
Considerations for the Alternate Assessment based on Modified Achievement Standards

Appendices

Appendix A: Individuals Involved in this Project

**Appendix B: List of Internet Resources for
Effective Curriculum and Instruction**

Appendix C: Tools for State Policymakers

Glossary

The contents of this publication were developed under cooperative agreement S283B050019 with the U. S. Department of Education. However, the contents do not necessarily represent the policy of the Department of Education, and you should not assume endorsement by the Federal Government.

For the full version of this white paper, visit:

http://nycomprehensivecenter.org/initiatives/inits_sp_assessment



University of the
State of New York
State Education
Department

APPENDIX A: INDIVIDUALS INVOLVED IN THIS PROJECT

Project Coordinator:

Larry Hirsch, NYCC

Project Assistant:

Laticha Sotero, NYCC

Project Manager/Editor:

Marianne Perie, Center for Assessment

Expert Panel:

Jamal Abedi, University of California, Davis

Chris Domaleski, Center for Assessment

Steve Dunbar, University of Iowa

Howard Everson, Fordham University

Claudia Flowers, University of North Carolina, Charlotte

Brian Gong, Center for Assessment

Meagan Karvonen, Western Carolina University

Suzanne Lane, University of Pittsburgh

Scott Marion, Center for Assessment

Jim Pellegrino, University of Illinois, Chicago

David Pugalee, University of North Carolina, Charlotte

Rachel Quenemoen, National Center on Educational Outcomes

Robert Rickelman, University of North Carolina, Charlotte

Katherine Ryan, University of Illinois, Urbana Champagne

Gerald Tindal, University of Oregon

Cathy Welch, University of Iowa

Phoebe Winter, Pacific Metrics

New York State Department of Education:

David Abrams, Assistant Commissioner for Standards, Assessment, and Reporting

Candy Shyer, Bureau Chief of Test Development, Office of State Assessment

Rebecca Cort, Deputy Commissioner, Office of Vocational and Educational Services for Students with Disabilities

APPENDIX B: LIST OF INTERNET RESOURCES FOR EFFECTIVE CURRICULUM AND INSTRUCTION

General Resources

K-8 Access Center (<http://www.k8accesscenter.org/index.php>)

Federally-funded project that has ended, but the Web site still hosts publications on access to curriculum in a variety of content areas as well as instructional issues.

National Center on Educational Outcomes (<http://nceo.info>)

“Provides national leadership in the participation of students with disabilities in national and state assessments, standards-setting efforts, and graduation requirements”

National Center on Accessing the General Curriculum (NCAC)
(<http://www.cast.org/policy/ncac/index.html>)

Web site “provides a vision of how new curricula, teaching practices, and policies can be combined to create practical approaches for improved access to the general curriculum by students with disabilities”

IEPs

NASDSE 2007 Standards-based IEP examples (available at projectforum.org)

Document that describes the 7-step process for creating standards-based IEPs, then applies those steps to two students. Sample IEPs for those students are provided.

Assistive Technologies

National Public Website on Assistive Technology (<http://www.assistivetech.net/>)

Hosted by the Center for Assistive Technology and Environmental Access at Georgia Institute of Technology. Provides access to the latest assistive technology.

ABLEDATA (<http://www.abledata.com/>)

NIDRR-sponsored project operated by ICF Macro. Provides a comprehensive database of assistive technology and rehabilitation devices, as well as publications and external links related to assistive technologies.

CBM, Rtl, and Progress Monitoring

National Center on Student Progress Monitoring (<http://www.studentprogress.org/>)

Web site that provides descriptions of what progress monitoring is, benefits/challenges, and provides a list of CBMs reviewed on site.

National Center on Response to Intervention (<http://www.rti4success.org/>)

Web site that provides information on what Response to Intervention is, how the tiered-system works, how Rtl can be used with different populations, and resources such as explaining the difference between curriculum-based measurement and curriculum-based assessment.

