

Fall 12-2017

Comparing Two Alternate Assessments: Dynamic Learning Maps and Multi-State Alternate Assessment

Dana N. Sir
dananbsir@gmail.com

Follow this and additional works at: <https://scholarship.shu.edu/dissertations>

 Part of the [Educational Assessment, Evaluation, and Research Commons](#), and the [Special Education and Teaching Commons](#)

Recommended Citation

Sir, Dana N., "Comparing Two Alternate Assessments: Dynamic Learning Maps and Multi-State Alternate Assessment" (2017). *Seton Hall University Dissertations and Theses (ETDs)*. 2421.
<https://scholarship.shu.edu/dissertations/2421>

Comparing Two Alternate Assessments: Dynamic Learning Maps
and Multi-State Alternate Assessment

by

Dana N. Sir

Submitted in partial fulfillment of the requirements for the degree

Doctor of Education

Department of Education Leadership, Management, and Policy

Seton Hall University

December 2017

© 2017 DANA N. SIR

SETON HALL UNIVERSITY
COLLEGE OF EDUCATION AND HUMAN SERVICES
OFFICE OF GRADUATE STUDIES

APPROVAL FOR SUCCESSFUL DEFENSE

Dana N. Sir, has successfully defended and made the required modifications to the text of the doctoral dissertation for the **Ed.D.** during this **Fall Semester 2017**.

DISSERTATION COMMITTEE
(please sign and date beside your name)

Mentor:
Dr. Luke J. Stedrak 

Committee Member:
Dr. Daniel Gutmore 

Committee Member:
Dr. Jan Furman 

The mentor and any other committee members who wish to review revisions will sign and date this document only when revisions have been completed. Please return this form to the Office of Graduate Studies, where it will be placed in the candidate's file and submit a copy with your final dissertation to be bound as page number two.

ACKNOWLEDGEMENTS

No words could ever express how grateful I am for my husband, Franklin, for his never-ending patience and support for me during the challenge of balancing family, work, and school on a daily basis. He always believes in me even on days that I have trouble believing in myself. I want to thank my children, Nicole, Christian, and Ryan, for being my cheerleaders who support any goal that I set. They help me focus on what is important in life. I truly appreciate my dad, Robert J. Breitenbach, and my sister-in-law, Jeannye Sir Petruzzi, who provided listening ears, assisted with my children when I needed to focus on my work, and appreciated and valued my educational goals. My grandmother, Loretta Nolan DeGeorge, always resonated the message that “education is never a waste”.

I am grateful to be surrounded by numerous friends and family, too many to name, who always gave me words of encouragement and reminded me that I was more than capable of achieving this great goal.

Thank you to the staff, faculty, Special Education Department, Child Study Team, Fran Orefice, and Gina McCormack at the Norwood Public School, who are my work family. They encourage me to grow, even if that means taking a different path. Every day they remind me through their actions on what it means to put students first. I thank them for their trust, faith, and believing in me.

This paper would not have come to fruition or completion without the encouragement and rallying from Dr. Courtney Carmichael and Dr. Catherine Stanzione. They were there to get me back on track when my wheels kept spinning.

Thank you to Dr. Luke J. Stedrak, my Seton Hall University committee chair, for his leadership in assisting me through this journey to the end. My gratitude is extended to my committee members, Dr. Jan Furman and Dr. Daniel Gutmore, for their guidance and reflection.

ABSTRACT

In 2001 under No Child Left Behind, states were required to create an alternate assessment for students with significant cognitive disabilities using alternate achievement standards. In 2003, all states had created an alternate assessment. All fifty states independently developed, implemented, and revised their alternate assessments. By 2014, Dynamic Learning Map (DLM) and Multi-State Alternate Assessments (MSAA) (formerly National Center and State Collaborative)—two alternate assessments developed through consortiums consisting of state departments, universities, and organizations using federal funding—were created. At the time of this study, the DLM and MSAA were used by approximately 49% of states for their alternate assessment. This study compared the DLM and MSAA in English language arts for students with significant cognitive disabilities in grades three through eight. The study focused on the DLM's and MSAA's measurement criterion and how they affect informational outcomes as well as how the alternate assessments are administered. The findings illustrated that the DLM and MSAA are primarily administered online to students with significant cognitive disabilities. Accessibility supports available through both alternate assessments are comparable. The DLM and MSAA may be administered on various devices, increasing their ability to individualize and accommodate to a student needs. Although there were differences in how the assessments were differentiated, the informational outcomes produced by both alternate assessments for students with significant cognitive disabilities were similar and, therefore, one alternate assessment could not be identified as superior to the other.

Keywords: alternate assessments, Dynamic Learning Maps, Multi-State Alternate Assessments, consortiums, significant cognitive disability, alternate achievement standards

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	iv
ABSTRACT.....	v
LIST OF TABLES.....	viii
CHAPTER 1: INTRODUCTION.....	1
Purpose of the Study.....	2
Delimitations.....	3
Limitations.....	4
Definitions of Terms.....	5
Organization of the Study.....	6
CHAPTER 2: REVIEW OF THE LITERATURE.....	8
Literature Review.....	8
Section One: National Center on Educational Outcomes.....	9
Section Two.....	29
Types of Assessments.....	29
Alternate Achievement Standards.....	31
Scoring Criteria.....	35
Validity of Alternate Assessments.....	38
Score Reporting.....	42
Summary.....	47
CHAPTER 3: METHODOLOGY.....	48
Study Design.....	48
CHAPTER 4: RESULTS.....	51

Types of Assessments.....	52
Criterion of Measurement.....	53
Informational Outcomes.....	59
CHAPTER 5: SUMMARY AND CONCLUSION.....	64
Research Questions.....	64
Significance of the Study.....	64
Implications of the Study.....	65
Topics for Further Research.....	68
Final Conclusions.....	69
REFERENCES.....	70

LIST OF TABLES

Table	Page
2-1 Alternate Assessment Standards.....	10
2-2 States Engaged in Various Alternate Assessment Activities.....	11
2-3 Alternate Assessment Approaches Selected by States.....	12
2-4 Estimated Percentages of All Students Whose Exposure to Content is Too Limited for Them to Participate in Regular Assessment.....	12
2-5 Stakeholders Involved in the Development of State Alternate Assessments.....	13
2-6 Alternate Assessment Approaches or Types.....	16
2-7 Alternate Assessment Scoring Criteria for Rubrics.....	18
2-8 Alternate Assessment Scorers.....	18
2-9 Alternate Assessment Approaches or Types.....	20
2-10 Scoring Criteria for Alternate Assessment Responses 2003-2005.....	22
2-11 Measured Outcomes from the Alternate Assessments.....	22
2-12 AA-AAS Formats	24
2-13 Measured Outcomes from the Alternate Assessments.....	25
2-14 State Contribution to Alternate Assessment Consortia (DLM, NCSC) Decision Making...	27
3-1 Organization of Data	50
4-1 DLM Supports.....	56
4-2 MSAA Accessibility Features.....	56

COMPARING TWO ALTERNATE ASSESSMENTS

CHAPTER I

INTRODUCTION

Students within the United States are required to participate in a statewide assessment system to show accountability to state and federal requirements (NCLB, 2001). For students with significant cognitive disabilities who are unable to participate in the statewide general education assessments even with accommodations, states are required to provide the option of taking an alternate assessment (IDEA 1997). The obligation of creating an alternate assessment gave the responsibility of each individual state to maintain their compliance by offering an alternate assessment that was reliable and demonstrated accountability. All states had at least one alternate assessment to offer by 2003 (Quenemoen, 2008, p. 5). In a 1999 National Center for Educational Outcomes (NCEO) survey of state special education directors, only twenty states were in the process of developing an alternate assessments, with only Kentucky and Maryland reporting that they had fully developed an alternate assessment (Quenemoen, 2008, p. 5). In 2001, No Child Left Behind (NCLB) was passed into law; it “specified that alternate assessments based on alternate achievement standards (AA-AAS) must be ‘aligned with the state’s academic content standards’” (Cho & Kingston, 2012, p. 162). Alternate achievement standards (AAS) are “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” (Maryland Department of Education, 2009). Under the reauthorization of IDEA in 2004, states were required to report their number of students with disabilities who participated in statewide assessments. States worked independently and in isolation to develop their alternate assessments. Educational organizations, consortiums of departments of education, and universities started to team up, using federal grants to develop

COMPARING TWO ALTERNATE ASSESSMENTS

alternate assessments that were more reliable and valid than in past years (Kingstom, Karvonen, Bechard, & Erickson, 2016; NCSC website). As of May 2017, current alternate assessments are created through partnerships and are based on research to meet federal requirements and to demonstrate state and local accountability. Before 1997, students with significant cognitive disabilities were rarely a priority in statewide assessments; current federal policy has placed more responsibility on states and districts to include this population in their accountability measurements.

Purpose of the Study

This study compared the alternate assessments provided by Dynamic Learning Maps (DLM) and Multi State Alternate Assessment (MSAA) in English language arts for students with significant cognitive disabilities in grades three through eight. From January to May of 2017, fifteen states administered the DLM (Alaska, Colorado, Illinois, Iowa, Kansas, Missouri, New Hampshire, New Jersey, New York, North Dakota, Oklahoma, Utah, West Virginia, and Wisconsin) and eight states administered the MSAA (Arizona, Arkansas, Maine, Maryland, Montana, Rhode Island, South Dakota, and Tennessee). Twenty-seven states administered their own alternate assessments (Alabama, California, Connecticut, Delaware, Florida, Georgia, Hawaii, Indiana, Kentucky, Louisiana, Massachusetts, Michigan, Minnesota, Mississippi, Nebraska, Nevada, New Mexico, North Carolina, Ohio, Oregon, Pennsylvania, South Carolina, Texas, Vermont, Virginia, Washington, and Wyoming). Therefore, approximately 49% of states' departments of education selected the DLM or MSAA as their alternate assessments. Both alternate assessments were originally piloted within the last two to three years, with little research conducted about either alternate assessment. This study will focus on

COMPARING TWO ALTERNATE ASSESSMENTS

comparing the alternate assessments provided by Dynamic Learning Map (DLM) and Multi State Alternate Assessment (MSAA), which provide detailed information regarding academic achievement in the English language arts of students with significant cognitive disabilities in grades three through eight.

1. Do DLM and MSAA use the same criterion for measurements? If there are similarities between the alternate assessments, why are they necessary for informational outcomes?
2. If there are differences between the criterion for measurements, how do they affect informational outcomes?
3. How do the DLM and MSAA administer their assessments? How does the difference between their administrations affect informational outcomes?

Delimitations

This study focused on the DLM and MSAA administered from January to June of 2017 because they were the only two online assessments given through the collaboration of consortiums at the time of this study. The population members who take these assessments are students with significant cognitive disabilities. Students with significant cognitive disabilities represent a small percentage of students who are administered alternate assessments. This population's physical, emotional, social, and intellectual needs vary. This variety limits the ability to compare students' academic performance on alternate assessments. This population for alternate assessments was limited to grades three through eight for this study, which reflects the identical grades needed to take statewide assessments for the general student population. English language arts were selected as the focus for comparing these alternate assessments. English language arts, as a content area, is less linear than mathematics and allows for growth in different skills without mastery of skills in a particular area of knowledge. Only the End-of-the-Year

COMPARING TWO ALTERNATE ASSESSMENTS

(EOY) assessment for DLM and MSAA were the focus of this study because all states are required to have at least an EOY alternate assessment for their students with significant cognitive disabilities.

Limitations

In looking at only the DLM and the MSAA, the study was limited by only comparing two alternate assessments, even though various states have developed and implemented their own alternate assessments. It narrowed the focus of this study and the comparisons that could be made from the results. Students with significant cognitive disabilities who are eligible to take the DLM or MSAA exhibit a broad spectrum of disabilities within this one disability category. Students with extremely limited cognitive ability are unable to access an alternate assessment even with access to the maximum support of accessibility features and/or accommodations. In focusing on alternate assessments for students in grades three through eight, the informational outcomes developed from student's performance on the alternate assessment was shown to have a strong impact on a student's elementary and middle school programs with some recommendations for transitional IEP's and programs. Eliminating high school alternate assessments limits the study's recommendations for high school and 18-21 programs based upon the informational outcomes produced from a student's performance on the DLM or MSAA. In addition, using only English language arts as the focus for comparing the DLM and MSAA limited the lens from comparing these assessments. Mathematics and science are other content areas that may be assessed using the DLM or MSAA, but for the purpose of this study were not included.

COMPARING TWO ALTERNATE ASSESSMENTS

Definition Of Terms

Alternate Achievement Standards (AAS)—AAS are performance standards used for comparing the achievement of students with significant cognitive disabilities against grade-level standards. AAS differs in complexity from general education grade-level standards, but they are linked or aligned to these standards to allow students with significant cognitive disabilities access to the general curriculum. AASs vary across states and are a part of each state’s standards and assessment system (USDOE, 2005).

Alternate Assessment—A statewide alternative assessment measures academic performance for students who are unable to participate in general statewide assessments even with accommodations, and one must be made available by each state (IDEA, 1997). Alternate assessments must demonstrate reliability and validity (IDEA, 1997).

Dynamic Learning Map (DLM)—DLM is an alternate assessment created by multiple states and the University of Kansas. DLM’s assessment uses a computer-based system that is administered through the DLM website (<http://dynamiclearningmaps.org/>). It was funded in late 2010 using a \$22 million, five-year U.S. Department of Education Office of Special Education Programs grant (<https://cete.ku.edu/dynamic-learning-maps>). DLM was originally implemented in 2014 (Kingston, Karvonen, Bechard, & Erickson, 2016).

Individualized Education Program (IEP)—An IEP is an individualized legal document for a student who is eligible for special education and related services. IEPs are created through an IEP team and include what the child needs to learn, the services that the school will provide, and how progress will be measured.

COMPARING TWO ALTERNATE ASSESSMENTS

Multi State Alternate Assessment (MSAA)—MSAA consists of state partners who transition and carry forward the alternate assessment developed through the National Center and State Collaborative (NCSC). (<http://www.azed.gov/assessment/msaa/>).

National Center and State Collaborative (NCSC)—“NCSC is a project led by five centers and 24 states to build an alternate assessment based on alternate achievement standards (AA-AAS) for students with the most significant cognitive disabilities” (National Center and State Collaborative, 2013, <http://www.ncscpartners.org/>). The NCSC alternate assessment is a summative and computer-based assessment. An operational test was administered March 30 through May 15, 2015. NCSC was federally funded and transitioned to MSAA in the fall of 2016.

