Exceptional Children

Vol. 78, No. 1, pp. 58-72.

Capturing Implicit Policy From NCLB Test Type Assignments of Students With Disabilities

HYUN-JEONG CHO

NEAL KINGSTON

University of Kansas

ABSTRACT: *This study examined the learner characteristics and performance scores of students in the 2009 alternate assessment-modified achievement standard for one Midwestern state. Comparing performance differences by disability category for each content area from the students' 2008 test type assignments and performance scores facilitated examining the appropriateness of the 2009 test type assignment. The results raise concerns because some students with disabilities seemed to have been inappropriately assigned to test type. Students with intellectual disabilities had the lowest performances across grade level and subject area. Limitations and implications of this study are discussed and suggestions for future research are offered.*

The No Child Left Behind Act of 2001 (NCLB, 2006) and the Individuals With Disabilities Education Act (IDEA) have become crucial for holding schools accountable for the teaching and learning of all students, including students with disabilities. Most students with disabilities participate in Title I general assessments with or without accommodation, but some students with disabilities participate in alternate assessments. Alternate assessments are intended for students with disabilities who have been unable to participate in general assessments, even with accommodations (Kleinert & Thurlow, 2001). Students with the most significant cognitive disabilities take the alternate assessments based on alternate achievement standards (AA-AAS) which are linked to grade level content standards but differ in complexity and scope (U.S. Department of Education, 2005). States also have the option of developing an alternate assessment based on grade-level achievement standards (AA-GLAS) for students who are capable of performing on grade level but need a format other than the traditional multiple-choice test to demonstrate their knowledge and skills.

However, some students with disabilities fall into a gap between the general assessment and the AA-AAS; these students would not be able to

show proficiency on the general assessment and yet are not assessed appropriately by the AA-AAS. In 2007, revised NCLB regulations responded to this situation by giving states the option of providing an alternate assessment based on modified academic achievement standards (AA-MAS; 34 C.F.R. § 200.1[e]). When the AA-MAS option was first announced by the U.S. Department of Education in 2005, it provoked an immediate, prolonged, and sometimes heated debate over its role in student assessment. The debate centered on how an AA-MAS should be implemented, whether it would distance some students from the general education curriculum, and how to define which students should participate.

Some students with disabilities fall into a gap between the general assessment and the AA-AAS; these students would not be able to show proficiency on the general assessment and yet are not assessed appropriately by the AA-AAS.

Federal rules and guidelines have evolved in response to the controversy. Current regulations for participation in an AA-MAS are that the student must be identified as having a disability that precludes him or her from achieving the grade-level standard within the current year, and the student must have an individualized education program (IEP) that emphasizes grade-level content standards (34 C.F.R. Parts 200 and 300). The assessments are required to cover the same breadth and depth as the other grade-level assessments.

WHAT IS THE PURPOSE OF MODIFIED ACHIEVEMENT STANDARDS?

IDEA and NCLB require that students with disabilities have access to the general education curriculum and have their progress in the curriculum monitored; however, for adequate yearly progress (AYP) purposes, IDEA and NCLB impose a cap of 2% for students with disabilities out of the whole school population (calculated at the state and local education agency [LEA] levels) that may be categorized as proficient or above based on modified academic achievement standards. In addition, these Acts expect states to provide an appropriate assessment of what students know and what they can do across all content areas, not only for accountability purposes but also in order to provide information that could help guide instruction.

WHO IS ELIGIBLE TO PARTICIPATE IN ALTERNATE ASSESSMENTS BASED ON MODIFIED ACHIEVEMENT STANDARDS?

There are five rules to be used to decide a student's eligibility for an AA-MAS. First, the student's IEP team must decide the student's eligibility to participate in an AA-MAS. Even if significant growth occurs after appropriate instruction, the IEP team must be reasonably certain that the student will not achieve grade-level proficiency within the year covered by the student's IEP (34 C.F.R. § 200.1[e][2][ii][A]).

Second, states must provide school districts with clear and appropriate criteria to guide the IEP team's decision making about whether to assign a student to the AA-MAS (34 CFR § 200.1[f][1][i]). There are three criteria states must address (U.S. Department of Education, 2007). The first criterion is that the student's IEP goals are based on academic content standards for the grade in which the student is enrolled (U.S. Department of Education, 2007, p. 29). There are several ways this criterion may be met: by statements of IEP goals that are explicitly aligned to grade-level-content standards in reading, math, and science; by IEP requirements that the curriculum contain grade-level material; or by a statement from an IEP team member that the IEP goals and instructions align with grade-level content. The second criterion is that the student is unable to achieve grade-level proficiency due to a disability, as reflected in the student's performance on school district, state, or other assessments that place the student at the lowest proficiency level in academic achievement. The third criterion is that even if significant growth

occurs, it is apparent to the IEP team that the student cannot achieve grade-level proficiency within the timeframe stated in the IEP (34 C.F.R. § 200.1[e][2][ii][A]).