Easycbm.com (<http://www.easycbm.com>)

A Web site that provides free membership to have access to curriculum-based measurements, reports, and charts to track student progress.

RtI Resources (http://www.jimwrightonline.com/php/rti/rti_wire.php)

Web site containing information on RtI, how to choose interventions, how to use problem-solving teams, how to monitor student progress, and it provides graphs to monitor.

RtI Network (<http://www.RTInetwork.org>)

Web site providing information on RtI; how to develop and implement an RtI plan; and breaks down resources into Pre-K, K-5, middle school, high school, and parents/families.

Interventions and CBM (<http://www.interventioncentral.com>)

Web site providing information on interventions, progress monitoring, curriculum-based measurements, graphing data, and RtI.

Progress Monitoring (https://dibels.uoregon.edu/faq.php#faq_dib3)

Web site that provides free membership and provides information on progress monitoring, as well as access to probes for curriculum-based measurement.

APPENDIX C: TOOL FOR STATE POLICYMAKERS

This tool is simply a list of guiding questions for state policymakers considering the development of an AA-MAS. Beyond providing information to think about in deciding whether an AA-MAS fits well into a state’s current system, it also provides guidance during the design and development process. Each question is linked back to a section of the report for further information about topical considerations.

Topic	Guiding Question(s)	Relevant Chapter(s)/Section(s)
Appropriateness of Developing an AA-MAS	Are there students who cannot be appropriately assessed with the state’s current large-scale assessment system?	Chapter 2, pages 23–39 Chapter 4, pages 103–105 Chapter 10, pages 345–347
	How do we know that the problem is the format or design of the general assessment rather than a lack of opportunity to learn the material on the assessment?	Chapter 2, pages 32–39 Chapter 3, pages 52, 60
	Will this state reap more benefits from developing a new assessment targeted towards eligible students rather than focusing on their instruction through another means, such as professional development of teachers?	Chapter 3, pages 78–79 Chapter 5, pages 163–167
	What is your theory of action for how this assessment will improve student outcomes?	Chapter 9, pages 317–323
Identifying the Target Population	How do we identify the students who are eligible to take the AA-MAS?	Chapter 2, pages 30–39 Chapter 10, pages 343–347
	What are the characteristics of these students? (<i>This question may need to be answered for several different groups of students.</i>)	Chapter 2, pages 23–39 Chapter 10, pages 345; 356; 358
	What are your assumptions about these students’ ability to learn grade-level content and to show what they know?	Chapter 4, pages 103–105; 125–134
	Are these students different from students without disabilities who have performed poorly on the large-scale assessment? If so, how?	Chapter 2, pages 23–39 Chapter 4, pages 103–105

Appropriateness of current curriculum and Instruction	Do these students have standards-based IEPs that promote an opportunity to learn the standards-based curriculum? How do you know?	Chapter 3, pages 57–59; 61–63; 69–74
	What evidence exists to support a determination that these students will not achieve grade level proficiency in the current year because of the effects of their disability and not because of lack of opportunity-to-learn?	Chapter 2, pages 32–39 Chapter 3, pages 57–63 Chapter 9, pages 325–326
	What evidence exists to support the policy assumptions that these students are provided high quality access to the standards-based curriculum, through specialized instruction, services, and support?	Chapter 3, pages 59–69 Chapter 5, page 163
	What types of training and support are available for teachers of these students to improve participation and performance in the standards-based curriculum?	Chapter 3, pages 78–79 Chapter 5, pages 164–167
	What training, oversight, and monitoring processes are built into the system to ensure that IEP teams make high quality decisions about who participates in AA-MAS?	Chapter 3, pages 78–81
Appropriateness of AA-MAS for improving student outcomes	What is the nature of the barriers to these students’ participation on the general assessment?	Chapter 4, pages 103–105; 122–124; 135; 140 Chapter 5, pages 163–167 Chapter 6, pages 210–215
	How will this assessment provide a more accurate measure of the knowledge and skills of the participants compared with the general assessment?	Chapter 6, pages 224–232
	How will development of an AA-MAS yield more valid inferences about the students than other assessment approaches, such as improved general assessment design, appropriate accommodations, or development of an AA-GLAS?	Chapter 2, pages 27; 45–48 Chapter 3, pages 76–78; 82 Chapter 5, page 176 Chapter 6, pages 214–215 Chapter 8, pages 293–296; 300
	What are the relative costs and benefits of assessment development and implementation compared with other uses of resources, such as targeted staff	Chapter 6, page 197 Chapter 9, page 309 Chapter 10, page 342

