

Joint Legislative Committee on Performance Evaluation and Expenditure Review (PEER)

Report to
the Mississippi Legislature



A Review of the Accountability Standards of the Mississippi Department of Education

The Mississippi Department of Education's (MDE's) accountability standards were created to communicate how well Mississippi's schools and districts are performing, to identify schools and districts that need improvement, and to advise decisionmakers on necessary adjustments.

In determining accountability grades, MDE uses five different assessments. These assessments are administered at various grade levels throughout the elementary, middle school, and high school levels. After students take the assessments, MDE places each student's score into one of four achievement categories. MDE then applies a point system to determine the accountability grades, incorporating calculations for proficiency, growth (i.e., students' learning gains), and the graduation rate.

Regarding the effectiveness of the accountability standards in measuring performance, PEER determined that:

- the use of achievement categories obscures student score data;
- combining proficiency and growth to determine an accountability grade may not present the most accurate picture of actual student performance;
- the department's emphasis on growth fails to demonstrate actual performance; and,
- the assignment of weights to growth multipliers appears to be arbitrary.

Regarding the clarity and accuracy of the accountability standards' presentation of performance, PEER believes that:

- the use of "better of two years" and "pausing" adjustments yields accountability grades that do not accurately reflect current performance;
- accountability grades for six-component schools do not reflect those schools' own performance and growth; and,
- the method of creating assessment benchmarks and cut-points for the calculation of the accountability grades is not criterion-based.

August 11, 2015

PEER: The Mississippi Legislature's Oversight Agency

The Mississippi Legislature created the Joint Legislative Committee on Performance Evaluation and Expenditure Review (PEER Committee) by statute in 1973. A joint committee, the PEER Committee is composed of seven members of the House of Representatives appointed by the Speaker and seven members of the Senate appointed by the Lieutenant Governor. Appointments are made for four-year terms, with one Senator and one Representative appointed from each of the U. S. Congressional Districts and three at-large members appointed from each house. Committee officers are elected by the membership, with officers alternating annually between the two houses. All Committee actions by statute require a majority vote of four Representatives and four Senators voting in the affirmative.

Mississippi's constitution gives the Legislature broad power to conduct examinations and investigations. PEER is authorized by law to review any public entity, including contractors supported in whole or in part by public funds, and to address any issues that may require legislative action. PEER has statutory access to all state and local records and has subpoena power to compel testimony or the production of documents.

PEER provides a variety of services to the Legislature, including program evaluations, economy and efficiency reviews, financial audits, limited scope evaluations, fiscal notes, special investigations, briefings to individual legislators, testimony, and other governmental research and assistance. The Committee identifies inefficiency or ineffectiveness or a failure to accomplish legislative objectives, and makes recommendations for redefinition, redirection, redistribution and/or restructuring of Mississippi government. As directed by and subject to the prior approval of the PEER Committee, the Committee's professional staff executes audit and evaluation projects obtaining information and developing options for consideration by the Committee. The PEER Committee releases reports to the Legislature, Governor, Lieutenant Governor, and the agency examined.

The Committee assigns top priority to written requests from individual legislators and legislative committees. The Committee also considers PEER staff proposals and written requests from state officials and others.

PEER Committee
Post Office Box 1204
Jackson, MS 39215-1204

(Tel.) 601-359-1226
(Fax) 601-359-1420
(Website) <http://www.peer.state.ms.us>

The Mississippi Legislature

Joint Committee on Performance Evaluation and Expenditure Review

PEER Committee

SENATORS
THOMAS GOLLOTT
Vice Chair
SAMPSON JACKSON II
Secretary
KELVIN E. BUTLER
VIDET CARMICHAEL
NANCY ADAMS COLLINS
GARY JACKSON
PERRY LEE



REPRESENTATIVES
BECKY CURRIE
Chair
RICHARD BENNETT
KIMBERLY L. CAMPBELL
STEVE HORNE
MARGARET ELLIS ROGERS
RAY ROGERS
PERCY W. WATSON

TELEPHONE:
(601) 359-1226

FAX:
(601) 359-1420

Post Office Box 1204
Jackson, Mississippi 39215-1204

James A. Barber
Executive Director

www.peer.state.ms.us

OFFICES:
Woolfolk Building, Suite 301-A
501 North West Street
Jackson, Mississippi 39201

August 11, 2015

Honorable Phil Bryant, Governor
Honorable Tate Reeves, Lieutenant Governor
Honorable Philip Gunn, Speaker of the House
Members of the Mississippi State Legislature

On August 11, 2015, the PEER Committee authorized release of the report entitled **A Review of the Accountability Standards of the Mississippi Department of Education.**

A handwritten signature in black ink that reads "Becky Currie".

Representative Becky Currie, Chair

This report does not recommend increased funding or additional staff.

Table of Contents

Letter of Transmittal	i
Executive Summary	vii
Introduction	1
Authority	1
Problem Statement.....	1
Purpose and Scope.....	2
Method	2
Background	3
Purpose of MDE’s Accountability Standards	3
Importance of Effective Accountability Standards	5
Key Players in Developing and Administering Mississippi’s Education Accountability Standards.....	6
Why MDE Recently Revised the Accountability Standards.....	8
Goals of the Accountability Standards.....	9
Characteristics and Components of a School’s or District’s Accountability Grade	10
Assessments Used as the Basis for the Accountability Standards	10
Achievement Categories	12
Components of a School’s or District’s Grade	12
Cut-Points	14
How MDE Calculates a School’s or District’s Grade	16
The Proficiency Calculation	16
The Growth Calculation	18
The Graduation Rate Calculation.....	26
The Effectiveness of MDE’s Accountability Standards in Measuring Schools’ and Districts’ Performance.....	27
How Achievement Categories Obscure Student Score Data.....	27
Problems with Combining Proficiency and Growth into a Single School Grade	30
Why MDE’s Emphasis on Growth Fails to Demonstrate Actual School or District Performance.....	31
MDE’s Growth Multipliers.....	34

Table of Contents (continued)

Clarity and Accuracy of the Accountability Standards' Presentation of Schools' and Districts' Performance36

 "Better of Two Years" and "Pausing" of Schools' and Districts' Grades36

 How MDE Determines Accountability Grades for Six-Component Schools.....39

 The Creation of Assessment Benchmarks and Cut-Points.....41

 Changes in Graduation Requirements44

Recommendations46

Appendix A: Composition of Commission on School Accreditation, Accountability Task Force, Technical Review Committee, State Board of Education, and Technical Advisory Committee49

Appendix B: Glossary of Terms and Concepts Related to MDE's Accountability Standards52

Appendix C: Process to Determine Total Number of Points for Schools and Districts58

Appendix D: MDE's Graduation Options.....63

PEER's Response to the Department of Education's Response67

Agency Response69

List of Exhibits

1. Assessments Utilized by MDE as the Basis of the Accountability Standards and Grade Level at which Assessments are Administered, as of 2014-2015 Assessment Year	11
2. Scale Math Scores for MCT2 Test (May 2014 Administration)	12
3. Components of a School’s or District’s Accountability Grade, as of 2013-2014 Assessment Year	13
4. MDE Cut-Points for Schools and Districts, as of 2013-2014 Assessment Year	14
5. Method of Determining Reading Proficiency Score with MDE’s Accountability Standards, as of 2013-2014 Assessment Year	17
6. Example of a Proficiency Calculation, as of 2013-2014 Assessment Year.....	18
7. Possible Learning Gains and Respective Multipliers Utilized by MDE to Calculate Growth Components, as of 2013-2014 Assessment Year	18
8. Illustration of Types of Growth and Respective Weights Utilized with MDE’s Accountability Standards, as of 2013-2014 Assessment Year.....	21
9. Method of Determining Reading All-Growth Score with MDE’s Accountability Standards, as of 2013-2014 Assessment Year	22
10. Reading or Math All-Growth Calculation Example, as of 2013-2014 Assessment Year	24
11. Method of Determining Reading Low 25% Growth Score with MDE’s Accountability Standards, as of 2013-2014 Assessment Year	25
12. Graduation Rate Calculation for a Hypothetical School or District	26
13. May 2014 MCT2 Test Score Summary Showing Raw Score and Scale Score Ranges for Each Achievement Category	29
14. How Schools or Districts with Objectively Different Scale Scores Could Receive the Same Number of Proficiency Component Points.....	30
15. Illustration of Counting Growth upon Growth Utilizing MDE’s Accountability Standards, as of 2013-2014 Assessment Year	33
16. How a School or District Could Appear to Make Substantial Growth Gains	34
17. Hypothetical Example of “Better of Two Years” and “Pausing” Adjustments	38
18. Illustration of How MDE Determines Grades for Six-Component Schools Using its Accountability Standards, as of 2013-2014 Assessment Year	40

A Review of the Accountability Standards of the Mississippi Department of Education

Executive Summary

Introduction

In response to a legislative request, PEER conducted a review of the Mississippi Department of Education's accountability standards to address concerns of whether the standards adequately measure school performance.