Third, each subject area must be considered separately. Fourth, the decision to have the student participate in the AA-MAS must be made annually based on multiple valid and objective measures of the student's achievement. Fifth, the student must be receiving instruction based on his or her grade level general curriculum; students in an AA-MAS should not be involved in "out-of-level" tests (U.S. Department of Education, 2007, p. 25).

A National Center on Educational Outcomes project reviewed nine states' AA-MAS participation guidelines and synthesized fifteen categories of criteria (Altman et al., 2008). Altman and colleagues found that the most frequently occurring criteria were based on those specified in the federal regulations, for example, that the student must have an IEP and that the decision must not be based on the student's disability label. Some states also provided exclusion criteria, such as the decision not being based on placement setting, the student not being eligible for the AA-AAS, and not having a pattern of past performance attributable to excessive absences or other non-instructional factors (e.g., cultural, language, economic).

THE CURRENT STUDY

This study examined the characteristics of elementary school students with disabilities taking an AA-MAS in reading or math in the 2009 school year in one Midwestern state, their performance scores, and their disability category. We also examined their 2008 assessments with respect to test type (AA-MAS, General, AA-AAS) and performance scores in order to gauge compliance with state and federal laws.

In our study, the population of an AA-MAS for a given content area was students with disabilities who did not meet eligibility criteria for the AA-AAS or for the general assessment for that content area. We also excluded students with low performance scores due to noninstruction factors such as high mobility or socioeconomic or cultural factors. This state provides a flowchart to guide the IEP team in its decision about participation in AA-MAS (e.g., Is the student multiple years behind grade level expectation? Does the student need significant changes in the complexity and scope of the general standards to show progress in the curriculum? Does the student need support to significantly reduce the complexity or breadth of assessment items?). The state provides IEP teams with this flowchart to determine the appropriate population. AA-MAS have been designed for use with a specific population in mind: students who should have been taking the general assessment but are unlikely to perform well and are not eligible for the AA-AAS. Therefore, they have been recommended to participate in the AA-MAS.

An analysis of the 2008 Federal Peer Review of the five states with an AA-MAS option at that time found that all of the states met some of the federal requirements under NCLB or IDEA, but that none met all of the requirements (Filbin, 2008; Lazarus, Thompson, & Thurlow, 2006). One problem that all the states had in common was in defining the guidelines to ensure that appropriate students were identified (Filbin, 2008; Lazarus et al. 2006; Thurlow, 2008); inappropriate assignment to AA-MAS is a threat to the validity of AA-MAS.

The National Research Council's Committee on the Foundations of Assessment developed the framework for the concept of the *assessment triangle* (Pellegrino, Chudowsky, & Glaser, 2001). The triangle is a theoretical framework for conceptualizing assessments. The three vertices of the assessment triangle are cognition, observation, and interpretation, what Pellegrino and colleagues described as the three pillars on which every assessment must stand: (a) *cognition* refers to the model of student cognition and how a student develops competence, (b) *observation* refers to a means for the student to demonstrate knowledge and skills, and (c) *interpretation* refers to drawing inferences from the performance evidence (p. 44). The present study focused on the cognition and observation vertices of the assessment triangle and was designed to answer the following research questions:

- Were there differences in performance among the disability categories in the state's AA-MAS reading and math assessments in 2009?
- What type of tests did students with disabilities who took an AA-MAS reading or AA-MAS math in 2009 take in 2008, and what were their performance scores? Were student assignments to the 2009 AA-MAS assessments appropriate based on their 2008 performance score and test type? If inappropriate, what was their 2009 AA-MAS performance?
- What were the disability categories of students who took the 2008 AA-AAS reading and math assessment and subsequently took the 2009 AA-MAS reading and math assessment?

METHOD

We extracted data from a state database of student performance, assessment information, and demographics and selected participants from 4,209 elementary students who took the 2009 AA-MAS reading assessment and 3,489 elementary students who took the 2009 AA-MAS math assessment; we included only those students for whom assessment data also were available from the 2008 school year. Thus, this sample comprised students who were third, fourth, and fifth graders in 2008 and fourth, fifth, and sixth graders in 2009, with approximately equal numbers in each grade. The selected sample numbered 3,260 student participants in the reading assessment and 2,743 in the math assessment.

Demographic Information

Table 1 presents demographics for participants in the reading and mathematics AA-MAS. For each of the tests, more than half of the students were classified as White. Participants were predominantly male, and the most common primary disability category was learning disability (LD). Table 2 separately presents National School Lunch Program status; this information was not available for all participants.

Data Analysis

We used descriptive statistics to summarize demographic information, to describe disability category differences in accommodations provided, and to describe the distribution of students in the various assessment types in 2008 (i.e., general, AA-MAS, AA-AAS) to the AA-MAS assessment type in 2009. We conducted chi-square tests of independence to examine reading and math proficiency levels by assessment type in 2008 and to compare students' 2008 general, AA-MAS, and the AA-AAS math and reading proficiency levels with their 2009 AA-MAS math and reading proficiency levels.