	development on instructional and curricular interventions for teachers of struggling learners?	
	How will the inclusion of the AA-MAS as part of the state's assessment system lead to better instructional and curricular opportunities for these participating students?	Chapter 3, pages 59–69
Modified achievement standards	How do the performance expectations of the AA-MAS relate to those in the general assessment and the AA-AAS? Is Proficient on the AA-MAS similar in nature to Proficient on the general assessment? Is it closer to Basic? Or is it somewhere in between?	Chapter 7, pages 240–251 Chapter 8, pages 272–273; 276
	Is there an expectation that the AA-MAS may provide a stepping stone for students to reach Proficient on the general assessment? Or, is the expectation that students taking the AA-MAS are a unique population that will always need the modifications provided? Is a student who scores Advanced on the AA-MAS prepared to take the general assessment or an AA-GLAS or are they simply exceeding the criterion on their own assessment?	Chapter 2, pages 30–39, 45–47 Chapter 7, pages 242–243
Test Design	How will you carry your philosophy regarding the description of the students and their barriers to participation in a general assessment to your design of the AA-MAS?	Chapter 5, pages 186–192 Chapter 6, pages 208–224 Chapter 7, pages 240–246
	What type of assessment best fits your philosophy—a modification of your general assessment? AA-GLAS or a modification of the AA-GLAS?	Chapter 2, pages 45–48 Chapter 6, page 252
	If you choose to modify your general assessment, which types of modifications best match your philosophy regarding the students' barriers to participation in the general assessment?	Chapter 4, pages 140–142 Chapter 5, pages 186–192 Chapter 6, pages 208–224

	How do you intend to maintain the depth and breadth of the assessment but reduce the difficulty?	Chapter 4, page 135 Chapter 5, pages 161–162; 180–192 Chapter 6, pages 206–222
	How will you measure and demonstrate the degree of comparability between the AA-MAS and the general assessment?	Chapter 8
Documenting the technical quality or validating the AA-MAS	What are the important features of technical quality and validity that should be evaluated and documented throughout this process?	Chapter 9
Incorporating an AA-MAS into an existing assessment and accountability system	In which subjects/grades should we develop an AA-MAS?	Chapter 10, pages 340–342
	How does the AA-MAS fit between the AA-AAS and the general assessment? Do we expect to see smooth transitions from one assessment to the next?	Chapter 2, pages 48–49 Chapter 3, pages 52–53, 64–65, 85–86 Chapter 7, pages 246–247 Chapter 10, page 346
	How will you report results in a manner that will provide maximum information to teachers and parents?	Chapter 4, page 138 Chapter 10, pages 359–360

GLOSSARY

Achievement standard: A definition of a level of performance including both a minimum cut score and a written description that distinguishes the level of performance from other defined levels.

Accommodation: Changes in the administration of an assessment, such as setting, scheduling, timing, presentation format, response mode, or others, to provide better access to the assessment in a manner that does not change the construct intended to be measured by the assessment or the meaning of the resulting scores.

Accountability: The systematic use of assessment data and other information to evaluate the effectiveness of a program, such as an education system, for the purpose of rewarding desired outcomes and sanctioning undesirable outcomes.

