Introduction to the Special Issue on Alternate Assessments Based on Modified Academic Achievement Standards: New Policy, New Practices, and Persistent Challenges

Ryan J. Kettler and Stephen N. Elliott

Learning Sciences Institute, Peabody College at Vanderbilt University

This introduction to the special issue titled Alternate Assessments Based on Modified Academic Achievement Standards: New Policy, New Practices, and Persistent Challenges addresses the federal policy introducing the new alternate assessment for students with persistent academic difficulties, as well as related implementation issues that will be more thoroughly considered throughout the journal. Three guidelines are identified within the policy for alternate assessments based on modified academic achievement standards (AA-MASs), including that (a) a state’s grade-level academic content standards cannot be modified for an AA-MAS, (b) a state’s general test can be modified for an AA-MAS, and (c) a state’s achievement standards can be modified for an AA-MAS so long as they remain on grade level. This article introduces key issues including identification of students eligible for an AA-MAS, the degree of modification that can be applied to develop an AA-MAS, and the current state of AA-MAS development across the nation. The article concludes with overviews of each contribution in the journal.

The final regulations for the No Child Left Behind (NCLB) Act of 2001 indicate that a small group of students with disabilities are allowed to show proficiency through an alternate assessment based on modified academic achievement standards (AA-MASs; U.S. Department of Education [ED], 2007a, 2007b). These students, who may constitute up to 2% of all who are reported proficient within a school, can take a version of the regular assessment test with modifications. Modifications are changes to a test’s content or item format. The regulations strongly emphasize that although modifications may make a test easier, out-of-level (i.e., below grade level) testing is not acceptable, leaving developers and users of these AA-MASs to determine at which point a test has become too easy to be within the intended level. Modifications to large-scale achievement tests, like testing accommodations, are intended to facilitate access to the assessment for students with special needs, so that their scores can be meaningfully compared with the scores of students who take the standard test. If this can be accomplished, better assessment and accountability for students with disabilities would appear to be the result.

Correspondence should be sent to Ryan J. Kettler, 410E Wyatt Center, Peabody #59, 230 Appleton Place, Nashville, TN 37203-5721. E-mail: r.j.kettler@vanderbilt.edu
Challenges related to this new policy include (a) identifying participants who qualify, (b) determining the degree of modification to achievement standards that is appropriate, and (c) meeting the financial and technical burden associated with designing, validating, and implementing these new AA-MASs. This theme issue, titled Alternate Assessments Based on Modified Academic Achievement Standards: New Policy, New Practices, and Persistent Challenges, brings together assessment researchers, test developers, policy analysts, and education leaders from the government to address these challenges. Collectively, across the six articles and the two commentaries, these authors probe federal policy, test design and implementation issues, and anticipated consequences associated with the new AA-MAS component of the NCLB Act. In the remainder of this article, we start the examination of key elements of an AA-MAS and provide the big picture on how each of the invited authors contribute to a detailed elaboration of the challenges associated with these elements.

**KEY ELEMENTS OF AN AA-MAS SYSTEM**

In 2007, the ED published *Modified Academic Achievement Standards: Non-Regulatory Guidance* (2007a). This document “provides States with detailed information about how to use and implement modified academic achievement standards” (p. 8). The document uses a question-and-answer format to advance understanding and implementation of AA-MAS policies. One of the most useful features of the document is its appendix, where a table titled “Characteristics of Alternate Assessments” is provided. This table compares alternate assessments based on alternative academic achievement standards, modified academic achievement standards, and grade-level academic achievement standards. We have reproduced this table as a resource (see Table 1).

To facilitate an efficient and focused understanding of the ED guidance document as it pertains to building an AA-MAS, we quote from it and organize the information around three key elements of a standards-based assessment system: academic content standards, alternate tests, and modified achievement standards. Please note we have added all the italics in these quotes for emphasis. With regard to academic content standards—the subject matter content and skills teachers are expected to teach, students are expected to learn, and tests are intended to measure—the ED guidance document makes two important points:

- “The expectations of content mastery are modified, not the grade-level content standards themselves.” (ED, 2007a, p. 9, Introduction)
- “The primary reason for requiring Individualized Education Program (IEP) goals based on grade-level academic content standards is to ensure that students who participate in an assessment based on modified academic achievement standards receive instruction in grade-level content standards so that they can make progress towards meeting grade-level proficiency.” (ED, 2007a, p. 28, E-2)

With regard to alternate tests, the ED guidance document makes the following points:

