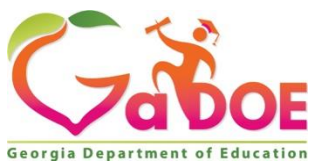


Georgia's Innovative Assessment Pilot Update

The Georgia Department of Education's Update on the
Innovative Assessment Pilot

SB 362, 2018 Legislative Session



2021-2022



Section 1: Background

State Innovative Assessment Pilot

Senate Bill 362 in 2018 established an Innovative Assessment Pilot Program that allowed up to 10 school districts or groups of districts to develop alternate assessment and accountability systems aligned with state academic content standards. To select the innovative assessments that would be part of the program, the State Board of Education (SBOE) held a competition in summer of 2018, with two application deadlines of August 1, 2018, and September 1, 2018.

Innovative Assessment Pilot Program Application Announcement:

- https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/Flexibility/Innovative_Assessment_Pilot_Program_Application_Announcement-2018-07-17.pdf

Innovative Assessment Pilot Application:

- https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/Flexibility/Innovative_Assessment_Pilot_Application-2018-07-17.docx

The SBOE reviewed the applications and supporting evidence from all submitted applications, ultimately approving three applications for participation in the pilot: Cobb County School District (Cobb Teaching and Learning System Assess platform), Georgia MAP Assessment Partnership (MAP Growth for Georgia), and the Putnam Consortium (Navy).

Federal Innovative Assessment Demonstration Authority (IADA)

To pursue maximum federal flexibility for the state innovative assessment pilot, Georgia applied to participate in the federal Innovative Assessment Demonstration Authority (IADA) under Section 1204 of the Elementary and Secondary Education Act (ESEA), as amended by the Every Student Succeeds Act (ESSA), on December 17, 2018.

Georgia's Application for the Innovative Assessment Demonstration Authority:

- https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/Flexibility/Georgia_IADA_Application.pdf

After receiving feedback from the U.S. Department of Education (ED), along with peer review notes, GaDOE provided additional information to ED demonstrating how Georgia's IADA application would meet all associated requirements. On July 12, 2019, Georgia received approval from ED to implement two of the innovative assessment models – Georgia MAP Assessment Partnership and the Putnam Consortium. As part of their approval, ED required that the Putnam Consortium's assessment model produce a measure for the literacy (Lexile) indicator of CCRPI before being

implemented. The Putnam Consortium submitted their plan for producing a literacy measure to ED in October 2019. Their plan was approved in February 2020.

U.S. Department of Education Interim Feedback Letter:

- <https://www2.ed.gov/admins/lead/account/iada/gaiadainterimfeedback03082019.pdf>

IADA Peer Review Notes:

- <https://www2.ed.gov/admins/lead/account/iada/gaiadapeernotes2019.pdf>

Georgia's Response to the Interim Feedback Letter:

- https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/Flexibility/Georgia_IADA_Addendum.pdf

IADA Approval Letter:

- https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/Flexibility/GA_IADA_approval_final_letter.pdf

Putnam Consortium Literacy Measure:

- https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/Flexibility/Putnam_Consortium_Navy_Literacy_Measure_10_28.pdf

Participating Consortia

Georgia MAP Assessment Partnership

The Georgia MAP Assessment Partnership is developing MAP Growth for Georgia in partnership with NWEA. MAP Growth for Georgia is a through-year assessment that leverages adaptive interim assessments to provide timely insights on students' command of grade-level standards, measure academic growth, provide norm-referenced test results, and produce summative proficiency scores.

Features of the assessment system include:

- computer adaptive
- measures student learning relative to grade-level expectations and adapts within, below, or above grade level based on student performance
- provides growth and norm-referenced scores
- provides interactive online reporting

MAP Growth for Georgia will be administered three times per year, in fall, winter, and spring.

Putnam Consortium

The Putnam Consortium is developing Navvy in partnership with Pearson Assessments.¹ Navvy is an on-demand assessment system that leverages cutting-edge data science to provide real-time diagnostic data.

Features of the assessment system include:

- assesses students on individual standards
- teachers provide assessments on demand throughout the school year
- includes multiple opportunities for students to demonstrate mastery of each standard
- provides interactive online reporting

Navvy is administered on demand throughout the school year. Each standard includes approximately 8 questions and students can be retested on each standard up to two additional times after the initial administration.