We conducted a series of univariate analyses of variance to explore student AA-MAS performance score differences among student disability categories for each content area separately (see Table 3). Each analysis consisted of an omnibus test of between-groups differences which, when found to be statistically significant, was followed by Bonferroni-adjusted, pairwise comparisons. We also performed descriptive statistics, cross tabulation, and chi-square analyses to compare performance scores, test types, subject areas, and years.

RESULTS

2009 AA-MAS Reading Performance Differences Among Elementary Student Disability Categories

An analysis of variance determined differences among disability categories with regard to 2009 AA-MAS reading performance scores. Reading performance means ranged from 2.81 (SD = 0.85) to 3.75 (SD = 1.17). Analysis of variance found a significant effect of disability category on reading performance category scores, F (10, 3249) = 16.78, $p < .001$, $\hat{\eta}^2 = .05$. Bonferroni adjusted pairwise comparisons showed that students with LD, emotional or behavioral disorders (EBD), speech/language impairment (SLI), and other health impairment (OHI) earned significantly higher reading scores than students with intellectual disabilities (ID). No other disability category pairs were statistically different from each other in reading performance scores.

TABLE 1

Characteristics of Participants in 2009 Alternate Assessments Based on Modified Achievement Standards (AA-MAS)

	Students Participating in Reading AA-MAS (N = 3,260)				*Students Participating in Mathematics AA-MAS (N = 2,743)*			
Student Demographics (%)	*Grade 4 (n = 1,102)*	*Grade 5 (n = 1,146)*	*Grade 6 (n = 1,012)*	*Total*	*Grade 4 (n = 912)*	*Grade 5 (n = 952)*	*Grade 6 (n = 879)*	*Total*
Gender								
Female	37.9	36.5	33.9	34.9	40.8	40.5	36.7	39.4
Male	62.1	63.5	66.1	63.8	59.2	59.5	63.2	60.6
Race/ethnicity								
Native American	2.0	2.4	1.5	1.9	1.6	2.5	1.5	1.9
Asian	1.5	1.1	1.9	1.4	1.2	1.4	1.7	1.4
Black	17.4	18.0	16.4	17.3	21.1	19.2	16.7	19.0
Hispanic	16.9	13.8	16.6	15.7	17.0	13.6	15.4	15.3
White	57.5	58.6	59.7	58.6	54.3	57.5	61.4	56.9
Multiethnic or missing	4.7	6.1	4.0	5.0	4.8	5.9	5.6	5.4
Disability status								
Autism	3.8	1.7	3.6	3.0	3.5	2.1	3.9	3.1
Developmental delay	5.4	.0	.0	2.2	5.9	1.1	.1	2.4
Emotional or behavioral disorder	3.0	.0	3.9	3.7	3.5	5.8	4.4	4.6
Hearing impairment	.9	.6	1.5	1.0	1.1	.6	1.7	1.1
Learning disability	60.4	63.0	63.2	62.2	54.5	57.1	55.7	56.2
Intellectual disability	9.3	11.4	12.1	10.9	11.6	13.2	13.7	12.8
Multiple disabilities	.0	.0	.0	.0	.0	.1	.2	.1
Other health impairment	13.5	13.7	13.2	13.5	15.9	16.1	16.3	16.1
Orthopedic impairment	.4	.0	.3	.2	.2	.2	.2	.2
Speech/language impairment	3.1	3.6	1.7	2.8	3.3	3.0	1.7	2.7
Traumatic brain injury	.0	.4	.2	.2	.1	.3	.2	.9
Visual impairment	.0	.2	.2	.1	.2	.3	.2	.3

TABLE 2

National School Lunch Program Status of Participants in 2009 Alternate Assessments Based on Modified Achievement Standards (AA-MAS)

	Students Participating in Reading AA-MAS (N = 2,297)				*Students Participating in Mathematics AA-MAS (N = 2,743)*			
Student Demographics (%)	*Grade 4 (803)*	*Grade 5 (791)*	*Grade 6 (703)*	*Total*	*Grade 4 (912)*	*Grade 5 (952)*	*Grade 6 (879)*	*Total*
Reduced-price lunch	12.3	10.9	13.8	12.2	12.3	10.3	13.8	12.1
Free lunch	60.6	58.1	55.6	58.2	58.8	58.0	55.5	57.5
No lunch assistance	27.1	31.0	30.5	29.5	28.9	31.7	30.7	30.5

TABLE 3

2009 Alternate Assessment Based on Modified Achievement Standards (AA-MAS) Reading and Math Performance Differences Among Elementary Student Disability Categories

	Reading AA-MAS			*Mathematics AA-MAS*		
Participants	N	*Mean*	SD	N	*Mean*	SD
Total	3,260	3.33	.97	2,743	3.03	.82
Disability category						
Autism	97	3.16	.98	86	2.86	.64
Developmental delay	73	3.08	1.01	65	2.88	.82
Emotional or behavioral disorder	121	3.06	.99	126	2.83	.85
Hearing impairment	33	2.88	1.11	31	3.06	.81
Learning disability	2,028	3.45	.96	1,541	3.15	.79
Multiple disabilities	0	0	0	3	3.00	.00
Intellectual disability	356	2.81	.85	352	2.59	.78
Other health impairment	440	3.35	.91	441	3.03	.85
Orthopedic impairment	8	3.63	1.19	6	3.00	.89
Speech/language impairment	92	3.30	.95	75	3.08	.88
Traumatic brain injury	8	3.75	1.17	7	3.57	.98
Visual impairment	4	3.00	.00	9	3.44	1.01