Significant Cognitive Disability—A student who is classified as SCD is defined as a student whose intellectual abilities limit his or her ability to accomplish grade-level academic standards, even with “systematic instruction” (National Alternate Assessment Center, 2010). Significant cognitive disabilities are noted in IDEA (2004).

Organization Of The Study

Chapter 1 explained the framework of federal legislation that shaped alternate assessments for students with significant cognitive disabilities. This included federal legislation passed from 1997 through May of 2017. Chapter 2 will describe the literature available to paint a picture of the literature, research, and policy available for states to read and process when selecting an alternate assessment for students with significant cognitive disabilities. The chapter is divided into national reports and research on alternate assessments. In Chapter 3, the methods of data collection and analysis needed to compare the DLM and MSAA within the study will be discussed. This will include important components of the alternate assessments that need to be

COMPARING TWO ALTERNATE ASSESSMENTS

considered in the comparison. Chapter 4 will present and analyze the data between the two different types of alternate assessments: the DLM and the MSAA. This information will to help decipher the differentiation between the two alternate assessments.

CHAPTER 2

REVIEW OF THE LITERATURE

Literature Review

This chapter reviews the literature, research, and policy relating to the topic of alternate assessments for students with significant cognitive disabilities. The chapter is divided into two sections. Section one presents the reports published from 1996 through 2014 by the National Center on Educational Outcomes (NCEO). The NCEO is an organization that “provides national leadership in assisting state and local education agencies in their development of policies and practices that encourage and support the participation of students with disabilities in accountability systems and data collection efforts” (Shyyan, Lazarus, & Thurlow, 2015). The reports document the progress of alternate assessments for students with significant cognitive disabilities from data gathered directly through surveys of state directors. The second section will discuss types of alternate assessments, alternate achievement standards (AAS), scoring criteria, the validity of alternate assessments, and score reporting, which are reported throughout the national reports in section one. Types of alternate assessments, alternate achievement standards (AAS), scoring criteria, the validity of alternate assessments, and score reporting all provide the framework around which alternate assessments are developed and implemented by each state. These topics led to the research questions comparing the DLM and MSAA. This literature review resembles the research and topics sought out by state departments of education when deciding between the DLM and MSAA.

COMPARING TWO ALTERNATE ASSESSMENTS

Section One

National Center on Educational Outcomes Reports

The National Center on Educational Outcomes (NCEO), in conjunction with the National Association of State Directors of Special Education (NASDSE) and the Council of Chief State School Officers (CCSSO), “provides national leadership in assisting state and local education agencies in their development of policies and practices that encourage and support the participation of students with disabilities in accountability systems and data collection efforts” (Shyyan, Lazarus, & Thurlow, 2015). NCEO is an affiliated center of the Institute on Community Integration, which can be found at the College of Education and Human Development at the University of Minnesota. The organization’s website (www.nceo.info) states that it is primarily supported through “a Cooperative Agreement (#H326G160001) with the Research to Practice Division, Office of Special Education Programs, U.S. Department of Education.” Since 1992, the organization has been tracking and analyzing state policies and data on assessment participation and accommodations in the fifty United States and, in subsequent years, included some or all of eleven unique states (American Samoa, Bureau of Indian Education, Commonwealth of Northern Mariana Islands, U.S. Department of Defense Education Activities, District of Columbia, Federated States of Micronesia, Guam, Palau, Puerto Rico, Republic of the Marshall Islands, and U.S. Virgin Islands) (Shyyan, Lazarus, & Thurlow, 2015, p. vii). NCEO began tracking alternate assessments in 1996 even though states were first required to implement them in 2000. Although these reports include data about assessments and all students with disabilities, for the purpose of this study, the information shared focused only on information pertaining to alternate assessments and students with significant cognitive disabilities (Thompson & Thurlow, 2012).

COMPARING TWO ALTERNATE ASSESSMENTS

1996 Alternate Assessments for Students with Disabilities

In this report, it was found that only three states (Kentucky, Maryland, and Texas) were developing or had developed an alternate assessment for students who were unable to participate in the general state assessment (Thurlow, Olsen, Elliott, Ysseldyke, Erickson, & Aherarn, 1996). At the time of this report, Kentucky had implemented an alternate portfolio assessment, Maryland had developed and was field-testing an alternate assessment, and Texas was in the early stage of developing an alternate assessment. It was stated that “there is not much experience on which to build” and presented to the reader the three questions that should be asked when developing an alternate assessments (p. 2). The three questions are, “Who is to take the alternate assessment?,” “What should be assessed?,” and “How should the alternate assessment be integrated into the accountability system?” (p. 3). This report was the first of many to chart the states’ progress in alternate assessments.

1999 State Special Education Outcomes

In this report, based upon surveys returned to the NCEO, only forty-three states participated (Thompson & Thurlow). States were questioned about the standards that they used for alternate assessments (see Table 2-1), which were reproduced from the report (Table 8, p. 16).

Table 2-1 Alternate Assessment Standards

Alternate Assessment Standards	Number of States
Identical to those applied to general education	6
General education standards with some additions	1
Subset of those applied to general education	14
Independently developed for students needing alternate assessments	8
Uncertain at this time	14

COMPARING TWO ALTERNATE ASSESSMENTS

Of the forty-three states, the majority used the same standards as those used for statewide general education assessments or were uncertain at the time of this study. It appears that 18% of the states' standards were independently developed. As shown in Table 2-2, states were asked about where they were in the development of their alternate assessments (Table 9, p. 17). The majority of states were focused on identifying standards, establishing guidelines, and creating their alternate assessment systems.

Table 2-2 States Engaged in Various Alternate Assessment Activities

State Activity	Number of States
Identifying standards	34
Establishing eligibility guidelines	36
Identifying/creating instrument	32
Training on alternate assessment	12
Establishing proficiency levels	22
Determining reporting procedures	23
Determining inclusion in high stakes	18

In Table 2-3, states were asked about where they were in the development of their alternate assessments (Table 9, p. 17). The majority of states were using observations, portfolios, or performance assessments. It should be noted that this inquiry was given to forty-three states, and only twenty-seven states responded with an approach or type; therefore, sixteen (or 37%) of the states had not selected an approach or type for their alternate assessment at the time of this report.

COMPARING TWO ALTERNATE ASSESSMENTS

Table 2-3 Alternate Assessment Approaches Selected by States* (p. 17)

<i>Approach or Type</i>	<i>Number of States</i>
Observation (direct, video, or other)	8
Student portfolio	4
Performance assessment	4
Survey (mail or other) or Interview	3
Review of progress	3
Adapted regular state assessment	3
Adaptive behavior scale	2

The forty-three states were asked to estimate the limited exposure to the general education curriculum that would warrant the decision to have a student not take the general assessment. Only twenty-nine state directors responded to this question. Their estimates are presented below, in Table 2-4 (p. 17-18).

Table 2-4 Estimated Percentages of All Students Whose Exposure to Content is Too Limited for Them to Participate in Regular Assessment

<i>Percentage</i>	<i>States</i>
<1-1%	Delaware*, Kansas, Kentucky, Maryland, Minnesota, Nebraska, Vermont
>1-2%	California, Colorado, Idaho, Hawaii, Indiana, Florida*, Louisiana, Nevada, Oregon, Rhode Island, Virginia
>2-3%	Arkansas*, Connecticut, Massachusetts, Missouri, New Hampshire, New Mexico, Utah, Washington, Wisconsin
>4%	Mississippi, Ohio, South Dakota, Tennessee, Texas*, West Virginia

*State provided percentage of students with disabilities was transformed to a percentage of all students using the special education rate.

The information from this report illustrated that most states were in the beginning stages of development for their alternate assessments. The number of states in the development process

COMPARING TWO ALTERNATE ASSESSMENTS

is a tremendous increase from the 1996 report, which noted that only three states had begun the process. Despite this increase, there was great variety in the development reported. There was no reference to collaboration among the states or organizations that would expedite or enhance the development of alternate assessments. In 1996, each state was working independently to create the same assessment.

*2001 State Special Education Outcomes: A Report on State Activities
at the Beginning of the Decade*

The data within this report were taken four years after the mandated due date for each state to develop an alternate assessment for their students with significant cognitive disabilities. Within the survey, the NECO did not directly ask the states whether they had an alternate assessment; rather, they assumed from the responses from the states that most were working on some facet of their alternate assessments (Thompson and Thurlow, p. 11). Table 2-5 was modified from the original (Table 2, p. 11) to illustrate what the stakeholders implemented in the development of alternate assessments as reported by the districts. The majority of states' directors involved state and local personnel and parents to develop their alternate assessments.

Table 2-5 Stakeholders Involved in the Development of State Alternate Assessments

<i>Stakeholders</i>	<i>Number of States</i>
States special education personnel	50
Local special educators	50
State assessment personnel	49
Parents	44
Local school administrators	44
Local related service personnel	41

COMPARING TWO ALTERNATE ASSESSMENTS

Local assessment coordinators	39
Advocates	34
Local general educators	31
Adults with disabilities	8
Students	6
Other	8

The standards assessed by alternate assessments, for most states, are the same or related to the academic standards of general assessments. At the time of this report, Georgia, Mississippi, Nebraska, and Ohio only used functional skills on their alternate assessments. Local standards were used in Iowa, and two states—Texas and Wisconsin—deferred to individual IEP teams to determine what the student's alternate assessments would assess (p. 12).

Alternate assessment approaches, or types, were documented from the states (p. 13). In 2000, twenty-eight states used “portfolio/body of evidence,” four states used a “checklist,” five states used “IEP Analysis,” six states noted “other,” and seven states used “state has not decided.” In 2001, twenty-four states used “portfolio/body of evidence,” nine states used a “checklist,” three states used “IEP Analysis,” twelve states noted “other,” and two states had “not decided.” In comparison to the previous report, over three times the number of states was implementing a portfolio/body of evidence approach or type of alternate assessment. As of 2001, there were states still undecided on the approach or type of alternate assessment that they were going to implement.

This report examined how states measured student performance and system performance. Student performance refers to the performance of students on an assessment (pp. 13-14). It may include a measurement of an assessment section, task, or overall performance. In looking at student performance measures, forty states measured “skill/competence,” about thirty states

COMPARING TWO ALTERNATE ASSESSMENTS

measured “independence,” about twenty-four states measured “progress,” about eighteen states measured “ability to generalize,” and about six states reported “other” (p. 14). System performance refers to variables that have a direct and indirect effect on student performance. In looking at system performance measures, about twenty states measured “variety of settings,” twenty states measured “staff support,” twenty states measured “appropriateness,” twelve states measured “general education participation,” nine states measured “parent satisfaction,” and nine states measured “no system measures.” Appropriateness was defined as “age appropriate and challenging for students (p. 13). Levels of staff support, variety of instructional settings, and appropriateness were variables that were measured the most frequently among states.

Alternate Assessment performance descriptors were gathered from the states. Thirty-eight percent of the states used “different from general assessment,” 36% of the states used the “same as general assessment,” 18% of states had not decided, 6% of the states were “same + different,” and 2% of the states had “no performance levels” (p. 14-15). The majority of states used the same performance descriptors or different performance descriptors as the general assessment.

To score alternate assessments, about twenty-two states used the students’ teachers, about fourteen states used “teachers from other districts,” about twelve states used a “test contractor,” ten states reported “other,” and about four states reported “undecided” (p. 15). Almost three-quarters of the states used teachers when scoring a student’s alternate assessment.

Documentation, in this report, demonstrated that most states were at some stage of developing alternate assessment systems. In comparison to the previous report, more states participated. State directors were capable of sharing information about stakeholders in the development process, the standards assessed, assessment approaches, performance measures, and

COMPARING TWO ALTERNATE ASSESSMENTS

performance descriptors. The information given included more details, but the answers varied among the states, confirming variation among the alternate assessments used for students with significant cognitive disabilities across the United States.

2003 State Special Education Outcomes

This report included responses from all fifty states. At the time of this report, three-quarters of the states had an alternate assessment option for students with significant cognitive disabilities. Fewer states were using a portfolio or body of evidence as their alternate assessment approach or type, and using a rating scale or checklist increased. Only three states were developing or revising their approach or type of alternate assessment at the time of the report (Thompson & Thurlow, p. 13). These data are shown in Table 2-6.

Table 2-6 Alternate Assessment Approaches or Types

<i>Year</i>	<i>Portfolio or Body of Evidence</i>	<i>Rating Scale or Checklist</i>	<i>IEP Analysis</i>	<i>Other</i>	<i>In Development/ Revision</i>
1999	28 states	4 states	5 states	6 states	7 states
2001	24 states	9 states	3 states	12 states	2 states
2003	23 states	15 states	4 states	5 states	3 states

Alternate Assessment content was surveyed from all fifty states. The authors noted that the option responses increased from 2003. Seventy-two percent of states were using “grade level or expanded state/district academic content standards,” 8% states were using “combination of state standards and functional skills,” 4% of states were using “functional skills (no alignment to standards),” 6% of states used “IEP teams determine alternate assessment content,” 4% states were “currently developing/revising,” and 6% of states reported using “other” (p. 14). There was a dramatic shift in states’ use of academic content standards from the prior report. This shift

COMPARING TWO ALTERNATE ASSESSMENTS

demonstrated the thought process of assessing students with significant cognitive disabilities based on academic standards rather than functional skills.

Alternate Assessment performance descriptors were gathered from the states. Thirty-two percent of the states used the “different as general assessment,” 62% of the states used the “same as general assessment,” and 27% of the states were “currently developing/revising” (pp. 14-15). States reported their student's assessment results using level descriptors (i.e., below proficiency, proficient, advanced proficient), and examples of these descriptors were noted to show variation. Arkansas used “independent, functional independent, supported independent, emergent, and non-evident” assessments, while Colorado used “inconclusive, exploring, emerging, developing, novice,” Georgia used “initial, emerging, proficient, functional,” Indiana used “not evident, emerging, developing, demonstrated,” Illinois used “attaining, progressing, emerging, attempting,” West Virginia used “awareness, progressing, competent, generalized,” and Wyoming used “beginner, partially skilled, skilled” (p. 15).

Forty states scored their alternate assessment using a rubric, eight states used a rating scale with points, five states scored correct items, and two states scored “reading rate or accuracy” (p. 16). In this report, a great variation of scoring criteria among the states was found. In the prior report, the criterion for measuring student performance and system performance were separated. In the 2003 report, all criteria were combined (Tables 2-7). The greatest response from states regarded independence, which is the level of support and skill/competence that a student needed to complete a task.

