Alignment Evaluation for the GAA 2.0 Assessments – Executive Summary

April 2021

Heather Peltier, Ph.D.
Alexandra Snipes, M.A.
Kevin Konort, M.Ed.
Ellen Forte, Ph.D.

Executive Summary

Introduction

Georgia's Alternate Assessment 2.0 (GAA 2.0) is the summative accountability assessment for Georgia students with significant cognitive disabilities and is based on the Extended Content Standards (ECSs) aligned to the Georgia Standards of Excellence (GSEs). This evaluation of the GAA 2.0 includes assessments of English language arts (ELA) and mathematics in grades 3 through 8 and high school, science in grades 5, 8, and high school, and social studies in grades 8 and high school.

The Georgia Department of Education (GaDOE) commissioned edCount, LLC (edCount), to conduct an independent evaluation of the alignment quality of the GAA 2.0 assessments with the ECSs. This report documents the methodology for and results of this independent alignment evaluation. The GaDOE intends to use the information gained via this evaluation to inform decisions about future item and assessment development and for federal peer review purposes.

Evaluation Methodology

edCount's approach to evaluating alignment quality between the aforementioned GAA 2.0 assessments and the ECSs encompasses the collection and evaluation of a comprehensive body of evidence. This evidence aligns with the demands of both the federal peer review criteria for alignment and the *Standards for Educational and Psychological Testing* (*The Standards;* AERA, APA, & NCME, 2014). Specifically, our alignment evaluation method addresses several key 'translation points' that connect standards to assessment scores. These translation points are where one component, such as a set of standards, is translated into the next component, such as the measurement targets, in the process chain that leads to tests and test scores. Our alignment framework, which reflects validity expectations in *The Standards* (AERA, APA, & NCME, 2014), takes considerably more information into account than any other approach to alignment evaluation. In particular, this framework considers the Achievement Level Descriptors (ALDs) and the blueprints to be critical components to an aligned system; no other approach to alignment evaluation includes these components and, as a result, no other method addresses the full spectrum of alignment concerns noted in *The Standards*, the federal peer review guidance, or the original statements from which all alignment methods have evolved (Forte, 2017a, 2017b, 2017c).

Because the GAA 2.0 assessments are alternate assessments based on alternate academic achievement standards, this framework also encompasses elements of the rigorous *Links for Academic Learning (LAL): An Alignment Protocol for Alternate Assessments based on Alternate Achievement Standards (AA-AAS)* method. The LAL methodology was developed by faculty at the University of North Carolina at Charlotte to support the evaluation of alignment between AA-AAS and the standards those assessments are intended to measure (Flowers, Wakeman, Browder, & Karvonen, 2007). In developing the LAL methodology, Flowers, Wakeman, Browder, and Karvonen (2007) first considered the policy and practice requirements for evaluating alignment quality for the AA-AAS. Namely, these assessments must[1]:

[1] Per *Improving the Academic Achievement of the Disadvantaged* (2003).

- Be aligned with a state's academic content standards;
- Promote access to the general curriculum; and
- Reflect the highest achievement standards possible.

The alternate expectation for achievement may focus on prerequisite skills or some partial attainment of the grade level, but students should still have the opportunity to meet high academic and performance expectations, to demonstrate a range of complexity, to achieve within their symbolic communication level, and to show growth across grade levels or grade bands.

Alignment evaluations provide primarily formative information; every test form is a new instance of the test, so a summative judgment is ungeneralizable and, thus, not helpful. Rather, the ideal outcome is that the agency that requires the assessment uses information about gaps or weaknesses to improve its practice. We encourage the GaDOE to use the findings and recommendations from this alignment evaluation to improve the meaningful use and interpretation of students' GAA 2.0 test results.

edCount evaluated the quality of alignment of the GAA 2.0 assessments in relation to the ECSs using the following evaluation questions:

1. How well do the claims and targets reflect the ECSs?
2. How well do the blueprints reflect the claims and targets?
3. How well do the ALDs capture the knowledge and skills expressed in the claims and targets?
4. How well do the items and test forms reflect the claims and targets?
5. How well do the assessments reflect fidelity with grade-level content and performance?
6. Is there a change in emphasis of age-appropriate content across grade levels?
7. Is the content accessible to students with varying levels of communicative competence?

Evaluation Findings and Recommendations

The alignment evaluation outlined in this report was designed with two key aims in mind: (1) satisfying federal peer review guidelines; and (2) providing meaningful and actionable feedback that contributes to the ongoing improvement of the GAA 2.0 assessments. Below, we highlight the results and recommendations of edCount's alignment evaluation.