- “Alternate assessments based on modified academic achievement standards are intended to be challenging for a limited group of students whose disability has prevented them from attaining grade-level proficiency. These students must have access to curriculum based
TABLE 1
Characteristics of Alternate Assessments

<table>
<thead>
<tr>
<th>Achievement standard</th>
<th>Alternate Assessment Based on Alternate Academic Achievement Standards (1%)</th>
<th>Alternate Assessment Based on Modified Academic Achievement Standards (2%)</th>
<th>Alternate Assessment Based on Grade-Level Academic Achievement Standards</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• An alternate academic achievement standard is an expectation of performance that differs in complexity from a grade-level achievement standard, usually based on a very limited sample of content that is linked to but does not fully represent grade-level content.</td>
<td>• A modified academic achievement standard is aligned to grade-level content standards for the grade in which a student is enrolled and challenging for eligible students, but may be less difficult than grade-level achievement standards.</td>
<td>• A grade-level academic achievement standard defines a level of &quot;proficient&quot; performance equivalent to grade-level achievement on the State's regular assessment.</td>
</tr>
<tr>
<td></td>
<td>• May be defined for grade clusters, e.g., 3–5.</td>
<td>• Achievement standards must include 3 levels of performance, cut scores that distinguish one level from another, and descriptions of the content-based competencies associated with each level.</td>
<td>• Achievement standards must include 3 levels of performance, cut scores that distinguish one level from another, and descriptions of the content-based competencies associated with each level.</td>
</tr>
</tbody>
</table>

Setting standards

• Requires a "documented and validated standard setting process." A detailed description of the procedures used, the qualifications of panelists (which must include persons knowledgeable about the State’s content standards and experienced in standards setting and special educators who are most knowledgeable about students with disabilities), the final cut scores, and performance level descriptors must be submitted for peer review.

• Requires a "documented and validated standard setting process." A detailed description of the procedures used, the qualifications of panelists (which must include persons knowledgeable about the State’s content standards and experienced in standards setting and special educators who are most knowledgeable about students with disabilities), the final cut scores, and performance level descriptors must be submitted for peer review.

• Requires a "documented and validated standard setting process." A detailed description of the procedures used, the qualifications of panelists (which must include persons knowledgeable about the State’s content standards and experienced in standards setting and special educators who are most knowledgeable about students with disabilities), the final cut scores, and performance level descriptors must be submitted for peer review.

(Continued on next page)
**TABLE 1**  
Characteristics of Alternate Assessments *(Continued)*

<table>
<thead>
<tr>
<th>Content standards on which the test is based</th>
<th>Alternate Assessment Based on Alternate Academic Achievement Standards (1%)</th>
<th>Alternate Assessment Based on Modified Academic Achievement Standards (2%)</th>
<th>Alternate Assessment Based on Grade-Level Academic Achievement Standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cap</td>
<td>• “Extended” standards may include substantially simplified content, including prerequisite skills.</td>
<td>• Grade-level</td>
<td>• Grade-level</td>
</tr>
<tr>
<td>Out-of-Level Assessments</td>
<td>• May include reduced coverage and/or simplification of grade-level content, based on “extended” standards. Format may permit variation in test content for individual students if results can be aggregated.</td>
<td>• Built on grade-level content but with easier items.</td>
<td>• Grade-level content</td>
</tr>
<tr>
<td>IEP</td>
<td>• State and LEA</td>
<td>• State and LEA</td>
<td>N/A</td>
</tr>
<tr>
<td>State guidelines define who is eligible</td>
<td>• Permitted only if consistent with the regulation, i.e. documented and validated standards-setting process employed.</td>
<td>• Not permitted because out-of-level assessments do not assess grade-level content.</td>
<td>• Not permitted because out-of-level assessments do not assess grade-level content.</td>
</tr>
<tr>
<td></td>
<td>• Must include annual measurable IEP goals and benchmarks or short term objectives.</td>
<td>• Must include annual measurable IEP goals that are based on grade-level content standards.</td>
<td>• Must include annual measurable IEP goals.</td>
</tr>
<tr>
<td></td>
<td>• Student with the most significant cognitive disabilities.</td>
<td>• Student whose disability has precluded the student from achieving proficiency, as demonstrated by objective evidence of the student’s performance and whose progress is such that, even if significant growth occurs, the student’s IEP team is reasonably certain that the student will not achieve grade-level proficiency within the year covered by the IEP.</td>
<td>• Student with a disability who cannot take the regular assessment with accommodations.</td>
</tr>
<tr>
<td></td>
<td>• IEP Team makes the decision regarding the appropriate assessment.</td>
<td>• IEP Team makes the decision regarding the appropriate assessment.</td>
<td>• IEP Team makes the decision regarding the appropriate assessment.</td>
</tr>
</tbody>
</table>

on grade-level content standards and, therefore, must be assessed with a measure that is also based on grade-level content standards, although the assessment may be less difficult than the general assessment. The content standards are not modified, but the achievement expectations are less difficult than those on the general test. This means that the same content is covered in the test, but with less difficult questions overall.” (ED, 2007a, p. 20, D-1)

