

Session 3B: Lessons Learned About the Implementation of IADA

Panelists: Thomas Lambert (thomas.lambert@la.gov), Kinge Mbella (kinge.mbella@dpi.nc.gov), Allison Timberlake (atimberlake@doe.k12.ga.us)

Moderators: Scott Marion, Carla Evans

[Session 3B Lessons Learned About the Implementation of IADA slides 6–10]

Dr. Evans opened the session by providing a timeline of IADA state approvals and withdrawals. All of the applications are available on [ED's website](#). Dr. Marion briefly explained a few items in and resulting from the statutorily required IES report, also available on [ED's website](#). (1) IADA starts upon approval of the application, so consider implementing a planning process before applying or be prepared to start once approved. (2) Very few states produced a summative determination as a result of their IADA. The COVID-19 pandemic added challenges to implementing a big program.

Dr. Marion then introduced the panelists and provided the questions that they were asked to respond to:

1. What motivated your state to pursue assessment innovation?
2. Why did you feel as if you needed the IADA to do this?
3. What have been your most important successes thus far?
4. What have been your biggest challenges?
5. What do you know now that you wish you knew when you were applying?

Thomas Lambert, Assistant Superintendent for Assessment, Accountability, and Analytics at the Louisiana Department of Education

What motivated your state to pursue assessment innovation? [Slide 3]

Mr. Lambert explained that Louisiana completes ELA, math, science, and social studies assessments yearly. Louisiana wanted to maintain the social studies assessment and reduce the testing burden.

Why did you feel as if you needed the IADA to do this? [Slide 4]

Louisiana was motivated to create multiple Louisiana Educational Assessment Program (LEAP) 2024 assessment formats and give greater flexibility on the form that best meets its needs. Louisiana wanted to focus on mastering broad domains of knowledge in state standards and be more relevant to classroom instruction.

In 2018, Louisiana conceptualized streamlining state testing through a comprehensive humanities assessment to build a cohesive English/language arts and social studies assessment that would (1) measure what students had learned by including passages that students had read (hot reads), as well as similar (warm reads) and new materials (cold reads); (2) allow students to take shorter tests throughout the year, versus one long test at the end of the school year; and (3) maintain local control of curricula. There would be multiple forms of the same test with slightly different construction. Districts could test with either the traditional stand-alone

summative assessment in each separate subject/course or an interdisciplinary model of interim text-based assessments, in combination with a shorter summative assessment at the end of the year.

What have been your most important successes thus far? [Slide 6]

Louisiana has successfully administered the tests and achieved strong comparability of results. Results are strongly predictive of the next year's assessment scores. Teachers report that they like the new assessments and that students are more confident.

What have been your biggest challenges? [slides 7–8]

Mr. Lambert reported several challenges. First, measuring ELA and social studies together was more challenging than anticipated. As a result, the plans were amended; the test transitioned to a full ELA assessment.

There is tension between innovation and comparability. If testing different constructs, should the results be comparable?

Reporting school year data in a way that is useful to educators and accurately using classical approaches is also a challenge. There's always a debate on what to release and at what level and what can be done. There's also a good behavioral shift among educators of looking at data to determine what students need—but there is a question about validity of the data. If midyear data are not completely accurate, could that do damage? Recruiting schools and systems to “try something new” that could negatively affect their letter grades is very difficult to do, particularly during a pandemic when scores are going down.

There can be a fundamental challenge in comparability between the new tests and traditional tests when you remove something that research tells you is a barrier to student achievement, such as background knowledge. Louisiana found that students achieved similar scores because the tests were designed for comparability by anchoring items to the traditional tests. Ironically, stakeholders want a test that is the same length or shorter yet also want more data.

What do you know now that you wish you knew when you were applying? [Slide 10]

Mr. Lambert replied that starting with math would have most likely been easier than starting with ELA. Louisiana has some exciting things underway with math. The standards for ELA and social studies were misaligned. Additionally, the standards updates were on different schedules. Pay as much attention to the delivery model and the educational theory as you do to the psychometric approach. Make sure that the psychometric model does exactly what you need it to do and that it accounts for students gaining mastery throughout the school year. Juggling multiple assessment programs requires much more capacity than previously assumed, and through-course assessments add a significant monitoring burden. Mr. Lambert was unsure what Louisiana might have done differently but recommends approaching it thoughtfully. The IADA does not come with specific financing. To date, Louisiana's innovative assessment pilot has been funded by partners and CGSA grant funding. This model is probably not sustainable for all states.

