



Establishing a Validity Argument for
the
Georgia Alternate Assessment:
Putting the Pieces Together

Reidy Interactive Lecture Series
Portsmouth, NH
September 25-26, 2008

Towards a Technically Robust AA-AAS

- Our journey began with
 - Development of a new alternate assessment
 - Collaboration between Assessment, Special Education, and Curriculum
 - Significant investment in training of educators surrounding curriculum access
 - Ongoing documentation of decisions surrounding development and implementation
 - Involvement of Georgia's Technical Advisory Committee
 - Augmented with an AA-AAS expert

AA-AAS: Technical Quality

- Our philosophy –
 - Test development and technical documentation is a journey, not a destination (there is no end)
 - Validity must be thoughtfully attended to from the onset
 - Required our contractor to propose validity studies immediately and vetted proposal with our TAC
 - The GAA is as important as any other assessment we develop and administer

Core Belief and Guiding Philosophy

- The purpose of the Georgia Student Assessment Program is to
 - measure the level of student achievement of the state-adopted curriculum,
 - identify students failing to achieve mastery of content,
 - provide teachers with diagnostic information, and
 - to assist school systems in identifying strengths and weakness in order to establish priorities in planning educational programs.

Core Belief and Guiding Philosophy

- Georgia believes all students, including students with significant cognitive disabilities, can learn when provided with access to instruction predicated on the state curriculum
- Student achievement is best enhanced through the integration of high-quality standards-based instruction and assessment

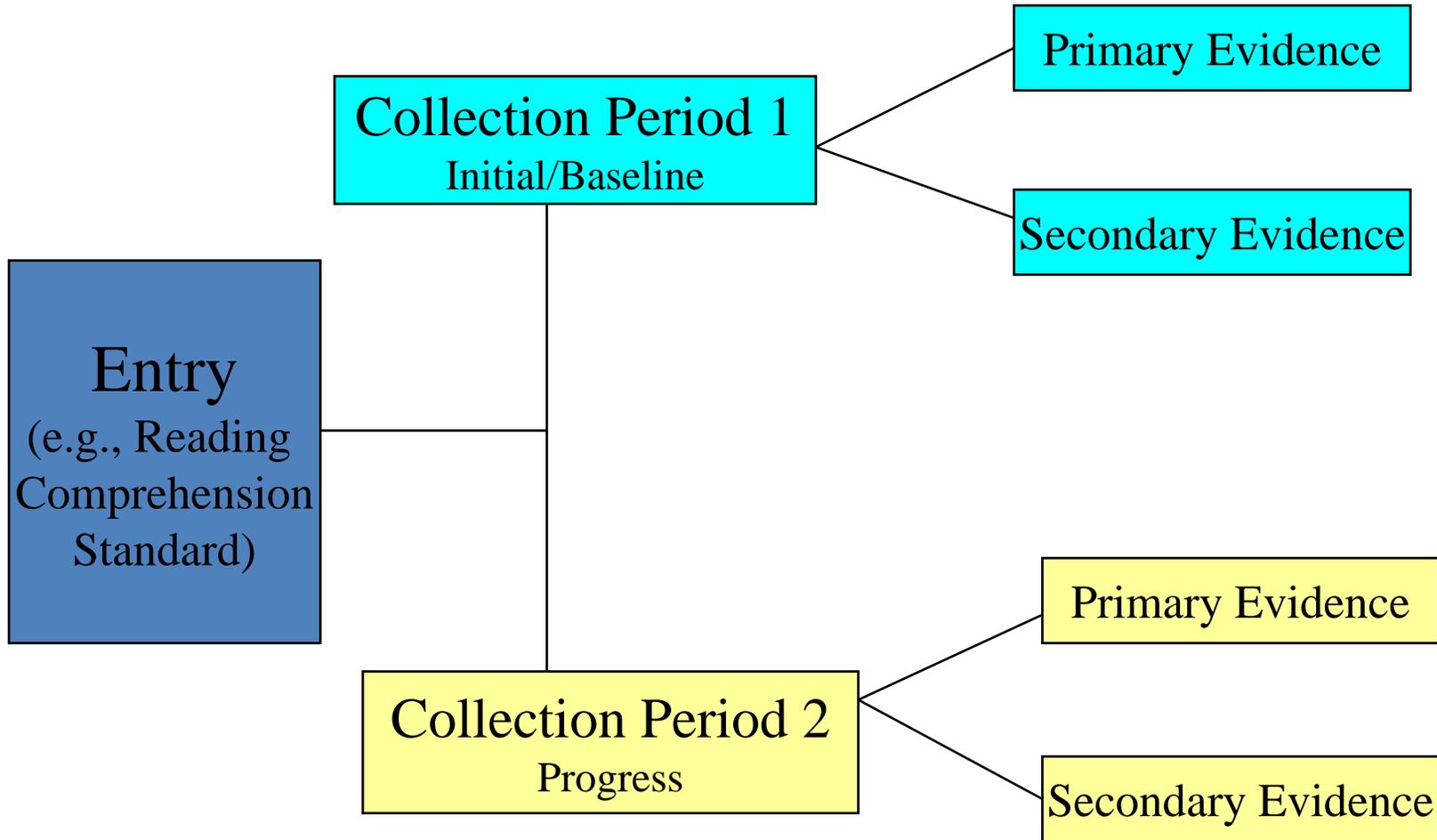
Georgia Alternate Assessment

- The GAA is a portfolio of student work provided as evidence that a student is making progress toward grade-level academic standards.
- Evidence provided must show instructional activities and student work that are aligned to specific grade-level standards.

