

# Inclusive, Innovative Assessments

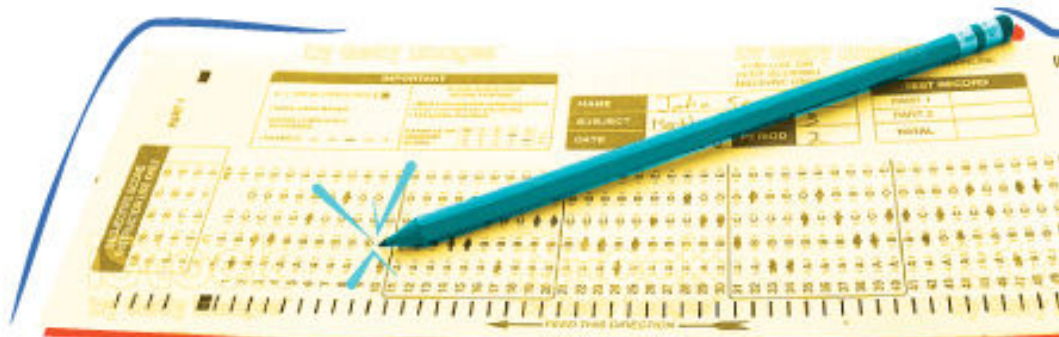
for Students With  
Learning Disabilities

*Lindsay Kubatzky & Meg Benner*



# Table of Contents

|       |  |    |
|-------|--|----|
| I.    | Introduction and Summary . . . . .   | 3  |
| II.   | The Road to Equity in and Access to Assessment . . . . .                             | 4  |
| III.  | Public Perceptions of Current Standardized Assessments . . . . .                     | 7  |
|       | 1. Growing Concerns Over Standardized Testing and the COVID-19<br>Pandemic. . . . .  | 7  |
|       | 2. Perceptions of Educators, Caregivers, and Students With<br>Disabilities . . . . . | 8  |
| IV.   | Growing Momentum to Adapt and Innovate Assessments. . . . .                          | 14 |
|       | 1. A Federal Push for Innovation in Assessment . . . . .                             | 14 |
|       | 2. Core Principles for Innovative Assessment. . . . .                                | 17 |
|       | 3. New Approaches to Assessment . . . . .  | 19 |
| V.    | Policy Recommendations . . . . .   | 22 |
| VI.   | Conclusion. . . . .  | 24 |
| VII.  | Acknowledgments. . . . .   | 25 |
| VIII. | Endnotes . . . . .   | 26 |





## Introduction and Summary

**The COVID-19 pandemic sparked a national conversation about the role of public education, what should be taught, and what makes a high-quality education.**

The mental, emotional, and physical health of students was a pillar in the debate between remote or in-person schooling. Remote learning also disrupted instruction and assessment. As a result, measuring and addressing the “instructional loss” due to the pandemic became a priority for families, educators, and policymakers. The 2022 National Assessment of Educational Progress (NAEP) results, often referred to as the Nation’s Report Card, showed a [dramatic drop](#) in proficiency scores across the board – but especially for students who are historically marginalized, such as those with disabilities.

The debate about how we measure student progress and how schools are supporting students is not a new one. The National Center for Learning Disabilities (NCLD), along with our disability and civil rights partners, have long advocated for including the performance of students with disabilities, students of color, English learners, and students impacted by poverty on statewide summative assessments when determining how well a school is meeting the needs of students. For our communities, these assessments have historically been viewed as a tool to identify opportunity gaps as they provide annual, comparative data on student progress. For others, though, the assessments are seen as taking away too much valuable instructional time and not providing actionable information. Because of the ongoing tension around the value of summative assessments, NCLD sought to discover what’s working and not working for students with disabilities in the current assessment system and to forge a path forward that’s more inclusive and equitable.

NCLD surveyed and conducted focus groups with educators, caregivers, and students to understand their perceptions toward statewide summative assessments. In addition, NCLD interviewed various assessment and disability rights experts to identify trends in innovative assessments as well as the benefits and risks for all learners – especially students with disabilities.

This paper includes principles that policymakers and assessment developers should consider when creating new assessments, an overview of current proposals, and policy recommendations to realize equity within assessment systems for all learners.



# The Road to Equity in and Access to Assessment

**D**ue to hard-fought advocacy and many reauthorizations of the Individuals with Disabilities Education Act (IDEA) and the Elementary and Secondary Education Act (ESEA), [federal law](#) requires that all students take assessments to measure performance in math and English language arts (ELA).<sup>1</sup> Students must be assessed yearly in grades 3–8 and once in high school.<sup>2</sup> These standardized summative assessments and the federal requirement to report performance by student subgroups ([see Box 1](#)) shed light on

## Box 1

### Current Federal Subgroup Reporting Requirements

The [Every Student Succeeds Act](#) — the most recent reauthorization of the Elementary and Secondary Education Act — requires that schools report academic performance and graduation rates for the following student subgroups:

- Students impacted by poverty
- Students from major racial and ethnic groups
- Students with disabilities
- Students who are English learners
- Students experiencing homelessness
- Students in foster care
- Students with a parent who is a member of the Armed Forces

States are not required to disaggregate data — or report data for certain subgroups — if the number of students in any subgroup is insufficient to produce statistically reliable information or if reporting may reveal individual student data. States may set the reporting requirements for disaggregation.

differences in performance among demographic profiles and allow stakeholders to compare how schools support various student subgroups. There are many types of assessments that serve many purposes ([see Box 2](#)). The tests we discuss here typically take place toward the end of the school year.

Congress didn't pass these requirements in one fell swoop. Rather, the current requirements for standardized summative assessments were slowly layered upon each other.

These requirements have been put in place to ensure that all students receive a high-quality education and to prepare all students for college, career, and beyond.

These data can be powerful in the aggregate to hold schools accountable and drive resources



to schools in greatest need.

Before these requirements were in place, some student subgroups – especially students with disabilities – were excluded from grade-level curriculum and, in some instances, their academic performance was hidden from families and decision makers.