Questions and Comments

An attendee asked about teachers' positive reactions and how Louisiana teased out preference for social studies versus ELA. Mr. Lambert responded that Louisiana actually did not combine testing the two subjects. From the teachers' response, Louisiana found that teachers liked getting information more quickly and felt they had a greater impact on achievement. Mr. Lambert is concerned that eventually teachers will not like test score data that are based on a snapshot before the school year's learning is done.

Dr. Banks asked Mr. Lambert to discuss use of ELA guidebooks, survey data from teachers and students, and how they feel about the curricula and instruction that they get in school. Mr. Lambert responded that most Louisiana schools choose to use locally developed ELA guidebooks, selecting different units from a set. His department then conforms the assessment to the selections. As a result of the assessment, Louisiana can give teachers information on student performance. Students report feeling more prepared, but the data so far does not indicate that they are better prepared. There may be a psychological effect at play.

An attendee asked whether there were any unintended consequences with the hot reads and teaching to the test. Mr. Lambert responded that teachers are supposed to teach material for the hot read—for example, *The Giver*. The assessment's warm read would then work with similar dystopian themes, with the hope that a better understanding of the context would remove the need for background knowledge and level the playing field. Regarding the hot read, analysis of the assessment results revealed a greater incidence of plagiarism, which could be because of lessons on specific items. More work needs to be done to determine whether this is cheating or replication of lessons.

Dr. Timberlake asked about next steps. Mr. Lambert responded that Louisiana is thinking a lot about lessons learned and what the next version will look like. He is uncertain whether curriculum-aligned assessments are sustainable in the long term. Updates to curricula require updates to tests, which are very expensive to build. Mr. Lambert wonders whether Louisiana will eventually develop its own "homegrown curriculum" and the default could be the traditional end-of-year assessment.

Kinge Mbella, Psychometrician at the North Carolina Department of Public Instruction

What motivated your state to pursue assessment innovation? [*Slides 3B_IADA NCPAT State Conference DC 2023 slides 3–8*]

Dr. Mbella explained that long before IADA, North Carolina had been on the leading edge of exploring innovative ways to approach standardized testing. However, educators had expressed concerns about the summative system in place (and still used by most of the state) and its inability to generate actionable data to inform classroom practices. North Carolina began to address these concerns in 2014. After considerable effort, in SY 2015–16, the North Carolina Department of Public Instruction (NCDPI) conducted a proof-of-concept pilot to explore the feasibility of implementing a through-grade system. Results from the proof of concept led to the

development of interim assessments, known as NC Check-Ins, and now there are NC Check-Ins 2.0. The main purpose of NC Check-Ins is to provide educators, students, and stakeholders with immediate and detailed feedback on student performance when it comes to grade-level-specific content standards so classroom instruction can be tailored to individual students' needs. In 2018–19, NC Check-Ins were administered to approximately 50 percent of students across the state as a formative model focusing on giving immediate feedback on student performance to teachers. (More information on NC Check-Ins is available on [NCDPI's website](#).)

IADA offered the opportunity to expand this concept to develop a comprehensive assessment system that would support the use of a through-grade model in lieu of the current summative assessment. Dr. Mbella reported that in addition to reporting progress to ED, North Carolina is required to issue progress reports to the North Carolina General Assembly, which passed a law in 2019 stating that North Carolina should “move toward a through-grade assessment model, in which the State-mandated assessments are administered in multiple short testing events throughout the school year rather than a single long testing event at the end of the year.”

Why did you feel as if you needed the IADA to do this? [slides 8–11]

Dr. Mbella stated that IADA enabled the flexibility during the pilot phase to concurrently report student achievement from both the innovative system and the traditional system without double-testing.

What have been your most important successes thus far? [slides 12–13]

Dr. Mbella reported that NC Check-Ins 2.0 continue to receive positive ratings from educators. Feedback from educators shows strong support for maintaining the formative purpose of these interim assessments. Feedback also shows that educators would not support a system in which NC Check-Ins 2.0 also become the main source of accountability data on student performance and growth. Participation in NC Check-Ins 2.0 continues to be voluntary for schools across the state. The state intends to scale up next year.