Section 2: 2021-2022 Year 3 Update

Participation

Each year, the consortia may amend the list of districts participating in the innovative assessment pilot program by 1) submitting a redlined version of their Consortium Membership as provided in their approved State Board of Education application and approved IADA application, 2) ensuring that the demographics of the updated consortium continue to be representative of the state, and 3) submitting a signed *Memorandum of Understanding* and *IADA Application Assurances* for each new consortium member.

In 2021-2022, 20 districts participated in the GMAP Consortium (an increase from 14 districts the previous year) and 12 districts participated in the Putnam Consortium (a decrease from 18 districts the previous year).

	Year 1 2019-2020	Year 2 2020-2021	Year 3 2021-2022	Year 4 2022-2023
GMAP	9	14	20	18
Putnam	12	18	12	10

GMAP participating districts in 2021-2022 include: Barrow County, Clayton County, Dalton City, Floyd County, Haralson County, Jackson County, Jasper County, Marietta City, Calhoun City (affiliate), Chattooga County (affiliate), Colquitt County (affiliate),

¹ Developed by Navvy Education LLC, the Navvy assessment system was acquired by Pearson Assessments in spring 2022.

Evans County (affiliate), Houston County (affiliate), Oglethorpe County (affiliate), Seminole County (affiliate), Treutlen County (affiliate), Trion City (affiliate), Chattahoochee County (participating), Georgia Cyber Academy (participating), and Elbert County (participating).

Putnam participating districts in 2021-2022 include: Dougherty County, Fayette County, Floyd County, Putnam County, Ben Hill County, Candler County, Chattooga County, Emanuel County, Scintilla Charter Academy, Statesboro STEAM Academy, Troup County, and Vidalia City.

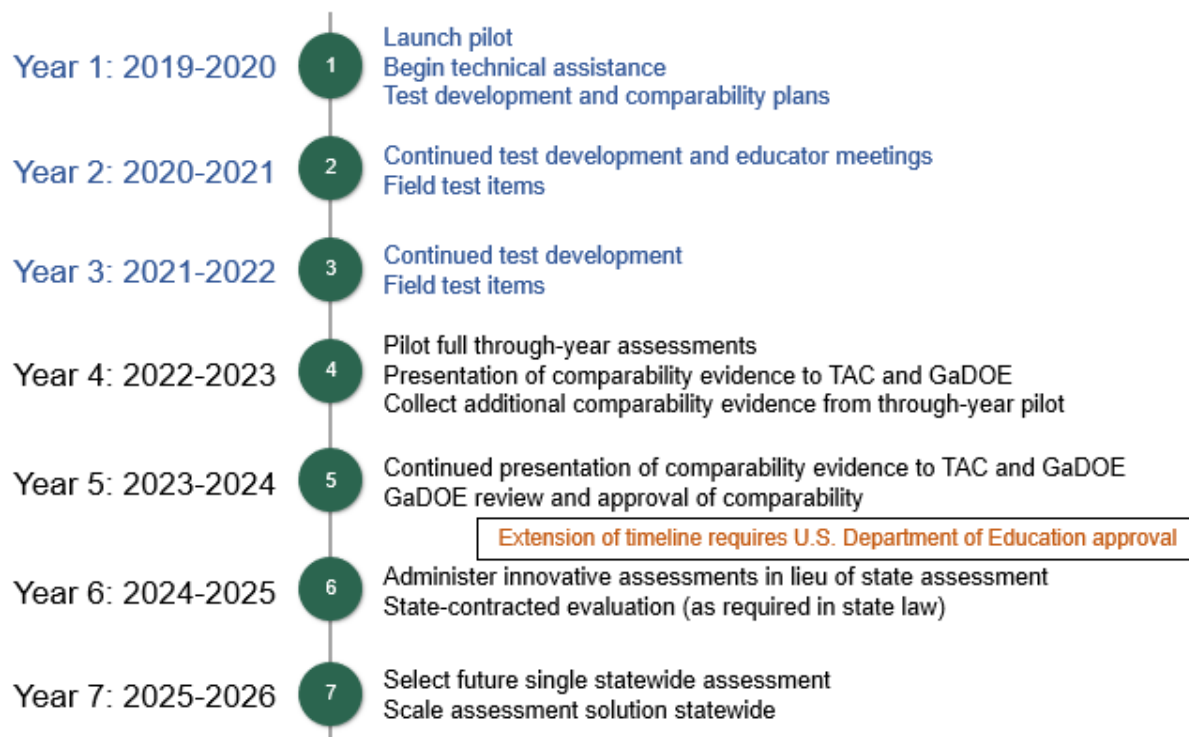
As Year 4 of the IADA begins (2022-2023), the GMAP consortium will include 18 districts and the Putnam Consortium will include 10 districts.

