# Evaluating the Federal Innovative Assessment Demonstration Authority:

Early Implementation and Progress of State Efforts to Develop New Statewide Academic Assessments

*A Publication of the National Center for Education Evaluation and Regional Assistance*

**U.S. Department of Education**
Miguel Cardona
*Secretary*

**Institute of Education Sciences**
Mark Schneider
*Director*

**National Center for Education Evaluation and Regional Assistance**

| | |
|---|---|
| Matthew Soldner | Marsha Silverberg |
| *Commissioner* | *Associate Commissioner* |
| | |
| Elizabeth Ty Wilde | Thomas Wei |
| *Project Officer* | *K-12 Branch Chief* |

The Institute of Education Sciences (IES) is the independent, nonpartisan statistics, research, and evaluation arm of the U.S. Department of Education. The IES mission is to provide scientific evidence on which to ground education practice and policy and to share this information in formats that are useful and accessible to educators, parents, policymakers, researchers, and the public.

We strive to make our products available in a variety of formats and in language that is appropriate for a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other IES product or report, we would like to hear from you. Please direct your comments to ncee.feedback@ed.gov.

This report was prepared for IES under Contract No. 91990019C0059 by Westat and HumRRO. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

April 2023

Evaluating the Federal Innovative Assessment Demonstration Authority: Early Implementation and Progress of State Efforts to Develop New Statewide Academic Assessments

**April 2023**

**Patricia Troppe**
**Michelle Osowski**
**Mary Wolfson**
**Liam Ristow**
**Erin Lomax**
Westat

**Arthur Thacker**
**Sheila Schultz**
HumRRO

Education officials have long hoped that the statewide academic assessments most students take each year could be used not only for accountability but also to guide instruction. Congress established the Innovative Assessment Demonstration Authority (IADA) program in 2015 to help address this goal, offering up to seven states temporary flexibility from federal testing requirements so that they may more easily make progress toward replacing their current assessments with more innovative ones. The key incentive to participate in IADA is that students trying out the innovative assessment are not required to also take the state's current assessment. However, states approved for IADA must still show that their innovative assessments meet most requirements for federal accountability, and they are expected to implement the new assessments statewide within 5 years. This report describes the progress of the first five assessment systems approved under IADA in order to help policymakers consider expanding the program to more states. The report is primarily based on an analysis of states' IADA applications and performance reports to the U.S. Department of Education through the 2020-21 school year and is part of a broader evaluation of IADA required by Congress.

## KEY FINDINGS

- **All five IADA systems sought to increase the usefulness of assessment data for classroom teaching, but few were ready to try out their assessments within a year of starting IADA—both program goals.** While specific approaches varied, all systems were working to include interim assessments to provide more frequent information about student progress and cover less, and often more targeted, content per test to better inform instruction. Only one of the five assessment systems began IADA in a position to soon administer the expected "operational assessment," a trial of all IADA testing components for a given grade and subject that results in a valid determination of whether students are proficient against the state's academic standards.

- **After 2 or 3 years of participation, the IADA systems had made limited progress and may not be on track to meet the program's 5-year statewide scale-up goal.** As of the end of the 2020-21 school year, a few systems still had to complete at least some initial planning and preparation activities. One system had administered an operational assessment in 15 districts but had not made progress expanding to additional districts.

- **States reported challenges hampering assessment development and implementation activities, with the COVID-19 pandemic causing major disruptions.** Pandemic-induced cancellation of statewide testing and limitations on in-person activities were universally cited challenges. But other challenges not specific to the pandemic were also common, such as persuading districts to adopt the new systems, coordinating roles and processes, and developing technical specifications and supporting materials for the assessments.

## INTRODUCTION

Despite three decades of federal law requiring states to assess students' academic achievement,[1] calls to reform state testing continue.[2] Ideally, these assessments serve multiple purposes: to measure individual students' progress toward meeting state standards, to guide classroom instruction, and to hold schools accountable for the academic learning of their students. There is wide consensus, however, that the one-time end-of-year assessments that states use to meet the Elementary and Secondary Education Act (ESEA) testing requirement may not be well-suited for all these purposes.[3]

**Common concerns about traditional state assessments**

- Limited frequency and end-of-year timing means information cannot be easily used to guide instruction in the same school year
- Summative measurement requires broad coverage of academic material, with insufficient detail to pinpoint students' needs

To broaden the possibilities for state assessments, Congress established the Innovative Assessment Demonstration Authority (IADA) in the 2015 reauthorization of the ESEA. IADA encourages up to seven states to further develop and administer assessments that better serve their multiple potential purposes and to do so in ways that may be less burdensome for students, teachers, and schools.[4,5]

**Key IADA program goals**

- Improve assessments' ability to inform instruction
- Increase assessment efficiency
- Reduce burden

IADA addresses these goals by easing some of the federal requirements for state assessment systems.[6] For example, IADA allows a greater range of assessment types and allows students to take the innovative assessment in lieu of the traditional statewide assessment (i.e., "no double testing" of students). In exchange for this specified flexibility, IADA states must meet many other legislated requirements for the development and administration of the assessment system during the demonstration period. These requirements relate to stakeholder input, assessment quality, supports provided to schools and students, reporting of assessment results, and what it means to scale up the program (see Appendix A for more detail about IADA program requirements). IADA's requirements reflect its aim to help states replace their current traditional accountability assessments with the innovative assessments. Thus, the innovative assessments must meet many of the specifications outlined in ESEA for state accountability and be comparable to the traditional assessment.[7]

This congressionally mandated report provides an initial look at the progress of the first five IADA assessment systems to help policymakers consider whether and how to expand the program to more states.[8] The report describes key features of the assessment systems that states proposed (see Appendix A for more details about each system)– their "readiness" to implement soon after they applied to IADA, how far they were in carrying out key IADA activities at the end of the 2020-21 school year, and the challenges they encountered including those related to the COVID-19 pandemic.

The progress analysis relies primarily on documents that states submitted to the U.S. Department of Education, as Congress required for this first report.

These documents include states' applications to the IADA program and their annual performance reports (APRs). APRs are typically submitted in the fall and provide information about the status of assessment development and implementation activities conducted in the prior school year. The APR describing activities in the 2020-21 school year was the most recent data available at the time this report's analysis was being conducted. These documents vary in the content and level of detail provided, requiring careful extraction of information that should be interpreted with some caution. To ensure as much consistency as possible, the study team submitted questions to states after reviewing the APR submissions to get clarification as needed and provided states with an opportunity to respond with additional information. Interviews with officials of participating IADA systems were also conducted to gather information not found in the APRs. These activities were conducted during the disruptions caused by the COVID-19 pandemic, which, as noted later, delayed aspects of IADA efforts and the ongoing evaluation. More details about how the report's analysis was carried out are in Box 1 below and in Appendix B.