What have been your biggest challenges? [Slide 14]

Dr. Mbella stated that change is really hard when happening in real time and that designing an innovative system while implementing it at the same time makes for a challenging communications plan. There is also a huge difference between what is technically possible and what is practical. Finding the balance between these has been key to North Carolina's design, particularly for a state with local control of curricula. Dr. Mbella also cautioned attendees to set appropriate expectations and avoid overpromising.

Questions and Comments

Dr. Mbella noted that if your state is thinking about doing something innovative, IADA is a good framework to do it in. An attendee asked about the NC Check-Ins. Is there a base form for students feeding into the system without a previous Check-In? Dr. Mbella responded that for students who have not previously participated, including those who have moved to the state

and are not part of the system, the middle-level Check-In is almost like traditional assessment and the natural default route.

Dr. Evans asked whether the interim assessments are modular or mini-summative in design. Dr. Mbella answered that math is based on content standards. For reading, North Carolina adjusts the same test to assess the same reading standards. An attendee asked whether all five achievement levels are available for students who take the spring summative test. Dr. Mbella answered that there are no restrictions. The levels are designed psychometrically to allow reflection of achievement.

Dr. Evans asked about scaling and the percentage of LEAs in the pilot last year. Dr. Mbella responded that about 8 percent of students in seven districts across the state from 189 schools participated in the pilot. An unidentified staff member from North Carolina added that students can participate in the interim assessment, independent of the pilot. About 50 percent of students across grade levels have taken the interim assessments, and the percentage is expected to reach 60 percent when the interim assessments are offered statewide next year.

[Slides 3B Georgia's IADA Experience slides 2–7]

Dr. Timberlake explained that Georgia has left IADA and that its path to IADA was a bit unconventional. In 2018, the state passed a law that established a state innovation assessment pilot, allowing up to 10 districts or consortia to develop their own alternate accountability systems. IADA provided the flexibility to follow federal law while working with the districts. A Georgia State Board of Education competition first selected three districts for the pilot, all of which were included on the IADA application. IADA approved two of the districts for the program—the Georgia Measures of Academic Progress (MAP) Assessment Partnership and its MAP Growth, a through-level assessment, and the Putnam Consortium's Navvy, an on-demand assessment. As part of the state competition, only one of the assessments would eventually be chosen for statewide implementation. (Slide 2 contains a QR code to access Georgia's IADA resources., which can be found at <https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Pages/Assessment-Innovation-and-Flexibility.aspx>)

IADA was needed to keep the state law for innovation assessment pilots to remain compliant with federal law. The pilots were driven by the consortia of districts, who were in charge of designing, developing, and implementing their assessments. The state law that created the pilot granted authority to the consortia with oversight from the Georgia State Board of Education; the state law intentionally cut the Georgia Department of Education (GaDOE) from the work. The IADA allowed GaDOE to be involved in the pilots because the department managed the IADA.

Dr. Timberlake explained that the greatest success was the level of stakeholder engagement attained through grassroots efforts; both consortia actively engaged educators. Additionally, at the start of the COVID-19 pandemic, the two consortia quickly developed pandemic-related resources to support educators and students. With the involvement of educators in the member

districts, both consortia were also successful with item development, resource development, and professional learning. Both consortia worked successfully on validity research and were working toward establishing comparability when the state left the IADA pilot.

Dr. Timberlake explained that there were several challenges—mostly because of the nature of the pilot’s structure, as well as the effort to turn formative assessments into summative assessments. The competitive nature of the pilot deterred the consortia from collaborating, possibly preventing them from illuminating duplication of efforts. The Georgia law constrained GaDOE’s ability to guide the work; for example, GaDOE staff members could not share their expertise in developing and implementing state assessments to pass peer review. The GaDOE had no involvement in test design or decision-making other than making sure the work met IADA requirements on comparability. With vendors collaborating on the pilots, if adopted, the pilots would have put Georgia in new territory with respect to ownership and procurement. Georgia currently owns all of its assessments; in the pilots, the vendors own the testing. The pandemic delayed timelines and increased costs; by Year 4 of the pilot, neither of the consortia had implemented a full three-year pilot in any frame or content area. Both consortia’s vendors stopped work at or around Year 3. Georgia adopted new state academic content standards this year. State assessments are currently being aligned to address the standards. This would have been a lot of work for the consortia. Both consortia struggled with how to calculate and validate a single summative score that would roll up all the testing occasions. The most pressing challenge was using formative assessments for summative purposes. If either one had become the statewide assessment system, it would have a marked significant increase in high-stakes testing.