COMPARING TWO ALTERNATE ASSESSMENTS

Table 2-7 Alternate Assessment Scoring Criteria for Rubrics (p. 17)

<i>Measured Outcomes</i>	<i>Percent of States</i>
Independence	66%
Skill/competence	54%
Variety of settings	46%
Progress	42%
Ability of generalize	42%
Appropriateness (age, challenge, authenticity)	38%
Alignment to standards	34%
Access to support	24%
Self determination	24%
Staff support	22%
Social relationships	18%
General education participation	12%
Parent satisfaction	2%
Other	4%

Table 2-8 shows who was used to score the alternate assessments. The majority of states used the students' teachers or teachers from other districts. These results were similar to those found in the prior report.

Table 2-8 Alternate Assessment Scorers (p. 18)

<i>Scorers</i>	<i>Percentage of States</i>
Student's teacher of IEP member	36%
Teachers from other districts	32%
Test contractor	24%

COMPARING TWO ALTERNATE ASSESSMENTS

State education agency	12%
Teachers within a student's district	6%
Currently developing/revising	8%
Other	16%

States were surveyed to determine if they were practicing a standards-setting process for their alternate assessments. In their responses, it was found that 52% of the states used a process, 14% did not have a process, 16% replied “other,” 10% replied “don’t know,” and 8% used an “informal process” (p. 20). For those states that did use a standard-setting process, fifteen (Alaska, Arkansas, Connecticut, Delaware, Illinois, Louisiana, Massachusetts, Michigan, Minnesota, New Hampshire, New York, North Dakota, Rhode Island, Virginia, and West Virginia) used “Reasoned Judgment,” six states (Alabama, California, Colorado, New Mexico, Pennsylvania, and South Carolina) used “Bookmarking or Item Mapping,” three states (Arkansas, North Carolina, and Pennsylvania) used “Contrasting Groups,” two states (Arkansas and Tennessee) used “Body of Work,” and one state (Kansas) used “Judgmental Policy Capturing.”

This was the second report to include responses from all fifty states. One quarter of the states still did not have fully implemented alternate assessments, and there was a shift in the approach or type of alternate assessments used from a portfolio to a rating scale or checklist. The alternate assessments focused more on academic standards than functional skills, but the academic standards used were the same as the general assessment. Therefore, the students with significant cognitive disabilities were given the option of using alternate assessments due to their inability to participate in the general assessment, but the content was linked to the same standards. Although a variety of choices evolved among the states’ responses regarding

COMPARING TWO ALTERNATE ASSESSMENTS

measured outcomes, the most selected outcomes were still consistent across the states. A formal standard-setting process was only used by 52% of the states, reflecting that the states were still working on refining how to find validity and reliability among each state's alternate assessments.

2005 State Special Education Outcomes

All fifty states participated in this report. The results from the surveys revealed a discrepancy in how states reported their participation rates for students with disabilities and how each state defined the population of students with significant cognitive disabilities (Thompson, Johnstone, Thurlow, & Altman, 2005). Each state documented its students' use of accommodations. Three years of data from evaluations provided states the knowledge required to note that there was an increase in the number of students who earned a "proficiency" on state accountability assessments, which coincides with the fact that students were receiving instruction to meet grade-level expectations. At the time of the report, forty-five states offered AA-AAS. The types of alternate assessments used by states were presented in Table 2-9, which is modified below from the original (Table 3, 2000-2005, p. 11). There was a reduction in the number of students who used a rating scale or checklist and IEP analysis.

Table 2-9 Alternate Assessment Approaches or Types

<i>Year</i>	<i>Portfolio or Body of Evidence</i>	<i>Rating Scale or Checklist</i>	<i>IEP Analysis</i>	<i>Other</i>	<i>In Development/ Revision</i>
1999	28 states	4 states	5 states	6 states	7 states
2001	24 states	9 states	3 states	12 states	2 states
2003	23 states	15 states	4 states	5 states	3 states
2005*	25 states**	7 states***	2 states	7 states	8 states

*One state has not developed any statewide alternate assessment approaches.

**Of these 25 states, 13 use a standardized set of performance/events/tasks/skills.

***Of these 7 states, 3 require the submission of student work.

COMPARING TWO ALTERNATE ASSESSMENTS

Alternate Assessment content was surveyed from all fifty states. The authors noted that the option responses increased from 2003. Twenty-one states were using “extended/expanded state/district academic content standards,” ten were using “grade level state/district academic content standards or grade level expectations,” one was using “state/district academic standards at one or more grade levels below the student’s current grade,” one used “a combination of state/district academic standards and functional skills not aligned to standards,” one used “IEP teams determine alternate assessment content for each student,” six reported using “other,” and ten reported that they were “currently revising.” In how states reported their student's assessment results using level descriptors (i.e., *below proficiency*, *proficient*, *advanced proficient*), examples of these descriptors were noted. Arkansas used “independent, functional independent, supported independent, emergent, and non-evident”; Colorado used “novice, developing, emerging, explorative, exploring, and inconclusive”; Georgia used “initial, engineering, progressing, and functional”; Illinois used “attempting, emerging, progressing, and attaining”; and New Mexico used “insufficient data, beginning step, nearly proficient, proficient, and advanced” (p. 23). States responded to questions about their AA scoring criteria with the most frequent response being rubrics. Their responses are provided in Table 2-10, which was modified from the original (Table 5. Scoring Criteria for Alternate Assessment Responses 2003-2005, p. 13).

COMPARING TWO ALTERNATE ASSESSMENTS

Table 2-10 Scoring Criteria for Alternate Assessment Responses 2003-2005 (p. 13)

Year	Rubric	Points Assigned on a Rating Scale	Number of Items Correct	Reading Rate or Accuracy
2003	40 states	8 states	5 states	2 states
2005	37 states	16 states	12 states	4 states

Within these rubrics, the measured outcomes from the alternate assessments were reported. The results are presented below (p. 14). Skill/competence and level of assistance were the top measured outcomes on the alternate assessment, followed by degree of progress. The top measured outcomes focused on student performance.

Table 2-11 Measured Outcomes from the Alternate Assessments (p. 14)

Measured Outcomes	Number of States
Skill/competence	25
Level of assistance	25
Degree of progress	23
Number/variety of settings	20
Alignment with academic content standards	18
Ability of generalize	15
Appropriateness	13
Staff support	10
Social relationships	10
Self determination	9
Participation in general education settings	7
Support	7

COMPARING TWO ALTERNATE ASSESSMENTS

States were surveyed to identify their formal standard-setting processes for their alternate assessments. In their reports, it was found that twenty-six states used “formal standard-setting process,” three states used an “informal standard-setting process,” five states were “developing a standard-setting process,” thirteen states were “developing/revising alternate assessments,” two states reported “no standard-setting process,” and one state reported “other” (p. 14). States that used a formal standard-setting approach were further questioned about the techniques that they used. It should be noted that some states used more than one technique. Eleven states (Alabama, Delaware, Kentucky, Massachusetts, North Carolina, New Mexico, Nevada, Ohio, South Dakota, Virginia, and Vermont) used a “body of work”; Ten states (Arizona, Delaware, Hawaii, Illinois, Indiana, Massachusetts, Mississippi, South Dakota, Wisconsin, and West Virginia) used “reasoned judgment”; nine states (Arizona, California, Colorado, Indiana, Kentucky, Maryland, Montana, Oregon, and South Carolina) used “bookmarking or item mapping”; five states (Alaska, Massachusetts, Maryland, Nevada, and South Dakota) used “contrasting groups”; and five states (Kentucky, Massachusetts, Nevada, South Dakota, and Vermont) used “judgmental policy capturing” (p. 15). There was still variation among states regarding if and how they set their assessment standards.

There was more evidence that states were invested in their alternate assessments. Feedback from the surveys illustrated various ways that the states were meeting their requirements. Scoring and the interpretation of the scoring still varied, but states were consistently scored by the same group of professionals. About half of the states still lacked formal standards setting, which is needed to demonstrate validity and reliability.

COMPARING TWO ALTERNATE ASSESSMENTS

2009 State Policies on Assessment Participation and Accommodations

This report further documented the continual growth and changes to alternate assessments across the United States. By the time of this report, all fifty states aligned their AA-AAS with “grade-level or with extended (or expanded) academic content standards,” and Hawaii, Idaho, Mississippi, Nebraska, New Hampshire, Nevada, and Utah were in the process of revising their AA-AAS (Christensen, Braam, Scullin, & Thurlow, 2011, p. 20).

Alternate assessment approaches or types were revisited for the states. The results, presented in Table 2-12 (Table 3. AA-AAS Test Formats, p. 20), confirmed that “portfolio or body of evidence” continued to be the most popular format. The second most popular answer—“standardized set of performance tasks”—was not an option in prior reports. Both the “portfolio or body of evidence” and “standardized set of performance tasks” were used by over three-quarters of the states.

Table 2-12 AA-AAS Formats

<i>Format</i>	<i>Number of States</i>
Portfolio or Body of Evidence	20*
Standardized Set of Performance Tasks	18**
Multiple Choice Test	8
IEP Analysis	0
Other	2
Currently in revision	7

*Of these 20 states, 8 used a standardized set of performance tasks

**Of these 18 states, 8 required the submission of evidence

In the revisited AA-AAS scoring methods, rubrics were still the most frequently used method, but the number of states (33) that used rubrics decreased compared to previous reports. Nineteen states marked alternate assessments by the “number of items correct,” eight states

COMPARING TWO ALTERNATE ASSESSMENTS

scored “points assigned on a rating scale,” two scored “Reading rate or accuracy,” and seven were currently in revision at the time of this report (p. 21).

Measured outcomes responses transitioned into more academic measures, with less focus on non-academic measures such as social relationships and self-determination. This is reflected in the “skills/competence” and “alignment with academic content standards” as common outcomes. It was noted by the authors that “skill/competence” was considered an identical category for states that reported “accuracy” (p. 20). All of the results are presented below in Table 2-13 (Figure 18, Outcomes Measured by Rubrics, p. 20).

Table 2-13 Measured Outcomes from the Alternate Assessments

Measured Outcomes	Number of States
Skill/competence	25
Level of assistance	24
Degree of progress	6
Number/variety of settings	10
Alignment with academic content standards	15
Ability of generalize	10
Appropriateness	8
Staff support	5
Social relationships	3
Self determination	4
Participation in general education settings	5
Support	6
Parent satisfaction	1

COMPARING TWO ALTERNATE ASSESSMENTS

In looking at who was scoring the AA-AAS, the top response from previous reports was dramatically overturned. Test company contractors were used by twenty-one states, followed by local personnel. These data suggest that there was an increase in the use of test companies by states. There was also a dramatic decrease in the use of state agencies or state personnel.

For determining achievement levels, fifteen states used a “body of work,” twelve states used “bookmarking,” eight states used “contrasting groups,” seven states used “reasoned judgment,” five states used “Modified Angoff,” and four states used “other method” (p. 23). In these data, some states used more than one method. “Body of work” and “bookmarking” appeared to be the most frequently-used methods to determine achievement levels.

2012 Survey of States

This report was written twelve years after states were mandated to develop an alternate assessment for students with significant cognitive disabilities. All fifty states were represented in the report. It included the acknowledgement of alternate assessment consortiums. This was the first report to name specific alternate assessments—the DLM and the NCSC—that were not developed by one state, but through a consortium of states and organizations. States were surveyed to note the participation of their various staff in education, which helped in the development of guidelines and policies through an alternate assessment consortia. Their responses are provided in Table 2-14 (Ricke, Lazarus, Thurlow, & Dominguez, p. 13)

COMPARING TWO ALTERNATE ASSESSMENTS

Table 2-14 State Contribution to Alternate Assessment Consortia (DLM, NCSC) Decision Making*

	Participation Guidelines	Assessment Claims	Technology Decision making	Assessment Scoring Policy	Accommodations Guidelines	Performance level Descriptors	Item Development	Reporting Formats
Senior Administration Staff	11	10	10	7	10	7	9	6
Curriculum and Instruction Staff	4	5	3	5	5	6	7	3
Special Education Staff	25	17	17	18	26	19	20	15
Assessment Staff	24	20	17	17	23	17	16	17
Technology Staff	0	0	8	1	1	0	1	2

*only included states involved in an alternate assessment consortia.

With regard to alternate assessments based on alternate achievement standards, thirteen states had made major revisions to their alternate assessments, and thirty-two states had not made any major revisions to their alternate assessments since the 2009 NCEO report.

This report had dramatically reduced in data compared to previous reports about the states' alternate assessments. It represented a plateau of data about initial development and the implementation of alternate assessment systems. The information shared focused on the growth of consortiums, especially the DLM and the NCSC. No other alternate assessment was specially mentioned.

COMPARING TWO ALTERNATE ASSESSMENTS

2014 Survey of States

This report acknowledged the fact that many states reported participation in an assessment consortia through a “a variety of consortium-led activities, including pilot tests, field tests, and special accessibility studies” (Shyyan, Lazarus, & Thurlow, 2015, p. 4). As of 2014, eighteen states reported participation in the DLM consortia and eighteen reported participation in the NCSC (or MSAA) consortia. The report stated that “most” states “revised their alternate assessment based on alternate achievement standards (AA-AAS) between 2012-2014” (p. 23). Changes to forty-three states’ AA-AAS were documented, with twenty-eight of these states “making major revisions” (p. 23). AA-AAS for end-of-course assessments were incorporated in state assessments provided by seven states, with only six additional states noting that they administered AA-AAS for some of their courses. Some of these courses included English language arts, mathematics, science, social studies, and biology (p. 23).

This report mirrored the 2012 report, lacking the same information about alternate assessments for students with significant cognitive disabilities. Data about alternate assessment formats, measured outcomes, scorers, and performance indicators were non-existent. This was the second report to note data on the DLM and the NCSC without mentioning any other specific alternate assessment.

When the NCEO began inquiring, in 1996, about each state’s alternate assessment for students with significant cognitive disabilities, only three states had information to contribute. The report concentrated on three questions for states to apply to their development process, illustrating the infancy of alternate assessment systems development for at least forty-seven states. From 1999 through 2009, NCEO’s reports documented the slow progression of alternate assessment development by the states. The information presented revealed variation among the

COMPARING TWO ALTERNATE ASSESSMENTS

states' alternate assessments. The data unveiled how each state was working individually and in isolation to create an alternate assessment for students with significant cognitive disabilities. As the reports continued over the years, so did the increase of refined responses to the surveys. Within these ten years, the types of alternate assessments, alternate achievement standards, scoring criteria, scoring reports, and the beginning stages of standard settings were the focused topics. From 2012 through 2014, the reports dropped the prior topics and dramatically shifted to alternate assessment consortiums, specifically focusing on the DLM and NCSC (later known as the MSAA). This shift demonstrated the evolution of many state departments to collectively working together to develop and implement alternate assessments. The information shared about these consortiums was limited in comparison to the volume of information that was shared in prior reports, demonstrating the limited availability of information on these two alternate assessment consortiums.