## BOX 1. OVERVIEW OF REPORT'S ANALYTIC APPROACH

**What questions did the study address?**

- What were the key objectives and features of the IADA systems?

- How "ready" were the IADA systems at the start of the demonstration to meet early program expectations?

- How far along were the IADA systems at the end of the 2020-21 school year toward implementing key program activities?

- What challenges to developing and implementing assessments did IADA systems report?

**What data were collected?**

- *State IADA application materials.* This publicly available information includes not only the applications but also responses to U.S. Department of Education (the Department) reviewer comments. Louisiana and New Hampshire submitted their applications in 2018, while Georgia and North Carolina applied in 2019.

- *State IADA APRs.* Each state is required to report annually to the Department on several of their IADA system activities.[9] For example, states must report their progress toward their intended implementation timeline; the performance of students in participating schools on the IADA assessments; feedback from educators, parents, and other stakeholders on their satisfaction with the IADA system; and if the system is not yet implemented statewide, school demographic information for participating schools and progress in scaling up to additional districts and schools. Starting with the 2019-20 APR, the form also provided space for states to describe how the COVID-19 pandemic may have affected the development and implementation of their IADA systems. The APR form allows states to provide narrative descriptions of their progress or submit supporting evidence. States provided additional information in response to questions posed by the study team to clarify descriptions in the APRs.[10] APRs were collected for 2019-20 and 2020-21 from all four early IADA states.[11] (Georgia submitted a single APR each year that included reports for the GMAP through-year and Navvy assessment systems.[12]) Louisiana and New Hampshire's 2018-19 APRs were also collected.

- *Interviews with IADA system officials.* Semi-structured telephone interviews were conducted in November and December 2020 with officials who were most familiar with the development and early implementation of the IADA systems, including challenges during the IADA application period and efforts through the 2019-20 school year.

## IADA SYSTEMS USE DIFFERENT APPROACHES, BUT ALL TARGET KEY PROGRAM GOAL: INCREASE TIMELINESS AND RELEVANCE FOR INSTRUCTION

Although all IADA systems were expected to satisfy the goals and specific requirements of the demonstration authority, the legislation was purposely broad in allowing a wide range of assessment types. The program included this flexibility so that states could propose assessments that best meet their state's specific objectives, with the law citing options such as competency-based, instructionally embedded, interim, and performance-based assessments.[15] Applicants were expected to choose assessment features that would help them improve assessment quality and relevance for instruction while continuing to generate annual summative results about individual students and schools. Accomplishing these key program goals required states to provide educators with more timely access to assessment scores and to be attentive to testing burden.

- ***All five IADA systems intended to increase timely access to data during the school year.*** Each includes an interim assessment, which could be administered one or more times within a school year (see Exhibit 1). Thus, information about student progress is expected to be available more frequently than is typically the case with only an end-of-year assessment. The more frequent assessments are to be executed in different ways across state systems. For example, teachers are expected to play a large role in determining when assessments are administered under NH PACE because the system includes performance assessment tasks that are embedded throughout the curriculum, and teachers decide how to order and pace the curriculum. Under Navvy, districts can decide the assessment windows or delegate the decision to their schools or teachers, whereas the remaining systems (LEAP 2025 Humanities, NC PAT,[16] and GMAP through-year) have preset windows for administering interim assessments during the school year. Except for LEAP 2025 Humanities, the systems also expect to report nearly immediate student performance information either because the assessments are scored by the teacher who administered them (NH PACE) or are computer-scored right away.[17] Because some LEAP 2025 Humanities items will be hand-scored by a vendor, results will not be immediate; however, the results are meant to be timelier than those from the traditional ESEA summative assessment.[18]

- ***Features of all five IADA systems aim to make their assessment data more meaningful and relevant to classroom teaching.*** The IADA systems expect to provide educators different information than the states' traditional ESEA summative assessments. In choosing to use interim assessments covering less, but more targeted, content per test than the traditional summative test, all systems have the potential to better inform instruction (see Exhibit 1). All systems have tied the content of their assessments more closely to instruction happening near the time of the assessment. Under NH PACE, for example, teachers can use the curriculum-embedded assessments when they teach the related content, leading to performance assessment that is directly tied to the content and timing of classroom instruction. Navvy districts have flexibility to choose which state academic standards to assess during a particular testing period, allowing them to align the assessments to their planned instructional pacing. In addition, Navvy teachers have the option to administer assessments that address a single standard and, since students have multiple attempts to demonstrate proficiency on the standard, teachers can capture several data points on student performance.[19] Because the GMAP through-year assessments are [adaptive](#) (i.e., content and level of difficulty are based on each student's prior performance), they are expected to produce more meaningful estimates of learning for particularly low- and high-performing students even though the assessments are not directly linked to materials recently taught.[20]

**Exhibit 1. Key features of traditional state assessments and IADA assessments: 2020–21**

| Planned assessment features | Typical traditional assessment system | IADA assessment system | | | | |
|---|---|---|---|---|---|---|
| | | NH PACE | LEAP 2025 Humanities | NC PAT | GMAP Through-Year Assessment | Navvy |
| How many summative assessments are administered? | 1 | 0 | 1 | 1 | 0 | 0 |
| How many interim assessments are administered? | 0 | Number can vary | 3 | 3 | 3 | Number can vary |
| When are assessments administered? | Preset assessment window, typically in spring | Flexible, teacher-determined as part of curriculum | Preset assessment windows | Preset assessment windows | Preset assessment windows | Flexible, district-determined |
| What types of items are in the assessment? | Multiple choice and constructed-response | Performance tasks (e.g., project portfolio, presentation) | Multiple-choice/ multiple-select, constructed-response, extended essay, and technology-enhanced | Multiple-choice/ multiple-select, constructed-response, and technology-enhanced | Multiple-choice/ multiple-select, constructed-response, and technology-enhanced | Multiple-choice/ multiple-select |
| How are assessment items selected? | From a common bank | From common and locally developed task banks | From a common bank | From a common bank | From a common bank | From a common bank |
| How are assessments administered? | Computer administration and/or paper and pencil | Computer administration, paper-and-pencil components, or other mode as determined by teacher | Computer administration | Computer administration | Computer administration | Computer administration |
| | Not adaptive | Not adaptive | Not adaptive | Adaptive | Adaptive | Not adaptive |
| How is the overall assessment score reported? | Content area score (e.g., grade 9 math) | Content area proficiency determination | Content area score | Content area score | Content area score | Content area proficiency determination |
| What additional detail about the score is reported to guide instruction? | Sub-score (e.g., algebra, geometry) | Competency score for one or more content or performance standards | Sub-score | Sub-score | Sub-score | Proficiency determination by specific content standard |

| Planned assessment features | Typical traditional assessment system | IADA assessment system | | | | |
|---|---|---|---|---|---|---|
| | | NH PACE | LEAP 2025 Humanities | NC PAT | GMAP Through-Year Assessment | Navvy |
| Which test scores contribute to accountability? | Single summative score | All interim proficiency determinations | Interim and summative scores | Single summative score | All interim scores | All interim proficiency determinations |

SOURCE: Application materials and annual performance reports for New Hampshire, Louisiana, North Carolina, and Georgia through 2020–21.