Dr. Timberlake had several recommendations for those considering innovating or IADA. Start with assessment literacy and ensure that all stakeholders have a shared understanding of assessment types and uses of data; in all likelihood, they will be used for state and federal accountability purposes. It is also important to build broadscale buy-in and commitment and to work with everyone (including advocates) and make sure everyone understands the plans and TOA. She advised states to ensure that they have the time and resources needed; an innovative assessment takes more of both. States should also consider practical and operational challenges. Plan to study the impact of the innovative assessment on teaching and learning; this is the root of why you are doing this. In Georgia, students and teachers currently love the formative nature of existing assessments. Dr. Timberlake wonders how they would feel if the formative assessments became summative. In closing, Dr. Timberlake said that for those considering a through-year assessment solution, a [National Center for the Improvement of Educational Assessment paper](#) on the topic is essential reading. If Georgia had been able to have statewide discussions on these issues first and had better attended to challenges on the front end, the pilot might have gone differently.

Questions and Comments

Dr. Marion asked whether Georgia understands why participation among the districts dropped for SY 2022–23. Dr. Timberlake said “no” and surmised that it was related to increased

understanding of what it would mean if the formative assessments, which the state valued so much, became the statewide summative. An audience member asked whether any of the consortia's work or lessons learned informed the statewide assessment. Dr. Marion responded that the entire panel would answer this question before the session's conclusion. Dr. Mbella asked what caused Georgia to stop participating in IADA. Dr. Timberlake responded that a combination of things contributed to its withdrawal from IADA. With the APR, it became apparent that Georgia was not making progress and did not have a plan to address that.

Dr. Evans asked how work or lessons learned informs statewide assessment plans. Dr. Timberlake responded that it does—but not specifically the assessments that were under development. Georgia has been working toward a more balanced approach to assessment, shortening the footprint of the summative and providing more resources throughout the year, making it optional for districts so that they can get tests set earlier and potentially offer one for instructional purposes and another for assessment. Mr. Lambert responded that Louisiana has learned a lot and will do some things differently, such as having a clearer theory of action and trying to have something more coherent, at least on ELA and math. Dr. Mbella replied that North Carolina is moving forward; right now, the state is making a final decision about whether to make the transition.

Janette Kirk, Chief of Federal Programs in Maine, inquired about structures for implementing IADA. Dr. Timberlake replied that just she and a colleague managed IADA. The two consortia were responsible for everything else. Georgia had a combination of district resources, vendor contributions, and external grants. Regardless, implementation was not cheap. Mr. Lambert said that Louisiana has one project manager, which is not enough. The psychometric and assessment and administration teams were already working on other tests. Dr. Marion said vendors have been brought in but added that it gets expensive. Dr. Evans asked about North Carolina having a testing department. Dr. Mbella replied that North Carolina partners with North Carolina State University to develop the assessment platform. North Carolina also works with the University of North Carolina at Greensboro on psychometrics. Dr. Mbella's department has a full slate of testing staff members. They have directed resources from the main assessment to IADA.

Dr. Marian asked about the 95 percent participation requirement, what counts, and how it is scored. Dr. Timberlake replied that Georgia foresaw issues with meeting peer review standards, and she didn't see a path forward. Mr. Lambert responded that Louisiana doesn't generally have a participation problem; if a student doesn't participate, the student is given a zero in the accountability system. As a result, people are finding ways for students to sit for tests. If a student has a doctor's note excusing the student from testing, the student is moved from the innovative test to the traditional summative test. The same applies for students moving into a system giving the innovative test. Every student sits for a test.

Dr. Marion asked the panelists to give one quick piece of advice to colleagues who are curious about IADA. Dr. Timberlake said to treat IADA like a tool for developing something new; IADA is not the something new. Dr. Mbella replied to be systematic. Take one step at a time. Innovation takes time. Mr. Lambert said to think about what you are trying to make better and whether you

can do it in your current construct. If your test is so fundamentally broken that you don't have time to make it better, you don't have time to do something new. Approach IADA with caution. It has many constraints to keep compliant with federal law; it is designed to have you do something in a box.