Scores for individual students also can be important for families. Caregivers can compare their child's individual performance on an assessment to grade-level expectations and to their peers. Families with a child who has a disability can use these data to help drive conversations about needed services and accommodations to help a child meet grade-level expectations.

While not required by federal law, some states link scores on statewide summative assessments to high-stakes decisions for individual students, including [grade retention](#) and high school graduation. This creates unnecessary stress on the students being assessed and can lead to negative consequences for individual students.

### Box 2

## Standardized vs. Formative Assessment

There are many different types of assessment. This report focuses on standardized versus formative assessments.

Standardized assessments are uniform and, generally, used for accountability or policymaking purposes – to inform instruction, guide allocation of resources, and hold schools accountable for their support of subgroups and individual students. This report focuses on state standardized, summative assessments.

Formative assessments can be more informal and monitor student learning so educators can inform their instruction.

A comprehensive assessment system should include assessments of many different types, including summative assessments, diagnostic, interim, formative, etc. Different assessments can serve different purposes, and each can be a smaller piece of a comprehensive assessment system. When selecting and administering the assessment, it's important to recognize the primary goal of each assessment. This will influence the way the assessment should be developed, executed, and translated for stakeholders.

## Power of Assessment to Drive Learning

Assessment helps measure performance in order to inform future decisions, but a [growing body of research](#) demonstrates that tests also help students learn. [Retrieval](#), or the act of recalling previously learned information, is one of the most powerful tools for learning. Testing can further embed concepts into long-term memory. This is often referred to as the “testing effect.”



# Alternate Assessments for Students With the Most Significant Cognitive Disabilities

**IDEA requires that all children with disabilities be included in all state- and district-wide assessment administrations. Most students with disabilities take the general assessment with or without accommodations. A small number of students with the most significant cognitive disabilities need an alternate assessment.**

The 2015 reauthorization of ESEA includes a 1% cap on the total number of students in the state participating in the alternate assessment based on alternate academic achievement standards (AA-AAAS). This requirement underscores that only a very small percentage of students should be assessed on alternate academic achievement standards. AA-AAAS must be aligned to alternate standards based on the state's challenging academic content standards for the enrolled grade (which are the same for all students) and must promote the involvement and progress of students with the most significant cognitive disabilities in the general education curriculum.

Despite this, many states have more than 1% of students taking an AA-AAAS. Nearly half of all states — 23 states in the 2021–22 school year — continue to request waivers to exceed the 1% cap. In some of those states, a disproportionate number of students of color are taking the alternate assessment, suggesting a clear need for revised state participation policies, additional training for Individualized Education Program (IEP) team members on assessment participation decision making, and monitoring of participation decisions by both the state and federal government.

There are both short- and long-term consequences for students when they take alternate assessments rather than the same assessments as their peers. Research shows the benefits of inclusion for all students with disabilities, including those with the most significant cognitive disabilities. Only 3% of students who take an AA-AAAS are educated in the general education classroom 80% or more of the time, as compared to 64.8% of all students with disabilities. Long-term consequences include access to and possible effects on high school graduation (including diploma received), postsecondary training, and employment opportunities.

For other assessments, many test developers generally have not created aligned interim alternative assessments. As a result, students with the most significant cognitive disabilities have historically often been excluded from participation, even though federal laws require their participation in all state- and district-wide assessment administrations. More research, professional development, and funding are needed to ensure that states and districts have the resources to assess all students, including those with the most significant cognitive disabilities.





# Public Perceptions of Current Standardized Assessments

## Growing Concerns Over Standardized Testing and the COVID-19 Pandemic

**D**espite the benefits of federal assessment requirements, including promoting inclusion of students with disabilities, the assessment requirements in federal law have challenges and critics. The requirement in the recent reauthorization of ESEA is intended as [one measure](#) of school quality and represents just one prong of several required criteria in a state and district accountability system. Yet some [families](#) and [many educators](#) criticize statewide summative assessments for detracting from instructional time or narrowing curriculum. The assessments require anywhere from hours to days to administer — a period that may be even longer for students with accommodations such as extended time. Pointedly, the results of the assessments often are not available until the following school year. As a result, educators may be unable to use the results of the assessments to drive instruction for the students in their classrooms because by the time educators get data, they have a new group of students who may perform differently or require different pedagogical approaches.

In addition, while the quality of many statewide summative assessments has significantly improved in the past 20 years,<sup>3</sup> experts and educators say that these assessments do not adequately measure the most critical higher-order thinking skills. Most statewide summative assessments default to a majority of multiple-choice questions that are more efficient to score but that range in quality.

There have been significant efforts to reduce bias built into standardized assessments, but there are lingering concerns about current statewide summative assessments.<sup>4</sup> Historically, some assessment items have included language or contexts that are unfamiliar to certain student groups, which limits a student's ability to answer the question correctly even if they have mastered the underlying concept. For example, a test question could include terms that are less familiar to students who are impacted by poverty than those from wealthier



backgrounds (see Box 3). A student from a well-resourced family would be more likely to be given an expensive telescope, so would be more likely to know the term. Moreover, test items could reinforce a negative stereotype about a race, culture, or gender, and can contribute to testing anxiety.<sup>5</sup>

## Perceptions of Educators, Caregivers, and Students With Disabilities

**To better understand attitudes toward assessments among educators and caregivers, NCLD conducted a survey of educators and caregivers of students with disabilities in the spring of 2022, followed by focus groups of young adults with disabilities and caregivers of students with disabilities in the summer of 2022.**

The survey showed that caregivers have a more positive perspective on assessments' ability to measure their child's academic proficiency, evaluate school quality, and support student learning compared to educators. Both groups believe accommodations, as provided to students with disabilities, are essential for students to demonstrate their full potential on assessments.

### Box 3

Suppose you wanted to determine if another planet had rivers or mountains on it.

*Which of the tools below would best help you find out?*

- A Camera
- B Microscope
- C Binoculars
- D Telescope

*Retrieved from "Assessment Bias: How to Banish It"*

## SURVEY METHODOLOGY

Using QuestionPro, NCLD's survey in the spring of 2022 collected 800 total responses — 400 responses from educators of students with disabilities and 400 responses from caregivers who have a child with a disability.