Another feature, to be adopted by LEAP 2025 Humanities, seeks to improve the validity of assessment results and therefore its relevance to instruction. The state will assess students using the same texts they study in their classrooms and limit the use of texts to which students have not been exposed. Unfamiliar contexts or frames of reference in assessment items can cause confusion, leading students to fail to demonstrate their understanding of underlying concepts being tested.[21] For example, consider a group of assessment items that assess students' reading comprehension using a detailed text about a musical composer's selection of notes and chords to elicit an emotional response from their audience. Those test items could potentially favor students who play in a band or orchestra, and disadvantage those who do not, irrespective of their underlying skill in reading comprehension. Removing the potential barriers associated with the unknown context may allow students to better demonstrate their abilities, yielding more accurate and fairer information to guide ongoing instruction.

## AT START OF DEMONSTRATION, SOME SYSTEMS MORE READY THAN OTHERS TO ADMINISTER OPERATIONAL ASSESSMENT, AN EARLY PROGRAM EXPECTATION

States granted demonstration authority were expected to administer a fully operational IADA assessment in at least one school and in at least one grade or subject required under ESEA within their first year.[22] Because the process of making an assessment fully operational can take many years, the design of the IADA program effectively required states to have development of their assessment systems well underway at the time of application.

To prepare for an operational assessment, the IADA program expected systems to complete a variety of assessment planning, design, and development activities, and then put assessment administration procedures in place. Steps included developing the full set of tests (e.g., interim and summative); defining performance levels (i.e., determining what constitutes proficiency); and

ensuring that teachers and school leaders were ready to implement the assessment system, prepared to provide appropriate testing accommodations for students who needed them, and aware of how to use assessment results as intended. After administering a fully operational assessment, the IADA program expected that systems would ensure the technical quality of the scores, including their comparability to the traditional ESEA statewide assessment; produce and use score information at the classroom, school, district, and program levels; and engage in activities to improve the system for the next administration of the assessment. The extent to which a system had prior experience completing these steps could influence how likely the system was to meet program expectations in (and beyond) the first year.

- ***Only one system clearly started IADA ready to administer an operational assessment within a year.*** New Hampshire's application materials[23] suggested that it had completed nearly all the major activities expected for assessment planning, preparation, administration, post-administration, and continuous improvement for NH PACE prior to IADA (see Exhibit 2). New Hampshire had been working to develop its NH PACE assessment long before the state received the IADA authority and had administered NH PACE operationally (generating scores that could be used for accountability) in the state as early as 2014-15 under a special U.S. Department of Education waiver from federal testing rules.[24]

- ***The other systems needed more work to achieve operational readiness, and the extent of the additional work varied considerably.*** At the start of their demonstration periods, the other four systems still needed to complete major activities to make their assessments operational (see Exhibit 2). Navvy's application materials suggested that it was close to completing all activities needed to administer an operational assessment. However, it still needed to develop procedures to combine student results from its interim assessments into a final score, and to conduct analyses to determine performance levels and calculate annual determinations of student proficiency. In addition, Navvy's application indicated a need to familiarize students and parents with the assessment and possibly a need for training school staff on test security and protecting student data.[25]

States' responses to the IADA application's section on prior experience suggested that each of the three other systems had to complete foundational assessment planning and development work and other preparation activity before they could administer an operational assessment. For example, while the GMAP through-year assessments planned to draw on the item pool developed for NWEA's Measures of Academic Progress assessment where appropriate, it still needed to establish test requirements (i.e., specifications for test development[26]) and ensure that the assessments covered the full scope of Georgia's academic standards. North Carolina and Louisiana had more assessment design and development work to complete because neither had sufficient items to draw on. Louisiana had not yet created an assessment framework to ensure coverage of the state standards, established item requirements (e.g., item type and level of

9

**Exhibit 2. Status of IADA systems on major assessment implementation activities at the start of IADA**

| Major activity and sub-activity | System status at the start of IADA | | | | |
|---|---|---|---|---|---|
| | NH PACE | LEAP 2025 Humanities | NC PAT | GMAP Through-Year Assessment | Navvy |
| **Initial planning of assessment** | | | | | |
| Test and item requirements established | ● | ○ | ○ | ⊘ | ● |
| Other initial planning completed | ⊘ | ● | ⊘ | ⊘! | ⊘! |
| **Prepare for first operational assessment** | | | | | |
| Initial items developed | ● | ○ | ○ | ⊘ | ● |
| Items piloted | ● | ○ | ○ | ⊘ | ● |
| Appropriate testing accommodations ensured | ● | ●! | ● | ⊘ | ● |
| Assessment administration procedures developed | ● | ⊘ | ●! | ○! | ● |
| Professional development conducted | ● | ○! | ○ | ⊘ | ⊘! |
| **Operational assessment administered** | ● | ○ | ○ | ○ | ○ |
| **Post-administration activities** | | | | | |
| Technical quality of scores ensured | ● | ○ | ○ | ○ | ○ |
| Reports produced | ● | ○ | ○ | ○ | ○ |
| Data from operational assessment used | ● | ○ | ○ | ○ | ○ |
| **Continuous improvement activities** | | | | | |
| Feedback on the system obtained | ⊘ | ○ | ○ | ○ | ○ |
| Annual evaluation conducted | ⊘ | ○ | ○ | ○ | ○ |
| **Progress toward scaling up the system made** | ○ | ○ | ○ | ○ | ○ |

● = conducted activity; ⊘ = conducted part of activity; ○ = no report of activity or not yet applicable

! Interpret with caution, as application materials were unclear about these activities.

SOURCE: IADA applications and application addenda for New Hampshire, Louisiana, North Carolina, and Georgia.

difficulty), or developed and piloted new assessment items for LEAP 2025 Humanities. North Carolina used its existing, voluntary "NC Check-Ins" interim assessments as the inspiration for its IADA assessment NC PAT, but it still needed to convene panels to

develop and finalize test requirements for each of the three NC PAT interim tests and to develop and pilot test items. Finally, state applications also suggested that all three systems had more to do in developing and conducting professional development for teachers and school leaders on key aspects of the assessment.