2009 AA-MAS Math Performance Differences Among Elementary Student Disability Categories

An analysis of variance determined differences among disability categories with regard to 2009 AA-MAS math performance scores. Math performance means ranged from 2.59 (SD = 0.78) to 3.44 (SD = 1.01). The analysis of variance omnibus test found a significant effect of disability on math scores, $F(12, 2730) = 13.60$, $p = .001$, $\acute{\eta}^2 = .06$. Pairwise comparisons using a Bonferroni adjustment for multiple comparisons showed that students with LD, EBD, SLI, OHI, and visual impairment (VI) had significantly higher math scores than students with intellectual disabilities (ID). No other disability category pairs were statistically different from each other in math performance scores.

Which Type of Reading Assessments Did Study Participants Take in 2008?

Frequency tabulations showed that 26.5% (n = 864) of the 2009 AA-MAS reading students took the general education reading test in 2008. The majority of the 2009 AA-MAS reading students, 71.8% (n = 2,338), took the AA-MAS reading test in 2008, and 1.7% (n = 54) took the AA-AAS reading test in 2008.

How Well Did Participants Perform on Their 2008 Reading Assessments?

For this analysis, we divided students into two groups, (a) less than proficient and (b) proficient or above, based on the level of performance on their 2008 reading tests. A chi-square test of independence testing the relationship of the dichotomized student performance with test type (i.e., general, AA-MAS, AA-AAS) found a significant dependency of performance on test type, $\chi^2 (2) = 1{,}097.63$, $p < .001$. Of those who took the 2008 general reading test, 20.4% were proficient or above, compared with the 82.5% of the students who took the 2008 AA-MAS reading test who were proficient or above, and the 87.0% of 2008 AA-AAS students who were proficient or above (see Table 4). All of these students were subsequently assigned to take the 2009 AA-MAS reading assessment.

TABLE 4

2008 Reading Assessment Performance of Students Who Took the 2009 Reading AA-MAS (N = 3,256)

Assessment Type		*Less Than Proficient*	*Proficient or Above*
General	*n*	688	176
	Percentage	79.6	20.4
AA-MAS	*n*	409	1,929
	Percentage	17.5	82.5
AA-AAS	*n*	7	47
	Percentage	13.0	87.0
Total	*n*	1,104	2,152
	Percentage	33.9	66.1

Note. Excludes study participants who did not take a reading assessment in 2008 ($n = 4$). AA-MAS = alternate assessment based on modified academic achievement standards; AA-AAS = alternate assessment based on alternate achievement standards.

How Did Students' Performance on the 2008 Reading Assessments Compare With Their 2009 AA-MAS?

For this analysis, performance scores were again dichotomized: (a) less than proficient and (b) proficient or above (see Table 5). A chi-square test of independence found a dependency between performance on the 2008 general reading test and performance on the 2009 AA-MAS reading test, $\chi^2(1) = 14.144$, $p < .001$. Of those who were proficient or above in their 2008 general reading tests, 98.3% ($n = 173$) were also proficient or above in their 2009 AA-MAS reading assessment, whereas 89.2% ($n = 614$) of those who were less than proficient in their 2008 general reading assessment were proficient or above in their 2009 AA-MAS reading test.

A chi-square test of independence also found a dependency between performance on the 2008 AA-MAS reading test and performance on the 2009 AA-MAS reading test, $\chi^2(1) = 304.391$, $p < .001$. Of those who were proficient or above on their 2008 AA-MAS tests, 91.7% ($n = 1,769$) also were proficient or above on their 2009 AA-MAS performance; 58.7% ($n = 240$) of those who were less than proficient on their 2008 AA-MAS reading assessment were proficient or above on their 2009 AA-MAS reading performance.

A chi-square test of independence found that students' performance on the AA-MAS reading test in 2009 was independent of their performance in the AA-AAS reading performance in 2008. It should be noted that 50% of the cells in this two-by-two design had expected values less than 5, so statistical power to demonstrate dependency, if there is a dependency, was lacking in this data.

Which Type of Math Assessment Did Study Participants Take in 2008?

Frequency tabulations showed that 29.7% ($n = 815$) of the 2009 AA-MAS math students took the general education math test in 2008. The majority of the 2009 AA-MAS math students (68.3%, $n = 1,874$) took the 2008 AA-MAS math test, and 1.5% ($n = 42$) of the students who took the 2009 AA-MAS math test took the 2008 AA-AAS math test.

How Well Did Participants Perform on Their 2008 Math Assessments?