- “. . . alignment with grade-level content standards is the foundation of an alternate assessment based on modified academic achievement standards. An alternate assessment based on modified academic achievement standards must cover the same grade-level content as the general assessment.” (ED, 2007a, p. 21, D-2)
- “A State may modify an existing assessment or develop a new assessment.” (ED, 2007a, p. 24, D-8)
- “Some States have suggested replacing the most difficult items on the general test with simpler items appropriate for the grade level, while retaining the same coverage of the content standards. Others have suggested modifying the same items that appear on the grade-level assessment by simplifying the language of the item or eliminating a distracter in multiple-choice items (e.g., having 3 options to choose from, instead of 4).” (ED, 2007a, p. 25, D-12)
- “. . . an alternate assessment based on modified academic achievement standards must be aligned with State content standards in the same manner as the general assessment . . .” (ED, 2007a, p. 26, D-12)
- “Built on grade-level content but with easier items.” (ED, 2007a, p. 51, Appendix A)

Finally, with regard to the topic of modified achievement standards the ED guidance document makes the following points:

- “A modified academic achievement standard is an expectation of performance that is challenging for eligible students, but is less difficult than a grade-level academic achievement standard. Modified academic achievement standards must be aligned with a State’s academic content standards for the grade in which a student is enrolled. Thus, only the academic achievement standards are modified, not the content standards on which those modified academic achievement standards are based. Although the assessment and modified academic achievement standards for a particular grade must be challenging for eligible students, they may be less difficult when compared with the general test and grade-level academic achievement standards . . . . The characteristics of modified academic achievement standards are the same as those described in the Title I assessment regulations for grade-level achievement standards. That is, they must be aligned with the State’s academic content standards; describe at least three levels of achievement; include descriptions of the competencies associated with each achievement level; and include assessment scores (cut scores) that differentiate among the achievement levels and describe the rationale and procedures used to determine each achievement level.” (ED, p. 14, B-1)
- “Modified academic achievement standards are not based on academic content standards that have been modified or restricted. Alternate academic achievement standards, on the other hand, are based on a very limited sample of content that is linked to
grade-level content standards but may not fully represent grade-level content and may include substantially simplified content.” (ED, 2007a, p. 15, B-2)

- “Proficient performance on an alternate assessment based on modified academic achievement standards . . . is expected to represent understanding of grade-level content based on a less rigorous assessment.” (ED, 2007a p. 21, D-2)

- “An alternate assessment based on modified academic achievement standards must have at least three achievement levels. If a State’s general assessment has six achievement levels, the alternate assessment based on modified academic achievement standards does not need to have all six achievement levels. In such cases, decreasing the number of achievement levels to three, instead of six, would allow the design of a test with fewer items, while covering the same grade-level content standards as the general assessment.” (ED, 2007a, p. 23, D-5)

- “…an accurate and meaningful measure of their achievement will require a different definition of proficiency in the form of a more appropriate test and related academic achievement standards. Setting a lower cut score on the general assessment does nothing to make the test more accessible or understandable.” (ED, p. 24, D-9)

We believe this collective guidance leads to three important conclusions from the standpoint of an alternate assessment accountability system:

1. You cannot modify a state’s grade-level academic content standards. In other words, the same content standards must be used to guide instruction and the development of an AA-MAS.
2. You can modify a state’s grade-level achievement test for students who qualify for an AA-MAS.
3. You must modify a state’s grade-level academic achievement standards for any modified alternate assessment that is used.

Figure 1 provides an integrated summary of the key elements of an AA-MAS assessment and accountability system and reiterates our main conclusions about key elements. Let’s now turn to several of the implementation issues that currently are challenging educators and test developers.

WHICH STUDENTS WILL BENEFIT FROM AN AA-MAS?