Participants came from 45 states. Educators and caregivers of students with disabilities were diverse in terms of race/ethnicity, income level, gender, age, and grade level represented.

The survey asked caregivers and educators to identify the value of standardized assessments, desired frequency, and how the results of the assessments should be used. The full survey is available upon request.



## Key Findings

The survey identified trends across the perception of educators and caregivers toward statewide summative assessments. Those include:

- Caregivers see more value than educators in statewide summative assessments as a measure of school quality and as data for educators and families to support student learning.** 73% of caregivers “agree” or “somewhat agree” that these assessments provide important information to compare their child’s performance in school (Fig. 1). Only 52% of educators believe the same (Fig. 2).

Figure 1

**Caregivers: “State assessments provide important information to determine how my child is performing in school.”**

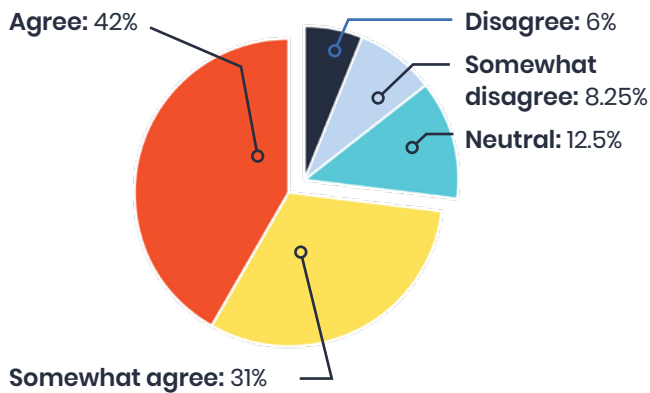
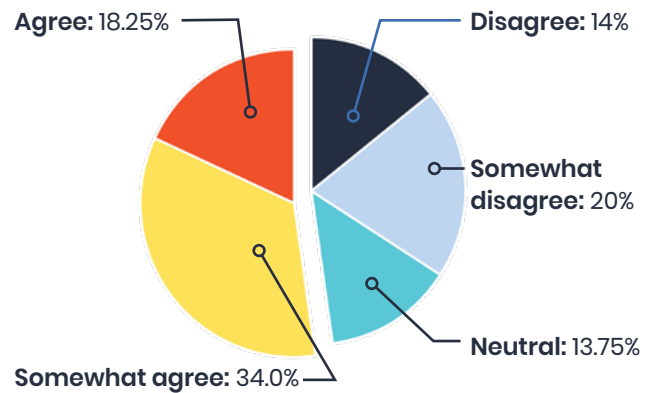


Figure 2

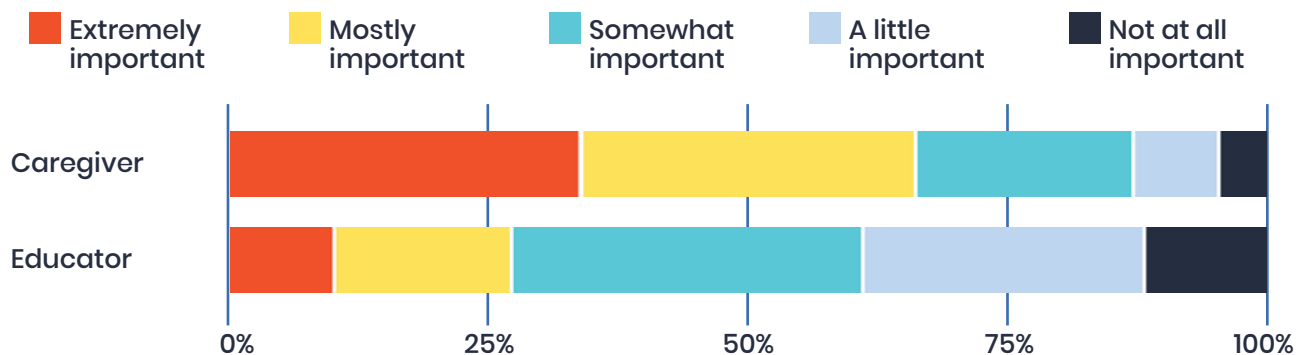
**Educators: “State assessments provide important information to determine how my students are performing in school.”**



The responses on a similar question were more divergent. Among caregivers, **66%** reported that these assessments are “extremely important” or “mostly important” to measure student success, while only **27%** of educators considered the same (Fig. 3). A greater percentage of caregivers as compared to educators also reported that assessments provide supportive data for families, educators, and school districts.

Figure 3

**State Test Scores as a Measure of Success**

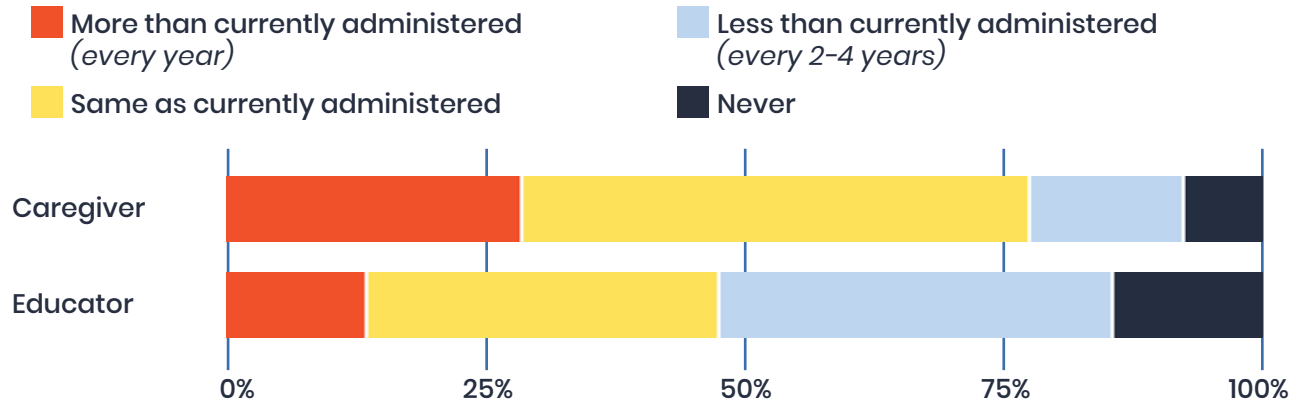


**2. Caregivers believe statewide summative assessments should be administered at least annually (as required by current law), but educators have mixed opinions.**

78% of caregivers reported believing that the federal government should request that states administer assessments at least once per year, compared to 48% of educators.