Adaptation: A generalized term that describes a change made in the presentation, setting, response, or timing or scheduling of an assessment that may or may not change the construct of the assessment.

Adequate Yearly Progress (AYP): Under the *No Child Left Behind* Act, the minimum level of performance that states, school districts, and schools must demonstrate each year as measured by the proportion of students classified as Proficient or better to reach 100% Proficiency by 2014.

Alternate achievement standards: Cut scores and performance-level descriptors differentiating achievement on tests of content linked to grade level curriculum appropriate for students with the most significant cognitive disabilities.

Alternate assessment: An instrument used in gathering information on the performance and progress of students whose disabilities preclude them from valid and reliable participation in the general state assessment. Alternate assessments may be developed to measure alternate achievement standards, modified achievement standards, or grade-level achievement standards.

Annual Measurable Objective (AMO): a set of federally-required state-established benchmarks serving as targets for performance among and across student subgroups, schools, and districts.

Assessment: Any systematic method of obtaining evidence to draw inferences about people or programs. Assessment may include both formal methods, such as large-scale state assessments, or less formal classroom-based procedures, such as quizzes, class projects, and teacher questioning.

Bias. In a statistical context, a systematic error in a test score. In discussing test fairness, bias may refer to construct under-representation or construct-irrelevant components of test scores that differentially affect the performance of different groups of test takers.

Classification Errors: (aka, Type I/Type II errors). Errors made when the application of a cut score or other determinant results in “failing” a student/school/district when they should have passed (Type I error) or “passing” someone who should have failed (Type II error).

Cognition: How students represent knowledge and develop competence in a subject domain.

Cognitive architecture: The information processing system that determines the flow of information and how it is acquired, stored, represented, revised, and accessed in the mind.

Cognitive complexity: An individual psychological characteristic related to the type of thinking a student would need to do in order to correctly answer an item or task, including the number of mental structures a student would have to use, how abstract the item structures were, and how elaborately the structures interacted with each other.

Comparability: The degree to which similar inferences can be made from the outcomes of two or more assessments.

Construct: As applied to assessment, the complete set of knowledge, skills, abilities, or traits representing a particular domain of knowledge, such as American history, reading comprehension, study skills, writing ability, logical reasoning, honesty, intelligence, and so forth.

Content domain: The set of behaviors, knowledge, and skills to be measured by a test, represented in a detailed specification and often organized into categories by which items are classified.

Content standards: Statements of the knowledge and skills that students are expected to learn. Content standards should drive instruction and test construction.

Curriculum: The knowledge and skills in subject matter areas that teachers are supposed to teach and students are supposed to learn including a scope or breadth of content in a given subject area and a sequence for learning.

Curriculum-Based Measurement (CBM): A method teachers use to determine how students are progressing in basic academic areas such as math, reading, writing, and spelling by testing students weekly using a short measure that is then graphed and analyzed to see if the progress is sufficient to meet the target.

Cut score: A point on a score scale at or above which test takers are classified in one way and below which they are classified in a different way. For example, if a cut score is set at 60, then people who score 60 and above may be classified as “passing” and people who score 59 and below classified as “failing.”

Decision consistency: A measure of the reliability of the classification decision. Decision consistency estimates the extent to which, if an examinee were administered a test on two separate occasions, the same classification decision (whether pass or fail) would be made.

Declarative knowledge: Information about “the way the world is.”

Depth-of-knowledge: Degree of depth or performance complexity required to understand/perform academic content/process found in content standards or assessment items; a description of different ways students interact with content measured by how deeply students must understand the content in order to respond.

Difficulty: In assessment, the proportion of respondents answering the item correctly. Conceptually, it is based on underlying knowledge and cognitive processes required to answer an item correctly.

Differential Item Functioning (DIF): A statistical property of a test items in which different groups of test takers who have the same total test score have different performance on particular items.