Timeline

Figure 1 presents a current timeline for the IADA, though detailed timelines vary by grade/content area and consortia. For example, both consortia are currently focused on developing English language arts and mathematics assessments in grades 3-8 and will begin developing science, social studies, and high school assessments at a later date.

The COVID-19 pandemic impacted both consortia and delayed some of their test development and implementation plans. These delays are not unique to Georgia’s IADA project and were experienced by many states. Any possible extension of the IADA timeline will require approval from the U.S. Department of Education.

Figure 1: IADA Implementation Timeline



Technical Assistance

Upon approval of Georgia's IADA application, the Georgia Department of Education (GaDOE) began providing technical assistance to the two participating consortia through a contract (procured through an RFP process) with WestEd. Through this contract, WestEd provides technical assistance hours to both consortia. These hours may include consultation time or other work (such as research or analyses) performed by WestEd, not to exceed the number of hours allocated to each consortium. Each consortium, in consultation with WestEd, determines how to utilize their hours. For the 2021-2022 school year (October 1, 2021 – September 30, 2022), each consortium had 12 technical assistance hours available. The Georgia MAP Assessment Partnership utilized 8.5 of their available hours. The Putnam Consortium utilized 3.5 of their available hours.

In addition to the technical assistance hours paid for by GaDOE, WestEd also convenes, coordinates, and facilitates two, two-day technical advisory committee (TAC) meetings annually to provide impartial advice and support to both consortia. The TAC will also make recommendations to GaDOE regarding comparability evidence and other technical issues associated with Georgia's demonstration authority. Two TAC meetings were held in 2021-2022, one in December 2021 and the second, which was split into two meetings, in March 2022 and June 2022. The second meeting was split to better support the consortia's timelines and availability of data.

TAC members include:

- **Dr. Wayne Camara**, Distinguished Scientist for Measurement Innovation at the Law School Admissions Council, is an expert in college and career readiness, psychometrics, assessment validation, and policy research. He has served on state technical advisory panels, including in Texas. In addition, Dr. Camara is past president of the National Council on Measurement in Education.
- **Dr. Gregory Cizek**, Guy B. Phillips Distinguished Professor of Educational Measurement and Evaluation at the University of North Carolina (UNC) at Chapel Hill, is an expert on standard setting, testing policy, classroom assessment, and detecting cheating on tests. He serves on multiple state assessment TACs, and prior to joining the UNC faculty, he managed national licensure and certification testing programs for American College Testing; served as a test development specialist for a statewide assessment program; and taught elementary school in Michigan.
- **Dr. Stuart Kahl**, founder and former CEO of Measured Progress, has more than 35 years of experience designing and implementing state programs that include innovative performance components, such as Vermont's portfolios, Kentucky's KIRIS program, Rhode Island's Distinguished Merit Program, and California's Golden State Examinations.
- **Ms. Lillian Pace**, Vice President of Policy and Advocacy at KnowledgeWorks, has extensive experience working with federal policymakers to create flexible policy environments that support personalized learning at scale. She has helped state and district leaders leverage federal policy to advance their vision for next-

generation education reform, and she has authored several national publications on competency education, innovative assessments, and high school redesign. Ms. Pace joined KnowledgeWorks after directing the U.S. House Subcommittee on Early Childhood, Elementary and Secondary Education. She spent nearly a decade on Capitol Hill advising policymakers on K–12 and higher education policy.

- **Dr. Stanley Rabinowitz**, Senior Technical Advisor at EdMetric, has over 30 years of experience in large-scale assessment, including work as a former assessment director. He has consulted extensively on standards, assessment, and school/educator accountability issues with researchers, policymakers, and assessment staff at national, state, and district levels in the United States and elsewhere globally. Dr. Rabinowitz has served on more than a dozen state and national technical advisory committees. He has directed the National Center for Standards and Assessment Implementation and served as the Project Management Partner for the Smarter Balanced Assessment Consortium. His expertise includes computer adaptive testing, developing and scoring technology-enhanced items, and assessing students across languages.
- **Dr. Stephen Sireci**, Distinguished University Professor in the Psychometrics Program, Director of the Center for Educational Assessment at the University of Massachusetts Amherst, and President of Sireci Psychometric Services, is a national expert in educational test development and evaluation, particularly issues of cross-lingual assessment, standard setting, and computer-based testing. Dr. Sireci serves on several advisory committees, including committees for Texas and Puerto Rico, and he is the former president of the National Council on Measurement in Education.