## SYSTEMS ESTABLISHED SOME FOUNDATION FOR OPERATIONALIZING ASSESSMENTS, BUT MADE LIMITED PROGRESS IN FIRST YEARS OF DEMONSTRATION

The IADA program requires participating states to scale up their IADA assessment system to statewide use and demonstrate suitability as a replacement to the traditional statewide assessment system within 5 years, with the possibility of a 2-year extension.[27] To meet this requirement, IADA systems must complete many activities both leading up to and following the administering of an operational assessment expected within the first year. Depending on where each system started from when they entered the demonstration, they must make headway on pre-operational assessment activities such as assessment planning, or post-operational assessment activities such as verifying the assessments' technical quality, producing (and using) assessment data, and undertaking various continuous improvement activities. The IADA program, however, does not specify timelines or milestones for completing these specific activities, perhaps in part because some previously completed activities may need to be revisited (e.g., trying out new assessment items). Still, it is important to examine the status of each of the five systems' efforts at the end of the first few years of the demonstration authority (2020-21) and, to the extent possible, the progress they made while under that authority.

- *Some systems still had to complete initial planning and preparation activities 2 to 3 years after starting IADA.* While all systems met at least some indicators for initial planning activities in 2020-21, the GMAP through-year and Navvy assessment systems did not appear to have completed the critical activities of specifying what their IADA assessments will measure (test specifications) and how (item specifications) (see Exhibit 3 and Appendix Exhibit C.1 for indicators by activity). Lack of fully established test and item specifications hampers these systems' ability to complete item development, leaving them with potentially considerable work to do before they can administer an operational assessment.

All systems met at least some indicators related to assessment preparation activities, but there was variation by system, and no system met indicators related to conducting professional development (see Exhibit 3 and Appendix Exhibit C.2). For example, all systems met most indicators related to developing test items for 2020-21 (see Appendix Exhibit C.2). However, there were two item development indicators that seemed difficult

| Major activity and sub-activity | System status at the end of 2020-21 | | | | |
|---|---|---|---|---|---|
| | NH PACE | LEAP 2025 Humanities | NC PAT | GMAP Through-Year Assessment | Navvy |
| **Initial planning of assessment** | | | | | |
| Test and item requirements established | ● | ● | ● | ○ | ○ |
| Other initial planning completed | ● | ● | ⊘ | ● | ● |
| **Prepare for first operational assessment** | | | | | |
| Initial items developed | ⊘ | ● | ⊘ | ⊘ | ⊘ |
| Items piloted | ● | ○ | ● | ● | ● |
| Appropriate testing accommodations ensured | ● | ● | ○ | ○ | ● |
| Assessment administration procedures developed | ● | ● | ○ | ○ | ⊘ |
| Professional development conducted | ○ | ○ | ○ | ○ | ○ |
| **Operational assessment administered** | ● | ○ | ○ | ○ | ○ |
| **Post-administration activities** | | | | | |
| Technical quality of scores ensured | ⊘ | ○ | ○ | ○ | ○ |
| Reports produced | ⊘ | ○ | ○ | ○ | ○ |
| Data from operational assessment used | ○ | ○ | ○ | ○ | ○ |
| **Continuous improvement activities** | | | | | |
| Feedback on the system obtained | ● | ○ | ○ | ○ | ○ |
| Annual evaluation conducted | ● | ○ | ○ | ○ | ○ |
| **Progress toward scaling up the system made** | ○ | ○ | ○ | ○ | ○ |

● = met all indicators for activity; ⊘ = met at least one indicator for activity; ○ = did not meet any indicators or indicators for activity not yet applicable

SOURCE: Derived from the 2020-21 IADA annual performance reports (APRs) and clarifying materials for Louisiana, North Carolina, and Georgia. For New Hampshire, the ratings are derived from the 2018-19 and 2019-20 APRs because New Hampshire received approval to pause their IADA timeline for 2020-21 and did not administer the NH PACE assessment that school year.

for some systems to meet: documenting how items and the item pool were selected and documenting how the assessment will address all major content areas from the state standards. All systems except Louisiana met the indicator for item piloting, sufficiently
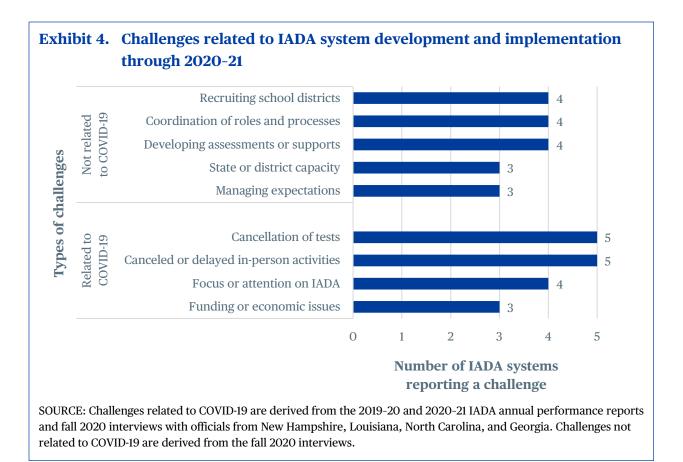
documenting that each assessment item allows students to demonstrate the specific content knowledge as intended. New Hampshire, Louisiana, and Navvy met or partially met indicators related to developing test accommodations and standardized administration procedures to ensure test security, a fair and reliable assessment, and the delivery of reports on individual student performance that protect personally identifiable information. Although each system reported providing at least some training opportunities for educators, none met indicators demonstrating that the majority of educators were trained on key topics, including general information about the system, test security, accommodations, and data use.

- ***Only one IADA system administered an operational assessment and completed at least some post-administration and improvement activities, except for scale-up.*** New Hampshire was the only system to administer an operational assessment by the end of 2020-21 (see Exhibit 3 and Appendix Exhibit C.3). As a result, New Hampshire was the only system with the opportunity to meet indicators of the kinds of post-administration, continuous improvement, and scale-up activities that IADA require to be completed within 5 years. New Hampshire administered an operational assessment and engaged in these post-administration activities during the 2018-19 school year. While New Hampshire met most indicators of post-administration activities and all indicators of continuous improvement, it did not increase the number of districts participating in NH PACE between 2018-19 and 2019-20 (see Appendix Exhibits C.4 to C.6).[28] None of the other systems administered an operational assessment by the end of 2020-21.[29]

- ***It appears that most systems were not yet advancing through major assessment implementation activities.*** Comparisons over time suggest that systems were largely working to complete the same major planning and preparation activities as they were at the outset of IADA (see Exhibits 2 and 3).[30] Systems may have made incremental progress on some sub-activities, but overall, most still had more work to do to advance to administering an operational assessment. In some cases, it may appear that systems have completed fewer activities or completed activities less fully a few years into the demonstration than they had when they began. However, readers should interpret the changes over time, both potential progress and backsliding, with caution. Some of the changes could be, at least partly, an artifact of relying only on IADA applications to assess states' status at the start of the demonstration. The applications may have included more aspirational language, for example, as compared to the actual performance reports used to assess states' status in subsequent years.