Disability Category: Assignments that qualify a child for special education and related services; different from a medical diagnosis. Federal law (IDEA 2004, Part B) has 13 disability categories that States must use to determine if students, ages 3-21, are eligible to receive special education and related services: Autism, Deaf-blindness, Deafness, Emotional disturbance, Hearing impairment, Mental retardation, Orthopedic impairment, Other health impairment, Specific learning disability, Speech or language impairment, Traumatic brain injury, Visual impairment including blindness, or Multiple disabilities.

Distractor: An incorrect option presented to an examinee in a multiple choice item.

Domain Sampling. The process of selecting test items to represent a specified universe of performance.

Dynamic Evaluation: As used in the context of validity, dynamic evaluation refers to the notion that evaluative judgments will be updated as new information about the assessment system is presented. In other words, dynamic evaluation refers to the idea that the evaluation continues to move (or adjust) as new information is gathered.

General Assessment: Assessments given to the majority of students at each grade level such as the state end of year tests.

Grade-level achievement standard: A minimum cut score and written description that provide an expectation for a level of performance aligned to the grade level in which a student is enrolled or that matches his biological age.

Guiding Philosophy: The fundamental beliefs or set of assumptions that guide the conception, development, implementation, and continuous improvement of an approach, program, practice, or policy.

Individualized Education Program (IEP): A written plan and legal document designed to meet the unique educational needs of one child, as defined by federal regulations under the Individuals with Disabilities Education Act (IDEA). An IEP describes a child's present level of functioning; specific areas that need special services; annual goals; short-term objectives; services to be provided; and the method of evaluation to be implemented for children 3 to 21 years of age who have been determined eligible for special education.

Instruction: The methods of teaching and the learning activities used to help students master the content and objectives specified by a curriculum and encompasses the activities of both teachers and students.

Interim flexibility (aka 2% proxy): The practice of allowing states that meet certain criteria to count as proficient for purposes of AYP a portion of the students with disabilities; the portion is determined by dividing 2 percent by the percent of students with disabilities in the state.

Interpretative argument: A plan specifying the proposed interpretations and uses of test results by laying out the network of inferences and assumptions leading to the observed performances to the conclusions and decisions based on the performances.

Item format: The variety of test item structures or types that can be used to measure examinees' knowledge, skills, and abilities, typically including multiple-choice or selected-response, open-ended or constructed-response, essay, or performance task.

Learning progression: Description of successively more sophisticated ways of reasoning within a content domain that follow one another as students learn.

Measurement error: The differences between observed scores and the theoretical true score; the amount of uncertainty in reporting scores; the degree of inherent imprecision based on test content, administration, scoring, or examinee conditions within the measurement process that produce errors in the interpretation of student achievement.

Metacognition: The set of skills and processes that allow one to reflect on, monitor, adjust and direct one's own thinking and learning.

Modified achievement standard: A minimum cut score and written description that provide an expectation for a level of performance aligned to grade level content standards but less rigorous than a grade-level achievement standard.

Modification: Changes made in both instructional and assessment situations that are individualized to student needs. In the context of assessment, changes are made to the content, format, and/or administrative procedures of a test in order to accommodate test takers who are unable to take the original test under standard test conditions. Unlike *accommodations*, modifications may directly or indirectly compromise the validity of the content standard by changing the construct. Modifications include a much wider range of supports and instructional scaffolding than do accommodations but can be effectively used in combination with accommodations in instructional and assessment situations when individualized to the student's strengths and needs. Modifications are intended to allow for meaningful participation and enhanced learning.

No Child Left Behind (NCLB): The 2001 reauthorization of the Elementary and Secondary Education Act that added new requirements for annual student testing and annual measurable objectives with a focus on improving achievement of all students and reducing the achievement gap.

Opportunity to learn: The provision of learning conditions, such as curriculum, courses, and instruction, including suitable adjustments, to maximize a student's chances of attaining the desired learning outcomes, such as the mastery of content standards.