The final regulations allow that an AA-MAS may be used for a subset of students with disabilities for whom such an assessment may be more appropriate. The modified standards are intended to help meaningfully include in the assessment and accountability system those students for whom the regular assessment is too difficult but for whom the alternate assessment based on alternate academic achievement standards is too easy. These are students whose disabilities have kept them from reaching proficiency, and who are unlikely to reach proficiency within the same timeframe as students who take the regular assessment (ED, 2007a). These are students with disabilities who can make significant progress, but have not reached grade-level achievement standards even after exposure to quality instruction, and who have a reliable record of below-level performance on achievement tests even with appropriate accommodations. They learn at a slower rate than most students and are considered to have persistent academic difficulties.

Students determined to meet these criteria may take the AA-MAS, so long as no more than 2% of the students within a state show that they are proficient via this test. An exception allows that as
many as an additional 1% of students in a state could show proficiency, as long as less than 1% of a state’s student population shows proficiency on the alternate assessment of alternate academic achievement standards. Even with this additional consideration, the law may not allow a large enough proportion to account for all students who have disabilities that are very likely to keep them from attaining proficiency in a state. Based on a survey of representatives of departments of education in five states, the second author of this article determined that between 3% and 9% of the total student population might meet these criteria. Educators, legislators, and researchers must seek a clearer picture of who the 2% of students eligible for this AA-MAS should be.

**HOW MODIFIED CAN THE TEST BE?**

The AA-MAS is expected to be easier than the regular assessment, but should be challenging for students who are eligible, and should reflect grade-level content (ED, 2007a). Because the
standards must reflect grade-level content, out-of-level testing is not appropriate. Out-of-level testing is the practice of assessing a child with a test that is designed for a lower grade level than the one in which he or she receives instruction. During the 2003–2004 academic year, 17 states used out-of-level testing for some students for whom the regular assessment was considered too difficult (Minnema, Thurlow, & VanGetson, 2004). The requirement that the AA-MAS cannot result in out-of-level testing has put test designers in the position of having to determine how much easier an assessment can be made before it can no longer be considered grade level. A number of technical issues generated by this new policy confront test developers and state assessment directors alike. These include concerns about (a) a test’s content alignment with that of the content taught and the content in state standards, (b) the depth of knowledge assessed, and (c) evidence for the comparability of the constructs assessed.

Testing accommodations and modifications are alterations made to a test or testing situation that make the test more accessible and appropriate for a group of students. According to Hollenbeck, Rozek-Tedesco, and Finzel (2000), appropriate accommodations are identifiable by four attributes:

1. **Unchanged constructs**—The alterations must not alter the construct being measured.
2. **Individual need**—The alterations must be individualized, rather than applied to groups.
3. **Differential effects**—The alterations must cause differential effects, such that the students who need them benefit in ways that other students would not.
4. **Sameness of inference**—The alterations must result in test scores that can be interpreted similarly to the interpretations of the regular assessment.

To the degree that a test alteration reflects these four attributes, it is likely to be considered an accommodation. If a test alteration fails to reflect one or more of these attributes, it is likely to be considered a modification. Examples of modifications include removing one of three distracters from a multiple-choice test, reducing the reading load of an item by simplifying language, adding visual cues that make an item easier to understand, or retesting students and using the better of multiple results. The field has not yet reached consensus on terminology to refer to an alteration that reflects the first, third, and fourth attributes but is applied to a group rather than individualized. Such terminology may soon be necessary, as states make alterations to items and tests for the group of students that qualify for an AA-MAS, in hopes of measuring unchanged constructs, evidenced by differential effects, which are useful for making the same inferences.

**WHAT IS THE CURRENT STATE OF AA-MASS?**

As of June 2007, three states (Kansas, Louisiana, and North Carolina) already had AA-MASs in place. Twenty of 36 state leaders who responded to a National Association of State Directors of Special Education survey indicated that they were going to adopt an alternate assessment for 2% of students, as allowed by the new regulations, and 12 state leaders indicated that they were undecided. In addition to determining the degree of modification that is necessary, and who should qualify for the modified alternate assessment, these states face a number of practical and technical challenges. States must train IEP teams to implement decisions about eligibility for the new assessment, must determine how to handle repeated testing within a single school year, and
must develop plans for management and reporting of additional data and results. States also have to develop and validate the new AA-MASs, write performance level descriptors, and implement standard setting procedures for the modified tests.

At the time of submission, the authors of this special issue were aware of one unpublished study of large-scale testing, item modifications, and students with disabilities. This study was conducted by Famularo and Russell (2006), and their findings were not conclusive about the influence of item modifications on the test scores of students with disabilities. Much larger bodies of research exist in the related areas of testing accommodations, universal design, and number of multiple-choice item response options. These areas, appropriately conceptualized as issues to be addressed when making assessments more accessible for all students, can also be viewed as guides in the process to develop AA-MASs.