# CHALLENGES HAMPERED IADA SYSTEM DEVELOPMENT AND IMPLEMENTATION, WITH PANDEMIC A MAJOR DISRUPTION

Under normal circumstances, states may encounter challenges as they transition to a new assessment system. The COVID-19 pandemic, which led to early school closures in 2019-20, cancellation of state assessments–including IADA assessments–in spring 2020, and disruptions to the learning environment into 2020-21, only added to the complexity of states' work. These difficulties, taken together, provide a signal about the likely pace of progress in later years of the demonstration.

- *Recruiting and preparing districts to adopt the new system, coordination issues, and development of assessments and program supports were challenges reported by most systems.* Four systems reported recruitment challenges either leading up to their IADA application (NC PAT, Navvy, and the GMAP through-year assessments) or upon attempted system scale-up (New Hampshire) (see Exhibit 4). These challenges seemed to be unrelated to the COVID-19 pandemic and seemed to stem from how information about the assessment was disseminated and district concerns about the extra burden or readiness to implement. For example, New Hampshire's IADA officials reported that some districts do not feel ready to adopt the new practices and to set aside the extra time needed for NH PACE activities, which hampered efforts to expand the assessment across the state. All systems except Navvy discussed difficulties with coordination and with development of the assessments and their supports. For example, North Carolina reported challenges in determining the timing and content of interim assessments when instructional pacing varies across districts. The GMAP through-year assessment system noted misunderstandings about who would lead certain portions of the work (districts or assessment vendor) and challenges providing signing support or a braille format for an online assessment.

- *Limited state or district capacity and managing expectations of stakeholders were also common challenges.* Three systems reported capacity challenges (see Exhibit 4). North Carolina and New Hampshire officials reported challenges in developing or deploying educator professional development, which they worked to resolve by leveraging support from vendors or participating districts. Louisiana officials noted issues with developing assessment items on a shorter schedule than usual for assessment work. They resolved this issue by focusing on less complex assessments first (e.g., middle school instead of high school). Three systems discussed the importance of managing stakeholder expectations to support system adoption, including familiarizing school-, district-, or state-level stakeholders with the new assessment approach or clarifying where some aspects of the new system are unchanged from the traditional assessment.

**Exhibit 4. Challenges related to IADA system development and implementation through 2020-21**



SOURCE: Challenges related to COVID-19 are derived from the 2019-20 and 2020-21 IADA annual performance reports and fall 2020 interviews with officials from New Hampshire, Louisiana, North Carolina, and Georgia. Challenges not related to COVID-19 are derived from the fall 2020 interviews.

- *Cancellation of traditional ESEA statewide testing and limitations on in-person activities due to the pandemic were universal challenges.* All systems reported assessment development and administration challenges related to test cancellations (see Exhibit 4). In 2019-20, spring testing cancellations meant that three systems (New Hampshire, Louisiana, and Navvy) could not complete critical activities, which may have included plans to administer an operational IADA assessment, generate annual determinations, or conduct post-administration activities. North Carolina did not expect to administer NC PAT in 2019-20, but instead to pilot new items by embedding them in traditional ESEA end-of-grade tests. It was unable to do so. These pandemic-related disruptions had a subsequent effect in 2020-21, as none of the systems ended up administering an operational IADA assessment that year. All systems reported that the pandemic also made in-person and person-intensive activities more challenging for participating districts and schools. In 2019-20, New Hampshire and Louisiana deferred or canceled educator training on their IADA system. North Carolina and the GMAP through-year assessment systems canceled or delayed meetings with technical advisers.

The pandemic drew attention away from IADA in four systems (see Exhibit 4), affecting activities such as recruitment, professional learning, and teacher engagement. For example, Navvy reported that the pandemic slowed expansion of the system in 2019-20

because districts were focused on pandemic response, and noted a drop in districts using the system in 2020-21. The GMAP through-year assessment system reported that the spring 2020 school closures prompted refocusing of professional learning activities in the following year. North Carolina reported that district personnel and teachers were occupied with the transition to remote learning in spring 2020 and could not attend a planned webinar to offer feedback on the NC PAT system. Louisiana limited the scope of new programs like the IADA system to prioritize student learning recovery and acceleration.

Additionally, officials in three states reported that financial challenges related to the pandemic affected available resources for the IADA work (see Exhibit 4). For example, New Hampshire reported that a data management contractor went out of business due to the pandemic, hampering the state's work in 2019-20. In Georgia, pandemic-related budget cuts reduced technical assistance available to Navvy and the GMAP through-year assessment systems.

## LOOKING AHEAD

The IADA program offers some flexibility from key federal testing requirements to incentivize states to replace their current statewide assessments with "innovative" assessments that better serve the needs of students. Congress stipulated IADA begin as a demonstration program for a handful of states for the purpose of learning from their experiences, with the intention of subsequently expanding to more states across the country. This report reveals the extent to which participating states did–or did not–meet key IADA program expectations, including administering fully operational assessments by the end of their first year and, in later years, making progress toward scaling up the assessments. In doing so, the report raises several questions for policymakers working to achieve the program's intended goals.

- **How are states weighing the benefits and costs of participation in IADA?** States wishing to develop new assessments to replace those already in use are not required to participate in IADA. The main benefit of participation is that, under IADA, a state's students are not required to take both current and new assessments during the trial phase. But fulfilling the requirements that come with IADA participation, including having to scale up the newer assessment to statewide use within 5 years (see Exhibits 2 and 3), is a complex undertaking. As this report outlines, most states that began participating in IADA did not appear to be very far along in assessment development at the start, and they experienced a variety of challenges, not all pandemic-related, in moving toward the program's scale-up expectations. These experiences may be a signal to both current participants and other states of difficulties that lie ahead; it is one hypothesis for why New Hampshire (in 2022) and Georgia (in 2023) withdrew their assessments from the

program, and IADA has so far not reached the capacity Congress allowed. At the same time, there are also signs that at least some of the current participants are forging ahead post-pandemic, with Louisiana and North Carolina reporting plans to administer an operational assessment in at least some schools during the 2022–23 school year.[31]

- **To what extent is IADA truly facilitating state innovation?** The IADA legislation is purposefully broad when defining the types of assessments that could be considered innovative. Some features of initial participants' systems (e.g., interim, competency-based, and instructionally embedded assessments) are far from new, but they are rarely used for accountability.[32] In this regard, their intended adoption for statewide testing may reflect meaningful innovation. But ultimately, there is little in the IADA legislation that explicitly "ensures" state innovation in terms of developing the next generation of assessments that are truly different from what has been done before. Participating states commonly reported struggling to gain district buy-in fast enough to meet the requirement to scale statewide within 5 years–a finding that other states considering new ways to assess students may view as a cautionary tale.