Figure 4

**Standardized Math and Literacy Test Administration Rates**



**3. Caregivers, more than educators, believe that disability status has an impact on a student’s attitude toward taking assessments.**

58% of caregivers reported that disability status “strongly” or “mostly” influences their children’s attitude toward assessment (Fig. 5), while only 29% of educators reported that disability status “strongly” or “mostly” influences their students’ attitude toward assessment (Fig. 6).

Figure 5

**Child’s Disability Status Influence on Test Taking Attitude – Caregivers**

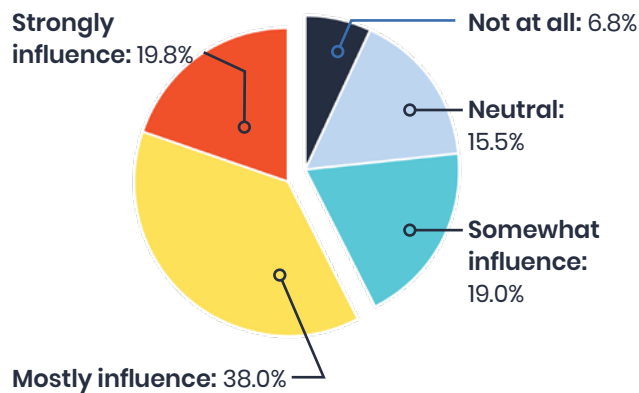
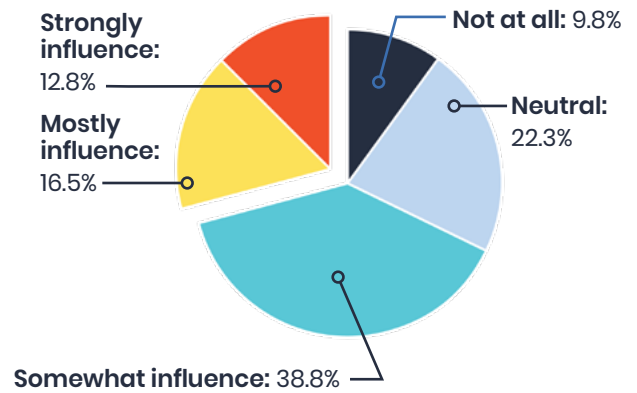


Figure 6

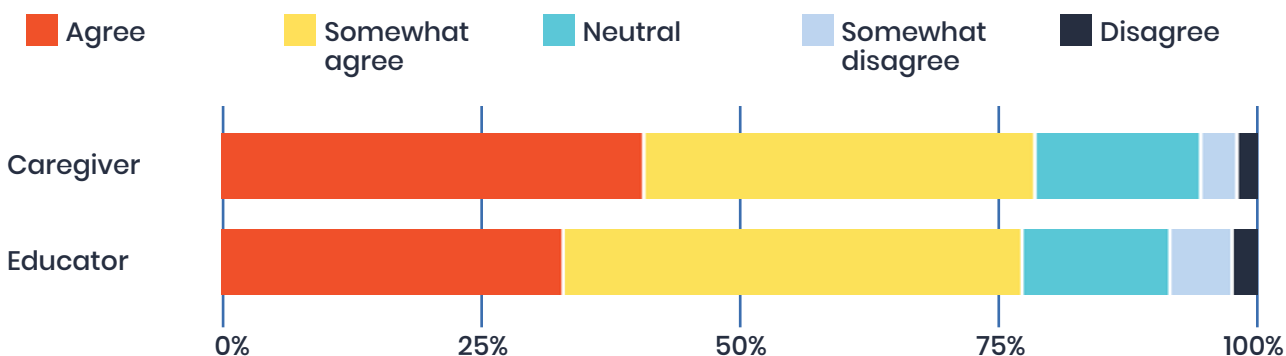
**Student’s Disability Status Influence on Test Taking Attitude – Educators**



**4. Caregivers and educators agree that accommodations are essential for children to show their potential on assessments. 79% of caregivers and 77% of educators reported they “agree” or “somewhat agree” that accommodations on assessments provide necessary support for children to fully access assessments.**

Figure 7

**Accommodations on State Assessments as Necessary Supports for Students to Show Potential**



These findings provide insights into the perspectives educators and families have on statewide summative assessments. Overall, families placed more value on them and were inclined to maintain or increase the frequency. Many factors might contribute to this. Specifically, educators are likely to understand that these assessments don’t necessarily have immediate implications for their pedagogical strategies, take away from instructional time, and may still impact their performance reviews. In addition, educators have daily experience that provides real-time, formative data on students. As a result, educators may feel that they’re in a better position to measure student proficiency. Families, on the other hand, may not receive frequent or consistent information about how well their child is doing in school and don’t have any context or ability to compare their child’s progress to that of their peers.

NCLD also conducted virtual 60-minute focus groups with young adults with disabilities and parents of students with disabilities. <sup>6</sup> Key trends from these focus groups include:

- **The purpose of statewide summative assessments needs to be better communicated to families.** Many parents and young adults broadly understood the purpose of standardized assessments to measure a person against others or against particular standards. But focus group participants were unaware that the assessment results are primarily used to hold schools accountable for student subgroup performance and to provide additional supports to schools when necessary. Parents and young adults held the misconception that a student’s scores could impact the individual student

(e.g., grade promotion). Therefore, parents and young adults suggested that schools be more explicit about the purpose of statewide summative assessments and how a student's individual score would impact the student, the teacher, and/or the school.