Finally, WestEd provides GaDOE with an annual written report summarizing the technical assistance needs addressed at the TAC meetings and through technical assistance hours, lessons learned, and recommendations for future pilot program activities.

WestEd Year 3 Annual Report:

- <https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/Flexibility/GA-IAPP-Annual-Report-110722.pdf>

Year 3 Annual Performance Report (APR)

Georgia is required to submit an annual performance report to the U.S. Department of Education at the conclusion of each year of the five-year IADA period. Georgia's third annual performance report was submitted on September 30, 2022.

IADA Annual Performance Report, Year 3: 2021-2022:

- <https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Documents/Flexibility/Georgia-Year3APR-September2022.pdf>

Progress Made in Year 3

In 2021-2022, both consortia made progress building out resources and training materials for their participating districts, including professional learning on using reported scores to guide instruction.

Validity research (how well the assessments measure the intended constructs and can support their intended uses and interpretations) was also a focus for both groups this year. The GMAP Consortium completed a testing engine evaluation study to determine if their specified computer algorithm adequately meets blueprint demands for a summative score in addition to other reported scores. The Putnam Consortium examined participation, fidelity, and sample representativeness within an item-level analysis to examine potential summative score approaches, given the prevalence of missing data. GMAP likewise considered sample representativeness, and both consortia moved forward with plans to assess classification accuracy and consistency as part of the broader requirements for establishing comparability.

Both groups worked with stakeholders to review and revise their assessment reports.

As mentioned previously, the Navvy assessment system was also acquired by Pearson Assessments (Pearson PLC), gaining access to extended expertise and resources.

Challenges in Year 3

Timelines

Both consortia have made considerable progress in developing their assessments. However, and understandably, the COVID-19 pandemic and other challenges specific to each consortium have delayed their original timelines. By the end of Year 3, both consortia were still working toward the necessary steps required to implement a full through-year pilot with summative score reporting for individual students.

The GMAP Consortium originally planned its through-year pilot in ELA and mathematics for Year 3, anticipating approval to implement their assessment in lieu of Georgia Milestones in Year 4. Science was planned to follow one year behind ELA and mathematics, while social studies had not yet been addressed. Due to the challenges referenced above, GMAP's updated timeline had the consortium piloting their through-year assessment in grades 3-8 ELA and mathematics in Year 4, one year behind the original timeline. Furthermore, the consortium's testing vendor, NWEA, made the decision to stop work on the GMAP project. Given this indefinite pause in work, the earliest GMAP can pilot their through-year assessment in grades 3-8 ELA and mathematics is Year 5 (2023-2024) if work resumes by summer 2023.

The Putnam Consortium originally indicated the Navvy assessment would be ready for use in ELA and mathematics for accountability purposes in 2018-2019 in Putnam County, upon the acceptance of their application to the State Board of Education, and in

other member districts in 2019-2020. Science was planned to be operational in 2021-2022. Putnam's updated timeline has the consortium piloting their through-year assessment in grades 3-8 ELA and mathematics in Year 4 (2022-2023).

Both consortia have prioritized the development of English language arts and mathematics assessments in grades 3-8; a timeline for high schools, science, and social studies assessments has not been provided.

Costs

Both consortia outlined expected development costs and funding mechanisms in Georgia's IADA application. In addition to the consortia's funding mechanisms, the Georgia legislature provided \$250,000 to each consortium in 2021. The Georgia Department of Education has similarly funded IADA technical assistance, despite overall cuts to the agency's testing budget, by providing \$174,691 in 2019-2020, \$105,908 in 2020-2021, \$120,083 in 2021-2022, and \$174,013 in 2022-2023.

At the end of Year 3, however, both consortia have indicated that the costs necessary to continue building and piloting their assessments are greater than their currently available funds. Work on the GMAP assessment has been placed on hold indefinitely by the vendor, NWEA, unless additional funding is secured. The Navy work continues, but the Putnam Consortium has also indicated that funding is a significant concern.