Background

State accountability standards must be designed in such a way that they effectively demonstrate actual school performance. If standards do not reflect actual student performance, education stakeholders and decisionmakers cannot make the appropriate decisions or necessary adjustments to improve schools' and districts' performance.

MDE's accountability standards were created in order to communicate how well Mississippi's schools and districts are performing, to identify schools and districts that need improvement, and to advise decisionmakers on necessary adjustments. Although college and career readiness was not included in the original purpose of the standards, as Mississippi shifts toward what will likely be more rigorous standards, college and career readiness will begin to shape the overall purpose of the state accountability standards.

The Accountability Standards Task Force, the membership of which is approved by the Mississippi Board of Education, makes accountability standards recommendations to the Commission on School Accreditation. Once recommendations are approved by the commission, the Board of Education provides the final approval before new standards or changes in standards go into effect. Selected staff at the department provide information necessary for the task force to make accountability standards recommendations.

According to the Mississippi Department of Education, changes in Mississippi state law, federal requirements, and the desire to make the accountability standards equitable for all schools and districts and easier to understand led to the adoption of MDE's current accountability standards.

Characteristics and Components of a School's Accountability Grade

MDE uses five different assessments to determine schools' and districts' accountability grades. These assessments are administered at various grade levels within schools and districts. After students take the assessments, MDE uses each student's *scale score* to determine that student's placement within one of four achievement categories (advanced, proficient, basic, or minimal).^A

MDE then uses the percentage of students that a school or district has in the top two achievement categories (i. e., advanced and proficient) to determine that school's or district's accountability grade. MDE uses seven components (i. e., 700 possible points) to determine a grade for a school with no twelfth grade^B or a district with no high school and nine components (i. e., 900 possible points) to determine a grade for a school with a twelfth grade or a district with a high school. The components and their possible points are illustrated in Exhibit A, page ix.

MDE then uses cut-points established by the Accountability Standards Task Force to determine what total score must be achieved for a school to receive an A, B, C, D, or F accountability grade. MDE's current cut-points are shown in Exhibit B, page ix.

How MDE Calculates a School's Grade

The components of each school's or district's accountability grade contain three types of calculations: proficiency, growth, and graduation rate (see Exhibit A, page ix). Although MDE uses a graduation rate calculated in accordance with federal requirements (see page 26 of the report), the department has its own methods for calculating proficiency and growth, as follows:

- *proficiency*-MDE determines proficiency by calculating the percentage of students who performed at or above the "proficient" achievement category on assessments. In other words, this is the percentage of students whose assessment score placed them in the proficient or advanced achievement category.

^A *Scaling* refers to the process of converting a student's raw test score to a common score that allows for comparison between students.

^BAn example of a school with no twelfth grade would be an elementary school.

Exhibit A: Components of a School's or District's Accountability Grade, as of 2013-2014 Assessment Year

Components	Without 12 th Grade	With 12 th Grade
	700 Possible Points	900 Possible Points
Reading Proficiency	100	100
Reading Growth-All Students	100	100
Reading Growth-Low 25% of Students	100	100
Math Proficiency	100	100
Math Growth-All Students	100	100
Math Growth-Low 25% of Students	100	100
Science Proficiency	100	50
U.S. History Proficiency		50
Graduation Rate-All Students*		200

*MDE uses a federally approved four-year graduation rate calculation (MISS. CODE ANN. Section 37-17-6 [1972]). See page 26 of the report.

NOTE: MDE does not currently use “college and career readiness” and “acceleration” to calculate a school’s or district’s grade. However, according to MDE, these components will be included beginning with school year 2015-2016 results. See pages 52-53 of the report for more information on these components.

SOURCE: MDE.

Exhibit B: MDE Cut-Points for Schools and Districts, as of 2013-2014 Assessment Year

Letter Grade	Cut-Point Range	
	Without 12 th grade	With 12 th grade
A	518 or higher	695 or higher
B	455-517	623-694
C	400-454	540-622
D	325-399	422-539
F	324 or lower	421 or lower

SOURCE: MDE.

- *growth*--MDE defines growth as the percentage of students who made “learning gains.” The department considers two areas of growth when determining a school’s or district’s accountability grade:
 - growth of all students, which refers to the percentage of students who made learning gains from one year’s assessment to the next year’s assessment; and,
 - growth of the lowest twenty-five percent of students, which refers to the percentage of students who scored in the low 25% of their class the previous testing year

who made learning gains between the previous year's assessment and the current year's assessment.

MDE uses the growth components only for math and reading/language arts because math and reading/language arts are tested every year in grades three through eight and once in high school.

- *graduation rate*--MDE calculates this by determining the percentage of students who graduated in four years with a "regular high school diploma" (i. e., the standard high school diploma that is fully aligned with the state's academic content standards). MDE uses the number of students who graduated in four years from a school or district with a "regular high school diploma" as the numerator and the number of students who entered four years earlier as first-time ninth graders (with adjustments for deaths and transfers in and out) as the denominator. The method of calculating the graduation rate is prescribed by federal regulation.

Conclusions

Because of the way in which Mississippi's accountability standards are currently calculated, the standards do not provide stakeholders and the public with a clear picture of how Mississippi schools and districts are performing. Not only does the calculation of the current standards make it impossible to compare one school or district to another, but also to compare a school or district to itself over time. Mississippi's standardized tests are carefully constructed to ensure that a student has mastered a certain level of competency; those tests alone should provide the criterion/standard for measuring school performance.

The Effectiveness of MDE's Accountability Standards in Measuring School Performance

Achievement Categories Obscure Student Score Data

MDE's use of achievement categories obscures actual student test score data because all scores in an achievement category are basically considered to be equal, despite the wide range of scores within a category. Determining proficiency by calculating the percentage of students whose scores are in the top two achievement categories, described in MISS. CODE ANN. Section 37-17-6 (5) (c) (i) (1972), compounds the problem because the range of scores deemed "proficient" is even wider, indicating an insensitive measurement instrument.

Combining Proficiency and Growth into a Single School Grade

Due to the way MDE's accountability grade components are structured, combining proficiency and growth to determine a

school's or district's accountability grade may not present the most accurate picture of actual student performance. PEER believes that growth is a very important factor in school performance, but if the way growth is calculated affects a school's or district's grade in such a way that it no longer demonstrates true student performance at that school or district, MDE's overall purpose of the accountability standards is not being fulfilled. If the purpose of the accountability standards is to improve student achievement and increase the level of accountability of schools and districts, then more emphasis should be placed on proficiency--how a student actually performs on the assessments.

Emphasis on Growth Fails to Demonstrate Actual School or District Performance

MDE emphasizes growth in order to ensure that lower performing schools or districts that are improving positively contribute to their school's or district's accountability grades and, as required by state statute, to emphasize the progress of the lowest twenty-five percent of students in the school or district.

However, because of the way MDE has structured its accountability standards, in certain situations (such as a student whose score places them in the low 25% of scores), a student's growth from one achievement category to another could be counted up to three times in the determination of the school's or district's accountability grade for a given year. Additionally, a school or district could appear to have made substantial growth gains, which might actually be inaccurate.

If proficiency scores are accurate, comparing proficiency scores from one year to the next or reporting scale scores divided by total possible scale score points would be other ways of showing whether a school or district improved from one year to the next.

Growth Multipliers Appear to be Arbitrary

MDE's assignment of weights for learning gains appears to be arbitrary and results in the obfuscation of data, which impedes MDE from reaching its goal of improving student achievement and increasing school and district accountability.

While it might seem beneficial to provide incentives for schools and districts to encourage them to reach a higher level of achievement, if those incentives obfuscate data regarding actual student performance, the ultimate goal of improving student achievement and increasing school and district accountability has not been reached.