For example, one popular item-modification strategy with empirical support involves the reduction of the number of answer options to a multiple-choice question. By reducing the number of response options, it is assumed that you have reduced the complexity of the decisions to be made and reduced the amount of reading, if not the difficulty of the task. Rodriguez’s (2005) meta-analysis of 27 studies addressed the question, “What is the optimal number of response options for a multiple-choice test?” Using the psychometric criteria of item difficulty, item discrimination, and test score reliability, the researcher concluded,

Three options are optimal for MC items in most settings. Moving from 5-option items to 4-option items reduces item difficulty by .02, reduces item discrimination by .04, and reduces reliability by .035 on average. Moving from 5- to 3-option items reduces item difficulty by .07, does not affect item discrimination, and does not affect reliability on average . . . . Moving from 4- to 3-option items reduces item difficulty by .04, increases item discrimination by .03, and increases reliability slightly by .02. (p. 10)

Although many states are using this strategy for item modification for students with disabilities, Rodriguez’s (2005) findings indicate that reducing the number of distracters does not harm the psychometric properties of the test within the general population but does theoretically reduce the reading load of the entire test. Modifications such as this are truly in the spirit of making assessments more accessible to students with disabilities.

AN INTRODUCTION TO OUR COLLEAGUES’ ARTICLES

The authors invited to contribute to Alternate Assessments Based on Modified Academic Achievement Standards: New Policy, New Practices, and Persistent Challenges were identified from a group of presenters at the introductory meeting in Washington, DC, on January 16, 2008, for Principal Investigators for General Supervision Enhancement Grants focusing on AA-MASs. They are educators and researchers who are motivated to address the aforementioned issues evoked by the new regulations regarding the AA-MAS.

In the first article, “The ‘Two Percent Students’: Considerations and Consequences of Eligibility Decisions,” Naomi Zigmund and Amanda Kloo expand upon their presentation at the introductory meeting. The presentation provided a critical look at the new policy and the impact that it might have on individual, state, and national levels.
In the second article, “The Changing Landscape of Alternate Assessments Based on Modified Academic Achievement Standards: An Analysis of Early Adopters of AA-MASs,” Sheryl Lazarus and Martha Thurlow discuss the current policies that states have in place for AA-MASs. Their article includes information on the number of states designing new tests, the methods that are being used, and the numbers of students likely to be affected.

In our third article, Andrew Roach and associates examine instructional aspects of an AA-MAS. Their article, “Opportunities and Options for Facilitating and Evaluating Access to the General Curriculum for Students With Disabilities,” reviews research on potential indicators of access to the general curriculum and opportunity to learn.

In the fourth article, “Modifying Achievement Test Items: A Theory-Guided and Data-Based Approach for Better Measurement of What Students With Disabilities Know,” along with Peter A. Beddow, we describe a systematic process of item modification. This process includes a foundation in theory, individual item analysis through statistics and a modification tool, and pilot testing in a cognitive lab.

The fifth article by Karla Egan, Steve Ferrara, Christina Schneider, and Karen Barton, titled “Writing Performance Level Descriptors and Setting Performance Standards for Assessments of Modified Achievement Standards: The Role of Innovation and Importance of Following Conventional Practice,” addresses critical aspects of developing performance level descriptors and achievement standards for a modified assessment. As they explain in detail, the process for making these alterations will have to incorporate consideration of the new test, as well as the student population for which it is designed.

The final article, “State Perspectives on Implementing, or Choosing Not to Implement, an Alternate Assessment Based on Modified Academic Achievement Standards,” by Porter Palmer, is a report of a national survey of state special education and assessment directors. The article highlights key issues on which state leaders are focusing to make decisions about this optional assessment for students with disabilities.

Susan Weigert, a Program Officer at the ED, and Michael Rodriguez, a psychometrician at the University of Minnesota, agreed to read all six articles in this issue and provide comments and perspectives. Their comments conclude the issue while concurrently providing a springboard for further examination of assessment policies, practice, and research.

Although this issue on AA-MAS will likely raise as many or more questions as it will answer, it represents a step toward determining how these assessments should be used to meaningfully include students with disabilities in the educational accountability framework, and ultimately in the classroom. We hope you find the articles informative with regard to the new alternate assessment, changes in practices, and issues that will likely challenge developers and users of AA-MASs.

REFERENCES