Overview of the GAA

- There are two collection periods over the course of the school year.
- For each entry, teachers collect four pieces of evidence of student performance on tasks aligned to a specific content standard.
- This evidence should show the student's academic progress toward those standards.

Anatomy of a GAA Entry



GAA Purpose

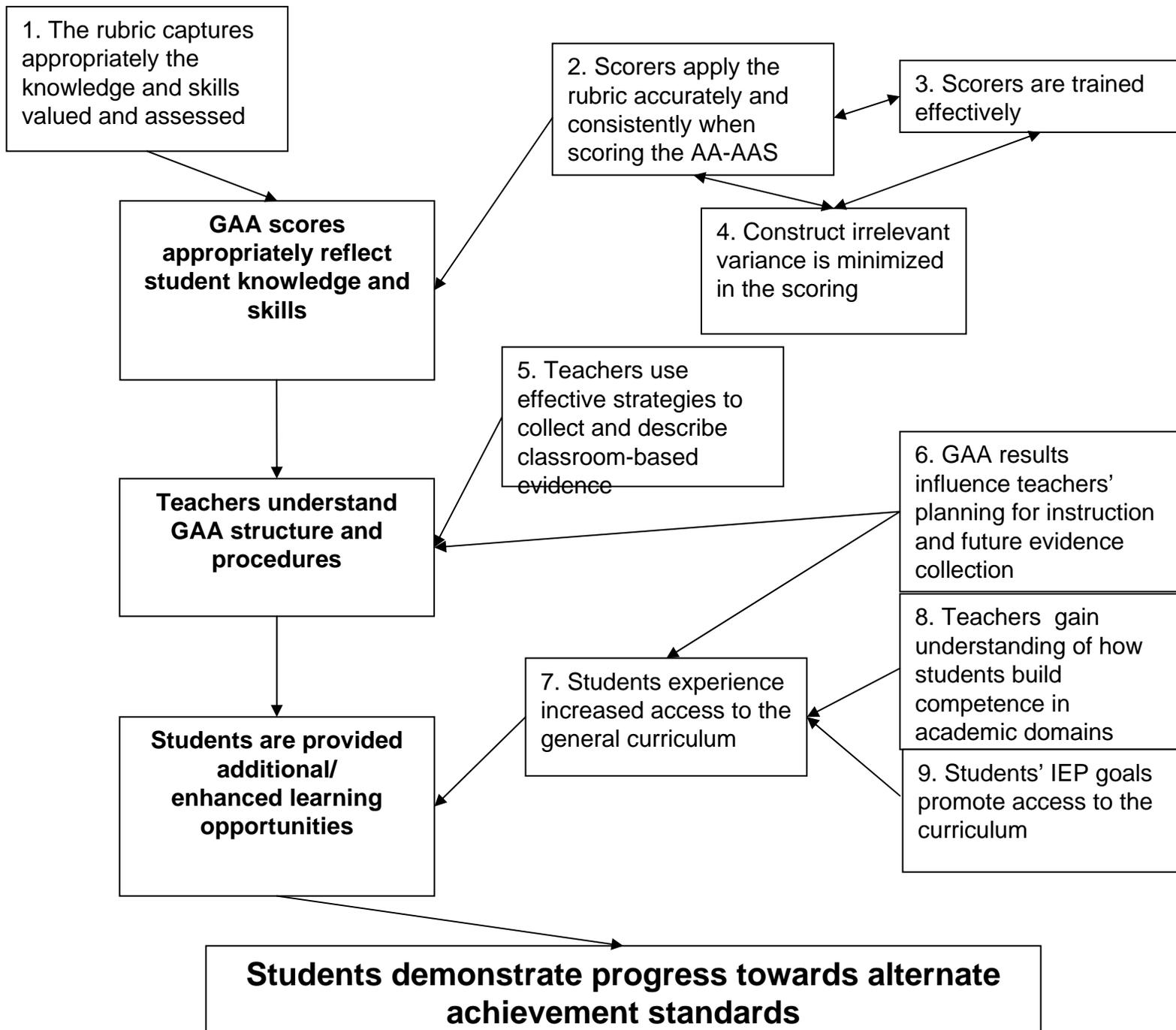
- To ensure all students, including students with significant cognitive disabilities, are provided access to the state curriculum
- To ensure all students, including students with significant cognitive disabilities, are given the opportunity to demonstrate their progress in learning and achieving high academic standards

Desired Outcomes of GAA

- Provides on-going documentation of student skills with a progressive history;
- Integrates instructional and assessment activities;
- Allows students to demonstrate strengths, knowledge, skills, and independence; and
- Provides meaningful ways to review student progress with parents.