The Clarity and Accuracy of the Accountability Standards’ Presentation of Schools’ and Districts’ Performance

“Better of Two Years” and “Pausing” of Schools’ and Districts’ Grades

Although MDE developed its current accountability standards for use in the 2013-2014 assessment year, because of the implementation of college-and career-readiness standards that year, MDE has used “better of two years” or “pausing” adjustments to schools’ and districts’ accountability grades.

The “better of two years” adjustment meant that after having calculated the actual accountability grades for each school and district, MDE could decide, for each school and district, to apply the calculated grade based on the 2013-2014 assessment results or to retain the previous year’s grade. “Pausing” means that rather than calculating actual accountability grades for each school and district for that assessment year and choosing the “better of two years,” if approved by the U. S. Department of Education, MDE may automatically apply the previous year’s accountability grade.

These practices obscure the actual performance of students on assessments, therefore preventing MDE from making accurate comparisons among schools or districts to each other or to themselves over time. Further, accountability grades could reflect the accountability standards as they were calculated in a previous year rather than as they should be calculated in the current year.

How MDE Determines Accountability Grades for Six-Component Schools

Rather than determining cut-points for the accountability grades for schools without a twelfth grade and without a science assessment (i. e., six-component schools), MDE determines these accountability grades based on the actual distribution of grades for seven-component schools. MDE takes the A-F distribution of the actual grades of schools that have seven components and applies that distribution to the six-component schools. MDE then applies, or “links,” that distribution (i. e., the percentages for each A-F grade) to the six-component schools. This method forces the six-component schools into the seven-component distribution, reflecting the performance and growth of those schools rather than their own performance and growth.

The staff at MDE is aware of this problem and according to MDE, in May 2015 the Board of Education approved a rule that would address this problem.

The Method of Creating Assessment Benchmarks and Cut-Points is Not Criterion-Based

MDE's current process for determining accountability grades is not being driven by student performance; rather, a Mississippi teachers' group determines the benchmarks for student performance. MDE, the task force, and the Technical Review Committee, with the help of a consultant, determine the cut-points for establishing the accountability grades each year, maintaining significant control over the outcome of accountability grades.

Thus the processes used to determine achievement category benchmarks, A-F cut-points, and the number of possible points for each accountability component are subjective rather than criterion-based. Moreover, the placement of benchmarks and cut-points can affect the magnitude of trends, possibly giving some schools and districts an advantage in their accountability grades.

Changes in Graduation Requirements

In 2013, the federal government began requiring that the graduation component had to account for twenty percent of a school's or district's accountability grade. At that time, a student could not graduate high school in Mississippi unless he or she passed each subject area test (i. e., English II, Algebra I, U. S. History, and Biology I).

In January 2014, the State Board of Education voted to allow students to graduate if they failed one or more of their subject area tests but met certain other requirements. In March 2015, the board amended this action to allow additional options. The perception is that MDE has made graduation more easily attainable, thus allowing schools and districts to have better graduation rates.

Recommendations

1. In order for a school's or district's student proficiency to be represented accurately by its accountability grade, MDE should report performance grades that reflect student assessment score data as closely as possible. This could be done by:
 - eliminating the use of the four achievement categories (minimal, basic, proficient, and advanced); or,
 - reporting scale scores divided by total possible scale score points (in the form of a percentage).

To accomplish this, the Legislature should amend MISS. CODE ANN. 37-17-6 (5) (c) (i) (1972).

(Note: When *proficiency* is referenced in other recommendations in this report, it is with the assumption that an accurate proficiency measure will be utilized.)

2. In order to communicate and report student proficiency and student growth accurately and to prevent either proficiency or growth from greatly affecting a school's or district's accountability grade, MDE should separate proficiency and growth into two separate grades.

MDE could do so by assigning a letter grade (A thru F) for proficiency, followed by another indicator to represent growth. The department could use a letter grade to demonstrate proficiency and an arrow that indicates direction to reflect whether a school has made adequate learning gains. For example, a school that made learning gains and earned a B in proficiency would have a grade of B↑. However, a school that earned a B in proficiency, but did not make adequate learning gains, would have a grade of B↓.^c

For the separation of scores to take place, the Legislature would need to amend MISS. CODE ANN. Section 37-17-6 (4) (g) (1972) to allow for separate proficiency and growth indicators.

3. To ensure that a school's or district's growth is represented accurately in its accountability grade, MDE should indicate growth by a student's improvement from one year to the next in the accurate proficiency grade. MDE uses growth multipliers of 1, 1.2, or 1.25 to indicate greater growth, but any multiplier or incentive that alters an original score takes a rating farther away from accurately demonstrating true performance.
4. To ensure that a school's or district's grade for a given year is a direct representation of that school's or district's performance for that year, MDE should instruct schools and districts to report and publicize not only their official grade, but also their "paused" or "waived" grades in any school year that is considered a transitional year. Allowing schools and districts the opportunity to publicize the better grade of two years, or an outdated school grade, does not provide a clear picture of current performance.

Further, to ensure that schools' and districts' grades can be reliably compared to those of other schools or districts for that year and that a single school or district can analyze its performance over a period of time, MDE should report schools' and districts' grades using the

^cThe growth component is not a measure of performance; it seeks to communicate where a school or district stands relative to current performance. An A school or district that earns an A in proficiency would not have much (if any) room for growth and would not necessarily have an arrow indicator following the school's or district's grade.

same accountability standards (as opposed to a previous year's standards or a previous year's grades).

5. To ensure that the A through F cut-points and assessment benchmarks are directly related to student mastery over material, MDE should develop a defensible criterion for being "proficient."
6. To ensure that the accountability standards accomplish what they are designed to accomplish, MDE should ensure that task force recommendations support the purpose of the accountability standards so that appropriate changes, where necessary, can be made.
7. In the best interest of the students and to acknowledge the distinct honor of successfully completing high school, MDE should develop a method to ensure that the changes made to the graduation options are equivalent and comparable to a standard/regular high school diploma.
8. The Legislature should enact legislation requiring that the Mississippi Department of Education submit any proposed changes to the school accountability standards to the appropriations and education committees of the House and Senate and to the Executive Director of the Legislative Budget Office one year before those standards would become effective. Such submission should also include a statement of estimated economic impact detailing how the proposed changes could impact the development of recommendations for the funding of the adequate education program. This is important because school districts' accountability grades are figured into the MAEP formula^D and any changes in the way that a "successful" district (currently, a district receiving a C accountability grade) is defined will affect the calculation of the MAEP funding formula and thereby affect the amount of funding requested by MDE and ultimately the amount of funding received by school districts.

^DComponents of the MAEP funding formula process are defined in MISS. CODE ANN. Section 37-151-5 (1972). Currently, districts receiving a grade of C are classified as "successful" and if other statistical requirements are met, their expenditures form the base of the MAEP funding formula. Expenditures from districts receiving higher grades (A or B) or lower grades (D or F) impact the statistical calculations used in the MAEP formula, but expenditures from these districts do not otherwise impact the funding formula. The MAEP funding formula is calculated every four years, with adjustments for inflation during the intervening years. The most recent recalculation was for FY 2015. (A full recalculation of the MAEP funding formula will be completed for FY 2019.)

As noted above, MDE uses the MAEP formula to determine the amount of funding necessary to fund all schools at the funding level of the schools used in the formula that met the "successful" level of student performance. However, if the classification of student performance is flawed, as is illustrated in this report, the assumptions underlying the selection of schools to be used in the computation of funding are also flawed from a performance perspective.

For More Information or Clarification, Contact:

PEER Committee
P.O. Box 1204
Jackson, MS 39215-1204
(601) 359-1226
<http://www.peer.state.ms.us>

Representative Becky Currie, Chair
Brookhaven, MS

Senator Thomas Gollott, Vice Chair
Biloxi, MS

Senator Sampson Jackson, Secretary
DeKalb, MS

A Review of the Accountability Standards of the Mississippi Department of Education

Introduction

Authority

The PEER Committee reviewed the Mississippi Department of Education's accountability standards. The Committee acted in accordance with MISS. CODE ANN. Section 5-3-51 et seq. (1972).

Problem Statement

During the 2013-2014 school year, the Mississippi Department of Education (MDE) began using new accountability standards. However, concerns arose that the standards still did not adequately measure school performance.

The PEER Committee identified the following concerns and public perceptions related to the current accountability standards.