- ***Opinions about the reasons for test anxiety differ, perhaps due to experiences with accommodations.*** Given the lack of information from schools on the purpose and impact of statewide summative assessments, some parents noted that their student has anxiety around taking assessments, thinking it was high stakes for the individual. A majority of young adults also mentioned testing anxiety, but most said accommodations provided during the assessments reduced their anxiety. However, others shared that accommodations, such as additional time or a quieter location to take the assessment, were stigmatizing and sometimes embarrassing.
- ***Assessments should be used to evaluate schools but should not be the sole measure of school performance.*** All parents and young adults reported that measures other than statewide summative assessments are critical to fully capture how well schools serve students. Parents identified attendance, teacher qualifications, and teacher:student ratios as indicators they valued when assessing school quality. Young adults, on the other hand, suggested potential metrics of school climate, including forms of support such as access to counselors and opportunities to participate in advanced coursework.



- **Multiple opportunities to demonstrate progress — including assessments that assess recently taught standards — allow for more accurate depiction of student learning.**

Families and young adults highlighted that statewide summative assessments may not be the most effective way to measure learning. Instead, they preferred methods that allowed students to demonstrate knowledge on recently taught subject areas throughout the course of the year. Participants also identified the potential value of demonstrating mastery through multiple ways, such as an open-ended task.

### ***Limitations of the Findings***

It is important to note that a few elements of the methodology may limit the generalizability of these findings. First, the survey and focus groups had a small sample size and likely had population bias. Specifically, families, educators, and students who had more vocal views around assessment may have been more likely to participate. In addition, the survey and focus groups were limited to participants who could read English and had internet access.

For the survey, caregivers of color were slightly underrepresented. The survey also asked caregivers with multiple children to select one child and complete the survey with only that child in mind. This is important because different types of assessments are administered at various grade levels, which may impact attitudes toward assessment. But caregivers with multiple children in different grades or with various disability statuses could have conflated their experience across students.



For the focus groups, most of the young adult participants were over 18 and were asked to reflect back several years on their experiences in K–12. It's unclear how their opinions would differ from those of current public school students. Furthermore, three of the focus group participants were members of NCLD's Young Adult Leadership Council, which consists of young adults focused on advocating on behalf of their communities. These participants' answers may have differed from those of other young adults who did not have a prior relationship with NCLD. Finally, focus group size varied from one to five, which could have impacted how freely participants felt like they could answer.

# Growing Momentum to Adapt and Innovate Assessments

## A Federal Push for Innovation in Assessment

**A** groundswell of support for changes to federally required annual assessments to make them less time-consuming and more useful to students and educators has prompted policymakers, practitioners, and test developers to seek innovations in assessment. New models of assessment have spurred changes to assessment structures in states across the country. These innovative approaches must be designed and accessible to all learners, including students with disabilities and English learners. Historically, alternate assessments to support inclusion of students with the most significant cognitive disabilities also have not been included in these innovations, but must be.

Two federal programs, the Innovation Assessment Demonstration Authority (IADA) and Competitive Grants for State Assessments (CGSA), seek to cultivate this innovation. In addition, various states, funded through current budgets or private funding, continue to refine their assessment policies and tools to better support student learning.

### *Innovative Assessment Demonstration Authority (IADA)*

The recent reauthorization of ESEA, called the Every Student Succeeds Act (ESSA), included the [Innovative Assessment Demonstration Authority](#) to allow up to seven state education agencies (SEAs) or consortium of SEAs to pilot different assessment designs or structures in a smaller scale for select districts over a five-year trial period. Innovative designs had to meet certain requirements:

- **Assessment quality.** The system must include high-quality assessments that create valid and reliable measures.
- **Comparability.** The pilot design must produce state-level determinations of student proficiency that are comparable to statewide assessment results.
- **Scaling statewide.** The SEA must have a logical, feasible plan to scale the innovative method statewide after the demonstration period.
- **Demographically representative.** The selected district must have student populations representative of the state population.



The IADA program does not require states to “double test” the students in the pilot. In other words, states do not need to administer the traditional statewide summative assessment to the students in the pilot as long as they can demonstrate that the pilot assessment is comparable to the traditional assessment. Demonstrating comparability is difficult to do, however, without double testing. As a result, most states double test to demonstrate comparability. Only three states – Georgia, North Carolina, and Louisiana – are participating in IADA as of September 2022. One additional state, New Hampshire, participated in IADA but withdrew from the program in 2020. The IADA program also does not provide additional funding to states. While the intent of the IADA program is positive, many have criticized the program because of the lack of additional funding to conduct the pilot and the stringent requirements.

## Implications of COVID-19 School Disruptions on Assessment

**School closures related to the COVID-19 pandemic disrupted traditional standardized assessment processes and grew demand for new types of assessment.** In the 2019–20 school year, schools stopped in-person learning for much – if not all – of the spring. As a result, most schools could not administer planned summative assessments that are required for school accountability systems.

In response, the U.S. Department of Education allowed states to apply for [waivers](#) for multiple elements of ESEA, including the annual requirement to administer and report student proficiency. All [50 states, the District of Columbia, Puerto Rico, and other territories](#) requested and received a waiver in 2020, but only one, the District of Columbia, received a waiver in 2021. Despite this, though, states and districts varied widely in their participation rates, with some assessing fewer than 10% of students.

Given the inability to obtain the summative assessment scores and the unprecedented [instructional loss](#) as a result of pandemic-related school disruptions, states and districts requested new methods to assess proficiency in the 2020–21 school year. Policymakers, practitioners, and families were eager to identify unfinished learnings and adapt instruction to meet student needs.

Test developers offered solutions, including interim assessments tied to summative expectations. Many of these assessments, however, were not designed to support all learners, as accommodations to support students with disabilities were not validated with the new assessments.



### ***Competitive Grants for State Assessments (CGSA)***

ESSA also created the Competitive Grants for State Assessment (CGSA) program to improve state assessments that measure student performance. CGSA is a competitive grant program that offers states funding to improve current assessments or to add new components. It does not waive any testing or reporting requirements under ESSA.