- **Could other federal programs support, or provide more support for, IADA efforts or expansion?** IADA is just one of several complementary federal programs that can be used to improve statewide assessments. Title I-B of the ESEA, which was appropriated $390 million in 2022,[33] provides funds to states via both formula grants and a competitive program called Competitive Grants for State Assessment (CGSA). States may use the formula funds to support development of innovative assessments before seeking IADA participation, while they are participating under IADA, or for assessments that would not qualify under IADA because the testing is not intended to be used for statewide accountability. In 2020, the Department experimented with incentivizing participation in IADA by prioritizing CGSA funds for that purpose and competing "development grants" to help states get to a stage where they are ready to apply for IADA.[34] While continuing that priority or setting aside CGSA exclusively to help states participate in IADA may be worth considering as a way to further expand IADA, its effect could be limited by the relatively small share of total Title I-B funds set aside for CGSA ($29 million in 2022).[35] Moreover, funds through CGSA may not fully address some underlying challenges reported by IADA systems and presented earlier in this report, such as recruiting districts and coordinating roles and processes. Other Department programs, such as the Office of Elementary and Secondary Education's Comprehensive Centers,[36] are intended to help address a variety of state education needs, although with any federal program, trade-offs are typically present in deciding what purposes to prioritize.

The next report for this evaluation, expected in 2025, will include further updates on systems' progress and summarize lessons learned through this effort.

# DEFINITIONS

An adaptive test provides a tailored assessment based on a student's performance on prior assessment questions (or prior assessment scores).

Competency-based assessments allow students to demonstrate that they have learned the expected knowledge and skills (i.e., a competency) needed to progress to the next academic content, grade, or level.

Competency scores reflect an aggregation of scores from performance tasks designed to measure specific knowledge or skills for a specific competency.

A constructed-response item requires a student to independently create a response, rather than select a response from a set of choice options. Examples of constructed-response items include fill-in-the-blank and short essay.

Cumulative year-end assessments evaluate whether students have learned what they were expected to learn by the end of the school year.

Instructionally embedded assessments are woven into instruction.

Interim assessments are administered multiple times throughout the year.

Items piloted includes small- or large-scale administration of items or test forms to a group of students. States review the results of the pilot to determine whether to include an item in an operational assessment and whether the group of items as a whole performs as expected.

Multiple-choice/multiple-select items require students to select one or more than one correct response from a set of responses.

An operational IADA assessment requires that students participate in all IADA testing components (e.g., all interims and the summative) within a given grade and subject and that IADA results can be used to determine whether students are proficient against the state's academic standards.

Other initial planning includes establishing an implementation timeline and whether the system provided information to parents and students and training to staff to familiarize them with the IADA assessment system's purpose and major features to help them transition to the new system.

Performance-based assessments allow students to demonstrate mastery of academic content standards (i.e., what they should know and be able to do) by performing a specific task or series of tasks.

A proficiency determination is an indication of whether the student has met a proficiency threshold (e.g., Proficient, Not Proficient).

Technology-enhanced items are computer-administered and more complex than multiple-choice/multiple-select items. For example, students may drag and drop answers from a list, order items, or highlight text to identify a central idea.

Through-year assessments, as defined by GMAP, are a series of interim assessments taken throughout the year that produce data that can be aggregated to generate a final, end-of-year summative score for students.

# ENDNOTES

[1] Since the 1994 reauthorization of the Elementary and Secondary Education Act (ESEA) (under the Improving America's Schools Act), states have been required to establish statewide standards and aligned assessments in selected subjects and grades. Under the most recent ESEA reauthorization (under the 2015 Every Student Succeeds Act), states are required to test students in English language arts/reading and mathematics in each of grades 3 through 8 and at least once in high school, and in science at least once in grades 3-5, grades 6-9, and grades 10 through 12 for accountability and reporting under Title I, Part A.

[2] See for example, Elish-Piper et al. (2013), Faxon-Mills et al. (2013), Jimenez and Boser (2021), Millitello et al. (2013), and Young and Kim (2010).

[3] See for example, Lake and Worthen (2021), O'Keefe et al. (2021), and Rennie Center Education Research & Policy (2021).

[4] The Secretary of Education may authorize a state or consortia of state educational agencies to establish IADA systems. Consortia may include no more than four states. The Secretary may authorize no more than seven states (including those participating in a consortia) for the initial demonstration authority (i.e., the first 3 years of the IADA program).

[5] In its December 2016 letter to Chief State School Officers, the U.S. Department of Education identified IADA among the new provisions in ESSA that can "support high-quality, fair assessments and transparent testing policies, consistent with the principles and practices outlined in the President's Testing Action Plan. These provisions may also help States and local education agencies (LEAs) improve assessment literacy and reduce unnecessary testing through efforts to evaluate and streamline existing assessments and to strive for continued improvement and innovation in assessment" (Office of Elementary and Secondary Education (OESE), 2016, p. 3). IADA, in particular, may "help States develop new and different ways to measure student knowledge and skills that provide more timely and useful feedback to students, parents, and families, and educators, and that may serve as a model to other States and LEAs" (OESE, 2016, p. 4).

[6] IADA does not provide funding to states for participating in IADA, although states may choose to use funds that they receive under Title I of the ESEA (Parts A and B) to help offset costs.

[7] IADA assessment systems must meet the requirements of section 1111(b)(2)(B) of ESEA, except that an innovative assessment (1) does not need to be administered to all students during the demonstration period if the assessment is initially administered to a subset of schools as long as the traditional assessment is given to students in nonparticipating schools; and (2) does not have to be administered in all grades and subjects required for federal accountability purposes as long as the traditional assessment is given in grades and subjects in which the IADA assessment is not administered.

[8] The U.S. Department of Education approved two assessment systems for Georgia to test under IADA– the Georgia MAP Partnership (GMAP) through-year assessment system and the Navvy assessment system administered by the Putnam County Consortium. At the end of its IADA period, Georgia planned to pick one of these assessment systems for statewide use (OESE, 2019). In addition, Massachusetts received approval for its IADA science assessment in 2020. Given that the focus of this report is on the early implementation of the IADA systems through the 2020-21 school year, the Massachusetts system is not included in this report.

[9] The U.S. Department of Education office that is responsible for IADA shared annual performance reports (APRs) for New Hampshire, Louisiana, Georgia, and North Carolina with the study team as states submitted them. All of the APRs are now available on the IADA program [website](#).

[10] In general, the APR invites states to provide narrative descriptions of their progress or submit supporting evidence. States are asked to provide counts on student participants and their performance, counts of the number of teachers eligible for trainings and the number who participated, and lists of participating LEAs and schools and the grades and subjects in which the IADA assessments were administered. In responding to APR reporting requirements and follow-up questions, states may have provided documentation including technical and training manuals that covered such issues as test administration, accommodations, test security, and scoring; alignment and validity study reports; assessment schedules; examples of district reports/report cards; and training and data use information.

[11] New Hampshire received approval from the U.S. Department of Education to pause its implementation of NH PACE during the 2020–21 school year. Although New Hampshire submitted an APR for the 2020–21 school year, it reported no NH PACE activities during that year. The study team drew on information from New Hampshire's prior APRs for the analyses in this report. New Hampshire subsequently withdrew from the IADA program in spring 2022 (OESE, 2022a).