- MDE's calculation of students' scores on assessments, upon which the accountability standards are based, does not offer a clear picture of how schools and districts are actually performing, nor does it show whether schools and districts have achieved what they are supposed to have achieved.
- The U. S. Department of Education's voluntary waiver that allows schools to choose the higher of two years' grades (for the 2012-2013 and 2013-2014 school years) does not make it possible to compare a school's accountability grade from one year to another or to compare the grades of several schools or districts over time.
- Certain multipliers included in a school's or district's accountability grade calculation result in inflation of schools' and districts' grades.
- Some of the MDE's Accountability Task Force's decisionmaking practices do not align with the overall purpose of the accountability standards.

- MDE's ability to create assessment benchmarks and cut-points for the accountability grades gives the department an inappropriate amount of control over these grades.

In response to a legislative request, PEER conducted a review of MDE's accountability standards in order to address these concerns.

Purpose and Scope

PEER reviewed whether MDE's current accountability standards adequately measure schools' and districts' performance and fulfill the stated purpose of the standards.

PEER's review addressed the following questions:

- Why was there a need for the recent revision of MDE's accountability standards?
- What are the current accountability standards?
- Do the current standards adequately measure school/district performance and do they follow established criteria for such standards?
- Does the public have a clear and accurate picture of how schools and districts are actually performing?

Method

In conducting this review, PEER:

- reviewed relevant sections of state law;
- interviewed selected staff of the Mississippi Department of Education;
- interviewed a consultant with Research in Action, Inc. (an MDE contractor that helped develop the current accountability standards);
- reviewed documents provided by the Mississippi Department of Education;
- attended MDE Accountability Task Force meetings;
- attended State Board of Education meetings; and,
- reviewed information promulgated by Isaac I. Bejar's 2008 article, "Standard Setting: What Is It? Why Is It Important?" pertaining to best practices for accountability standards.

Background

This chapter addresses:

- the purpose of MDE’s accountability standards;
- the importance of effective accountability standards;
- the key players in developing and administering Mississippi’s education accountability standards;
- why MDE recently revised the accountability standards; and,
- goals of the current accountability standards.

Purpose of MDE’s Accountability Standards

MDE’s accountability standards were created to communicate how well Mississippi’s schools and districts are performing, to identify schools and districts that need improvement, and to advise decisionmakers on necessary adjustments. Although college and career readiness was not included in the original purpose of the standards, as Mississippi shifts toward what will likely be more rigorous standards, college and career readiness will begin to shape the overall purpose of the state accountability standards.

According to MDE, the accountability standards are “designed to improve student achievement and increase the level of accountability for both school districts and individual schools.” The accountability standards provide an annual estimate of instructional effectiveness for each school district. This estimate is delivered through statewide assessments that are used to ultimately determine a school’s or district’s grade.

MISS. CODE ANN. Section 37-16-1 (1972) states:

The primary purpose of the statewide testing program is to provide information needed for state-level decisions. The program shall be designed to:

- a. Assist in the identification of educational needs at the state, district and school levels.*
- b. Assess how well districts and schools are meeting state goals and minimum performance standards.*
- c. Provide information to aid in the development of policy issues and concerns.*
- d. Provide a basis for comparisons among districts . . .and between districts, the state and the nation, where appropriate.*

- e. *Produce data which can be used to aid in the identification of exceptional educational programs or processes.*

Thus the purpose of MDE's accountability standards is to use results from statewide assessments, administered at certain grade levels, to communicate to schools, districts, education stakeholders, the Legislature, and the public how schools are performing. The standards are to be used to identify where schools need to improve and advise decisionmakers on necessary adjustments. PEER believes that since the accountability system has to provide a clear picture on how schools are performing, the system's process must be protected from any potential arbitrary score adjustments that might dilute actual student performance.

MDE's assessments provide the criteria to demonstrate school performance. Since MDE, districts, and schools rely heavily on the results of the assessments, schools' and districts' grade reports must reflect as closely as possible students' original test scores if results are to be associated with actual school performance.

Incorporation of College- and Career-Ready Standards

“College and career readiness” describes the shift to design education in such a way that it will prepare students to be successful whether they choose to pursue a higher education degree or enter the workforce.

According to MDE, the Mississippi College- and Career-Ready Standards Initiative is a state-led effort that established a single set of clear educational standards for kindergarten through twelfth grade in English language arts and mathematics. States voluntarily adopt these standards and they are designed to ensure that students graduating from high school are prepared to enter credit-bearing entry courses in two-year or four-year college programs or enter the workforce. The standards ensure that parents, teachers, and students have a clear understanding of the expectations in reading, writing, speaking and listening, language, and mathematics and they put students on a level playing field regardless of their zip code.

According to MDE, Mississippi decided to adopt college- and career-ready standards because:

- they provide a consistent, clear understanding of what students are expected to learn so that teachers and parents know what they need to do to help them;
- they are consistent standards, adopted by forty-five other states, and will provide appropriate academic benchmarks for all students at each grade level, regardless of where they live;
- they incorporate the best and highest of previous state standards in the U. S. and are internationally

benchmarked to the top-performing nations around the world;

- students will learn the skills and abilities demanded by the workforce of today and the future;
- they emphasize critical thinking, teamwork, and problem-solving skills; and,
- they are grounded in college and career readiness.

According to MDE, College and Career-Ready Standards have led MDE to increase the rigor in instruction and testing. MDE believes that these standards give Mississippi's children the ability to compete, not just with their classmates and other students in the state, but with students from across the country and world. Incorporating college and career ready standards into Mississippi's education framework means the use of new assessments that determine whether students have achieved a certain level of competency necessary to be successful in either a career or college.

Importance of Effective Accountability Standards

State accountability standards must be designed in such a way that they effectively demonstrate actual school performance. If standards do not reflect actual student performance, education stakeholders and decisionmakers cannot make the appropriate decisions or necessary adjustments to improve schools' and districts' performance.

Accountability standards, the standards that provide benchmarks that characterize a school's or district's performance, must be constructed in such a way that they adequately depict how schools and districts are performing. According to The Accountability Systems and Reporting State Collaborative in Assessment and Student Standards Project (ASR SCASS),¹ accountability standards should:

- identify and promote improved educational practices and results;
- inform stakeholders of the condition of education at school, district, and state levels;
- identify areas in which improvement is needed and success is being achieved;
- obtain the support of all stakeholders in making the changes needed to enable all students to achieve at high levels; and,
- inform policy decisions and actions by officials at the local, state, and federal levels; parents, students, and members of the community; and other interested individuals to improve

¹The Accountability Systems and Reporting State Collaborative in Assessment and Student Standards Project works to develop documents that will help state departments of education design, improve, or review their state accountability and reporting systems.

academic performance where needed and to reward it where appropriate.

One of the specific characteristics of accountability standards described by the ASR SCASS is that accountability standards should report school performance in relation to criteria or standards that are established by the state. The criteria should provide a credible operational system for evaluating and publicizing school performance results and assigning rewards, assistance, and sanctions.

Key Players in Developing and Administering Mississippi's Education Accountability Standards

The Accountability Standards Task Force, the membership of which is approved by the Mississippi Board of Education, makes accountability standards recommendations to the Commission on School Accreditation. Once recommendations are approved by the commission, the Board of Education provides the final approval before new standards or changes in standards go into effect. Selected staff at the department provide information necessary for the task force to make accountability standards recommendations.

The key players in developing and administering the accountability standards are:

- the Mississippi Board of Education;
- the Commission on School Accreditation;
- the Accountability Task Force; and,
- the Mississippi Department of Education.

The Mississippi Board of Education

The Mississippi Board of Education is made up of nine members. Each member is appointed according to the Mississippi Constitution and as defined in MISS. CODE ANN. Section 37-1-1 (1972). The Governor appoints five members of the board--one from each of the three Supreme Court districts, one who is an active and full-time school administrator, and one who is an active full-time teacher. The Lieutenant Governor appoints two at-large members and the Speaker of the House of Representatives appoints two at-large members.

The Board of Education appoints the State Superintendent of Education, sets public education policy, and oversees the Mississippi Department of Education. The Board of Education must approve any changes to the accountability system before they go into effect. See Appendix A, page 51, for a list of current board members.