Again, we divided the students into two groups, (a) less than proficient and (b) proficient or above, based on their scores on their 2008 math assessments. A chi-square test of independence testing the relationship of the dichotomized student performance with the type of test (i.e., general, AA-MAS, AA-AAS) found a significant dependency of performance with test type, $\chi^2(1) = 763.48$, $p < .001$. Of those who took the 2008 general math assessment, 21.0% were proficient or above; 77.1% of those who took the 2008 AA-MAS math test and 90.5% of those who took the 2008 AA-AAS math assessment were proficient or

TABLE 5

Dichotomous Grouping of 2008 Reading Performance Compared With Performance on the 2009 Reading AA-MAS

2008 Reading Assessment		*Performance on 2009 Reading AA-MAS*	
		Less Than Proficient	*Proficient or Above*
General Assessment (N = 864)			
Less than proficient	*n*	74	614
	Percentage	10.8	89.2
Proficient or above	*n*	3	173
	Percentage	1.7	98.3
Total	*n*	77	787
	Percentage	8.9	91.1
AA-MAS (N = 2,338)			
Less than proficient	*n*	169	240
	Percentage	41.3	58.7
Proficient or above	*n*	160	1,769
	Percentage	8.3	91.7
Total	*n*	329	2,009
	Percentage	14.1	85.9
AA-AAS (N = 54)			
Less than proficient	*n*	2	5
	Percentage	28.6	71.4
Proficient or above	*n*	19	28
	Percentage	40.4	59.6
Total	*n*	21	33
	Percentage	38.9	61.1

Note. AA-MAS = alternate assessment based on modified academic achievement standards; AA-AAS = alternate assessment based on alternate achievement standards.

above (see Table 6). All of these students were subsequently assigned to the 2009 AA-MAS math assessment.

How Did Students' Performance on the 2008 Math Assessments Compare With Their 2009 AA-MAS?

Using dichotomized performance scores, (a) less than proficient and (b) proficient or above, a chi-square test of independence found a dependency between the performance of students who took the 2009 AA-MAS math assessment and their prior performance on the 2008 general math assessment, $\chi^2(1) = 13.728$, $p < .001$. Of those who were proficient or above on their 2008 general math assessments, 95.9% (n = 164) were proficient or above on their 2009 AA-MAS math performance. Of those who were less than proficient on their 2008 general math assessments, 85.4% (n = 550) were proficient or above on their AA-MAS 2009 performance (see Table 7).

A chi-square test of independence found that the performance of students who took the 2009 AA-MAS math assessment also was dependent on their performance on the 2008 AA-MAS math assessment, $\chi^2(1) = 228.97$, $p < .001$. Of those who were proficient or above on their 2008 AA-MAS assessments, 81.4% (n = 1,175) were also proficient or above on their 2009 AA-MAS performance. Of those who were less than proficient on their 2008 AA-MAS assessments, 44.4% (n = 191) were proficient or above on their 2009 AA-MAS performance.

A chi-square test of independence found that students' performance on the 2009 AA-MAS math assessment was independent of their perfor-

TABLE 6

2008 Math Assessment Performance of Students Who Took the 2009 Math AA-MAS (N = 2,731)

2008 Assessment Type		Performance 2009	
		Less Than Proficient	*Proficient or Above*
General	*n*	644	171
	Percentage	79.0	21.0
AA-MAS	*n*	430	1,444
	Percentage	22.9	77.1
AA-AAS	*n*	4	38
	Percentage	9.5	90.5
Total	*n*	1,078	1,653
	Percentage	39.5	60.5

Note. Excludes study participants who did not take a math assessment in 2008 (n = 12). AA-MAS = alternate assessment based on modified academic achievement standards; AA-AAS = alternate assessment based on alternate achievement standards.

mance on the 2008 AA-AAS math assessment. With few AA-AAS students, 50% of the cells in this 2 by 2 design had expected values less than 5, so statistical power to demonstrate independence was lacking. Table 8 presents the disability categories of the students who took the reading or math AA-AAS in 2008 and took the corresponding (reading or math) AA-MAS in 2009.

TABLE 7

Dichotomous Grouping of 2008 Math Performance Compared With Performance on the 2009 Math AA-MAS

2008 Math Assessment		Performance on 2009 Math AA-MAS	
		Less Than Proficient	*Proficient or Above*
General Assessment (N = 815)			
Less than proficient	*n*	94	550
	Percentage	14.6	85.4
Proficient or above	*n*	7	164
	Percentage	4.1	95.9
Total	*n*	101	714
	Percentage	12.4	87.6
AA-MAS (N = 1,874)			
Less than proficient	*n*	239	191
	Percentage	55.6	44.4
Proficient or above	*n*	269	1,175
	Percentage	18.6	81.4
Total	*n*	508	1,366
	Percentage	27.1	72.9
AA-AAS (N = 42)			
Less than proficient	*n*	2	2
	Percentage	50.0	50.0
Proficient or above	*n*	23	15
	Percentage	60.5	39.5
Total	*n*	25	17
	Percentage	59.5	40.5

TABLE 8

Disability Categories of Students Who Took the 2008 Alternate Achievement Standards Assessment and Subsequently Took the 2009 Alternate Assessment Based on Modified Achievement Standards