Parallel Forms: Two or more assessments that provide similar outcomes (true scores) of the construct being measured.

Performance Index: A measure that weights scores at each performance level and awards a school partial credit for students whose achievement improves, even though they may not yet be proficient, and can be included in determining the adequate yearly progress (AYP) of the school.

Portfolio (assessment): An assessment comprising the collection and analysis of examinee work samples, typically consisting of performance tasks gathered over a specific period of time; often used to assess special populations who have difficulty with standard paper-and-pencil assessments.

Procedural knowledge: Information about “how things are done.”

Progress Monitoring: The process of collecting and evaluating data to make decisions about the adequacy of student progress toward a goal by evaluating the student’s actual rate of change compared to the expected rate of change.

Prompt. Any form of verbal, nonverbal, or physical cue to structure, pace, or signal a response to be made by the student. Examples include verbalisms like, ‘continue,’ next,’ now what,’ or reminders of each step; physical guidance is an example of a prompt.

Reliability: The characteristic of test scores of being dependable, generally conceptualized as stability or consistency over both time and items.

Response to Intervention (RTI): A comprehensive, multi-step process that closely monitors how the student is responding to different types of services and instruction.

Sampling Error: The error associated with observations from a sample instead of the whole population, used to quantify the expected range within which the true population value might be located relative to the sample data.

Scaffolding: An approach to enhancing items derived from supports provided during learning that are gradually removed when learning becomes solidified and/or the learner becomes more independent. Includes any type of structural assistance introduced to organize information or guide responses embedded in the presentation of the item or task. These supports are not intended to change the construct being measured.

Standard setting: An activity in which a procedure is applied systematically to gather and analyze human judgment for the purpose of deriving one or more cut scores for a test.

Standards-based IEP: An individualized education plan that specifically refers to instruction of the state’s academic standards for the student’s enrolled grade and focuses on aligning instruction of students with disabilities to the academic content that all students at that grade level should know and be able to do.

Student with disabilities (SWD): In the Individuals with Disabilities Act, a student with disabilities is defined as “a child evaluated in accordance with §§300.530-300.536 as having mental retardation, a hearing impairment including deafness, a speech or language impairment, a visual impairment including blindness, serious emotional disturbance (hereafter referred to as emotional disturbance), an orthopedic impairment, autism, traumatic brain injury, another health impairment, a specific learning disability, deaf-blindness, or multiple disabilities, and who, by reason thereof, needs special education and related services.”

Test domain: The portion of all knowledge and skill in a subject matter area that is selected to be assessed because there is consensus that it represents what is important for teachers to teach and for students to learn.

Test specifications: A detailed description for a test that specifies the number or proportion of items that assess each content and process/skill area. Aka, test blueprint.

Theory of Action: Originally drawn from sociology and organizational studies, theory of action is used in the education context to refer to higher level view of the interpretative argument. Essentially, it provides an overview of how the specific components of the testing/educational system are intended to work in concert to bring about the desired aims.

Universal design: The creation of products and environments meant to be usable by all people, to the greatest extent possible, without the need for adaptation or specialization.

Universal design for learning: A framework for designing educational environments that enables all learners to gain knowledge, skills, and enthusiasm for learning, by simultaneously reducing barriers to the curriculum and providing rich supports for learning.

Validity: The extent to which inferences and actions made on the basis of a set of scores are appropriate and justified by evidence. It is the most important aspect of the quality of a test. Validity refers to how the scores are used rather than to the test itself.

Validity argument: An evaluation of the completeness and coherence of proposed interpretations and uses of test results, based on both empirical evidence and logic, as specified by the interpretative argument.

Validity evaluation: The full set of activities related to evaluating the proposed interpretations and uses of test results and includes the interpretative and validity arguments as well as the validity studies plan and the actual studies themselves.

Working memory: A kind of cognitive energy level or “resource” that exists in limited amounts, with substantial individual variations.