The Commission on School Accreditation

The Commission on School Accreditation, addressed in MISS. CODE ANN. Section 37-17-3 (1972), is appointed by the Board of Education. This commission of fifteen members continually reviews the standards on accreditation and their enforcement and makes recommendations to the Board of Education. The commission must also initially approve recommendations to the accountability standards before they go to the Board of Education for final approval. See Appendix A, page 49, for a list of current commission members.

The Accountability Task Force

The Accountability Task Force was originally created by the Board of Education in May 2012 to look into the graduation component of the accountability standards. Task force recommendations first go to the commission for approval; this is followed by a vote for final approval by the board.

In August 2012, the board voted to expand the charge of the task force to include a study of the growth model (see page 18), as well as the graduation/dropout component and to make recommendations to the commission and the board regarding the Quality of Distribution Index (QDI) and QDI cut-point ranges. A new task force was created in April 2013, as minor changes to the original standards turned into an entirely new accountability standards process. According to MDE, task force meetings are open to the public.

See Appendix B, page 57, for a brief explanation of the QDI. See Appendix A, page 50, for lists of the original and current task force members.

The Mississippi Department of Education

The Mississippi Department of Education is the administrative arm of the Board of Education. MDE is responsible for implementing state and federal education laws, disbursing state and federal funds, holding schools and districts accountable for performance, and licensing all educators. MDE provides resources and technical support to Mississippi's public school system and functions as a resource for federal education requirements and funding. MDE staff provides information necessary for the task force to make decisions.

Why MDE Recently Revised the Accountability Standards

According to the Mississippi Department of Education, changes in Mississippi state law, federal requirements, and the desire to make the accountability standards equitable for all schools and districts and easier to understand led to the adoption of MDE's current accountability standards.

The federal No Child Left Behind Act of 2001 required states to develop and implement a single statewide accountability system (see Appendix B, page 54, for a summary explanation). MISS. CODE ANN. Section 37-17-6 (1972) calls for the establishment and implementation of a permanent performance-based accreditation system. According to this section, the annual performance standards will "measure the performance of each school against itself through the standard that has been set for it." The statute specifies that the performance-based accreditation system must include:

- high expectations for students and high standards for all schools;
- strong accountability for results;
- a process to implement accountability at both the school district level and the school level;
- holding individual schools accountable for student growth and performance;
- annual performance standards for each of the schools and measurement of the performance of each school against itself through the standard that has been set for it;
- determination of which schools exceed standards and a plan for providing recognition and rewards to those schools;
- determination of which schools fail to meet standards and means for intervention; and,
- development of a comprehensive student assessment system to implement these requirements.

Goals of the Accountability Standards

MISS. CODE ANN. Section 37-17-6 (1972) established three accountability goals: all students leaving third grade are reading on grade level, the dropout rate is reduced to thirteen percent, and sixty percent of students are scoring proficient or advanced on assessments.

According to MDE, the current accountability standards measure progress toward the following goals identified in MISS. CODE ANN. Section 37-17-6 (1972):

- to ensure that all students exit third grade reading on grade level by 2015;
- to reduce the dropout rate to thirteen percent by 2015; and,
- to have sixty percent of students scoring proficient and advanced on the assessments by 2016 (with 3% annual incremental increases thereafter).

According to MDE, the following are the key differences between the current accountability standards and the previous accountability standards:

- The current model emphasizes student growth, particularly the lowest performing 25% of students.
- The previous system calculated student growth using an equation that predicted growth. Now, students meet growth if their scores improve from one proficiency level to the next, or move sufficiently within the lower proficiency levels, or stay the same within the proficient or advanced level.
- The previous system included a graduation and dropout component for the twelfth grade score, which gave schools partial credit for GED completers and other types of non-traditional diplomas. These students do not accumulate credit in the current system, per state statute.

The current standards, described in detail in MDE's *Guide to Calculating School and District Grades*, outline what components are included and how data is calculated to determine a school's and district's grades.

Characteristics and Components of a School's or District's Accountability Grade

This chapter addresses:

- assessments used as the basis for the accountability standards;
- the achievement categories, based on assessment scores;
- components of a school's or district's grade; and,
- cut-points.

Assessments Used as the Basis for the Accountability Standards

Five different assessments are used to determine schools' and districts' accountability grades. These assessments are administered at various grade levels throughout the elementary, middle school, and high school levels within schools and districts.

The assessments used to determine schools' or districts' accountability grades are:

- The *Mississippi Curriculum Test*, Second Edition (MCT2) is administered in grades 3 through 8. According to MDE, the MCT2 consists of language arts and mathematics assessments that are fully aligned with the 2006 Mississippi Language Arts Framework-Revised and the 2007 Mississippi Mathematics Framework-Revised. Administering these assessments allows Mississippi to be in compliance with the requirements of the No Child Left Behind Act of 2001 (NCLB). (See Appendix B, page 54, for the requirements of NCLB.) A committee of Mississippi teachers, selected by the MDE, approved the items that appear on these tests.
- *The Mississippi Science Test* (MST2), administered in grades 5 and 8, is a science test that allows Mississippi to be in compliance with the requirements of NCLB. A committee of Mississippi teachers, selected by the MDE, approved the items that appear on these tests.
- *The Subject Area Testing Program*, Second Edition (SATP2) tests are required for graduation; however, a student can take the SATP2 assessments prior to even entering high school--i.e., as an eighth grader.² SATP2 tests are developed and administered by MDE and consist of four subject area tests: Algebra I, Biology I, U. S. History, and English II. Prior to the March 2015 Board of Education meeting, which changed graduation requirements, students earning a high

²Students who take SATP2 tests prior to entering the tenth grade have their scores "banked" until their tenth grade year to be used to determine the school's grade. See Appendix B, page 54, for more information on banking of scores.

school diploma had to pass all four subject area tests to meet graduation requirements. (See page 44 for further discussion of the new graduation requirements.) Unlike the MCT2 tests that are administered each year from third grade to eighth grades, the subject area tests are taken once but can be repeated if a student fails the initial test.

- *The Dynamic Learning Maps³* (DLM) assessment is for students with significant cognitive disabilities (SCD). This reading and math assessment replaced the Mississippi Alternate Assessment of Extended Curriculum Frameworks (MAAECF) and was first administered during the 2014-2015 school year.
- *The Mississippi Alternate Assessment of Extended Science Frameworks (MAAESF)* is the science assessment for the SCD student population. Students who take the alternate assessments do not take the MCT2 in elementary school or the SATP2 in high school. According to MDE, the MAAESF is designed to assess the educational performance of students with disabilities who cannot participate in the general education curriculum, even with accommodations. Students in third through eighth grade and eleventh and twelfth grade who meet the state’s three SCD criteria are eligible. See Appendix B, page 56, for an explanation of the state’s three SCD criteria.

Exhibit 1, below, shows the grade levels at which the assessments are administered.

Exhibit 1: Assessments Utilized by MDE as the Basis of the Accountability Standards and Grade Level at which Assessments are Administered, as of 2014-2015 Assessment Year

Grade Level	MCT2 Reading	MCT2 Math	MST (Science)	SATP2 Algebra I	SATP2 English II	SATP2 Biology I	SATP2 U.S. History	DLM Math	DLM Reading	MAAESF (Science)
K										
1										
2										
3	Y	Y						Y	Y	
4	Y	Y						Y	Y	
5	Y	Y	Y					Y	Y	Y
6	Y	Y						Y	Y	
7	Y	Y						Y	Y	
8	Y	Y	Y					Y	Y	Y
High School				Y	Y	Y	Y	Y	Y	Y

SOURCE: MDE.

³The MAAECF was the SCD assessment used prior to the DLM assessment. It was administered for the last time during the 2013-2014 school year. According to MDE, the DLM assessment was pilot tested in the 2013-2014 school year. Scores for that year’s DLM were not included in the calculation of a school’s grade, but were included in the 2014-2015 school year. A new alternate assessment will be administered in the 2015-2016 school year.

Achievement Categories

After students take the assessments, MDE places each student's score into one of four achievement categories (advanced, proficient, basic, or minimal).

After taking the assessments previously described, each student's score is placed into one of four categories:

- advanced;
- proficient;
- basic; or,
- minimal.

MDE uses a "scale" score to determine each student's placement within an achievement category. *Scaling* refers to the process of converting a student's raw test score to a common score that allows for comparison between students.

As an example, Exhibit 2, below, shows, by grade, the scale score ranges for the MCT2 math assessment in May 2014 for each of the achievement categories.