	Autism	*Developmental Disability*	*Emotional/ Behavioral Disorder*	*Hearing Impairment*	*Learning Disability*	*Intellectual Disability*	*Other Health Impairment*	*Orthopedic Impairment*
Reading ($N = 55$)								
n	5	2	3	5	19	14	6	1
Percentage	9.1	3.6	5.5	9.1	34.5	25.5	10.9	1.8
Math ($N = 42$)								
N	1	4	4	0	11	14	7	1
Percentage	2.4	9.5	9.5	.0	26.2	33.3	16.7	2.7

DISCUSSION

The purpose of this study was to describe the characteristics of students with disabilities, their AA-MAS performance scores, and their performance in the preceding year, in order to refine understanding of this population and improve practice. This study corroborates the findings of other recent studies and raises questions about eligibility. Similar to the findings of the NCEO survey (Altman et al., 2008), we found that the majority of AA-MAS students in this Midwestern state belonged to a few disability categories, primarily LD and ID. We found that students with ID had the lowest performance scores across subject areas and across grade levels. Traditionally, many of these students, including those with more mild intellectual disabilities, functioned academically below grade levels appropriate to their age groups. Perhaps the focus of their instruction was on content standards at lower grade levels and nonacademic content, such as life skills. Joseph and Eveleigh (2009) synthesized 20 years of research on self-monitoring but found no studies on self-monitoring instruction for students with ID. Failure to provide appropriate instruction to students with ID jeopardizes their academic achievement.

The present study found that 5.3% of students assigned to the 2009 AA-MAS reading assessment and 6.2% of those assigned to the 2009 AA-MAS math assessment scored proficient or above on the previous (2008) year's general reading and math assessments, respectively. If these students were not misassigned to the AA-MAS assessments in 2009 after demonstrating proficiency the previous year, what justifications might their IEP teams have for these placements? Perhaps the students were not performing well on other assessments, such as interim assessments or curriculum based assessments. Perhaps some students achieved proficiency but with little margin to spare, and the IEP teams felt the student would profit more by participating in the next year's AA-MAS. Another possibility is that teachers are under such great pressure that taking the extra time and resources to instruct students with disabilities detracts from their ability to properly prepare general assessment students without disabilities. Maybe teachers are under such great pressure for their classes to meet AYP that making things easier for the ones who are eligible, students with disabilities, seems prudent. Or perhaps some teachers simply have low expectations for their students with disabilities.

This study corroborates the findings of other recent studies and raises questions about eligibility.

In 1984, special education researcher Anne Donnellan wrote that

> The criterion of least dangerous assumption holds that in the absence of conclusive data,

> educational decisions ought to be based on assumptions which, if incorrect, will have the least dangerous effect on the likelihood that students will be able to function independently as adults. (p. 142)

Donnellan (1984) maintained that in the absence of evidence to the contrary, educators should assume that poor performance is due to instructional deficits rather than student deficits. Some students in the present study achieved advanced or exemplary scores in AA-MAS reading (15.5%, $N = 591$) or in AA-MAS math (9.4%, $N = 416$) 2 years in a row (2008 and 2009). According to the U.S. Department of Education (2007), "there will be students with disabilities who take an alternate assessment based on modified academic achievement standards one year, make considerable progress during the school year, and then take the general grade-level assessment the following year." The Department of Education guidelines do not specify a performance score criterion for shifting students from AA-MAS to the general assessment.

In the present study, approximately 7% of students achieved less than proficient scores on both their 2008 and 2009 AA-MAS reading assessments, and 13% were less than proficient on their 2008 and 2009 AA-MAS math assessments. Approximately 65% of these students qualified for the National School Lunch Program. Of the reading assessment group, 46% were students with LD, 25% were students with ID, and 10% were students with OHI. Of the math assessment group, 35% were students with LD, 25% were students with ID, and 15% were students with OHI. Research has shown that student characteristics related to disability, ethnicity, poverty, and other demographic categories are associated with a history of low expectations and limited opportunities to learn. Unless there is effective intervention to change the learning opportunities for students with these demographics, past performance cannot be used to predict which students will achieve academic proficiency. This is true for students with disabilities as well as for those without disabilities. More research is needed that accounts for the characteristics of students and determines whether low-achieving students did not respond well to intervention or simply did not have adequate opportunity to learn. Studies should consider the cognitive, interventional, and neurobiological perspectives of this subgroup.

The primary disabilities of the students who took part in the AA-AAS reading assessments in 2008 and then participated in the 2009 AA-MAS reading assessments, and of the students who took the 2008 AA-AAS math and then the 2009 AA-MAS math, were LD, hearing impairment (HI), EBD, and OHI. None of these primary disability categories involves intellectual impairment, one of the main eligibility criteria for inclusion in the state AA-AAS. Considering disability category would be desirable to establish exclusionary criteria so that students with noncognitive disabilities are not included in the AA-AAS.

Limitations of Findings

We selected student data from the 2009 assessment for analysis in this study based on whether there was assessment data for the same student in the 2008 school year. Many of the student records in the 2009 assessment had no matching record in the 2008 database.