For **Evaluation Question 1**, which addresses the relationship between the claims and targets and the Extended Content Standards (ECSs), edCount experts reviewed the claims and targets development process documentation provided by the GaDOE. Evaluators agreed that the process is **well documented** for all four content areas. Evaluators found that the claims and targets were developed to reflect the key concepts defined in the ECSs using a process that involved key content and assessment experts and were reviewed and revised as necessary until they were judged adequate by state experts. Evaluators commend the GaDOE for the review and revision process and the documentation that encompassed this portion of the development process, which used educator committees and state experts to judge the adequacy of the claims and targets.

For **Evaluation Question 2**, which addresses the blueprints and not test items, edCount evaluators judged the development of the blueprints for all GAA 2.0 content areas to be **well documented**. Evaluators noted that the documentation provided by the GaDOE details information on the

prioritizations of the ECSs that are represented in the blueprints, including information regarding the involvement of state content and assessment experts, as well as that of Georgia educators who work with students with disabilities and English learners.

Expert reviewers also evaluated the blueprints in terms of four aspects of alignment: Domain Concurrence, Range of Knowledge, Balance of Representation, and Complexity. Reviewers found all blueprints to **meet expectations** in each of these areas. For all four content areas, the blueprints do not reference any content that is not defined in the grade- or course-level ECSs, and the blueprints indicate how the Georgia Standards of Excellence (GSEs) will be sampled within the domains on each form. Evaluators commend the GaDOE on the representation of content domains within the blueprints, the thoughtfulness of the sampling of the ECSs represented in the blueprints, and the appropriateness of the vertical articulation of ECSs within the blueprints, across grade levels. While we commend the GaDOE on the development and use of a complexity framework for the GAA 2.0 that is content-specific and aligns clearly with the complexity framework within the ECSs, we recommend that, for purposes of streamlining documentation, the GaDOE consider including a description of form-level complexity expectations within the test specifications.

In relation to **Evaluation Question 3**, which targets ALD development, edCount evaluators found the ALD development process to be **well documented** for all four content areas. The GaDOE provided sufficient documentation to show that industry-standard practices were followed for the creation of ALDs.

Furthermore, across all content areas and grade levels, the ALDs meet expectations for Domain Concurrence and Differentiation. The ALDs for all assessments meet the criteria for **well aligned** for Domain Concurrence and **well differentiated** for Differentiation. Panelists had minor suggestions for edits to the language in order to strengthen the alignment between the claims and targets and the ALDs. edCount commends the GaDOE for developing ALDs that capture the key concepts from the content standards and reflect reasonable progressions in the sophistication of student knowledge, skills, and performance.

For **Evaluation Question 4**, which explores the relationship between test forms and the claims and targets, all forms meet the criteria for **well aligned** to the writer's intended GSE, which examines exact matches of items by panelists at the GSE- and claim level. We commend the GaDOE on the strong alignment of the assessment content with that intended by the test developers. For all four content areas, panelists aligned between 80% and 100% of items on each form to the GSE of record and 100% of items to the claim of record.

All forms across all grade levels and content areas meet the criteria for **well aligned** in terms of Domain Concurrence, Range of Knowledge, and Balance of Representation. All claims included on the blueprints are represented in these forms, no items reflect expectations not defined in the blueprints, and the GSEs within each claim are fully represented on the forms, reflecting the intended distribution outlined on the blueprints. We commend the GaDOE on the quality of these aspects of alignment for all forms included in this evaluation.

With the exception of the ELA grade 3 and science grade 5 test forms, all test forms evaluated meet the expectations of alignment to the intended complexity at the *item level*. In terms of alignment to the *form-level* complexity expectations within the technical documentation, all test forms except for the grade 5 science test meet alignment expectations. edCount recommends that the GaDOE consider

examining items aligned by panelists to different complexity levels than intended. We commend the GaDOE on the development and use of a standards-based, content-specific complexity framework for item development.

In terms of item alignment with ALDs, all forms **meet expectations**, with the exception of the grade 5 and high school science assessments, as these forms have between zero and three percent of items aligned by panelists to ALD Level 4. Given some unevenness in the representation of achievement at the higher or lower ends of the scale in a small number of test forms, the GaDOE may wish to incorporate ALD guidance within the test blueprints to ensure that every test form includes the range of opportunities that can recognize student achievement across the full score scale. We commend the GaDOE on the development of a majority of test forms that include a range of performance opportunities to appropriately distinguish student achievement across the full scale.

Evaluation Question 5 examines the degree to which the items from operational test forms reflect the intentions, as expressed through the ECSs, in terms of content and performance centrality. All test forms evaluated **meet the highest standards for alignment** in terms of fidelity with grade-level content and performance. We commend the GaDOE on the development of test forms that meet these alignment criteria for content and performance centrality.