Exhibit 2: Scale Math Scores* for MCT2 Test (May 2014 Administration)

	Minimal	Basic	Proficient	Advanced
Grade 3 Math	107-137	138-149	150-165	166-189
Grade 4 Math	112-140	141-149	150-164	165-187
Grade 5 Math	114-141	142-149	150-163	164-187
Grade 6 Math	114-141	142-149	150-163	164-186
Grade 7 Math	107-142	143-149	150-163	164-189
Grade 8 Math	117-142	143-149	150-163	164-184

*MDE defines scale scores as statistical conversions of raw scores that adjust for variations in the difficulty of items on different tests and permit valid comparison across all test administrations within a particular subject area or grade.

SOURCE: MDE.

Components of a School's or District's Grade

MDE created a point system that it utilizes to determine what accountability grade a school or district will receive.

MDE uses the percentage of students that a school or district has in the top two achievement categories to determine that school's or district's accountability grade. MDE uses seven components (i.e., 700 possible points) to determine a grade for

a school with no twelfth grade⁴ or a district with no high school and nine components (i. e., 900 possible points) to determine a grade for a school with a twelfth grade or a district with a high school. The components and their possible points are illustrated in Exhibit 3, below.

Exhibit 3: Components of a School’s or District’s Accountability Grade, as of 2013-2014 Assessment Year

Components	Without 12 th Grade	With 12 th Grade
	700 Possible Points	900 Possible Points
Reading Proficiency	100	100
Reading Growth-All Students	100	100
Reading Growth-Low 25% of Students	100	100
Math Proficiency	100	100
Math Growth-All Students	100	100
Math Growth-Low 25% of Students	100	100
Science Proficiency	100	50
U.S. History Proficiency		50
Graduation Rate-All Students*		200

*MDE uses a federally approved four-year graduation rate calculation (MISS. CODE ANN. Section 37-17-6 [1972]). See page 26.

NOTE: MDE does not currently use “college and career readiness” and “acceleration” to calculate a school’s or district’s grade. However, according to MDE, these components will be included beginning with School Year 2015-2016 results. See pages 52-53 for more information on these components.

SOURCE: MDE.

For schools and districts that neither administer a science assessment (for example, 6th and 7th grade schools) nor have a twelfth grade, MDE assigns the grade based on the 700-point system that it “equates” to a six-component system. (See explanation on page 15.)

MDE also has established rules that will give schools that do not take assessments (i. e., K-1 and K-2 schools) an accountability grade. To do this, MDE uses a process it refers to as “back mapping.” According to the *Mississippi Statewide Accountability System: Guide to Calculating School and District Grades*, effective for the 2013-2014 academic year, back mapping is used in any elementary or middle school that does not have reading or math scores because the school does not have the required grade level. MDE uses the scores from the students in the next higher grade in the tested subject within the same district and applies them to the student’s lower elementary school of origin. (For more information on back mapping see Appendix B, page 52.)

⁴An example of a school with no twelfth grade would be an elementary school.

Cut-Points

MDE's Accountability Task Force, with the guidance of the contractor Research in Action, Inc., created a set of cut-points that demonstrate how many points must be earned for a school or district to receive its accountability grade.

A *cut-point*, sometimes referred to as cut-score, is the score that serves to classify the students whose score is below the cut-point into one level and the students whose score is at or above the cut-point into the next level or a higher level. MDE uses established cut-points to determine what score needs to be achieved for a school to receive an A, B, C, D, or F accountability grade. MDE's current cut-points are shown in Exhibit 4, below.

Exhibit 4: MDE Cut-Points for Schools and Districts, as of 2013-2014 Assessment Year

Letter Grade	Cut-Point Range	
	Without 12 th grade	With 12 th grade
A	518 or higher	695 or higher
B	455-517	623-694
C	400-454	540-622
D	325-399	422-539
F	324 or lower	421 or lower

SOURCE: MDE.

How the Task Force Established Cut-Points

MDE, the Task Force, and the Technical Review Committee, with the help of a consultant from Research in Action, Inc., established the current cut-points for the seven- and nine-component schools and districts by adding each one's component points to determine a composite score. Then they ranked the composite scores and used each composite score's median as a starting point for discussions regarding whether a particular school or district "looked like" an A, B, or C school while keeping in mind original discussions on the characteristics of such schools.

To establish the current cut-points, MDE enlisted the help of a consultant from Research in Action, Inc. (RIA). In the fall of 2013, the consultant, along with MDE's Accountability Task Force, discussed what characteristics made an A school a B school, a B school a B school, etc. The National Strategic Planning and Analysis Research Center (nSPARC) at Mississippi State University provided MDE with actual student score data to determine each school's or district's individual component scores. RIA divided the data according to whether the school or district had seven components or nine components. Those

schools or districts that had six components were not used in the calculation for determining cut-points.

Research in Action performed calculations and had discussions with MDE and the task force for the seven- and nine-component schools and districts separately. For each school or district, RIA calculated the composite score by summing each one's component scores to give each one a total possible number of 700 points or 900 points, depending on the respective number of components. Then, for each of the two sets, RIA ranked the composite scores from highest to lowest and determined a median value of all scores for each set. This median value was determined to be the mid-point of a C for that set.

Then RIA determined the median value for each component. The Technical Review Committee and the task force, along with RIA, then used the median for each component as a starting point for discussions regarding whether a particular school "looked like" an A, B, or C school while keeping in mind the original discussion on the characteristics of A, B, and C schools.

Based on these calculations and discussions, RIA, the Technical Review Committee, and the task force determined initial cut-points for A through F for the nine-component schools or districts and the seven-component schools or districts. For the six-component schools, MDE staff applied the actual distribution of grades (A-F) from the seven-component schools and districts to the six-component schools rather than determining the six-component school cut-points in the same manner as the seven- and nine- component cut-points.

Even though the median will change with each year's calculation, theoretically, the cut-points are not meant to change each year. However, once the initial cut-points were determined, MDE, RIA, the Technical Review Committee, and the task force made certain changes to those cut-points. For example, according to MDE, in 2013, it was the federal government's intent that the graduation component account for twenty percent of a school's or district's grade. When this happened, MDE adjusted the cut-points to reflect the extra weight on this component. Also, according to MDE, when assessments change, MDE will adjust the cut-points.

How MDE Calculates a School's or District's Grade

This chapter addresses:

- the proficiency calculation;
- the growth calculation; and,
- the graduation rate calculation.

