year. Carolina Demography projects a 2025-26 private school enrollment between 145,000 and 170,000, which suggests between 61 and 71 percent of current 2025-26 private school students receive an Opportunity Scholarship.
- 2 A "public school unit" in North Carolina is a traditional public school, charter school, regional school, lab school, or school for the deaf and blind (Chapter 115C, Section 5 of the North Carolina General Statutes). Throughout the remainder of the report use of "public schools" pertains to these public school units.
- 3 EVAAS, developed by SAS Institute, is a group of tools that uses statistical modeling to measure student academic growth by analyzing individual test score data over multiple years. It estimates educators' and schools' contributions to student learning by modeling historical performance and predicting future outcomes. These growth metrics help inform teacher evaluations, guide instructional practices, and support accountability systems.
- 4 NCSEAA is the state agency that administers the OS Program.
- 5 OS are paid out twice annually – once in the fall semester and once in the spring semester. In December 2025, the NCSEAA reported fall semester Opportunity Scholarship expenditures of \$292,069,545. If spring funding remains the same, the total 2025-26 OS Program funding would be \$584,139,090.
- 6 Appendix C provides a 13-state summary of program types, eligibility criteria, and approaches to assessing student and school performance.
- 7 This report references both EOGs and EOCs because some middle school students take advanced courses that require EOCs. However, through-grade assessments are less common at the high school level, where EOCs are most frequently administered in North Carolina.
- 8 In 2012, Chingos (2012) reported that North Carolina's EOGs and EOCs assessments had the 3rd-lowest per-pupil costs in the United States, reporting an annual expenditure of \$8,969,794 in the 2009-10 school year. In 2023, the state signed a five-year contract for EOGs and EOCs at an annualized cost of \$9,432,117.
- 9 Assessments differ in many ways, including construct coverage, item formats, use of adaptive versus fixed forms, scaling and norming models, and reporting metrics. These variations collectively can undermine the validity of equating scores across different assessments. For more information, see, for example, National Research Council (1999).
- 10 Given the dynamic nature of these criteria, the development and implementation of a standardized annual review process would provide assessment companies with an opportunity to present evidence that meets these criteria or an appealing evaluation of their candidacy for inclusion. The research team proposes that NCSSEA lead and manage this process.
- 11 Assessment pricing can vary widely across vendors and educational settings due to factors such as development and maintenance costs, the level of customization and alignment with state standards and curriculum, the unique needs of a district or state, the scale and volume of testing, legal compliance requirements, and other considerations. To supplement publicly available cost information, the research team also sent informational requests to test vendors, although some did not reply.
- 12 For a comprehensive and thoughtful review of the topic, see work by Reardon, Kalogrides, and Ho (2021) published in the *Journal of Educational and Behavioral Statistics*, along with published commentaries by Bolt (2021), Davison (2021), Moses and Dorans (2021), and von Davier (2021). Ho, Reardon, and Kalogrides (2021) also wrote a rejoinder to the commentaries.
- 13 An example of a recent alignment study in North Carolina is Egan, K., Davidson, A., & Rabinowitz, S. (2023). *Alignment Study for North Carolina End-of-Grade and End-of-Course Assessments in Grades 3-8 and High School*. Retrieved from <https://www.dpi.nc.gov/documents/accountability/testing/technotes/technical-report-alignment-study-nc-eog-and-eoc-assessments/open>.
- 14 A "public school unit" in North Carolina is a traditional public school, charter school, regional school, lab school, or school for the deaf and blind (Chapter 115C, Section 5 of the North Carolina General Statutes). Throughout the remainder of the report use of "public schools" pertains to these public school units.