For the stated purposes, outcomes, and uses to be carried out, the following must be true:

- State standards reflect important skills and concepts students should learn
- Teachers understand standards
- Teachers understand assessment structure and how to design aligned tasks (orchestrate meaningful learning opportunities for students)
- Assessment design appropriately reflects and captures student learning
- Scoring rubric picks up on salient aspects of student learning and is applied accurately and consistently
- Information emanating from the assessment process is meaningful to users and impacts future educational planning
- Students are provided additional/enhanced learning opportunities as a result of assessment feedback
- Student achievement increases as a result of the additional opportunities



Validity Plan

- Studies arose directly from the assessment purpose and/or from stakeholder questions/concerns
- In general, we seek information...
 - What are scorers attending to during scoring?
 - What characteristics and traits do experts use to evaluate portfolios?
 - What are teachers attending to in the creation of the portfolios?
How do they use information from the assessment process?
 - What is happening with instruction and student achievement over time?

GAA Validity Study #1

- Using think-aloud protocols, investigate scorer decisions about evidence and application of the scoring rubric in evaluating GAA portfolios
 - What construct-relevant information do scorers attend to?
 - What construct-irrelevant information do scores attend to?
 - To what degree might the construct-irrelevant information influence scores?

GAA Validity Study #2

- Convene focus group of 6 experts in curriculum, instruction, and assessment for students with significant cognitive disabilities
 - Each expert will be asked to independently organize 18 portfolios into three categories: Basic / Proficient / Advanced
 - Identify salient features and traits that distinguished portfolios at each level and construct a rubric
 - Evaluate the GAA scoring rubric using the criteria and features experts identified

GAA Validity Study #3

- What are teachers attending to in the creation of the portfolios? How do they use information from the assessment process?
- A group of teachers ($n = 30$) will be asked to:
 - “Talk through” one of their students’ portfolios, providing context and background information on how evidence samples were created and why they were selected for inclusion;
 - Discuss the rubric dimensions and how they apply to each entry;
 - Explain their understanding of the GAA results, and what changes (if any) they intend to make in instruction and/or the GAA process based on these results

GAA Validity Study #4

- Analyze longitudinal data regarding students' curricular access, instructional programming, and academic achievement
 - Collect information for a cohort of students for 3 consecutive school years, beginning with 2007-2008
 - Analyze students' GAA stage of progress and dimension scores, types of primary and secondary evidence submitted, and the content standards the portfolio addressed
 - Evaluate the influence of various student and instruction variables on GAA performance
 - Are learning progressions discernable?
 - Do certain profiles emerge?

GAA Validity Evidence

Source of Validity Evidence	Description of Validity Evidence Collected or to be Collected
Evidence Based on Test Content	<ol style="list-style-type: none">1. Test blueprint mapped to state standards2. Independent alignment evaluation of entries3. LAL alignment study4. Comparison of scoring rubric dimensions to potential evaluative criteria identified in expert scoring study (Study #2)
Evidence Based on Relations to Other Variables	<ol style="list-style-type: none">1. Investigate the relationship of various predictor student and instructional predictor variables to students' GAA performance (Study #4)
Evidence Based on Internal Structure	<ol style="list-style-type: none">1. Correlational analyses across rubric dimensions within entries and within and across content areas2. Exploratory factor analysis examining the scoring dimensions across academic areas

GAA Validity Evidence

Source of Validity	Description of Validity Evidence Collected or to be Collected
Evidence Based on Consequences of Testing	<ol style="list-style-type: none">1. Survey of teachers (n = 120) and parents (n = 120) about the acceptability, utility, and meaningfulness of the GAA (Study #4)2. Longitudinal data regarding quality of students' (n = 120) curricular access, instructional programming , and achievement (Study #4)
Evidence Based on Response Process	<ol style="list-style-type: none">1. Think-aloud protocols and interviews with a group of scorers (n = 14) regarding their scoring and evaluation of a sample of GAA portfolios. (Study #1)
Evidence Based on Reliability Scoring	<ol style="list-style-type: none">1. Scorer inter-rater reliability statistics2. Generalizability study3. Decision accuracy/contrasting groups study4. Inter-rater agreement data for scorers (n =14) who participated in the think-aloud investigation of scorers' behaviors and cognitions (Study #1)5. Agreement between experts' (n = 6) overall performance levels assigned to 18 portfolios and the scores assigned to the same portfolios by vendor (Study #2)

Synthesizing

- Keeping in mind the purpose of the assessment...

I to what degree might we have succeeded?

I to what degree have we failed?

I what next?

Contact Information

Melissa Fincher

Director of Assessment Research & Development

Georgia Department of Education

mfincher@doe.k12.ga.us / 404.651.9405

Andy Roach, Ph.D.

Assistant Professor of Counseling & Psychological
Services

Georgia State University

aroach@gsu.edu / 404.413.8176