For **Evaluation Question 6**, the goal of this review is to ensure students are not being assessed on identical content from year to year and that there are appropriate changes in the academic content in the ECSs and the items across grade levels. In an alternate assessment, evaluators expect that prerequisite skills exist in lower grades, and new content is introduced as the ECSs and items progress across grade levels. edCount experts indicated that the ECSs and items build upon themselves as they progress across the grade bands in ELA, mathematics, science, and social studies. Evaluators generally noted that most of the content within the ECSs and items is clearly unique and different across the grade levels, and there are limited instances cited of identical content repeated across grades. edCount commends the GaDOE on the development of **assessments that offer students the opportunity to demonstrate the depth and breadth of their learning across grade levels** in the ELA, mathematics, science, and social studies content areas.

Evaluation Question 7 examines the degree to which the content of the GAA 2.0 is accessible to students with various levels of communicative competence. Across panels, panelists determined that the GAA 2.0 assessments have accommodations available for a wide range of students. All panels concluded that **the majority of students should be able to access the test using the accommodations available or with flexibility built into the tasks**. The panels generally indicated that they felt the accommodations and supports that can be used for the GAA 2.0 assessments were defined. Panelist suggestions for increased accessibility include more explicit definitions or descriptions (e.g., tactile enhancements, assistive communication devices, replacement objects, pictures, etc.) and additional clarity around the accommodations and supports, in order to mitigate differences in administration of the assessment.

A majority of the panels did express concerns for access by students who are deaf and/or blind. edCount recommends that the GaDOE provide additional language in their trainings and administration manuals to address the administration of the GAA 2.0 for students who are deaf and/or blind and have specialists in visual impairment and hearing impairment review the Test Administration Manual and the assessment items in conjunction with these evaluation results, to identify any further recommendations for specific steps that can be taken to ensure accessibility for these students.

Across the four content areas, panelists indicated that students lacking a clear, intentional communication system even at the nonsymbolic level may not yet be prepared to actively engage with the GAA 2.0 assessments.

edCount commends the GaDOE for the provision of accommodations and supports available for a wide range of students. All panels judged that these **accommodations are appropriately defined**.

Conclusion

Evaluators found the development process for claims and test blueprints to be well documented and commend the GaDOE on the substantial role of Georgia educators in this process. Evaluators found all claims and targets to be closely related to the ECSs, while blueprints for all four content areas were judged as well aligned in the areas of Domain Concurrence, Range of Knowledge, Balance of Representation, and Complexity. Panelists judged all Achievement Level Descriptors (ALDs) to meet expectations in the areas of Domain Concurrence and Differentiation. For the 19 GAA 2.0 test forms reviewed through this alignment evaluation, for every aspect of alignment, all forms meet expectations, with the exception of alignment with the intended complexity for ELA grade 3, alignment with the intended complexity and form-level complexity targets for science grade 5, and representation of items across the ALDs for the grade 5 science and high school science forms. Panelists found all test forms to be well aligned in terms of content and performance expectations. Both the ECSs and the test forms evaluated show a progression across grade levels, and panelists judged that the degree of support and accommodations available for the GAA 2.0 provides accessibility for the vast majority of test takers.

These findings are notable given the breadth and depth of the methodology used, which exceeds the requirements laid out for state assessments through federal peer review. A test form is the product of a complex, multi-faceted development process, and the levels of alignment for these GAA 2.0 forms are the outcome of a clear and standardized test development process. We commend the GaDOE on the development of test forms that meet the rigorous expectations of this alignment evaluation and encourage the GaDOE to use the findings and recommendations from this alignment evaluation to improve the meaningful use and interpretation of students' scores.

References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing. (2014). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Flowers, C., Wakeman, S. Y., Browder, D. M., & Karvonen, M. (2007). *Links for Academic Learning: An alignment protocol for alternate assessments based on alternate achievement standards*. Charlotte, NC: National Alternate Assessment Center, University of North Carolina at Charlotte.

Forte, E. (2017a). *Evaluating alignment in large-scale standards-based assessment systems*. Washington, DC: Technical Issues in Large Scale Assessment SCASS of CCSSO.

Forte, E. (2017b). *Applications of an alignment evaluation framework.* Workshop presented at the Annual Meeting of the Northeastern Educational Research Association, Trumbull, CT.

Forte, E. (2017c). *Development and implementation of a comprehensive alignment evaluation framework.* Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Antonio, TX.

U. S. Department of Education (2003). *Title 1: Improving the academic achievement of the disadvantaged; Final rule*, 68 Fed. Reg. 236.