More states have participated in CGSA compared to those that have participated in IADA. Specifically, [13 states](#) have received grants since the federal government established the program. Some states are seeking to fundamentally change their overall state assessment systems using this program, while others are adding new elements to their current system. For instance, in the most recent awards named in August 2022, the Montana Office of Public Instruction was awarded a grant to develop and implement a through-year assessment that could, in time, replace a summative assessment. Arkansas was awarded a grant to better leverage multiple measures of achievement to make better decisions for students who may be eligible to take the state’s alternate assessment.





Many states and districts have leveraged CGSA to address key challenges in order to make their assessment system more robust, instructive, and less burdensome. Common themes in the state applications include:

- adapting assessments to provide more regular, actionable feedback on individual and group student performances that can inform instruction;
- ensuring that the time to administer assessment does not significantly detract from instructional time;
- enabling assessments to capture more comprehensive views of student learning and deeper learning across subject areas.

## Core Principles for Innovative Assessment

**There have been tremendous shifts in the role of standardized assessments in school accountability systems over the past three decades.** Many states and districts may attempt to reimagine assessment systems due to a groundswell of support for new measures combined with advancements in technology and psychometrics.

New assessments used for school accountability purposes should conform to principles that are critical to effectively measuring subgroup performance, especially that of students with disabilities. These principles are specifically focused on statewide, standardized assessments as required by current federal law.

Statewide, standardized assessments should:

- **Assess students' proficiency to grade-level standards.** Assessments that allow educators to evaluate students' performance should be comprehensive in order to assess the specific level of proficiency across various domains while also assessing individual student progress toward grade-level standards. All students, and especially students with specific learning disabilities, may have uneven performance even across standards or concepts in the same subject area. Providing students the opportunity to demonstrate mastery on a wide range of grade-level standards ensures that schools have a comprehensive perspective on student performance and are effectively held accountable.
- **Facilitate the comparison of individual students to grade-level standards and to their peers.** ESSA continues to ensure that overall student performance within each state is reported by subgroups, which sheds critical light on how different populations of students within schools are performing. In the current system, each family can compare their child's score to grade-level standards. This provides families, educators, and schools an objective measure of how the child is performing compared to their peers.



- **Be accessible and culturally and linguistically responsive.** Strategically crafted Individualized Education Programs (IEPs) or 504 plans outline the conditions students need to effectively access grade-level curriculum. In many instances, accommodations for classroom instruction and assessments are a key component of a student's IEP. Accommodations can range from a no-distraction test environment for students with ADHD, to interoperable assistive technology (AT), to large print for students with low vision. To ensure equitable access, any new assessments must be compatible and administered with federally required accommodations and access to AT. Assessments should be carefully constructed, reviewed, and evaluated to eliminate bias and ensure validity for all students.

- **Be primarily used as one component of an accountability system and balanced assessment system that includes all students.** Families, students, and educators must have a clear understanding about how the assessment results will be used to identify gaps in student subgroup performance and hold school and district leaders accountable. The assessments should ultimately capture the data necessary to transform schools and ensure that students are being served well. This should not be the only measure of school success nor should statewide, standardized assessment results be used in a punitive manner. As a part of a comprehensive balanced assessment system, the standardized assessment should also include an alternative assessment that is based on alternate academic achievement standards. The number of students assessed using this alternate assessment must not exceed 1% of the total number of students assessed in the state.
- **Minimize instructional disruption among students and educators.** Assessments are critical tools for measuring progress and for accountability purposes. But they also take time away from classroom instruction. Assessment administration and scheduling should be carefully designed to minimize the time it takes for students to take tests.

For more information on how assessments should be used in accountability systems, download [A Better Approach to Accountability PDF here](#).

## New Approaches to Assessment

**Select states are already seeking ways to improve state assessment systems and revamp standardized assessments.** Different approaches to modify these assessments have different goals. Many of these approaches are not currently set up to address the principles outlined above and have significant issues in supporting subgroup performance and accountability for all students. This section outlines the most common types of new assessment models and compares them to the core principles of assessment for students with disabilities.

### ***Through-Year/Interim Assessments***

Through-year standardized assessments, sometimes called through-course or interim assessments, require schools to disseminate shorter tests throughout the school year, usually assessing segments of the grade-level standards. In theory, the scores of the assessments can be aggregated to create a summative score comparable to the previous end-of-year, summative score.

This approach seeks to increase the frequency that educators, families, and students receive feedback on student learning to inform future instruction throughout the year. When implementing through-year assessments in lieu of a complete summative assessment, policymakers seek to reduce the amount of time students sit for assessment at the end of the year, instead spreading the time throughout the school year.

#### **! Areas of Concern and Further Investigation**

There is not [sufficient evidence](#) to demonstrate that through-year assessments can be aggregated to achieve comparable scores as standardized summative assessments. In addition, accommodations or accessibility features – such as braille or speech-to-text – are [not available for many interim assessments](#). Without these accommodations, many students with disabilities cannot participate. Moreover, some through-year assessments are not effectively mapped to grade-level standards. For more on interim assessments, see [here](#) and [here](#).

### ***Performance-Based Assessments***

Performance-based assessments seek to allow students greater opportunities to demonstrate mastery of content in a way that is curriculum-embedded, facilitates critical thinking, and is less disruptive to instructional processes. These assessments consist of performance tasks designed to reflect if a student has the key knowledge,

skills, and dispositions in a specific discipline such as reading, math, or science. Tasks are typically aligned to rubrics educators use to gauge a student's level of mastery of the content being assessed. For example, on a math performance-based assessment, students can use visuals, equations, or reasoning skills to respond. On a reading performance assessment, students could show mastery in writing or through a presentation, a portfolio of artifacts, or a research investigation.

This approach seeks to capture a holistic view of student learning and higher-order thinking skills. For years, [researchers have suggested](#) that performance-based assessments measure more complex skills and how students apply knowledge and skills to new contexts unlike other standardized assessments.