[12] In February 2023, just prior to the publication of this report, Georgia informed the U.S. Department of Education that it intended to discontinue IADA participation for both of its assessment systems. This development does not affect any of the analyses in this report, which only include data through the 2020–21 school year.

[13] The six activities were based on major topics included in *Operational Best Practices for Statewide Large-scale Assessment Programs* (Council of Chief State School Officers & Association of Test Publishers, 2013).

[14] The framework for the indicators was based on the five areas of progress identified in the ESSA for this IADA Progress Report—the extent to which (1) the state has solicited feedback from key stakeholders about their satisfaction with the IADA system; (2) educators have demonstrated a commitment and capacity to implement or continue to implement the IADA system; (3) the IADA system meets technical requirements for assessments; (4) the same innovative assessment was used to measure the achievement of all participating students; and (5) there were similar student participation rates in the IADA and traditional state assessments. ESSA did not define specific indicators or thresholds for the five areas of progress. Thus, the study team developed indicators to operationalize them and measure whether the IADA system performed an activity or provided evidence that technical assessment requirements were met. The indicators and the thresholds states must meet were guided by legislative requirements, best practice, or specific guidance for the U.S. Department of Education's assessment peer review process. However, the study team did not conduct a full-scale peer review of each IADA assessment. The indicators were then grouped into six major assessment development and administration activities for presentation: initial planning, preparation for the first operational assessment, operational assessment administered, post-administration activities, continuous improvement activities, and scale-up. See Appendix B for more information.

[15] ESSA also lists [cumulative year-end assessments](#) as an option for the program. In addition, in the final IADA program regulations, the Department expanded the allowable assessment types to include an "other" assessment design as long as it satisfies other requirements for IADA assessments.

[16] The North Carolina Department of Public Instruction plans two assessment windows for NC PAT. It plans for one large assessment window (e.g., October 1-May 31) for the three interim assessments, giving teachers flexibility on the timing of administration. In contrast, the assessment window for the summative assessment is fixed to the last 10 days of the school year.

[17] The GMAP through-year assessment is likely to include constructed-response items on the fall and winter interims for formative purposes and will include constructed-response items on the spring interim for summative purposes. If included, the formative constructed-response items will be scored by teachers for their own use, while the summative constructed-response item will be scored by professional scoring services.

[18] LEAP 2025 Humanities includes hand-scored portions of the interim and summative assessments. All scoring during the demonstration period and beyond will be done by an assessment vendor.

[19] While students taking the Navvy assessments have multiple attempts to demonstrate proficiency on a standard, the results of their final attempt are included in summative scores for accountability purposes.

[20] The GMAP through-year assessment items are adaptive within, below, and above grade based on individual student performance. This means that students' performance on early test items influence which items they will be administered later in the tests (poor-performing students will be given easier items; high-performing students will be given harder items). An adaptive test allows the test to better estimate a student's ability, with fewer items, even if the ability estimate is above or below the student's grade level (Becker & Bergstrom, 2013).

[21] See: Kozminsky and Kozminsky (2001), and Recht and Leslie (1988).

[22] In its December 2016 letter to Chief State School Officers, the U.S. Department of Education noted that "only those States that wish to use the innovative assessment in place of the statewide assessment, including for the purposes of accountability and reporting under Title I, Part A, in at least one school, require innovative assessment demonstration authority" (OESE, 2016, p. 4). Moreover, the Notice Inviting Applications (NIA) defines a participating school as one in which the IADA assessment is administered "instead of, or in addition to," the traditional statewide assessment and "where the results of the school's students on the innovative assessment system are used by its State and LEA for the purposes of accountability and reporting" (Applications for New Authorities, 2018, p. 46927). While this information suggests that the IADA program office expected applicants to administer a fully operational IADA assessment within their first year, the program office approved applications where the system timeline indicated that it would use its first year for assessment planning, design, or development work rather than administering an operational assessment for accountability purposes.

[23] The analysis of a system's "readiness" at the start of IADA relied on state application materials, including responses to reviewer comments, and in particular the application section where the state described its prior experience, capacity, and stakeholder support for the IADA system (see NIA section V[1][b]). This section asked states to describe (1) the extent and depth of prior experience the state and districts have in developing and implementing the components of the innovative system; (2) the extent and depth of the state and district capacity to implement the system considering the availability of technological infrastructure, state and local laws, dedicated and sufficient staff, expertise, and resources, and other relevant factors; and (3) the extent and depth of state and local support for the IADA application (Applications for New Authorities, 2018, pp. 46928-46929).

[24] NH PACE was approved for use in certain grades and subjects for accountability purposes in a subset of schools and districts through a waiver from the U.S. Department of Education granted during the 2014-15 school year. New Hampshire received subsequent annual waivers thereafter (State of New Hampshire Department of Education, 2018, p. 4). New Hampshire administered the traditional statewide assessment for accountability in non-NH PACE grades, subjects, and schools.

[25] The Navvy application materials were not detailed enough to ascertain whether the system had conducted activities to familiarize students and parents about the major features of the assessment or whether the system previously developed and conducted training for teachers and school leaders on test security and protecting student data.

[26] The *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014) indicate that test specifications include documentation of the purpose and intended uses of the test; detailed decisions about content, format, test length, and psychometric characteristics of the items and test; delivery mode; administration; scoring; and score reporting. The Georgia application did not identify whether the GMAP through-year assessment system needed to develop all these specifications or some during the early years under IADA.

[27] After scaling up to statewide use, the state must submit evidence to the Department's peer-review process to determine the quality of the IADA assessment (Application for New Authorities, 2018, p. 46930). The scale-up requirement looks somewhat different for Georgia. Since the U.S. Department of Education approved two IADA assessments for Georgia, the state planned to pick one of these assessments for statewide use at the end of its demonstration period (OESE, 2019).

[28] As indicated in endnote 11, New Hampshire did not administer an operational assessment or engage in other activities in 2020-21.

[29] Georgia has an added requirement for the Navvy and GMAP through-year assessment systems that requires the systems to demonstrate comparability with the statewide assessment before it allows the systems to administer the assessment in lieu of the traditional statewide assessment for accountability purposes. As a result, the consortia will have to administer their IADA assessments for at least a year to generate evidence of comparability for the state before they can administer an operational assessment under IADA.

[30] The 2018-19 annual performance report for Louisiana revealed that the state decided to roll out its IADA assessment with grade 7 first instead of high school, as was discussed in its application materials, so the status of the activities referenced for the first operational assessment in Exhibits 2 and 3 may not be for the same assessment.

[31] Based on a March 2022 inquiry from the IADA program office to participating assessment systems asking: "Do you plan to administer the operational versions of [your innovative assessment] for some schools in the state, provide individual student reports, and use the results in state and local report cards and in the State's federal accountability system in place of the traditional state assessment for at least one grade and one subject area in 2022-23?"