We had insufficient information to determine why certain students with disabilities achieved at or above proficiency in reading and math in the general education assessment in 2008, yet were placed in the AA-MAS assessment in 2009. We were unable to determine why students achieving proficiency or above in AA-MAS assessments in 2008 were still taking AA-MAS assessments in 2009. We were also unable to learn the reasons students with the primary disabilities of LD, EBD, and OHI were included in the AA-AAS. We were unable to access student IEP documents, which would have been helpful in understanding teachers' roles in the process. Additionally, we were not able to determine why some of the students taking the AA-AAS or AA-MAS in 2008 performed better in 2009 than in 2008. Perhaps teachers' expectations were higher for their students when taking the AA-MAS, affecting their quality of instruction and their students' ability to achieve above proficient scores in reading and math in 2009. We were also unable to learn what types of intervention teachers used to improve their instruction, how teachers monitored student progress, which scientific evidence

was used to measure progress, and what effect the quality of teaching may have had on their students' performance scores.

Further, this study concerns one state's implementation of an AA-MAS. The extent to which the findings of this study are generalizable to other states is unknown. However, policy makers and related personnel in other states may find this information of use in constructing or adjusting their programs.

Recommendations for Future Research

Because the AA-MAS assigned a majority of students to high proficiency levels for two consecutive academic years, it brings into question the appropriateness and validity of the question items: They do not seem to accurately differentiate between student ability levels. Future research should evaluate assessment items used in the AA-MAS using tools such as the Test Accessibility and Modification Inventory (TAMI; Beddow, Kettler, & Elliott, 2008), influenced by universal design (e.g., Elliott, Kurz, Beddow, & Frey, 2009), test accessibility (e.g., Johnstone, Thurlow, Moore, & Altman, 2006), cognitive load theory (e.g., Clark, Nguyen, & Sweller, 2006; Elliott et al., 2009), and fairness. TAMI is used to enhance accessibility and improve measurement validity of test items and is especially meant for assessments designed for the AA-MAS population. Future research should also examine teachers' decision-making process when assigning a student to the AA-MAS and whether standards-based IEPs are used and, if so, how student progress is being monitored.

Future research should . . . examine teachers' decision-making process when assigning a student to the AA-MAS and whether standards-based IEPs are used and, if so, how student progress is being monitored.

Curriculum-based measurement (CBM) is of particular interest given its strong research base. Recently, Fuchs, Seethaler, Fuchs, and Hamlett (2008) proposed CBM as meeting the requirements for determining eligibility under the 2% regulation: That is, CBM progress monitoring can be used to provide the necessary database on (a) whether grade-level proficiency is expected, (b) whether appropriate instruction has been provided, and (c) whether progress in response to that instruction is appropriate. The effectiveness of CBM instruction/response to intervention is reflected in AA-MAS performance. Future research should investigate the consequential validity of this intervention. If students don't respond to those interventions, research should examine why they fail to respond.

Recommendations for Future Practice

Most importantly, the present study highlights challenges for state policy makers in considering how to refine the eligibility of students with disabilities to take the AA-MAS and in identifying students who might benefit most from these assessments. The state we studied should refine its eligibility guidelines for participation in the AA-AAS and AA-MAS. If participation guidelines are not well defined, students may be continually assigned to inappropriate test types.

First, disability category should be considered when deciding which students belong to the 1% of students expected to have significant cognitive disabilities. For example, students whose primary disability is LD, EBD, or OHI are not cognitively impaired and should be excluded from the AA-AAS. The rules for AA-AAS eligibility should be refined to help teachers in their decisions about their students with significant cognitive disabilities.

Second, when teachers decide which test type to assign their students, they need clear guidelines of what kind of standardized tests should be considered in this decision and how many test performances should be counted (i.e., curriculum-based assessment, districtwide assessment, commercially available standardized assessment, or statewide assessment).

Third, states should establish guidelines for transition of proficient students from the AA-MAS to the general assessment. Placement

decisions, by law, must be made annually, and these decisions should be based on validated assessments and standards-based IEPs. The use of multiple indicators will strengthen such decisions. For example, Domaleski (2009) suggests a profile approach that could be implemented to take advantage of several data sources. Under this system, several clusters of performance indicators would be defined (*profiles*), any of which would indicate a student's eligibility to advance to a different assessment type. An example of such a profile might be achieving a performance category score of 2 or 3 on the AA-MAS and also achieving some criterion score on a district assessment. The development of these profiles must be accomplished at the state level, and their use must be monitored and well documented.

Fourth, states should take advantage of the professional networking capabilities of current technology to enable teachers of students with disabilities to share experiences, curricular materials, and approaches and to provide links to information on assistive technology, accommodations, and data collection materials for alternate assessment. A model for this is the web site developed by the Georgia Department of Education (2005–2006) for teachers of students who take that state's AA-AAS.

Fifth, there should be recognition of the importance of effort and teamwork at the local level. States and LEAs should provide guidance for strengthening collaboration among IEP team members, administrators, related service personnel, parents, and general and special educators. Schools could adopt a philosophy that the whole school is responsible for all students' success.