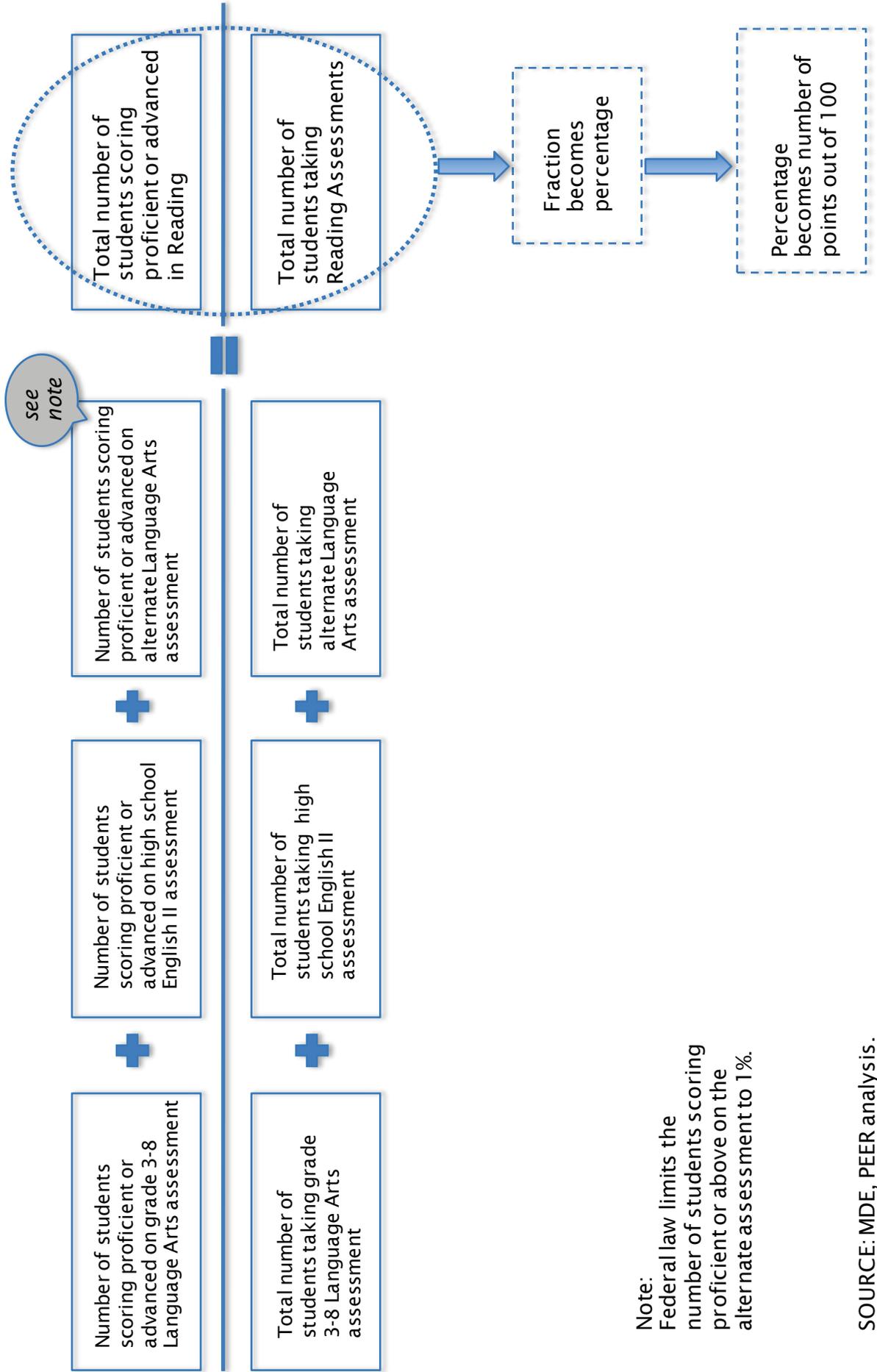
The Proficiency Calculation

For purposes of the school's or district's accountability grade, MDE determines proficiency by calculating the percentage of students who performed at or above "proficient" on assessments.

As previously noted, proficiency in reading, math, science, and history must be determined in order for MDE to calculate a school's or district's accountability grade. MDE determines proficiency in these subject areas based on the percentage of students who performed at or above "proficient." In other words, this is the percentage of students whose assessment score (assessments are described on page 10) placed them in the "proficient" or "advanced" achievement category (achievement categories are described on page 12).

For example, the assessments used to determine reading proficiency are the MCT2 LA (Language Arts), SATP2 English II, and the alternate assessment reading test. To determine a school's reading proficiency score, the number of students who scored proficient or advanced on each of these assessments is totaled and then divided by the total number of students who took the assessments. The fraction would then become a percentage and that percentage would be the number of points earned in the proficiency component. Exhibit 5, page 17, shows how MDE calculates student score data to arrive at a reading proficiency score for a school or district. Exhibit 6, page 18, illustrates how MDE would calculate the reading, math, U. S. history, or science proficiency rate for a school or district.

Exhibit 5: Method of Determining Reading Proficiency Score with MDE's Accountability Standards, as of 2013-2014 Assessment Year



Note: Federal law limits the number of students scoring proficient or above on the alternate assessment to 1%.

SOURCE: MDE, PEER analysis.

Exhibit 6: Example of a Proficiency Calculation, as of 2013-2014 Assessment Year

Number of Students	Achievement Category	Value
50	Minimal	
55	Basic	
60	Proficient	
35	Advanced	
Number of students scoring proficient or advanced →		95
Total number of students taking the assessment →		200
Percent of students scoring proficient or advanced →		47.5%
Number of points school or district would get for proficiency component →		47.5

SOURCE: PEER analysis.

The Growth Calculation

MDE defines growth as the percentage of students who made “learning gains.” MDE calculates growth for all students as well as for the lowest twenty-five percent of students.

MDE defines growth as whether a student made a “learning gain.” See Exhibit 7, below, regarding what constitutes a “learning gain.”

Exhibit 7: Possible Learning Gains and Respective Multipliers Utilized by MDE to Calculate Growth Components, as of 2013-2014 Assessment Year

Type 1 Learning Gain (1.0 Multiplier)	Type 2 Learning Gain (1.2 Multiplier)	Type 3 Learning Gain (1.25 Multiplier)
<ul style="list-style-type: none"> An increase in any performance level Staying at the same performance level that is at or above “proficient” from one year to the next An increase within the lowest two performance levels that crosses over the mid-point of the level (e. g., bottom half of “basic” to top half of “basic”) 	<ul style="list-style-type: none"> Increase of two or more levels 	<ul style="list-style-type: none"> Increase to the highest performance level (advanced)

NOTE: In the event that a student meets criteria in more than one scenario, the higher weight is applied.

SOURCE: MDE and PEER analysis.

MDE considers two areas of growth when determining a school's or district's accountability grade:

- growth of all students ("All-Growth"); and,
- growth of the lowest twenty-five percent of students ("Low 25%").

MDE uses the growth components only for math and reading/language arts because math and reading/language arts are tested every year in grades three through eight and once in high school, while science is only tested in grades five and eight and once in high school and history is only tested once in high school. See Exhibit 1, page 11, for an illustration of this explanation.

Calculating Growth for All Students

The growth of all students refers to the percentage of students who made learning gains from one year's assessment to the next year's assessment.

The All-Growth component refers to the percentage of students who were in the school or district for a full academic year for two assessment years that made learning gains from one year's assessment to the next year's assessment. (Refer to Appendix B, page 54, for a definition of full academic year.)

The assessments used for calculating all-growth are:

- reading--the MCT2 Language Arts test taken in grades three through eight, the SATP2 English II test taken once in high school, and the alternate assessment reading test taken in grades three through eight and once in high school; and,
- math--the MCT2 Math test taken in grades three through eight, the SATP2 Algebra I test taken once in high school, and the alternate assessment math test taken in grades three through eight and once in high school.

Refer to Appendix B, page 52, for exceptions when including the alternate high school assessment in the reading and math growth components.

For assessments that are not taken over two consecutive years, a student's achievement category in one year will be "banked" until that student takes the subsequent assessment. (Refer to Appendix B, page 52, for a more detailed explanation of banking scores.)

As described previously in the discussion on determining proficiency scores, each student's score on a given assessment is placed in an achievement category. For the purpose of determining growth, the bottom two categories (minimal and basic) are further divided into four sub-categories: bottom half of minimal, top half of minimal, bottom half of basic, and top half of basic. MDE determines growth (of all students and the low 25%) by whether a student made learning gains. Students can make a learning gain by satisfying a number of scenarios.

Students can also receive additional credit for making “additional learning gains.”

Exhibit 7, page 18, shows the additional weight a student’s score can receive based on the type of learning gain that that student makes.

To illustrate, Exhibit 8, page 21, shows what types of movement constitute growth, along with the multipliers associated with certain types of movement. Exhibit 9, page 22, illustrates how the reading growth of all students is calculated. Exhibit 10, page 24, illustrates how MDE calculates the All-Growth component for either reading or math.

Calculating Growth for the Lowest 25% of Students

The growth of the lowest 25% component refers to the percentage of students who scored in the low 25% of their class the previous testing year who made learning gains between the previous year’s assessment and the current year’s assessment.