Section Two

Types of Alternate Assessments

The requirement of creating an alternate assessment became the responsibility of each individual state, which had to maintain their compliance by offering an alternate assessment that was reliable and demonstrated accountability (IDEA 2004 612(a)(16)(c)). Alternate assessments required a collection of work samples to demonstrate a student's performance against a set of standards. All alternate assessments approaches that have been developed by states could be categorized as Portfolio Assessment, Performance Assessment, or Comprehensive Rating Scales of Achievement (Elliot & Roach, 2007, p. 301). These approaches included a collection of student work samples, a direct measure of student skills, and a checklist of skills (Quenemoen, Rigney, Thurlow (2002).

COMPARING TWO ALTERNATE ASSESSMENTS

Information was gathered about the alternate assessments for eight states (Delaware, Florida, Georgia, Indiana, Minnesota, North Dakota, Utah and West Virginia) (Burdette & Olsen, 2000). In looking at the types of alternate assessments chosen, six states (Delaware, Florida, Indiana, Minnesota, North Dakota, and West Virginia) used a portfolio format and two states (Georgia and Utah) used an IEP Progress format (p. 40). Despite the type of format used for each state's alternate assessment, all of the alternate assessments were individualized in some way through the student's IEP and all required the gathering of student performance evidence by the student's teacher. In contrast, all of the alternate assessments were different in their content, even within the same alternate assessment type.

Kingston, Karvonen, Bechard, and Erickson (2016) described the type of portfolios used for alternate assessments (p. 3). Portfolios "allow for a more flexible evaluation of students academic skills and more opportunity to choose academic content that fits best with the student's curricular priorities" (p. 3). There is a concern that alternate assessment portfolio scores are affected by the ability or skills of the teacher to construct the portfolio than the student's actual performance (p. 3). Flowers et al. (2007) reported that performance assessments are better aligned to alternate achievement standards. A drawback to the performance alternate assessments was that it prevented the use of accommodations due to the standardization of its administration.

Kohl, McLaughlin, and Nagle (2006) reported that ten states used portfolios as the primary tool for their alternate assessments. "Student products, anecdotal accounts, data sheets, data graphs, audio- and video tapes, rating scales, and photographs" were included as evidence in the portfolios (p. 116). Six states were reported as using performance tasks as their alternate assessment and two states used teacher checklists or inventories as alternate assessments. Three

COMPARING TWO ALTERNATE ASSESSMENTS

states reportedly allowed districts to “determine how individual students will be assessed and report the use if variety if instruments including student products, anecdotal records, and standardized commercially available assessments” (p. 116). Even among ten states, there was disparity on the types of alternate assessment approaches that were used. This was pattern that continued seven years later, when Wyse, Dean, and Viger (2013) found it hard to equate alternate assessments across the country due to the variety of alternate assessments, number of test items, and scoring practices. Using two scorers to assess student responses may “lead to oddly shaped score distributions” (p. 52).

Based on the research, there are three general types of alternate assessments that states tend to mandate as their statewide assessment for students with significant cognitive disabilities. Having a variety of assessment types makes it difficult to make comparative assumptions from state to state. Selecting an alternate assessment is also complicated by the fact that students with significant cognitive disabilities are a small population that is already so diverse that it clouds the ability to compare assessments. There is also a lack of literature on technology and its effect on the three types of alternate assessments. Therefore, selecting the ideal type of alternate assessment to adopt within a state is a complicated decision.

Alternate Achievement Standards (AAS)

Although IDEA 1997 provided the groundwork for requiring all students who received special education access to and progress in general education curriculum, NCLB, IDEA 2004, and ESSA confirmed the expectation that all students are held to state academic standards (Quenemoen, 2008, p. 10). “Alternate assessments are to be aligned to (or “linked to” in later terminology related to peer review) the state content standards in each grade” (p. 10). States had to demonstrate that their alternate assessments adequately measured a student’s performance

COMPARING TWO ALTERNATE ASSESSMENTS

against alternate academic standards. Under NCLB (2001), states were given the task of creating AAS in response to setting a standard to determine student performance (Kingston, Karvonen, Bechard, & Erickson, 2016, p. 3). There was great variation in what states determined as their AAS. Some states reflected on their general academic content standards, while others linked their state content standards but reduced them to allow students with significant cognitive disabilities access (p. 3). The results since 2001 have been slow and not very productive. This may be because many states have focused the majority of their attention on creating academic standards for the general education population, which represented approximately 91% (Almond & Case, 2004), resulting in little progress in the development of AAS.

In 2004, Browder et al. documented whether states focused their alternate assessments for students with significant cognitive disabilities on academic or functional performance. Kentucky was the first state to develop an alternate assessment and the first to shift its performance standards from a functional to an academic focus (p. 213). Within the study, the authors wanted to examine if the content from alternate assessments was aligned with academic standards or “functional life domains” (p. 213). The study included multiple groups as reviewers. Experts in the general content domains (i.e., language arts, mathematics), special education domains (i.e., disabilities specialists), school special education teachers and administrators, and national experts/researchers participated in reviews of the performance indicators for thirty-one states’ alternate assessments. Feedback from all of the reviewers found that performance indicators from three states were in true alignment with language arts and mathematics standards. Some states labeled mathematics or language arts standards that were truly functional skills. About one third of the thirty-one states used academic and functional performance indicators in their alternate assessments. It was specifically noted by the reviewers that Arizona’s alternate

COMPARING TWO ALTERNATE ASSESSMENTS

assessment included “well-aligned performance indicators in language arts” and “major functional life domains” (p. 218). Connecticut’s alternate assessment was noted as having “good access to the general curriculum using a functional approach, was also strong in representing the major life domains” (p. 218). The reviewers reflected that states exhibited a blended focus on academic and functional skills, which expanded the performance indicators, not restricting them. Overall, the study emphasized that functional skills are still a part of the performance indicators for many of the states’ alternate assessments. Even among thirty-one states, there is still variety regarding how each state uses their performance indicator.

In 2006, Kohl, McLaughlin, and Nagle reported how sixteen states established their alternate achievement standards. Nine out of sixteen states established levels of performance (i.e., proficient, advanced proficient, and basic) using a committee of stakeholders from the school community. State personnel determined levels of performance in four states. Another state used state personnel and a test contractor. One state depended upon their IEP teams, and four other states used a committee of stakeholders and “statistical techniques.” At the time of press, one state had not yet determined their levels due to the newness of their alternate assessment (p. 118). There was no norm for aligning academic standards during this time across the United States, demonstrating the disparity in the development in AAS.

In August of 2009, Cameto, Knokey, Nagle, Sanford, and Blackorby summarized the national findings from a report that profiled alternate assessments from all of the fifty states by the National Study on Alternate Assessments (NSAA). The original report was based on the 2006-2007 school year. States were surveyed about their alternate achievement standards (AAS) used in conjunction with their alternate assessment for students with significant cognitive disabilities. States were surveyed about the individuals who are involved in creating the AAS.

COMPARING TWO ALTERNATE ASSESSMENTS

The three most frequent responses were special education teachers (96%), content specialists (84%), and state special education staff (82%) (p. 23). The two most selected standard-setting methodologies used by states were body of work (31%) and bookmark or item mapping (24%) (p. 26). States focus on the teachers (who work directly with the students), the content experts, and the professionals who implement the assessments for the state (i.e., the state special education staff) when creating the standard for measuring student growth in the alternate assessments. This group of stakeholders, in the process, will have a strong impact on the creation of the AAS. The types of standard-setting methodologies used most often are usually used with a portfolio of body of evidence and items scored with a right/wrong answer or rubric. This, most likely, coincides with the most frequent types of alternate assessments used by the states.

In a survey of teacher perceptions about AAS by Restorff, Sharpe, Aberly, Rodriguez, and Kim (2012), 51% of teachers selected that the AAS “Better align special education with the state standards” and 50% selected that AAS did “Improve/help to develop IEP goals” (p. 190). Only 9% reported that AAS “Increase opportunities for Inclusion,” and 7% answered “Results in an increase of classroom materials” (p. 190). There was a range of between 7% and 18% of states that selected “Providing greater access to the general education curriculum” as a benefit of AA-AAS. Based upon this information, there appeared to be a little or no relationship between the AAS and its affect on increasing students’ performance from a teacher’s standpoint. It could be argued that teachers felt the pressure to include IEP goals that aligned with state standards. It is difficult to confirm that it directed instruction to increase student academic growth.

Kingston et al. (2016) noted that AAS are grounded on the fact that these standards are “unidimensional,” and it is an unfair measurement when they “cannot demonstrate their

COMPARING TWO ALTERNATE ASSESSMENTS

knowledge in the same way” (p. 4). Despite growth, Kingson et al. noted that there was “limited ability to measure growth,” “difficulty demonstrating the quality of assessment systems due to the diverse population, and “perceived irrelevance” of the results of the alternate assessment based upon the AAS. The recommendation included achievable AAS that were linked or aligned to general education academic standards.

Fifteen years since the requirement to develop AAS was mandated under NCLB, researchers are still requesting achievable AASs that align with or are linked to general education academic standards for students with significant cognitive disabilities. Studies reveal inconsistency among states in how they develop their AAS. Teachers are still waiting for an AAS that drives instruction and IEP goals that keep academic expectations high and that realistically matches the academic needs of these students. More research and development are needed to further expand our knowledge about academic standards and how they can be more accessible for the students who take alternate assessments.

Scoring Criteria

All assessments should evaluate a student’s performance against a set of academic standards that includes students with significant cognitive disabilities who take alternate assessments. Creating scoring criteria to define student outcomes on an alternate assessment is challenging. Assessment types diversify the scoring criteria because they are highly differentiated and are rarely item scored.

Eight states (Delaware, Florida, Georgia, Indiana, Minnesota, North Dakota, Utah, and West Virginia) were selected to compare their development of areas, such as scoring criteria, with their alternate assessment systems. Before addressing scoring, assessments approaches or type were shared. Delaware, Florida, Indiana, Minnesota, North Dakota, and West Virginia were

COMPARING TWO ALTERNATE ASSESSMENTS

using portfolios, also known by these states by datafolio, body of evidence, and electronic portfolio (p. 39). Georgia uses individual IEP goals and objectives and Utah used a functional framework. When comparing scoring, it is difficult when one state is scoring individualized goals and objectives and another state is correcting pen/paper items. The authors only identified three different scoring criteria: student performance, opportunities to learn, and both (p. 42). Florida, Georgia, Indiana, Minnesota, and North Dakota focus on student performance for their student criteria. Delaware and West Virginia used student performance and opportunities to learn as their scoring criteria (p. 42). The authors stated that they believed that developing an alternate assessment system is a “fluid process,” one that will take three to five years in order to fully develop scoring criteria and implementation (p. 3).

In a report (2003) prepared by Quenemoen, Thompson, and Thurlow, the scoring criteria for five alternate assessments were compared and contrasted. Scoring criteria were separated into student scoring criteria and system scoring criteria. Student criteria focused on student performance, and system criteria focused on variables that indirectly affected student performance. Arkansas uses a portfolio approach. Students are scored using skill/competence and level of independence. Kentucky also uses a portfolio approach, scoring students on skill/competence, degree of progress, level of independence, and ability to generalize. System scoring criteria included staff support, variety of settings, appropriateness (age appropriate, challenging, authentic), parent satisfaction, and participation in general education. Louisiana used performance assessment. Student performance was scored using skill/competence, level of independence, and other; no score was given for system performance. Oregon used a combination of approaches for their alternate assessment (checklist, performance assessment, and pencil/paper test). Students were scored using the number of correct items on the

COMPARING TWO ALTERNATE ASSESSMENTS

pencil/paper exam. System scoring criteria used a variety of settings and appropriateness (age appropriate, challenging, authentic) on performance assessment. Vermont used an evidence/IEP linkage alternate assessment. Students were scored using skill/competence and degree of progress, and system-scoring criteria used *other* (not described) (pp. 14-15). The common criteria across the five states are content standard linkage, independence, appropriateness, IEP linkage, and performance. Differences in criteria across the five states are system vs. student, mastery, progress, and single state criteria (criteria not used by other states). The report states that there was no right or wrong approach or criteria. Although there are similarities, there are many varieties among the similarities. For example, for the independence scoring criteria, what types of prompts are included? This study provided exposure to the scoring criteria implemented by other states while trying to find similarities and differences.

Using the alternate assessments for students with significant cognitive disabilities of sixteen states (California, Colorado, Delaware, Florida, Indiana, Kansas, Louisiana, Maryland, Massachusetts, Nebraska, New York, North Carolina, Ohio, South Carolina, Texas, and West Virginia), Kohl, McLaughlin, and Nagle (2006) conducted a descriptive investigation (p. 111). This investigation included how the states scored their alternate assessments. Ten of the states “addressed dimension of independence in their scoring procedure” (p. 118). Task independence was scored through standards in several states. Two states established “levels of functioning” (p. 118). This means that the states assessed using the level of support or lack of support given toward the performance (i.e., supported, independent). For scoring criteria, some states required teachers to document the level of support for each task. Independence is scored by about 69% of the sixteen states investigated. Scoring a student’s independence on a task reflects the evaluation

COMPARING TWO ALTERNATE ASSESSMENTS

of a student's performance. There appears to be variety in the criterion for how states evaluate "independence" in the alternate assessment.

Scoring criteria varied among the different state alternate assessments for students with significant cognitive disabilities. Even when states listed the same criteria, the interpretation or implementation of those criteria could vary greatly. Separating the student and system scoring criteria diversifies the alternate assessments. It also complicates comparing alternate assessments from student to student or state to state. Once states start to collaborate more on developing or revising their alternate assessment systems, there will be more consistency among scoring criteria.

Validity of Alternate Assessments

Under NCLB, state assessments are required to "be used for purposes for which such assessments are valid and reliable, and be consistent with relevant, nationally recognized professional and technical standards" (20 U.S.C. 6311 § 1111 (b)(3)(C)(iii)). The requirement for assessments was also duplicated in the passing of ESSA (Every Student Succeeds Act, 2015). The U.S. Department of Education published information on specific guidelines for demonstrating validity evidence for AA-AASs in January of 2009 (Office of Elementary and Secondary Education, 2009). Kohl, McLaughlin, and Nagle (2006) noted that, in nine states, state representatives shared that their states conducted studies for validity and alignment on their alternate assessments. Several states conducted these studies through their state education agencies and others used a test contractor (p. 117). In Rachel Quenemoen's report (2008), Quenemoen noted that validity still was an area of consideration for states (p. 25). At the time of her report, many states were participating in validity studies through various grants offered by the United States Department of Education's Office of Special Education Programs (p. 25).