[32] For example, the STAR Reading interim assessment program from Renaissance Learning has been available since at least 1996 for monitoring student achievement and progress (Piasta et al., 2012). ESEA allows traditional assessment systems to derive a summative score from either a series of interim assessments or a single summative assessment. However, a 2018 survey of all state education agencies

found that no states reported requiring districts to administer multiple statewide interim assessments instead of a single summative assessment in 2017-18 (Troppe et al., 2020, Exhibit 2.50a).

[33] See: Consolidated Appropriations Act (2022), p. 476.

[34] The Louisiana, Massachusetts, Nebraska, and Texas Departments of Education each received grants of approximately $3 million, and the Hawaii Department of Education received a $500,000 grant. (Applications for New Awards, 2020; OESE, 2022b).

[35] See: OESE (2022b).

[36] See: Comprehensive Centers [website](#).

# REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association. https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf

Applications for New Authorities; Innovative Assessment Demonstration Authority, 83 F.R. 46924. (2018, September 17). https://www.govinfo.gov/content/pkg/FR-2018-09-17/pdf/2018-20152.pdf

Applications for New Awards; Competitive Grants for State Assessments Program, 85 F.R. 25422. (2020, May 1). https://www.govinfo.gov/content/pkg/FR-2020-05-01/pdf/2020-09336.pdf

Becker, K.A., & Bergstrom, B.A. (2013). Test administration models. *Practical Assessment, Research, and Evaluation, 18*(14). https://doi.org/10.7275/pntr-yz21

Consolidated Appropriations Act, 2022, Pub. L. No 117-103, 136 Stat. 476 (2022). https://www.govinfo.gov/content/pkg/PLAW-117publ103/uslm/PLAW-117publ103.xml

Council of Chief State School Officers & Association of Test Publishers. (2013). *Operational best practices for statewide large-scale assessment programs*. Washington, DC: Author.

Elish-Piper, L., Matthews, M.W., & Risko, V.J. (2013). Invisibility: An unintended consequence standards, tests, and mandates. *Journal of Language and Literacy Education*, *9*(2), 4-23. https://files.eric.ed.gov/fulltext/EJ1023571.pdf

Faxon-Mills, S., Hamilton, L., Rudnick, M., & Stecher, B. (2013). *New assessments, better instruction? Designing assessment systems to promote instructional improvement*. RAND Corporation. https://www.rand.org/pubs/research_reports/RR354.html

Jimenez, L., & Boser, U. (2021). *Future of testing in education: The way forward for state standardized tests*. Center for American Progress. https://www.americanprogress.org/article/future-testing-education-way-forward-state-standardized-tests/

Kozminsky, E., & Kozminsky, L. (2001). How do general knowledge and reading strategies ability relate to reading comprehension of high school students at different educational levels? *Journal of Research in Reading*, *24*(2), 187-204. https://doi.org/10.1111/1467-9817.00141f

Lake, R., & Worthen, M. (2021). *State accountability systems in the COVID era and beyond*. Center on Reinventing Public Education. https://crpe.org/wp-content/uploads/v3_accountability_report_2021.pdf

Millitello, M., Bass, L., Jackson, K.T., & Wang, Y. (2013). How data are used and misused in schools: Perceptions from teachers and principals. *Education Sciences*, *3*(2), 98-120. https://doi.org/10.3390/educsci3020098

Office of Elementary and Secondary Education. (2016, December 7). *Dear colleague letter regarding Every Student Succeeds Act (ESSA) and assessments* (OESE-2016-29). U.S. Department of Education. https://oese.ed.gov/files/2020/07/essaassessmentdcltr1207.pdf

Office of Elementary and Secondary Education. (2019, July 12). *Georgia IADA approval letter*. U.S. Department of Education. https://www2.ed.gov/admins/lead/account/iada/gaiadaapproval2019.pdf

Office of Elementary and Secondary Education. (2022a, April 22). *Letter acknowledging New Hampshire's withdrawal from the IADA*. U.S. Department of Education. https://oese.ed.gov/files/2022/04/nhacknowledgementiadawithdraw4.22.22.pdf

Office of Elementary and Secondary Education. (2022b, August 26). *Awards: Competitive grants for state assessments*. U.S. Department of Education. https://oese.ed.gov/offices/office-of-formula-grants/school-support-and-accountability/competitive-grants-for-state-assessments/awards/

O'Keefe, B., Rotherham, A., & O'Neal Schiess, J. (2021, May). Reshaping assessments and accountability in 2021 and beyond. *The State Education Standard, 21*(2), 7-12. https://nasbe.nyc3.digitaloceanspaces.com/2021/05/OKeefe-Rotherham-Schiess_May-2021-Standard.pdf

Piasta, S.B., Justice, L.M., McGinty, A.S., & Kaderavek, J.N. (2012). Increasing young children's contact with print during shared reading: Longitudinal effects on literacy achievement. *Child Development, 83*(3), 810-820. https://doi.org/10.1111/j.1467-8624.2012.01754.x

Recht, D.R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology, 80*(1), 16-20. https://psycnet.apa.org/record/1988-24805-001

Rennie Center Education Research & Policy. (2021). *Reimagining the MCAS: The need for an accountability system that supports deeper learning*. https://www.renniecenter.org/sites/default/files/2021-06/Reimagining%20MCAS%20Policy%20Brief%20Final.pdf

State of New Hampshire Department of Education. (2018, March 29). *Applications for new authorities under the Innovative Assessment Demonstration Authority*. U.S. Department of Education, Office of Elementary and Secondary Education. https://www2.ed.gov/admins/lead/account/iada/nhapplication.pdf

Troppe, P., Isenberg, E., Milanowski, A., Garrison-Mogren, R., Rizzo, L., Gill, B.P., Ross, C., Dillon, E., & Li, A. (2020). *The transition to ESSA: State and district approaches to implementing Title I and Title II-A in 2017-18* (NCEE 2021-002sv). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/pubs/2021002/pdf/2021002_supplementalVolume.pdf

Young, V.M., & Kim, D.H. (2010). Using assessments for instructional improvement: A literature review. *Education Policy Analysis Archives, 18*(19), 1-40. https://doi.org/10.14507/epaa.v18n19.2010

## ACKNOWLEDGMENTS

## DISCLOSURE OF POTENTIAL CONFLICT OF INTEREST

The research team for this evaluation included staff from Westat and its subcontractors, HumRRO and Plus Alpha Research & Consulting. None of the research team members have financial interests that could be affected by findings from this evaluation. None of the five members of the Technical Working Group who participated in both meetings have financial interests that could be affected by findings from the evaluation. A sixth member of the Technical Working Group participated only in the first meeting and did not have financial interests that could be affected by the evaluation findings at the time of that meeting.