Finally, the ultimate goals for academic achievement are mandated by state and federal standards, but intermediate goals are created by teachers and others responsible for curriculum, assessment, and instruction. Determining these intermediate goals requires an understanding of how students learn and represent knowledge. Thus, teachers need to have an understanding of cognition and learning differences. Preservice and professional development should help teachers understand models of learning so they can recognize students' initial sense-making strategies.

REFERENCES

Altman, J. R., Lazarus, S. S., Thurlow, M. L., Quenemoen, R. F., Cuthbert, M., & Cormier, D. C. (2008). *2007 survey of states: Activities, changes, and challenges for special education.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Beddow, P. A., Kettler, R. J., & Elliott, S. N. (2008). *Test accessibility and modification inventory (TAMI).* Nashville, TN: Vanderbilt University. Retrieved from http://peabody.vanderbilt.edu/tami.xml

Clark, R., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load.* San Francisco, CA: Pfeiffer.

Domaleski, C. (2009). Operational and accountability issues. In M. Perie (Ed.), *Considerations for the Alternate Assessment based on Modified Achievement Standards (AA-MAS): Understanding the eligible population and applying that knowledge to their instruction and assessment. A white paper commissioned by the New York Comprehensive Center in collaboration with the New York State Education Department* (pp. 334–363). Retrieved from https://www.cehd.umn.edu/NCEO/AAMAS/AAMASwhitePaper.pdf

Donnellan, A. (1984). The criterion of the least dangerous assumption. *Behavior Disorders, 9,* 141–150.

Elliott, S. N., Kurz, A., Beddow, P., & Frey, J. (2009, February). *Cognitive load theory and universal design principles: Applications to test item development.* Paper presented at the National Association of School Psychologists (NASP) Annual Convention, Boston, MA.

Filbin, J. (2008). *Lessons from the initial peer review of alternate assessments based on modified achievement standards.* Washington, DC: U.S. Department of Education, Office of Elementary and Secondary Education Student Achievement and School Accountability Program. Retrieved from http://www.cehd.umn.edu/nceo/AAMAS/2percentLessonsAAMAS.pdf

Fuchs, L., Seethaler P. M., Fuchs, D., & Hamlett, C. L. (2008). Using curriculum-based measurement to identify the 2% population. *Journal of Disability Policy Studies, 19,* 153–161. doi:10.1177/1044207308327471

Georgia Department of Education. (2005–2006). *Georgia Performance Standards (GPS) for students with significant cognitive disabilities.* Retrieved from http://gadoe.georgiastandards.org/impairment.aspx

Johnstone, C., Thurlow, M., Moore, M., & Altman, J. (2006). *Using systematic item selection to improve universal design of assessments* (Policy Directions 18). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Joseph, L. M., & Eveleigh, E. L. (2009). A review of the effects of self-monitoring on reading performance of students with disabilities. *The Journal of Special Education, 20*, 1–11.

Kleinert, H., & Thurlow, M. (2001). An introduction to alternate assessment. In H. Kleinert & J. Kearns (Eds.), *Alternate assessment: Measuring outcomes and supports for students with disabilities.* Baltimore, MD: Paul Brookes.

Lazarus, S. S., Thompson, S. J., & Thurlow, M. L. (2006). *How students access accommodations in assessment and instruction: Results of a survey of special education teachers* (Issue Brief 7). College Park, MD: University of Maryland, Educational Policy Reform Research Institute.

NCLB Regulations, 34 C.F.R. § 200 *et seq.* (2007).

No Child Left Behind Act of 2001, 20 U.S.C. § 6301 et seq. (2006).

Pellegrino, J. W., Chudowsky, N. J., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academy Press.

Thurlow, M. L. (2008). Assessment and instructional implications of the alternate assessment based on modified academic achievement standards (AA-MAS). *Journal of Disability Policy Studies, 19*, 132–139. doi:10.1177/1044207308327473

U.S. Department of Education. (2005). *Raising achievement: Alternate assessments for students with disabilities.* Retrieved from http://www.ed.gov/print/policy/elsec/guid/raising/alt-assess-long.html

U.S. Department of Education. (2007, July 20), *Modified academic achievement standards: Non-regulatory guidance.* Washington, DC: Office of Elementary and Secondary Education, U.S. Department of Education. Retrieved from http://www2.ed.gov/policy/speced/guid/nclb/twopercent.doc

ABOUT THE AUTHORS

HYUN-JEONG CHO (Kansas CEC), Research Associate; and **NEAL KINGSTON** (Kansas CEC), Associate Professor, Center for Educational Testing and Evaluation, University of Kansas, Lawrence.

Correspondence concerning this article should be addressed to Hyun-Jeong Cho, Center for Educational Testing and Evaluation, University of Kansas, 1122 West Campus Road, Room 730, Lawrence, KS 66045 (e-mail: chohj@ku.edu).

The authors would like to acknowledge Rebecca Finney, Dr. Martha Thurlow, Gretchen Anderson, and Mickey Waxman for their assistance.

Manuscript received March 2010; accepted July 2010.