COMPARING TWO ALTERNATE ASSESSMENTS

Quenemoen stated that there is little experience and knowledge about large-scale alternate assessments and measuring competence in content areas over the past couple of decades. She felt that the results of these assessments must be “defensible” and valid. Quenemoen’s considerations for alternate assessments were a valiant effort in proclaiming that alternate assessments must be held to a strong standard to demonstrate their validity.

Johnson and Arnold (2004) were the first researchers to publish a study about the validity of a portfolio alternate assessment. Their study examined the Washington Alternate Assessment System (WAAS). Through their study, Johnson and Arnold concluded that that the results “indicated serious shortcomings in the evidence for content, response process, and structural validity,” which was due to the fact that the portfolio format did not accurately assess the student’s performance against an IEP skill or content standard (pp. 266, 273). This study highlighted the absence of validity, but it noted areas of improvement for future studies.

Roach, Elliot, and Webb (2005) focused their article on the Wisconsin Alternate Assessment (WAA). One of their research questions was, “Does the WAA adequately measure the concepts and skill areas represented in Wisconsin’s Model Academic Standards?” (p. 220). The WAA format includes a “behavior rating scale based on the state’s alternate performance indicators (APIs), a downward extension of the state’s academic standards,” and an “overall scoring continuum for each core subject area which allows student performance to be categorized” by proficiency levels that mirror the general assessment (p. 219).

In a study by Elliot, Compton, and Roach (2007), the Idaho Alternate Assessment (IAA) was matched with two norm-referenced teacher rating scales to gather information about the validity of the IAA. The IAA used a rating scale that was approved by the USDOE in 2006, assessing the content areas of reading, language arts, and mathematics (pp. 32, 34). The study

COMPARING TWO ALTERNATE ASSESSMENTS

intended to gather information about the IAA's construct measures (p. 40). The results demonstrated "meaningful amount of construct-irrelevant variance in the IAA scores, yet the scores are functioning rather well in the differentiating performances by known groups of students" (p. 42). The authors suggested that this validity will increase over time, with more professional development on alternate assessment administrations and with IEPs including more academic skills. The results from this study provided a platform for supporting the validity of rating scale assessments and identifying a model of an alternate assessment that demonstrates validity.

In 2009, Cameto et al. summarized the national findings from a report that profiled alternate assessments from all of the fifty states by the National Study on Alternate Assessments (NSAA). The original report was based on the 2006-2007 school year. States were surveyed about the validity of their alternate assessment for students with significant cognitive disabilities. When asked about the individuals who are involved in reviewing validity, the top three responses were outside experts (86%), state assessment staff (80%) and, tied for third, test vendor (60%) and special education teachers (60%) (pp. 32-33). Fifty-seven percent of states responded that they did not "claim or document the validity of the alternate assessment in terms of scoring and reporting structures consistent with the subdomain structures of its content standards" (p. 34). In documenting the validity of their alternate assessment "in terms of test and its scores related to internal or external variables as intended, 41% responded that a formal study was conducted and 47% responded that they did not conduct a formal study (p. 37). Eighty-six percent of the states used a correlational study to indicate validity on their alternate assessments (p. 39). How states document the validity of their alternate assessments has progressed. It appears that more states are delving deeper into the validity of multiple lenses of the assessment and including more

COMPARING TWO ALTERNATE ASSESSMENTS

perspectives, such as test vendors, in the process. More formal studies are lacking in many of the states, however, and most likely will continue to be put in place over time.

The alternate assessments from six states (Arizona, Hawaii, Idaho, Indiana, Mississippi, and Nevada) were used in a study conducted by Kettler et al. (2010). Their study derived from a call to answer questions about the “constructs measured and their relationship to other measures of achievement remain largely substantiated by rigorous research a validation studies” (p. 458). Published research on constructs measured by AA-AASs is miniscule due to the fact that the state populations of students with significant cognitive disabilities are not considered adequate sample sizes to conduct MTMM studies (p. 470). In this study, a rating scale AA-AAS was used by all six of the states. A multitrait-multimethod design was used to “determine the relationship among the AA-AASs, the state’s general achievement tests, and two established teacher-based rating scales” (p. 460). The study found that “in most states, the relationships among content areas (typically the correlation between reading ad mathematics) within the AA-AAS are in the range that would be acceptable reliability coefficient for a single, unitary construct” (p. 470). It is suggested from the results that a number of constructs are measured by alternate assessments. As with the characteristics of the population of students with significant cognitive disabilities, the number of students is small and heterogeneous, which complicates research by limiting generalizations.

Teachers were evenly split on their opinions about whether the items on the alternate assessment accurately reflected the alternate academic standards and their students’ actual performance (Restorff, Sharpe, Abery, Rodieriz, & Kim, 2012, p. 189). They argued that the population of students who met the criteria to take an alternate assessment were too diverse to show validity appropriately and represented a small percentage of the total student population.

COMPARING TWO ALTERNATE ASSESSMENTS

Students, even in the same classification category, vary greatly in their needs for accommodations. It was difficult to demonstrate validity among state alternate assessments since there was such disparity among the various types of alternate assessments (p. 52).

In looking at the research and studies on the validity of alternate assessments, it is understandable why validity is difficult to document. Most alternate assessments are not based on straight measurements to observe and calculate with, as with the test items found in large-scale assessments. To complicate matters, the populations of students who take the alternate assessment are as heterogeneous as the types of test items and student products. There appears to be very little scientific research and development on how states demonstrate the validity of their assessments. Federal policy dictates that validity is necessary, but there is little accountability for each state to publicly show how their alternate assessment is valid.

Score Reporting

Under NCLB and ESSA, districts are required to provide parents with information about the results of their child's performance on an alternate assessment. This information is meant to include the child's individual test results as well as supportive information to understand the results (Elledge, LeFloch, Taylor, & Anderson, 2009, p. 12). The outcomes published on a score report demonstrated how a state meets state and federal accountability requirements. In addition, the outcomes were intended to provide information about current instruction and its effectiveness on a student's ability to learn. This information about the student's learning is shared with students, families, schools, and states (Almond & Case, 2004, p. 7).

Burdette and Olsen (2000) highlighted the score reporting for eight states (Delaware, Florida, Georgia, Indiana, Minnesota, North Dakota, Utah, and West Virginia). Delaware was still emerging on their score reporting (p. 9). Score reports, modeled after the general assessment

COMPARING TWO ALTERNATE ASSESSMENTS

format, were given back to the schools, who forwarded the results to the student's parents (p. 9). Results from Florida's alternate assessments were used locally for "instructional programming and school-level accountability" and submitted to the state every three years (pp. 13-14). Georgia teachers prepared a "student reporting form" that recorded the final progress at the end of the year based on five priority goals and objectives in the student's IEP (p. 17). Based upon the final results, students were categorized on a rubric as "initial, emerging, progressing, or functional" (p. 17). At the time of the report, Georgia was revising its process. Indiana, at the time of this report, was aligning the score report of their alternate assessment based on a rubric, with the general education assessment. Indiana's rating scale will include five categories: advanced proficient, proficient, partially proficient, below proficient, prerequisite skills, and nonexistent (p. 21). Prerequisite skills were further broken down into the emergent level, the supported independence level, the functional independence level, and the independent level (p. 21). The nonexistent category included "no opportunity," "not applicable," and "not observed."

Minnesota teachers scored their alternate assessment portion for Reading, Writing, and Math using a scale from one to seven: "(1-2) Awareness, (3-5) Understanding, (6-7) Application" (p. 24). For the functional section, another seven-point scale was used: "(1-2) No Participation/Full support, (3-5) Moderate participation/Moderate support, (6-7) Full participation/No support" (p. 24). Local directors shared their scores with the State Department of Education, who organized and categorized the data for local schools to analysis. As of this report, "other data reporting issues have not been fully developed" (p. 24). The North Dakota alternate assessment uses a "body of evidence" format that is scored at a central location by a team that forwards the score to the state. Individual scores are sent to local schools and filed in the student's school records (p. 28). Utah uses two goals from each student's IEP as part of their alternate assessment. Once the

COMPARING TWO ALTERNATE ASSESSMENTS

final performance on the goals is charted, the information is shared with the State Office of Education, only if the school has more than ten students taking the alternate assessment, in order to protect the identify and confidentiality of the students (p. 31). West Virginia's Alternate Assessment Skill Inventory Rubric scored their students using the ratings: Awareness, Progressing, Competent, and Generalized (p. 35). Student performance was also rated on levels of accuracy and fluency, number of environments, intensity of instructional assistance, and number of varied demonstrations (p. 35). The teacher maintains the documentation of performance and shares the results through a parent conference. Afterward, the alternate assessment is sent to the State Department of Education. These eight states shared uniquely different scoring reports to represent a student's performance on an alternate assessment. Even when the type of alternate assessment was similar, the scoring was not identical. A few states, also, were in the development or revision stage of their score reporting. This may be due to the fact that they were still working on strengthening their alternate assessments.

Based on a 2005 study by Thompson, Johnstone, Thurlow, and Altman, school and individual score reports focused on one single measure (i.e., "proficient") to compare a student's performance against the alternate achievement standards, providing little-or-no information on schools, teachers, and parents that would drive instruction for a student. The information given was to satisfy the accountability responsibility of each state, as dictated by NCLB (20 U.S.C. 6311 § 1111 (b)(3)(C)).

In August of 2009, Cameto, Knokey, Nagle, Sanford, and Blackorby summarized the national findings from a report that profiled alternate assessments from all of the fifty states by the National Study on Alternate Assessments (NSAA). The original report was based on the 2006-2007 school year. States were surveyed about their score reports from the results of their

COMPARING TWO ALTERNATE ASSESSMENTS

alternate assessment for their students with significant cognitive disabilities. Individual score reports were given by 98% of the states to the parents, and 90% of the states shared them with schools and teachers (p. B173). Individual scores were expressed by state achievement standards by 88% of the states, scores by 88% of the states, and percentiles by 25% of the states (p. B175). Information included within the score reports included performance/achievement levels (by 92% of the states), scores (by 92% of the states), standard/strand breakouts (by 53% of the states), indicator/benchmark breakouts (by 20% of the states), performance/achievement level descriptors (by 63% of the states), and sample test items (by 6% of the states) (pp. B179-B181). The transparency within score reports has dramatically changed in comparison to years past. Based upon the responses from the states, their score reports provided more information than just a simple performance indicator. The data generated is shared with teachers, schools, and, most importantly, parents. Having alternate assessment data available to all IEP team members enriches decision making about instruction and individual student goals and objectives.

In a study that focused on teacher perceptions of the outcomes produced from the alternate assessment administered in three states, Restorff, Sharpe, Abery, Rodieriz, and Kim (2012) found that 55% of teachers surveyed felt that the outcomes presented on their students' alternate assessment results "helped them align their instruction to the states' alternate academic content standards" (p. 189). Although more teachers supported the administration of the alternate assessment for their significantly impaired students, 42% felt that it did not reflect their ability, and 40% indicated that "the assessment provides a inaccurate profile of their students' abilities" (p. 190). Although the process of administering the alternate assessment appeared to give more data to teachers than in years prior, the focus of the outcomes still centered around a single measure outcome, with little feedback about student performance (Cameto et al., 2009).

COMPARING TWO ALTERNATE ASSESSMENTS

Score reports for alternate assessments have evolved since the mandate for states to create an alternate assessment in IDEA (1997). The intent of score reports is to drive instruction and provide information about student growth, accountability, and documentation of participation. These reports, traditionally, have not had a significant effect on a student's daily academic life. As the development of alternate assessments for students with significant cognitive disabilities progressed, so have their score reports. Even though score reports are an important piece of alternate assessment systems, they can only be produced once an alternate assessment is established. These reports now include more information or data that is shared with the significant stakeholders in a student's education, leaving a more profound effect than in the past. Score reports are the one area that appears to have the least amount of variation among the states.

The populations of students who are eligible to take an alternate assessment has a profound effect upon the type, standards, scoring criteria, validity, and score reporting for alternate assessments. Due to the individuality of the students and their needs, there is a discrepancy in how standard protocols are applied to alternate assessments. As a result, comparing alternate assessments and their components has been complicated for researchers. Although there were a limited amount of types of alternate assessments used, a portfolio or body of evidence, the most popular type, was difficult to standardized in comparison to a multiple choice assessment administered to the general population. This includes scoring criteria, which included student performance and system performance. As alternate assessments grew in development, using alternate achievement standards based on or linked to general education academic standards increased the opportunity to compare alternate assessments among states. Score reports based on using alternate achievement standards also reflected more consistency

COMPARING TWO ALTERNATE ASSESSMENTS

within these assessments and mirrored the opportunity for more comparisons. The research on alternate assessments has been limited due to the slow development of alternate assessments and the flexibility needed to assess this diverse population of students. Students with significant cognitive disabilities who are administered alternate assessments present challenges to the development of statewide assessment types, standards, score criteria, validity, and score reporting.

Summary

Reports through the NCEO developed the story of alternate assessments for students with significant cognitive disabilities over the decades since the NCEO's first conception. Despite projected deadlines, states' alternate assessments have been in a dynamic state of development and revision. Only in recent years have collaborative groups grown to reflect fresh input and resources to produce and improve alternate assessments. Research focusing on alternate assessments has highlighted the variation in types, alternate achievement standards, scoring criteria, validity, and score reporting. There has been movement and growth over the years to reduce the amount of variation to produce a better alternate assessment, but at a very slow pace. Policy has defined broad expectations for how states develop alternate assessments, but the overall goal is the same: the valid and reliable measurement of growth for students with significant cognitive disabilities. Selecting an alternate assessment to implement in a state is complicated due to the variety and slow growth of progress.

CHAPTER 3
METHODOLOGY

Study Design

This two-part study was grounded in comparing two different alternate assessments: the DLM and the MSAA. The first part collected and organized the data. To answer the research questions that surrounded the comparison, data needed to be collected about the administration, criterion of measurement, and informational outcomes of the alternate assessments. These topics were based upon the information reflected in the NCEO reports and past research and literature. This information was gathered from the DLM and MSAA websites, websites of state departments of education that administered these alternate assessments, and direct contact with the organization that developed these alternate assessments. To organize the data, three coding agendas were developed with categories. The three coding agendas—types of alternate assessment, criterion of measurement, and informational outcomes—were necessary to disaggregate the DLM and the MSAA in order to make side-by-side comparisons. Within each coding agenda, categories were developed to further focus the data on the components needed to address the research questions (see Table 3-1). Mayring (2000) described this process as “organizing information by sorting content into categories in a systematic way, such as coding agendas, for the purpose of fitting the material into a mode of communication” (p. 1).