Technical Considerations

There are several technical steps required for U.S. Department of Education approval that both consortia have yet to complete, including:

- Methodology to categorize students into achievement levels based on their overall level of mastery
- External alignment study to determine if the assessments align to the depth and breadth of Georgia's academic content standards
- Statistical comparability analyses to determine if the innovative assessments are providing similar results to that of the state's existing assessment system
- Development of science, social studies, and high school assessments
- Updating mathematics assessments to align with the state's new mathematics content standards, which will be implemented in 2023-2024

In addition to the technical steps both consortia need to address, each consortium has unique technical steps to complete.

Georgia MAP Assessment Partnership:

- Conduct a full through-year field test
- Provide evidence that the GMAP-developed Range Achievement Level Descriptors (ALDs) are comparable to the Georgia Milestones ALDs and that comparisons of the knowledge and skills students have demonstrated are comparable between the two assessments

- Operationalize the through-year scoring model and demonstrate how the through-year results (fall, winter, and spring) or spring results will produce a final summative score based on assessment of the depth and breadth of Georgia’s academic content standards
- Provide evidence that the selected CAT (computer adaptive test) algorithm can adequately cover the range of knowledge and produce summative scores reliably

Putnam Consortium:

- Develop protocols for assessing students who are not present in one school/district for the full school year on the full depth and breadth of Georgia’s content standards
- Develop literacy and growth measures
- Develop a methodology for calculating the summative score based on assessment of the depth and breadth of Georgia’s academic content standards

Considerations for Year 4

Comparability

Participation in the IADA requires that the consortia demonstrate that their innovative assessments are comparable with the state’s assessment system, Georgia Milestones, prior to using the innovative assessments in lieu of Georgia Milestones during the IADA period. The project’s technical assistance provider, WestEd, developed comparability guidelines to support the consortia in their efforts to develop comparable assessments and collect necessary evidence. The guidelines were reviewed and approved by the IADA technical advisory committee (TAC).

The comparability guidelines are based on the original innovative assessment pilot application to the State Board of Education (SBOE), IADA regulations, and federal assessment peer review guidance. Acknowledging that the IADA is an opportunity to *pilot* new assessments, the comparability guidelines are less rigorous than the peer review requirements and focus solely on the elements required to allow the consortia to implement their assessments in lieu of Georgia Milestones during the IADA period. Comparability approval during the IADA period does not mean that development is complete, the assessment is ready to be scaled statewide, or the assessment is ready to be submitted for federal peer review. Rather, comparability approval means the consortia can implement their assessment in lieu of Georgia Milestones and results will be used for accountability purposes during the IADA period, while development is completed.

Both consortia planned to begin submitting comparability evidence to the TAC and GaDOE for review and approval in Year 4. However, both consortia have indicated that they are not prepared to submit evidence during the Fall 2022 review period. An additional review period is planned for Spring 2023.

New State Academic Content Standards

In accordance with state law, the SBOE conducted a review and revision of the mathematics and English Language Arts standards. The SBOE adopted new mathematics academic content standards in August 2021, to be implemented in 2023-2024, which coincides with Year 5 of the IADA. Both consortia have been building assessments aligned with the Georgia Standards of Excellence (GSE) in mathematics. However, at the beginning of the 2023-2024 school year, they will need to ensure their assessments are aligned with the new Georgia's K-12 Mathematics Standards in mathematics. Therefore, in addition to building and providing comparability evidence for their current mathematics assessments in Year 4, the consortia will also need to redevelop their mathematics assessments to align with the new standards, and provide evidence of the new alignment, prior to Year 5.

Similarly, the SBOE is expected to adopt new English Language Arts standards in 2022-2023, with implementation anticipated for the 2024-2025 school year. As with mathematics, both consortia will need to ensure their assessments are aligned with the new ELA standards by the beginning of the 2024-2025 school year.

GMAP Pause

As of June 2022, all GMAP development is paused indefinitely. There are currently no plans to continue work during Year 4. This pause raises questions about the path forward for the GMAP consortium and if work will resume or if development will be completed during the IADA period.

Putnam Comparability Evidence

The Putnam Consortium was scheduled to begin submitting comparability evidence to the TAC and GaDOE for review and approval in fall 2022. However, the consortium declined to submit evidence during this window. This delay raises questions about the path forward for the Putnam work, if development work and the full-year pilot is continuing, and when comparability evidence will be submitted.

Policy and Practical Considerations

The Year 2 Annual Report outlined several policy and practical considerations that are still relevant as Year 3 is complete and the project enters Year 4. The most pressing consideration, however, is how the formative nature of these through-year assessments can be preserved when they are reappropriated for summative uses.