### **! Areas of Concern and Future Investigation**

More research is needed to understand how students with disabilities respond to performance-based assessments of various forms. Further, it's not clear if this approach is as reliable and valid at measuring proficiency at scale compared to traditional summative assessments. While educators have experience and knowledge in several pedagogical approaches, they may have limited understanding of how to develop and use performance-based assessments in pursuit of student learning. To implement performance-based assessments at scale, school districts and states would likely need to provide significant investments in rubric development, educator supports, comparability studies, and additional accessibility features.

## **Matrix Sampling**

Matrix sampling occurs when students receive a limited number of test items representative of grade-level standards. As a result, no one student is assessed on all grade-level standards, but the scores across students can demonstrate how the overall population and subgroups of students perform compared to grade-level expectations. The scores are aggregated to create representations of performance on all grade-level standards.

This approach also seeks to reduce the amount of time students spend sitting for assessments. Because each student only receives a sampling of the overall assessment items, this will significantly reduce the time it will take to administer assessments. However, this also means that students do not receive individual scores, making it impossible to compare each individual to grade-level performance and their peers.

**! Areas of Concern and Future Investigation**

In its current form, matrix sampling cannot provide valid scores for individual students. While this method is valid for group accountability (and is currently employed for NAEP), it does not allow individual parents and educators to evaluate how a child is performing compared to peers and grade-level expectations. This system runs the risk of not providing any comprehensive measure of a specific student's ability to perform against grade-level standards. Parents, particularly parents of students with disabilities, have indicated that having access to these measures is important to them.

## **Computer Adaptive Assessments**

Computer adaptive assessments are administered on a computer and may estimate a student's ability with fewer questions than a traditional assessment. Based on a student's responses to previous questions, an algorithm adjusts which questions are presented to the student to make the next questions easier or more challenging. In other words, a child who continues to answer questions correctly will receive harder questions to assess more complex skills, often above grade level. If a student continues to answer questions incorrectly, the algorithm would present easier questions, sometimes below grade level.

This approach seeks to measure growth and proficiency. Computer adaptive tests assess standards above and below grade level, enabling schools and educational institutions to understand how students are progressing over time. This can encourage schools to maximize progress both for students who are below grade level and for those who are above grade level.

**! Areas of Concern and Future Investigation**

To ensure that all students are tested at grade-level standards, computer adaptive assessments must include [sufficient grade-level items](#), regardless of student responses to questions, to measure and allow for valid reporting against grade-level standards. Without this, these assessments may inadvertently erode the efforts to assess all students to grade-level standards.

In addition, many computer adaptive tests don't allow students to skip questions or move on without answering. This could impact students with learning disabilities or attention issues, as they may spend longer than expected on a single question, creating extreme test fatigue and undue stress or anxiety.



# Policy Recommendations

To support ongoing efforts to spur innovation in assessment and adhere to the principles mentioned before, the U.S. Congress and the U.S. Department of Education should consider the following recommendations:

## The federal government should:

- **Preserve requirements that states must develop and administer assessment systems to allow families, educators, and policymakers to compare both individual students and subgroups of students to grade-level performance targets.** Efforts to drive innovation in assessments to calculate summative scores should ensure that all students can fully access new modes of assessment and are measured against grade-level standards. Families, educators, and policymakers must also be able to compare students' performance to that of their peers.
- **Publish guidance or provide technical assistance to ensure that states and districts provide equitable access to standardized assessments, including the necessity of accommodations.** To ensure that states and testing companies prioritize developing and validating accommodations for new types of assessments, the U.S. Department of Education should reinforce the federal requirement that students have a right to accommodations. Validating accommodations can be costly and, as a result, it may be a lower priority when test developers and states are piloting approaches. However, until accommodations are developed for new testing approaches, it is not possible to evaluate the impact on critical student subgroups, like students with disabilities, who have historically been excluded from participating in standardized assessments.
- **Allocate additional funds toward assessment innovation, and prioritize development and research on models that apply to subgroups traditionally excluded from standardized assessment systems.** Developing and designing new assessments is costly and requires significant up-front investment. The federal government can facilitate greater innovation by increasing funds available for the CGSA program as well as by creating assessment-focused priorities in other research and development programs administered by the U.S. Department of Education. As a requirement for receiving funds, entities should be required to address how new innovations will support the needs of students with disabilities and English learners.
- **Incentivize cross-state, district, and test provider collaboration to drive innovation in assessments at scale.** Similar to the previous recommendation, the federal government should provide funding to reduce the cost of creating and validating innovative assessment practices.



In addition, states and districts should also continue to refine and implement comprehensive assessment systems to facilitate robust accountability systems and maximize learning for all.

## States and districts should:

- **Allocate sufficient funding to drive innovative assessment systems.** States in particular can apply for federal competitive programs, like the CGSA program, that provide supplemental funding for assessment design. States and districts should prioritize funding for assessments during budgeting processes in order to ensure that any new assessment system is implemented with fidelity.
- **Bring together stakeholders to refine assessment systems and make them more accessible and inclusive.** As states refine assessment systems, they should include and seek input from state officers who specialize in various student groups, including state directors for special education and English learners, at every stage of the design and implementation process. This collaboration is likely to increase accessibility for populations that have been historically excluded from standardized assessments. States and districts also should work with stakeholders to ensure that such assessment technologies are interoperable and compatible with assistive technology devices that comply with nationally recognized accessibility standards.
- **Apply the principles of Universal Design for Learning (UDL) to assessments of all kinds.** All assessments should follow UDL principles, or maximize instruction and learning by accommodating for differences in learners' abilities. [UDL is a scientifically valid framework](#) for guiding educational practice that provides flexibility in the way information is presented, in the way students respond or demonstrate their knowledge and skills, and in the ways students are engaged. This approach to assessment will increase the likelihood that students can effectively convey their true understanding of a topic.
- **Refine how assessment results are communicated and shared with families and educators.** Assessment data should arm educators and families with clear information about how subgroups and individual students perform compared to their peers. In collaboration with assessment vendors, states and districts must carefully design a system to communicate results with families in an accessible and culturally responsive way. In the results, families and educators should be able to understand their student's performance, how it compares to grade-level expectations, and how, if at all, it will impact the student's future instruction.