As stated in the literature review, since the development of alternate assessments, states have selected among portfolios, performance, and checklist or inventory formats. Throughout the research, the most popular format for alternate assessments has been portfolios. The coding agenda for types of alternate assessments included these three formats, as well as online formats. Although there was little or no research about the alternate assessment administered through an

COMPARING TWO ALTERNATE ASSESSMENTS

online format, this format is available for some alternate assessments.

The second category in organizing the data was criterion of measurement. Criterion of measurement defines the standard by which an outcome is measured on an alternate assessment. This category encompasses the prior topics of alternate achievement standards, scoring criteria, and validity. The coding agenda included how the performances of the skills that are being measured are linked to general education academic standards. Federal policy dictated that states must link these skills to general education academic standards. Scoring criteria for ELA for the DLM and MSAA was included in this coding agenda. This is important when comparing the two alternate assessments, especially for this population of students who are so diverse in their academic needs. With regard to the supports needed for many of the students with significant cognitive disabilities, the level of supports available within an alternate assessment was included in this coding agenda.

An informational outcome on an alternate assessment is produced from a student's performance on the alternate assessment. This category encompasses the prior topics of score criteria and reporting. Within this coding agenda, the number of performance levels was a category. Performance levels define how a student performs in comparison to the standards. These levels are reported to show a student's overall performance on alternate assessments. Another category reported within informational outcomes is the notation of a student's performance on an alternate assessment with various skills within a content area (i.e., reading comprehension in ELA). This type of information provides the school district, teacher, and IEP team with pertinent information that drives the goals and objectives of a student's IEP. Also included within this coding agenda is the level of supported used within the alternate assessments. Supports are used to level the playing field for students with significant cognitive

COMPARING TWO ALTERNATE ASSESSMENTS

disabilities. These supports can be noted in the informational outcome from the alternate assessment, which illustrates what supports a student needs to perform during the assessment.

Once the data were organized and coded, the researcher analyzed them through the guidance of the research questions to compare the DLM and the MSAA for part two. This content analysis paralleled the information from both alternate assessments in a spreadsheet using the coded information and the categories in this framework. Reliability was difficult with one reader when coding and categorizing information, but to increase reliability, the information given for each alternate assessment was confirmed with an employee from the DLM and the MSAA. Multiple readers increased the reliability of the coding and categories, which is recommended by Bengtsson (2016). The content analysis procedure analyzed the two alternate assessments comparatively. The results of the study may be found in Chapter 4.

Table 3-1 Organization of Data

<i>Coding Agendas</i>	<i>Definition</i>	<i>Categories for Coding</i>
Type of Assessment	The platform that is used to administer the alternate assessment to the student.	-Online -Portfolio -Paper/Pencil -Checklist/Inventories
Criterion of Measurement	The standard of how a measure is related to an outcome.	-Skills linked to standards -Scoring criteria for ELA -Level of supports available
Informational Outcomes	The information generated from a student's performance on the alternate assessment.	-Number of performance levels -Notation of student performance by assessment categories -Level of support used

COMPARING TWO ALTERNATE ASSESSMENTS

CHAPTER 4

RESULTS

The results were gathered to compare the alternate assessments provided by Dynamic Learning Map (DLM) and Multi State Alternate Assessment (MSAA) for the purpose of determining which assessment provided detailed information regarding academic achievement in English language arts for students with significant cognitive disabilities in grades three through eight. The research was guided by the following questions:

1. Do DLM and MSAA use the same criterion for measurements? If there are similarities between the alternate assessments, why are they necessary for informational outcomes?
2. If there are differences between the criterion for measurements, how do they affect informational outcomes?
3. How do the DLM and MSAA administer their assessment? How does the difference between their administrations affect informational outcomes?

Research was conducted between March and September of 2017. The information was gathered from the DLM and MSAA websites, the websites of state departments of education that administered these alternate assessments (Alaska, Arizona, Arkansas, Colorado, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Maine, Maryland, Michigan, Mississippi, Missouri, Montana, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Oklahoma, Rhode Island, South Carolina, South Dakota, Tennessee, Utah, Vermont, West Virginia, and Wisconsin), and direct contact with the organization that developed these alternate assessments. Before analyzing the data, the data were organized into categories and agendas, as described within in Chapter 3. To organize the data, three coding agendas were developed, with

COMPARING TWO ALTERNATE ASSESSMENTS

categories (see Table 4-1 in Chapter 3). The three coding agendas, types of alternate assessments, the criterion of measurement, and informational outcomes were used to disaggregate the DLM and the MSAA in order to make side-by-side comparisons. Within each coding agenda, categories were developed to further focus the data on the components needed to address the research questions. Under the agenda of Type of Assessment, the categories were consolidated since the DLM and MSSA were intended to be administered online. The Criterion of Measurement agenda for the DLM and MSAA was coded for the categories of “Skills Linked to Standards,” “Scoring Criteria for ELA,” and “Level of Supports Available.” Under the Informational Outcomes agenda, the categories for coding were “Number of Performance Levels,” “Notation of Student Performance by Assessment Categories,” and “Level of Supports Used.” For each category and agenda, the data were separated for the DLM and MSAA and included a summary of the results. The research questions were addressed after the organized data.

Type of Assessments

DLM

The DLM format is an online computer-based assessment (Wells-Moreaux, Bechard, & Karvonen, 2017, p. 7). The DLM alternate assessment system uses a learning map model based upon Universal Design principles for their alternate assessment system. DLM uses Essential Elements, which are “grade-level specific expectations about what the most significant cognitive disabilities should know and be able to do.” The Essential Elements are related to the general education standards that focus on being college and career ready. Essential Elements are linked to each state’s content standards (Wells-Moreaux, Bechard, & Karvonen, 2017, p. 8). Individual concepts and skills on the map are represented by “nodes.” There are slightly over 2,000 English

COMPARING TWO ALTERNATE ASSESSMENTS

Language Arts nodes within the learning map model, with approximately 5,000 connections among them (Wells-Moreaux, Bechard, & Karvonen, 2017, p. 1). A small collection of nodes are called linkage levels. In the DLM, assessment items are grouped together in a “testlet,” which includes an engagement activity and about three to nine items aligned to one or more essential elements (p. 8). Testlets are accessible at different levels of complexity at the linkage levels (p. 9). Although the format is based online, an alternate format may be given in Braille, but it limits a student's access to some sections of the assessment.

MSAA

The MSAA may be administered online, in a paper-pencil format, or a hybrid of the two (Multi-State Alternate Assessment, 2017, p. 19-20). The ELA section of the MSAA includes 30-40 items that are “mostly selected response” and a scaffolded writing prompt (p. 3).

Both alternate assessments were developed with the intent of administering the alternate assessment online. The DLM and the MSAA are supportive of having a test administrator input the answers online for students who are unable to input their answers online independently. The MSAA provides a paper-pencil format for students who need one. Most sections of the MSAA can be printed and given to the student. A test administrator then inputs answers into the system. The DLM offers a Braille format, but it is limited because not all of the sections are available in that format.

Criterion of Measurement

Skills Linked to Standards

DLM

The skills and knowledge assessed through the DLM are linked to standards. Within the DLM system, these standards are referred to as Essential Elements, which are “grade-level

COMPARING TWO ALTERNATE ASSESSMENTS

specific expectations about what the most significant cognitive disabilities should know and be able to do.” They are related to general education standards that focus on being college and career ready. Essential Elements are linked to each state’s content standards (Wells-Moreaux, Bechard, & Karvonen, 2017, p. 8). To determine a student’s appropriate skill level, the DLM uses the KITE system, which is a “special user interface,” to gather information about a student through survey questions about the student from a test administrator who is familiar with the student. The system survey gathers information on topics such as academic performance, expressive and receptive communication skills, and content specific skills (Wells-Moreaux, Bechard, & Karvonen 2017, p. 8). The level of difficulty of items is dynamic, depending on the accuracy of a student’s response and his/her position on the learning map (ETS, 2016, p. 22)

MSAA

Students are assessed using MSAA Core Content Connectors, which were developed from Common Core State Standards and Learning Progressions Frameworks (Nebelsick-Gullett, Towles-Reeves, Perkins, & Deters, 2015, p. 4). Learning Progression Frameworks are defined as “research-based pathways for learning. Learning Progression Frameworks are developed and refined using available research and evidence and have clear binding threads that articulate the essential core concepts and processes of a discipline” (National Center and State Collaborative, 2015). The MSAA includes three testing sessions for ELA. The first session administers numerous test items at various levels of complexity. Based upon the student’s responses, the student is then directed to the second session, which initiates one of the three versions of the test. All students are then given a writing prompt in session three (Multi-State Alternate Assessment, 2017 p. 11).

COMPARING TWO ALTERNATE ASSESSMENTS

Scoring Criteria for ELA

DLM

The DLM calculates “a student’s probability of mastering every node in the learning map; higher probabilities (e.g., 80 or higher) indicate a greater likelihood of mastery. In the DLM system, a threshold is applied to identify a probability that is high enough to be considered “mastery.” That information is then combined across nodes within a linkage level to determine whether a student has mastered the linkage level” (Kingston, Karvonen, Bechard, & Erickson, 2016, p. 19). Therefore, the DLM does not use raw scores, percentages, or scale scores.

MSAA

Most items are scored as correct or incorrect through the online test platform based upon programmed answers within the system. Constructed response items are scored by the test administrator and input as correct or incorrect within the test platform. Items that do not have a response are scored as a zero (MEA, 2016, p. 3). The writing prompt is inputted into the system for human scoring (MSAA, 2015, p. 42). Assessments are given a scale score to reflect the student’s performance.

Level of Supports Available

DLM

The supports or accessibility available for a student taking the DLM is determined by the KITE system. The system gathers information on topics such as communication, assistive technology devices, and motor and sensory impairments (Wells-Moreaux, Bechard, & Karvonen, 2017, p. 8). The information about the student is used to determine what additional tools or materials a student needs access to or should be provided by the test administrator (DLM, 2017, p. 20). The information given also determines an appropriate entry point that utilizes

COMPARING TWO ALTERNATE ASSESSMENTS

accessibility and challenge (p. 8). Through the survey information and the students' ongoing performance, the DLM determines "linkage levels of complexity are most accessible and appropriate for the student" (p. 8). Timing and setting are individualized for each student and are not defined by DLM. Other supports are divided into three categories (adapted from Table 4-1, p. 20):

Table 4-1 DLM Supports

Category 1	Category 2	Category 3
<i>Supports Provided in Kite System Through Survey Questions</i>	<i>Supports Requiring Additional Tools/Materials*</i>	<i>Supports Provided Outside the System</i>
<ul style="list-style-type: none"> -magnification -overlay color -invert color choice -contrast color -spoken audio 	<ul style="list-style-type: none"> -uncontracted Braille EBAE -uncontracted Braille UEB -single-switch system (access profile enabled) -two-switch system -individualized manipulatives -calculator -alternate form-visual impairment 	<ul style="list-style-type: none"> -human read aloud -sign interpretation of text -language translation of text -test administrator entering of responses for student -partner-assisted scanning (PAS)

* These supports require preplanning and sufficient time to set up.

MSAA

Timing is completely in control of the test administrator (MEA, 2016, p. 3). Table 4-2 presents accessibility and support features for a student taking the MSAA (MSAA, 2017, pp. 21-23).

Table 4-2 MSAA Accessibility Features

<i>Accessibility Features</i>	<i>Online Version</i>	<i>Paper Version</i>
Answer Masking	Electronically embedded.	A piece of paper or card may be used.

COMPARING TWO ALTERNATE ASSESSMENTS

Line Reader	Electronically embedded.	Two pieces of paper may be used to focus a student's attention on an item or section of the test.
Alternate Color Themes	Electronically embedded.	Color overlays may be used.
Audio Player Tool	Electronically embedded.	Not available.
Read aloud by test administrator	Directions, answer options, or passages from test may be read to the student when appropriate.	Directions, answer options, or passages from test may be read to the student when appropriate.
Alternate Text Read Aloud by test administrator	Alternate text may be read by the test administrator.	Alternate text may be read by the test administrator.
Increase Volume	Headphones or volume control on device.	Test administrator may adjust the volume of his/her voice.
Magnification	Electronically embedded.	Handheld magnification devices, which are regularly used by a student, may be used.
Increase/Decrease Size of Text and Graphics	Built-in tools on devices maybe used to zoom in or zoom out text and graphics. To increase the size of text and graphics, projection devices, video magnifiers, and Smart Boards may be used.	The text in a paper version may increased or decreased using projection devices or interactive white boards as needed.
Tactile Graphics	May be used if the student is already familiar with them prior to testing. It is the test administrator's responsibility of creating and administering them to the student when appropriate.	May be used if the student is already familiar with them prior to testing. It is the test administrator's responsibility of creating and administering them to the student when appropriate.
Tactile Symbols		May be used if the student is already familiar with them prior to testing. It is the test administrator's responsibility of creating and administering them to the student when appropriate.
Object Replacement		May be used if the student is already familiar with them prior to testing. It is the test

COMPARING TWO ALTERNATE ASSESSMENTS

		administrator's responsibility of creating and administering them to the student when appropriate.
--	--	--

The following assistive technology devices are supported by the MSAA system:

text-to-speech, alternate keyboards, switch-based navigation and answer selection, and eye-gaze.

Summary

Both alternate assessments systems' criterion for measuring student performance are linked to states' common standards. The MSAA also uses a second ingredient, Learning Progressions Frameworks, for developing their Core Content Connectors. This provides another perspective on the skills and knowledge that students should know. Scoring the student's performance is remarkably different between the two assessments. The DLM calculates probability based upon the indicated performances, whereas the MSAA uses scale scores. Both measurements use standard setting or cut scores to determine the performance level or label. Both systems also offer levels of support to students. The MSAA offers only two more options (line readers and answer masking) over the DLM. It should be noted that both systems are dependent on the device used to access some of the supports (i.e., volume, zoom). This includes other devices that are available within a school (i.e., SmartBoards, projection devices). On the contrary, having the flexibility to access different devices provides flexibility and adaptability to meet a student's needs to help ensure that that the student is demonstrating his/her abilities, not disabilities.

COMPARING TWO ALTERNATE ASSESSMENTS

Informational Outcomes

Number of Performance Levels

DLM

The DLM summative score report includes the overall performance categories of “emerging,” “approaching the target,” “at target,” or “advanced” (Kingston et al., 2016, p. 20) with “at target,” which is equivalent to “proficient” as used in many assessments. To create cut scores for performance levels, the DLM system uses diagnostic classification modeling or standard setting, based on the results, to identify patterns in performance (Wells-Moreaux, Bechard, & Karvonen, 2017, p. 1).