The greatest promise of both the GMAP and Navvy assessments is that they are *formative* assessments administered throughout the school year, thereby providing teachers with timely information to inform instruction while it is still occurring. However, these assessments would no longer be formative if they inform summative processes

like accountability. If either assessment becomes the statewide assessment system, they will become through-year *summative* assessments.

Many of the technical and administrative decisions being made to support the current formative nature of the assessments will conflict with the federal and state requirements for summative, high-stakes assessments used for accountability purposes. Conversely, when decisions are made to ensure the assessments meet the federal and state requirements for summative, high-stakes assessments used for accountability purposes, those decisions will conflict with the formative goals of the assessments.

For example, the Navvy assessment allows teachers to select which standards to administer to students and when, according to their instructional plans and students' readiness. While this is a key feature that maximizes the instructional value of the assessment, it creates challenges for maintaining test security and ensuring equivalence across student scores, both of which are necessary for high-stakes accountability assessments. Importantly, the Navvy assessment does not have any participation requirements related to the number of standards that teachers must assess. This protocol enables teachers to maximize the value of the assessment while not overburdening students with excessive testing. In an accountability context, however, students must be assessed on the full depth and breadth of the state content standards, necessitating that students receive a zero for standards not assessed. This necessary change in participation requirements would increase the amount of time spent on high-stakes summative testing, change how teachers utilize the assessment, and/or misrepresent achievement for students who do not complete all required assessments.

As another example, the GMAP assessment prioritizes its RIT (Rasch Unit) scale score, a national norm-referenced measure, as well as its ability to assess students below and above grade level. While these are key features that maximize the instructional value of the assessment, the limited number of on-grade-level test questions directly aligned to Georgia's standards and comprising the summative score may not be sufficiently valid or reliable and may not meet the technical quality requirements of federal peer review.

Additionally, when previously formative assessments become high-stakes in nature, it is possible that behaviors and attitudes toward the assessments will change. While teachers are currently willing to administer the assessments in the middle of instruction to gain feedback to guide next steps, that may no longer be the case when the results will be used for high-stakes purposes. Teachers may want to spend more time reviewing material before testing students, which could result in a significant increase in "test prep" given that these through-year summative assessments would be administered more frequently than the current end-of-year model. Given delays in the development and implementation timeline for both consortia, it is not likely that this concern will be examined during the IADA period as originally expected.

In recent years, Georgia's Governor, State Legislature, and State Superintendent have taken several steps to reduce high-stakes testing. The number of required high-stakes

assessments have been reduced by eliminating 13 tests since 2016 (science in grades 3, 4, 6, 7, and Physical Science; social studies in grades 3, 4, 5, 6, 7, and Economics; ELA in 9th Grade Literature; and mathematics in Geometry). The test design for the state's current assessment system, Georgia Milestones, has been updated to reduce total student time spent on high-stakes testing (on average, a 4 ½ hour reduction in testing time across all four content areas, and a 2 ½ hour reduction in testing time for ELA and math only). Adopting a through-year summative assessment statewide would represent a departure from previous actions on high-stakes testing, instead substantially increasing the amount of time spent on high-stakes summative testing as such testing events would occur either three times per year (GMAP) or year-round (Navy).

During the first three years of the pilot, both consortia have struggled to make technical and administrative decisions given the inherent conflict between formative and summative uses of assessment. With the inability to reconcile both purposes, it is likely not possible to create a single assessment that can adequately serve both purposes well.

Policymakers and educators will need to identify the goals and purposes required of the state assessment system and identify the role for both formative and summative assessment. While a single assessment likely cannot serve both purposes, a system of assessments can. A cohesive, balanced system of assessment can provide both through-year formative information to support instruction while maintaining an end-of-year summative component that meets state and federal requirements. For example, a through-year formative component can provide instructionally-useful information free of high-stakes decisions, while an end-of-year summative component can serve accountability purposes. Both components can maintain the technical and administrative decisions best suited to their purposes. Such a system of assessment would enable the state to meet both instructional and accountability needs by using the right tests for the right purposes.

Next Steps

In 2022-2023, the GMAP consortium's work is paused by NWEA, while the Putnam consortium is expected to implement a full through-year assessment pilot. The Putnam consortium will also begin submitting comparability evidence to the TAC for review and feedback, though the consortium declined to submit evidence as expected in Fall 2022.

Georgia is required to submit an annual performance report to the U.S. Department of Education at the conclusion of each of the five years of the IADA period. Georgia's fourth annual performance report is expected to be due August 31, 2023.