## Conclusion

**S**tandardized summative assessments serve an important purpose to building a high-quality education for all students. Decades of advocacy established the current assessment systems that drive school accountability systems. Even though the assessment system has improved over time, there is room to grow. More knowledge of assessment approaches, the science of learning, and advancement in technology will allow for new types of assessment. But innovations will not have the intended impact unless they provide access to all students.

We hope readers take away from this paper a key message: **Assessment is an important tool to support the learning of students with disabilities and to assess the effectiveness of our schools.** The disability and civil rights communities have won important victories regarding assessments that include all students and monitor student performance. What this paper argues is that states should be clear and transparent in building coherent systems of assessment that provide appropriate information to parents, students, educators, and advocates.

While assessment victories of the past help us improve how schools serve students with disabilities, we cannot yet be satisfied. We must build upon the recent positive assessment innovations. Together, we can advocate for balanced assessment systems that provide actionable information for parents and educators, inform students about their progress, and ensure that all students have access to a high-quality, equitable education.





# Acknowledgments

The authors – Lindsay Kubatzky and Meg Benner – would like to thank staff members Dr. Amelia Malone, Joey Hunziker, and Nicole Fuller for their thoughtful guidance. They would also like to acknowledge the following individuals for their insights, feedback, and review of this paper. Although this paper represents substantial feedback from these individuals, the views expressed in the paper do not necessarily reflect the positions of the organizations or individuals listed.

**Kathleen Airhart**

*Senior Advisor*  
Council of Chief State School  
Officers

**Candace Cortiella**

*Director*  
The Advocacy Institute

**Jennifer Denne**

*Consultant*  
Iowa Department of  
Education

**Roxanne Garza**

*Senior Policy Advisor*  
UnidosUS

**Dr. Andrew Ho**

*Charles William Eliot Professor  
of Education*  
Harvard Graduate School of  
Education

**Lindsay Jones**

*Chief Executive Officer*  
CAST

**Laura Kaloi**

*Partner*  
Stride Policy Solutions, LLC

**Kristin Kane**

*Founding Member*  
Decoding Dyslexia VA

**Dr. Stevan Kukic**

*Honorary Member,  
Professional Advisory Board*  
NCLD

**Dr. Sheryl Lazarus**

*Director*  
National Center on Educational  
Outcomes

**Lynn Makor**

*Consultant - Psychological  
Services*  
North Carolina Department of  
Public Instruction

**Nicholas Munyan-Penney**

*Assistant Director of P12 Policy*  
The Education Trust

**Dr. Nicole Ofiesh**

*Chief Innovation Officer and  
Co-Founder*  
Potentia Institute 21 and Brain  
Explained Products

**Ricki Sabia**

*Senior Education Policy Advisor*  
National Down Syndrome  
Congress

**Rebeca Shackelford**

*Director of Federal Government  
Relations*  
All4Ed

**Hwa Hee Sohn**

*Test Development Specialist*  
Hawai'i State Department of  
Education

**Dr. Martha Thurlow**

*Senior Research Associate*  
National Center on Educational  
Outcomes

**Sandra Warren**

*Independent Researcher*



# Endnotes

---

- 1 P.L. 114-95F requires that 95% of students take the annual assessment, and schools are held accountable if they fail to meet that threshold. Embedding participation rate into the accountability system ensures that schools test all students — even students who may not be meeting grade-level standards. Also, P.L. 108-446 requires that all children with disabilities be included in State and district-wide assessments with individual modifications and accommodations as needed.
- 2 Federal law also requires schools to administer an assessment in science in grade 10.
- 3 Batel, S., & Sargrad, S. (2016). *Better tests, fewer barriers: Advances in accessibility through PARCC and Smarter Balanced*. Center for American Progress.
- 4 Kim, K. H., & Zabelina, D. (2015). Cultural bias in assessment: Can creativity assessment help? *The International Journal of Critical Pedagogy*, 6(2).
- 5 Gordon, A. (2019). Don't remind me: Stereotype threat in high-stakes testing. *University of Baltimore Law Review*, 48(3), 3.
- 6 Seven young adults and seven parents of students with disabilities participated. Seven participants were White and seven participants were Black, Indigenous, or a person of color. Participants resided in 13 different states. The age-range requirement for the young adults was 17–26 years old. There was no age requirement for parents. NCLD recruited these individuals through email and social media marketing and through direct outreach to NCLD's Young Adult Leadership Council. Each participant conducted a pre-screening survey to determine eligibility and received a \$75 gift card for participating.

The moderator first asked participants to name some of the tests they or their child took in school and the purpose of those assessments. After participants answered, the moderator provided the federal definition of statewide summative assessments and the intended purpose.

After sharing this information, the moderator asked “Did you know that this was the purpose of these tests? If not, what did you think the purpose was?” Next, the moderator asked a series of questions about how participants felt about these assessments and how schools should be evaluated.

---

## Graphic Credits

Cover photo: [iStock.com/Caiaimage/Chris Ryan](https://www.iStock.com/Caiaimage/Chris%20Ryan)

Page 2: [iStock.com/ImagePixel](https://www.iStock.com/ImagePixel)

Page 6: [iStock.com/Hakase\\_](https://www.iStock.com/Hakase_)

Page 12: [unsplash.com/@gettyimages](https://unsplash.com/@gettyimages)

Page 13: [unsplash.com/@anniespratt](https://unsplash.com/@anniespratt)

Page 16: [unsplash.com/@cdc](https://unsplash.com/@cdc)

Page 18: [unsplash.com/@robowunderkind](https://unsplash.com/@robowunderkind)

Page 21: [unsplash.com/@thisisengineering](https://unsplash.com/@thisisengineering)

Page 24: [unsplash.com/@rockyhirajeta](https://unsplash.com/@rockyhirajeta)

Design and composites: Fil Vocasek Design, LLC