MSAA

Performance levels include scale scores and performance level descriptors. Performance level descriptors are numbered from 1-4, with 4 being the highest score. These levels are determined by ranges of scale scores, and grade levels and are separated by content. Scale scores are used not only to determine performance level descriptors but also to “make comparisons between groups of students, schools, and districts” (MEA, 2016, p. 4).

Notation of Student Performance by Assessment Categories

DLM

Within the DLM summative score report is a “brief narrative statement about the student’s mastery in each conceptual area” (Kingston et al., 2016, p. 20). Within the score report for the DLM, there is also a performance profile that provides a summary of the results for each content area. The performance profile includes the overall results, performance categories, and conceptual areas. The overall results section explains each student’s overall performance in the Essential Elements. Examples of Essential Elements and how that compares to grade level peers

COMPARING TWO ALTERNATE ASSESSMENTS

are also included within this section. The performance categories section explains the performance categories, focuses on the “student's highest level of mastery,” and might include a comparative guide of the performance indicators to the state’s performance indicators. Lastly, the conceptual areas section “summarizes the student’s performance in groups of related Essential Elements within the subject” focuses on mastery performance and lists the demonstrated skills. The results of this section are presented through a bar graph (Dynamic Learning Maps Consortium, 2016, p. 2)

MSAA

Individual Student Reports include a student’s scale score and performance level for each content area. The student's score is displayed among the spectrum of the performance-level descriptors using a bar graph. Text is included that describes “the performance level descriptor for the student’s performance level” and the skills and knowledge that a student is generally able to do at that student’s performance level (MEA, 2016, p. 11). Also included in the score report is a range of scale scores that the student would most likely score within if the assessment were taken again.

Level of Support Used

There is no notation on a student’s individual performance report about the level of support used by the student when administered the DLM or the MSAA.

Summary

Both assessments provide a symbol, word(s), or number as performance levels to determine the value of a student’s academic performance. Bar graphs and descriptors are included to provide a visual representation of the student’s performance. The DLM provides a breakdown of skills for each conceptual area in list format and the student’s mastery of each

COMPARING TWO ALTERNATE ASSESSMENTS

area. The MSAA provides a narrative of the skills that the student most likely is able to know and do in each conceptual area based on the student's performance score. Neither report document the supports used by the student during the administration of the assessment.

Conclusion

Both the DLM and the MSAA use criterion that are linked to state common core academic standards. It should be noted that the MSAA adds another perspective when creating their standards, but using that lens does not digress from the criteria used by states. Therefore, this similarity in the standards ensures that states are in compliance with standards dictated by federal and state legislatures. These standards are necessary to produce the required informational outcomes expected by federal, state, and local education departments. The contrast in the assessments is the format with which students are measured by that criterion. Both assessments administer surveys to gather preliminary data to distinguish the academic level and accessibilities that the students needs. The DLM uses information about the student to pinpoint a starting point for the alternate assessment. As the student progresses through the assessment, the system constantly reexamines and adjusts, when needed, the path on the learning map to ensure an appropriate level of complexity. The student is then given an individualized or customized assessment based upon his/her academic needs. The MSAA uses the preliminary surveys to administer one of three different versions of the assessment based upon the information given about the student. All students are given the same writing prompt. The informational outcomes will be different because the DLM has more detailed information to offer about a student's academic performance measured against the standards. The MSAA only has the potential to show a student's academic performance against one of three versions of an assessment based upon the standards. The lack of diversity in the differentiation of the

COMPARING TWO ALTERNATE ASSESSMENTS

assessment by the MSAA is not representative of the individualized needs of this diverse group of students.

The DLM and the MSAA are primarily online-based assessments. The MSAA provides more opportunity for other options in paper-pencil format. In reflecting on their formats, the DLM shares their technical format in great detail, whereas the MSAA does not have any technical information that sets it apart from an assessment that uses an online format to display questions and input answers. The online administration of the assessment provides a format that might be adapted for a student since the assessment may be accessible by multiple devices that are familiar to the student. The DLM format has the potential to provide a map or picture of where the student is academically, demonstrating more informational outcomes than the MSAA, which provides an assessment format to be administered online.

Based upon the organized data, it is difficult to determine whether one alternate assessment is superior to the other. Both alternate assessments are intended to be administered online to assess students with significant cognitive disabilities. The DLM and MSAA provide information about a student's performance measured by alternate achievement standards that are linked to the states' academic standards, which is required through state and federal policy. Informational outcomes include performance levels and summaries describing the skills and knowledge that a student with significant cognitive disabilities might attain based upon the performance on the alternate assessment. There is a significant difference between how those scores are determined, however. The DLM uses probability calculations and the MSAA uses scaled scores. Based upon these numbers, a student's performance level is determined. Despite the differentiation between calculating performance levels, performance levels do not have a direct impact on the daily academic performance of disabled students in comparison to their

COMPARING TWO ALTERNATE ASSESSMENTS

nondisabled peers. Performance levels do not assist us in aligning instruction to academic standards, improving or developing IEP goals, or providing opportunities for least restrictive environments for these students (Restorff et al., 2012, p. 190). Both the DLM and the MSAA provide an alternate assessment for students with significant cognitive disabilities that produce approximately similar informational outcomes based upon the performance on these alternate assessments.

CHAPTER 5

SUMMARY AND CONCLUSIONS

Research Questions

Comparing the alternate assessments provided by Dynamic Learning Map (DLM) and Multi State Alternate Assessment (MSAA), which assessment provides detailed information regarding academic achievement in English language arts of students with significant cognitive disabilities in grades three through eight. The research was guided by the questions:

1. Do DLM and MSAA use the same criterion for measurements? If there are similarities between the alternate assessments, why are they necessary for informational outcomes?
2. If there are differences between the criterion for measurements, how do they affect informational outcomes?
3. How do the DLM and MSAA administer their assessment? How does the difference between their administrations affect informational outcomes?

Significance of the Study

This study is significant because it appears to be the first to compare the DLM and the MSAA. The DLM and the MSAA were piloted within the last two to three years, but little research or reference in the literature is found about either alternate assessment. These two alternate assessments were created through a consortium made up of many organizations that applied federal funds through the development process. At the time of this study, approximately 49% of the states used the DLM or MSAA to provide an alternate assessment to students with significant cognitive disabilities. There is a potential for their continued growth through the

COMPARING TWO ALTERNATE ASSESSMENTS

recruitment of additional states to the consortiums. At the time of this study, there was no research or literature in print that compared these two alternate assessments.

Implications of the Study

Implications for Practitioners

Through a comparison of the DLM and MSAA, the research presented information side-by-side on both alternate assessments with the same lens. With the potential of more states to join each consortium, state departments of education that are interested in joining one of these consortiums will be seeking more information to assist in their decision-making process. For states that are currently providing their own alternate assessments to students with significant cognitive disabilities, this study might initiate reflection on revising their current alternate assessment, abandon it and joining a consortium, or recruiting a new consortium to develop a future alternate assessment. “The variation in alternate assessment practices across states, their ongoing development, and the limited research available to date has important implications for practitioners to become consumers and advocates in their states” (Browder et al., 2003, p. 57). These alternate assessments have the opportunity to expand the use of the informational outcomes and to directly affect the academic achievement of students with significant cognitive disabilities on a daily basis. In this study, the DLM appeared unable to produce a richer informational outcome when given the potential to extract more information about a student’s performance than the MSAA. This researcher speculates that this might be due to lack of sophistication and/or efficiency of the online system to extract the information used to readjust the level of difficulty during the process of the student’s performance. Another reason might be the broad spectrum of abilities of students with significant cognitive disabilities that the system is trying to accommodate when administering the alternate assessment. With such a vast range of

COMPARING TWO ALTERNATE ASSESSMENTS

complexity to accommodate all students with significant cognitive disabilities, the details of information might be just as general as the information produced when given the MSA.

Recommendations for Policy

Recommendations from this study focus on federal and state policy. Federal policy should provide encouragement for consortiums to increase in size and resources in order to further the development of alternate assessments for students with significant cognitive disabilities. Increasing and strengthening consortiums may be accomplished through the continuation of federal funds. Creating federal policy that supports more continuity in alternate assessments also encourages more validity and reliability across states when more states are working together in a consortium toward the same type of assessment. Supporting consortiums through federal policy is important to ensuring that current alternate assessments continue to be revised and updated as needed. This is especially important as technology continues to expand exponentially and the number of students who significantly depend on technology to execute daily life tasks increases. Holding states accountable through federal policy for the assessment of students with significant cognitive disabilities is necessary to ensuring that members of this population and those who work with them are kept to the same high standards that we hold for general education students.

“Alternate assessment has the potential to enhance expectations for students with significant cognitive disabilities and to increase consideration of this population’s needs in setting state and district policy” (Browder et al., 2003, p. 57). The state policy recommendations are suggested through the lens of an Individualized Education Program (IEP). In an IEP, the document is written by the IEP Team (teachers, parents, and case managers). Similar to an IEP Team, it is recommended that state policy include teachers and parents as stakeholders in the

COMPARING TWO ALTERNATE ASSESSMENTS

process of developing alternate assessments that produce valuable informational outcomes. Their feedback and input is important, especially that from teachers who administer the assessments to their students and work directly with the DLM and MSAA systems. All IEPs contain past and present levels of student performance. Data from a student's past performance in the alternate assessments should be included in the informational outcomes to demonstrate growth or lack of growth as defined by a state policy. This information can provide direction or recommendations to a student's program, goals, accommodations, modifications, or assistive technology to the IEP team. State policy about the use and availability of supports or accommodations for a student that are recorded in a student's IEP should be documented within the informational outcomes to portray the abilities of the student to provide current or future recommendations for each student. A student's IEP includes goals and objectives that drive appropriate instruction for that student. State policy should support the ways that informational outcomes or data from an alternate assessment drive the instructional decisions that lead students toward achieving alternate achievement standards that are included in the goals and objectives section of an IEP. This includes curriculum driven by the general education academic standards that is linked to policy and enforces the link between alternate achievement standards and the academic standards for general education expectations in order to keep the academic expectations high for students with significant cognitive disabilities. A state policy that drives revision of and reflection on a student's performance and supports on an alternate assessment, similar to the annual review of an IEP, ensures that districts and IEP teams continually value and utilize the informational outcomes produced by an alternate assessment. State policy needs to consider how informational outcomes or data are incorporated into the educational decisions about a student. Valuable data from these alternate assessments portray the abilities of these

COMPARING TWO ALTERNATE ASSESSMENTS

students. Browder et al. (2003) noted that throughout the NCEO reports in the late 1990s, it was suggested that students needed to participate in assessments to ensure their voice in policy decisions. The students' participation would also improve input into programs, improve opinions about people with disabilities, provide educational opportunities for these students to access the general education academic standards, and improve academic instruction within programs attended by students with disabilities (p. 46). These recommendations for federal and state policies could have an impact on how students with significant cognitive disabilities are assessed and on the use of the informational outcomes or data from these assessments on their education.

Topics for Further Research

This study ignited further topics for future research. User feedback from test administrators or teachers might provide information that would enhance the comparison between the two systems. Originally, this might be difficult because users are usually familiar with only one system, but the opportunity to use or view both would give valuable information when comparing these alternate assessment systems. Another topic to expand upon would be the technical aspect of the alternate assessment systems. Both systems rely heavily upon technology. Feedback about their ease of use on different devices would be beneficial, especially within a district where there are multiple users at the same time on the same infrastructure. Beyond the scope of this study, but an area worthy of further investigation, is the population of students whose cognitive abilities are so low that the student is unable to access an alternate assessment even with the maximum support of accessibility features and/or accommodations. Lastly, within both systems there is professional development and training for teachers who instruct the students with significant cognitive disabilities who take these

COMPARING TWO ALTERNATE ASSESSMENTS

assessments. Both systems believe that the goal of academic instruction should be to provide access to grade-level content that is aligned with their states' content standards. In order to properly administer these assessments, there should be significant training in how to administer the assessment so that it accurately reflects a student's academic performance. The DLM and the MSAA are the first consortium alternate assessments for students with significant cognitive disabilities that are administered online, initiating many areas for future research.

Final Conclusions

The DLM and the MSAA are alternate assessments for students with significant cognitive disabilities. They are unique in that they are the only alternate assessments developed and implemented by consortiums at the time of this study. Approximately 49% of states use these assessments. The DLM and MSAA are in compliance with federal laws that mandate that all students participate in assessments that measure student achievement using grade-level content standards. Both assessments are primarily administered online through most devices. Students are given accessibility supports through the alternate assessment systems and various devices. The DLM has developed a system that instantly individualizes a student's level of complexity based on that student's accuracy on test items. The MSAA provides three variations of their assessment. Student performances are scored through cut scores using probability (DLM) or scaled scores (MSAA) and are assigned a performance level. Individual score reports provide information on the students' academic performance using text and visual bar graphs. Overall, although both systems have similar characteristics in various areas, the DLM appears to provide more differentiation for the diversity of the students with significant cognitive disabilities. This allows the DLM system to provide a strong criterion for measurements that includes the potential to provide more informational outcomes.

COMPARING TWO ALTERNATE ASSESSMENTS

REFERENCES

- Alaska Department of Education. (2013). Assessments. Retrieved from <https://education.alaska.gov/tls/Assessments/alternate.html>
- Alternate Assessment Consortium. (2017, October 3). Dynamic Learning Maps (DLM). Retrieved from https://www.education.nh.gov/instruction/assessment/alt_assess/index.htm
- Arizona Department of Education. (2017). MSAA. Retrieved from <http://www.azed.gov/assessment/msaa/>
- Arkansas Department of Education. (2014). Multi-State Alternate Assessment. Retrieved from <http://www.arkansased.gov/divisions/learning-services/assessment/assessments-for-students-with-disabilities/multi-state-alternate-assessment>
- Bengtsson, M. (2016). How to plan and perform a qualitative study using content analysis. *Nursing Plus Open*, 2, 8–14. Retrieved from [http://www.nursingplusopen.com/article/S2352-9008\(16\)00002-9/fulltext-s0025h](http://www.nursingplusopen.com/article/S2352-9008(16)00002-9/fulltext-s0025h)
- Bouck, E. C. (2013). High stakes? Considering students with mild intellectual disability in accountability systems. *Education and Training in Autism and Developmental Disabilities*, 48(3), 320–331.
- Browder, D. M., Spooner, F., Algozzine, R., Ahlgrim-Delzell, L., Flowers, C., & Karvonen, M. (2003). What we know and need to know about alternate assessment. *Exceptional Children*, 70(1), 45–61.
- Browder, D., Flowers, C., Ahlgrim-Delzell, L., Karvonen, M., Spooner, F., & Algozzine, R. (2004). The alignment of alternate assessment content with academic and functional curricula. *The Journal of Special Education*, 37(4), 211–223.

COMPARING TWO ALTERNATE ASSESSMENTS

doi:10.1177/00224669040370040101

Browder, D. M., Wakeman, S., & Flowers, C. P. (2006). Assessment of progress in the general curriculum for students with disabilities. *Theory Into Practice, 45*(3), 249–259.

doi:10.1207/s15430421tip4503_7

Burdette, P., & Olsen, K. (n.d.). *Alternate assessments: A medley of alternate assessments* (Rep.). (ERIC Document Reproduction Service No. ED452642)

Dynamic Learning Maps Consortium. (2016). *Dynamic learning maps: What is a learning map model?* Lawrence, KS: The University of Kansas Center for Educational Testing and Evaluation

Dynamic Learning Maps Consortium (2016). About DLM tests. Retrieved from <http://dynamiclearningmaps.org/about/tests#accessibility>

Dynamic Learning Maps Consortium (2016, October 05). *Dynamic learning maps talking to parents/guardians about the DLM student reports, 2015-2016*. Lawrence, KS: The University of Kansas Center for Educational Testing and Evaluation

Dynamic Learning Maps Consortium. (2017, February 09). *Dynamic learning maps information about scoring, data files, and score reports, 2016-2017 school year*. Lawrence, KS: The University of Kansas Center for Educational Testing and Evaluation

Dynamic Learning Maps Consortium. (2017, August 01). *Dynamic learning maps test administration manual, 2017-2018* Lawrence, KS: The University of Kansas Center for Educational Testing and Evaluation

Elliott, S. N., & Roach, A. T. (2007). Alternate assessments of students with significant disabilities: Alternative approaches, common technical challenges. *Applied Measurement in Education, 20*(3), 301–333. doi:10.1080/08957340701431385

COMPARING TWO ALTERNATE ASSESSMENTS

Elliott, S. N., Compton, E., & Roach, A. T. (2007). Building validity evidence for scores on a state-wide alternate assessment: A contrasting groups, multimethod approach.

Educational Measurement: Issues and Practice, 26(2), 30–43.

Every Student Succeeds Act of 2015, Pub. L. No. 114-95 § 114 Stat. 1177 (2015-2016)

Flowers, C., Ahlgrim-Dezell, L., Browder, D., & Spooner, F. (2005). Teachers' perceptions of alternate assessments. *Research and Practice for Persons with Severe Disabilities*,

30(2), 81–92. doi:10.2511/rpsd.30.2.81

Gong, B., & Marion, S. (Eds.). (2006). *Dealing with flexibility in assessments for students with significant cognitive disabilities* (Rep.). Minneapolis, MN: National Center on Educational Outcomes. (ERIC Document Reproduction Service)

Idaho State Department of Education (SDE). (2017). Special Education Assessments. Retrieved from <https://www.sde.idaho.gov/assessment/sped/>

Illinois Department of Education. (2017). Retrieved from <https://www.isbe.net/Pages/DLM-AA.aspx>

Indiana Department of Education. (2017, October 6). Alternate Assessments. Retrieved from <https://www.doe.in.gov/assessment/alternate-assessments>

Individuals with Disabilities Education Act Amendments of 1997, 20 U.S.C. § 1400 (1997).

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004).

Iowa State Department of Education. (2017). Iowa's alternate assessments for students with significant cognitive disabilities. Retrieved from <https://www.educateiowa.gov/pk-12/special-education/assessment-testing/iowa-alternate-assessment-1-iaa>

Kansas Department of Education. (2017, August 25). Dynamic Learning Maps (DLM) & Essential Elements. Retrieved from <http://www.ksde.org/Agency/Division-of-Learning->

COMPARING TWO ALTERNATE ASSESSMENTS

Services/Career-Standards-and-Assessment-Services/CSAS-

Home/Assessments/Dynamic-Learning-Maps-DLM-Essential-Elements

Kettler, R. J., Elliott, S. N., Beddow, P. A., Compton, E., Mcgrath, D., Kaase, K. J., Bruen, C.,

Ford, L., & Hinton, K. (2010). What do alternate assessments of alternate academic achievement standards measure? A multitrait-multimethod analysis. *Exceptional Children, 76*(4), 457–474.

Kingston, N. M., Karvonen, M., Bechard, S., & Erickson, K. (2016). The philosophical underpinnings and key features of the dynamic learning maps alternate assessment.

Teachers College Record, 118(140312), 1–30.

Kohl, F. L., Mclaughlin, M. J., & Nagle, K. (2006). Alternate achievement standards and

assessments: A descriptive investigation of 16 states. *Exceptional Children, 73*(1), 107–123. doi:10.1177/001440290607300106

Kopriva, R. J., Thurlow, M. L., Perie, M., Lazarus, S. S., & Clark, A. (2016). Test takers and the validity of score interpretations. *Educational Psychologist, 51*(1), 108–128.

doi:10.1080/00461520.2016.1158111

Maine Department of Education. (2015). MEA Alternate Mathematics and English Language Arts/Literacy. Retrieved from <http://www.maine.gov/doe/alternate/>

Maine Department of Education. (2016). *Maine Department of Education MEA Alternate Mathematics and ELA/Literacy Multi-State Alternate Assessment (MSAA) 2016 Guide for Score Report Interpretation* [MEA Alternate Mathematics and ELA/Literacy Multi-State Alternate Assessment (MSAA) 2016 Guide for Score Report Interpretation].

Maryland Assessments. (2017, September 22). MSAA 2017-18. Retrieved from

<https://marylandassessments.org/msaa/>

COMPARING TWO ALTERNATE ASSESSMENTS

Division of Special Education/Early Intervention Services, & Division of Accountability and Assessment. (n.d.). Maryland's Differences Among Assessments Chart for Students Receiving Special Education Services (Maryland State Department of Education). Retrieved September 09, 2017, from www.hcpss.org/f/special/assessmentdiffer.pdf

Michigan Department of Education. (2017). MI-Access, Michigan's Alternate Assessment Program. Retrieved from http://www.michigan.gov/mde/0,1607,7-140-5234_6027_6494-36793--,00.html

Mississippi Department of Education. (2012). Mississippi Academic Assessment Program-Alternate (MAAP-A). Retrieved from [http://www.mde.k12.ms.us/OSA/SP/mississippi-assessment-program-alternate-\(map-a\)](http://www.mde.k12.ms.us/OSA/SP/mississippi-assessment-program-alternate-(map-a))

Missouri Department of Elementary & Secondary Education. (2017). MAP-A. Retrieved from <https://dese.mo.gov/search-mo-gov/alternate%2Bassessment>

Multi-State Alternate Assessment. (2015). *Multi-State Alternate Assessment Multi-State Alternate Assessment 2015-2016 Technical Report* [Multi-State Alternate Assessment 2015-2016 Technical Report]. (2015). Measured Progress, 100 Education Way, Dover, NH.

Multi-State Alternate Assessment (2017). *Multi-State Alternate Assessment Test Administrator Manual March 27-May 12, 2017* [Test Administrator Manual March 27-May 12, 2017].

National Center for Education Statistics. (2015, January 01). The condition of education - Participation in education - Elementary/secondary - children and youth with disabilities - Indicator May (2016). Retrieved from https://nces.ed.gov/programs/coe/indicator_cgg.asp

Nebelsick-Gullett, L., Towles-Reeves, E., Perkins, A., & Deters, L. (2015). *Evaluating the*

COMPARING TWO ALTERNATE ASSESSMENTS

quality and impact of items, products, and procedures: NCSC writing alternate assessment based on alternate achievement standards. A paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, 2015.

New Jersey Department of Education. (2014). Other State Assessments. Retrieved from <http://www.nj.gov/education/assessment/apa/dlm/>

New Mexico Department of Education. (2013). The Administration of the New Mexico Alternate Assessment. Retrieved from <http://cdd.unm.edu/sde/index.htm>

New York State Alternate Assessment (NYSAA). (2017). New York State Alternate Assessment (NYSAA). Retrieved from <http://www.nysed.gov/edtech/schools/new-york-state-alternate-assessment-nysaa>

No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319 (2002).

Nolet, V. & McLaughlin M.J. (2005). *Accessing the general curriculum: Including students with disabilities in standards-based reform* (pp. 50 –75). doi:10.4135/9781483329253.n5

North Dakota Department of Education. (2015). ConnectED. Retrieved from https://www.nd.gov/dpi/uploads/81/ConnectED_Newsletter_april_may_june2015.pdf

Office of Elementary and Secondary Education. (2009). *Standards and Assessments Peer Review Guidance: Information and Examples for Meeting Requirements of the No Child Left Behind Act of 2001*(pp. 1-74) (United States, Department of Education, Office of Elementary and Secondary Education). Washington, D.C.: U.S. Department of Education. Revised with technical edits January 12, 2009

Office of Public Instruction. (2017). MSAA Alternate Assessment. Retrieved from [http://montana.gov/search.aspx?q=alternate assessment&via=homepage&cx=013380590290877010950%3A3ubczas3i44&cof=FOR](http://montana.gov/search.aspx?q=alternate%20assessment&via=homepage&cx=013380590290877010950%3A3ubczas3i44&cof=FOR)

COMPARING TWO ALTERNATE ASSESSMENTS

ID%3A11&ie=UTF-8

Oklahoma State Department of Education. (2015). Oklahoma overview: Alternate assessment.

Retrieved from <http://sde.ok.gov/sde/alternate-assessment-overview>

Public Schools of North Carolina. (2017). North Carolina Alternate Assessments. Retrieved from

<http://www.ncpublicschools.org/accountability/testing/alternateassess/>

Quenemoen, R., Rigney, S., & Thurlow, M. (n.d.). *Use of alternate assessment results in reporting and accountability systems: Conditions for use based on research and practice* (pp. 1-17, Rep. No. 43). (ERIC Document Reproduction Service No. ED467720)

Quenemoen, R., Thompson, S., & Thurlow, M. (2003). *Measuring academic achievement of students with significant cognitive disabilities: Building understanding of alternate assessment scoring criteria* (Rep. No. 50). Minneapolis, MN: National Center on Educational Outcomes. (ERIC Document Reproduction Service)

Restorff, D., Sharpe, M., Abery, B., Rodriguez, M., & Kim, N. K. (2012). Teacher perceptions of alternate assessments based on alternate achievement standards: Results from a three-state survey. *Research and Practice for Persons with Severe Disabilities*, 37(3), 185–198. doi:10.2511/027494812804153570

Rhode Island State Department of Education. (2017). Dynamic Learning Maps (DLM) alternate assessments. Retrieved from

<http://www.ride.ri.gov/InstructionAssessment/Assessment/DLMAssessments.aspx>

Roach, A. T., Elliott, S. N., & Webb, N. L. (2005). Alignment of an alternate assessment with state academic standards. *The Journal of Special Education*, 38(4), 218–231.

Roach, A. T., Elliott, S. N., & Berndt, S. (2007). Teacher perceptions and the consequential validity of an alternate assessment for students with significant cognitive disabilities.

COMPARING TWO ALTERNATE ASSESSMENTS

Journal of Disability Policy Studies, 18(3), 168–175.

doi:10.1177/10442073070180030501

Saven, J. L., Anderson, D., Nese, J. F., Farley, D., & Tindal, G. (2016). Patterns of statewide test participation for students with significant cognitive disabilities. *The Journal of Special Education*, 49(4), 209–220. doi:10.1177/0022466915582213

South Carolina Department of Education. (2017). South Carolina Alternate Assessment (SC-Alt). Retrieved from <https://ed.sc.gov/tests/assessment-information/testing-swd/sc-alt/>

South Dakota Department of Education. (2017). Alternate Assessment. Retrieved from <http://doe.sd.gov/assessment/alternate.aspx>

State of Vermont Agency of Education. (2017). Vermont Alternate Assessments. Retrieved from <http://education.vermont.gov/student-learning/assessments/alternate-assessments>

Taylor, M. A., & Pastor, D. A. (2013). An application of generalizability theory to evaluate the technical quality of an alternate assessment. *Applied Measurement in Education*, 26(4), 279–297. doi:10.1080/08957347.2013.824450

Tennessee Department of Education. (2017). Tennessee Department of Education alternate assessment. Retrieved from <https://www.tn.gov/education/topic/alternate-assessment>

Thompson, S., & Thurlow, M. (2001). *2001 state special education outcomes: A report on state activities at the beginning of the new decade* (Rep.). Minneapolis, MN: University of Minnesota.

Thompson, S., & Thurlow, M. (2003). *2003 state special education outcomes: Marching on* (Rep.). Minneapolis, MN: University of Minnesota.

Thompson, S., Johnstone, C. J., Thurlow, M., & Altman, J. R. (2005). *2005 state special education outcomes: Steps forward in a decade of change* (Rep.). Minneapolis, MN:

COMPARING TWO ALTERNATE ASSESSMENTS

University of Minnesota.

Thompsons, S., & Thurlow, M. (n.d.). *State special education outcomes, 1999: A report on state activities at the end of the century* (Rep.). (ERIC Document Reproduction Service No. 440507)

Thurlow, M., Olsen, K., Ysseldyke J., Erickson, R., & Aherarn, E. (1996). *Alternate Assessments for Students with Disabilities* (NCEO Policy Direction). University of Minnesota, Institute on Community Integration, National Center on Educational Outcomes (NCEO). Retrieved from the University of Minnesota Digital Conservancy, <http://hdl.handle.net/11299/174083>.

Tindal, G., Nese, J. F., Farley, D., Saven, J. L., & Elliott, S. N. (2016). Documenting reading achievement and growth for students taking alternate assessments. *Exceptional Children*, 82(3), 321–336.

Utah State Board of Education. (2017). Utah Special Education Services (SES). Retrieved from <https://www.schools.utah.gov/specialeducation/resources/assessment>

Wells-Moreaux, S., Bechard, S., & Karvonen, M. (2017, August 01). *Dynamic learning maps accessibility manual 2017-2018*. Lawrence, KS: The University of Kansas Center for Educational Testing and Evaluation

West Virginia Department of Education. (2017). Alternate Summative Assessment Resources. Retrieved from <http://wvde.state.wv.us/assessment/DLM/resources.html>

Wisconsin State Department of Education. (2017, August 17). Assessment in Wisconsin. Retrieved from <https://dpi.wi.gov/assessment>

Wyse, A. E., Dean, V. J., Viger, S. G., & Vansickle, T. R. (2013). Considerations for equating alternate assessments: Two case studies of alternate assessments based on alternate

COMPARING TWO ALTERNATE ASSESSMENTS

achievement standards. *Applied Measurement in Education*, 26(1), 50–72.

doi:10.1080/08957347.2013.739460