MDE uses the following assessments for calculating low 25% growth:

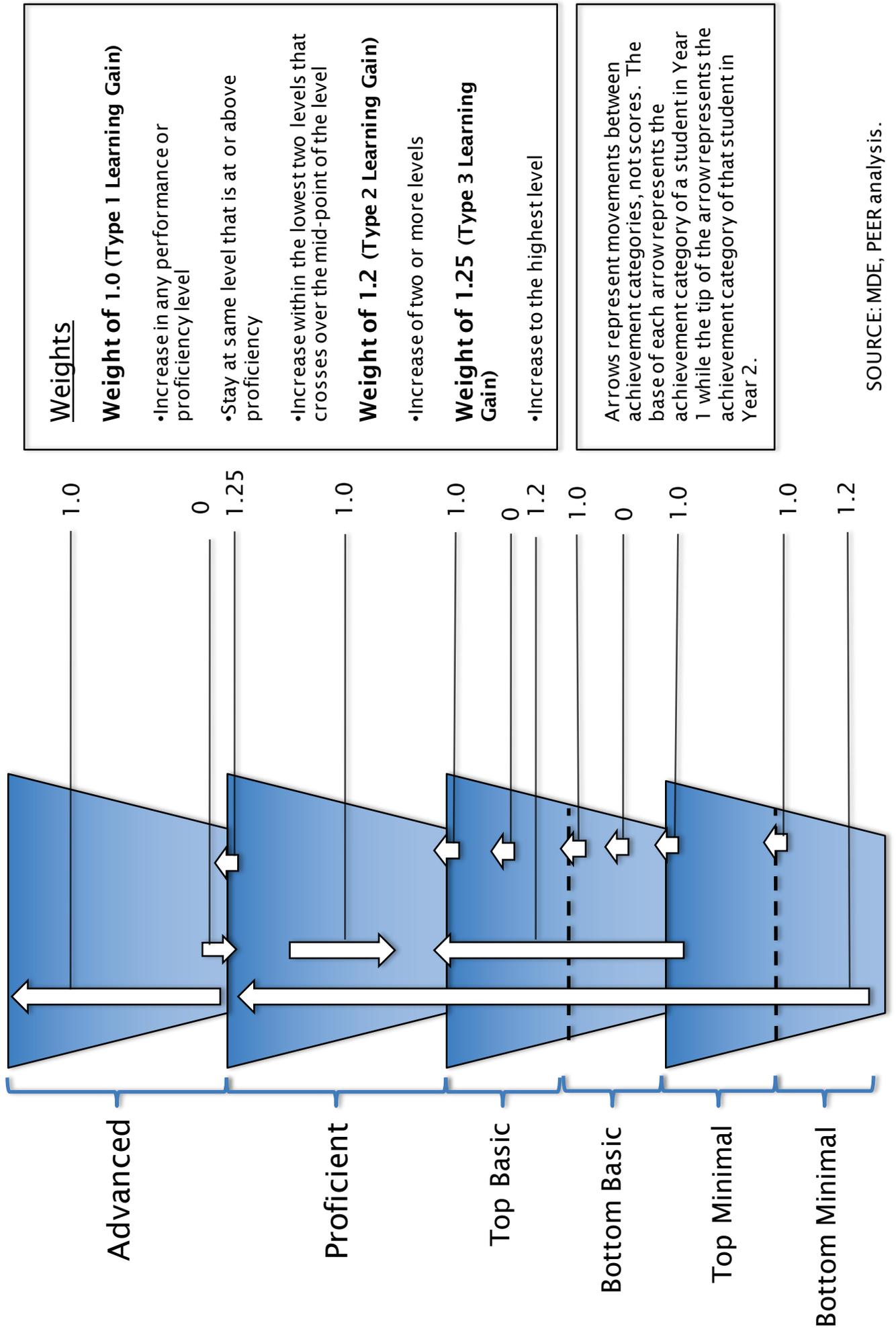
- reading--the MCT2 Language Arts test taken in grades three through eight and the SATP2 English II test taken once in high school; and,
- math--the MCT2 Math test taken in grades three through eight and the SATP2 Algebra I test taken once in high school.

Students who took the alternate assessments prior to the 2014-2015 school year were not included in the calculations for low 25% growth in either reading or math. According to MDE, beginning with the 2014-2015 school year, alternate assessments will be included in all components.

Refer to Appendix B, page 54, for an explanation of the exclusion of students with banked scores from the low 25% growth calculation.

The first step in the process of determining the low 25% growth group score is to identify which students made up the low 25% population in the previous year’s assessment. (This process is explained in detail in Appendix C on page 61.) This process is repeated for each current year’s grade level (4th graders through 8th graders, plus high school students) using each student’s scores on the previous year’s assessment (3rd grade through 8th grade tests). As illustrated in Exhibit 11 on page 25, the number of students identified through this process is summed to determine the total number of students who scored among the low 25% of their grade level’s scores on assessments the previous year. The number becomes the denominator for the low 25% growth calculation.

Exhibit 8: Illustration of Types of Growth and Respective Weights Utilized with MDE's Accountability Standards, as of 2013-2014 Assessment Year



Weights

Weight of 1.0 (Type 1 Learning Gain)

- Increase in any performance or proficiency level
- Stay at same level that is at or above proficiency
- Increase within the lowest two levels that crosses over the mid-point of the level

Weight of 1.2 (Type 2 Learning Gain)

- Increase of two or more levels

Weight of 1.25 (Type 3 Learning Gain)

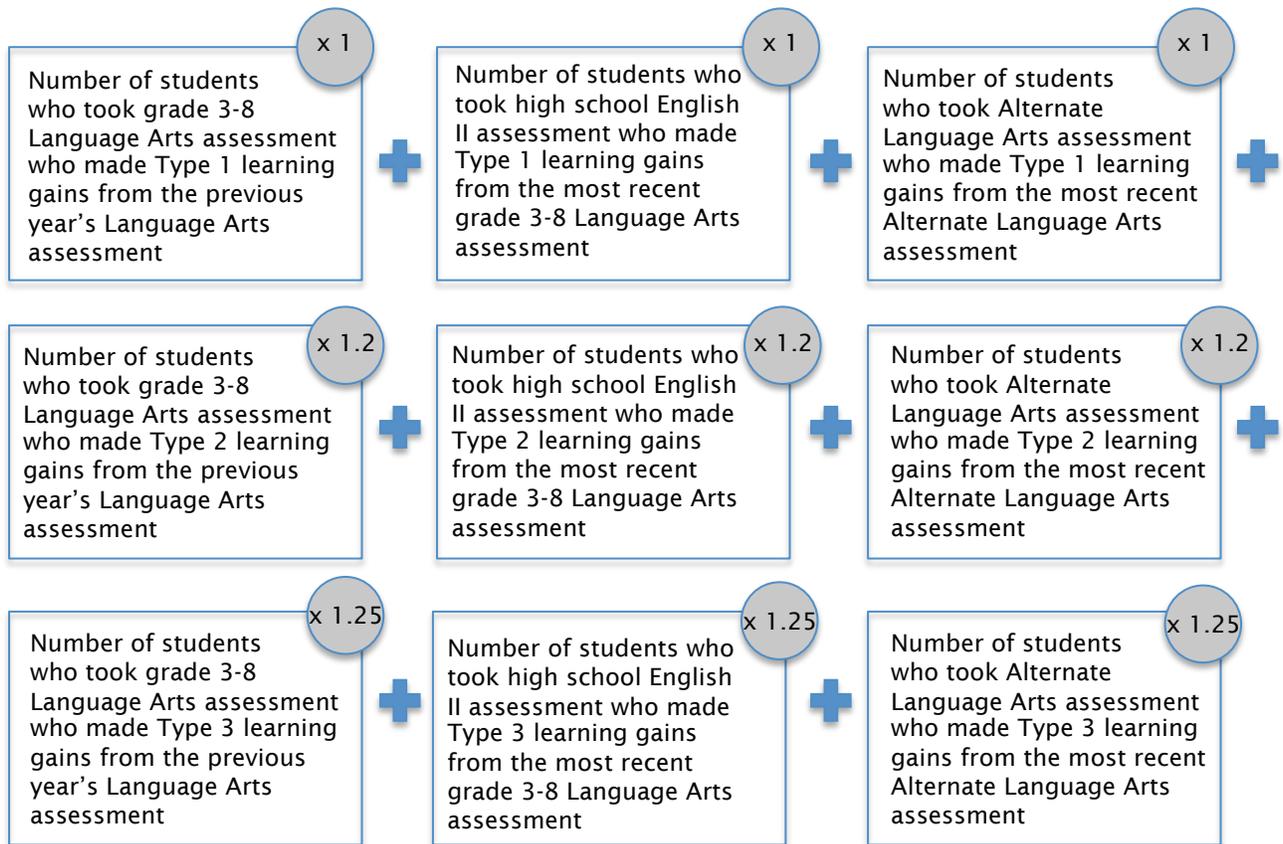
- Increase to the highest level

Arrows represent movements between achievement categories, not scores. The base of each arrow represents the achievement category of a student in Year 1 while the tip of the arrow represents the achievement category of that student in Year 2.

SOURCE: MDE, PEER analysis.

Exhibit 9: Method of Determining Reading All-Growth Score with MDE’s Accountability Standards, as of 2013-2014 Assessment Year

Calculation of Numerator



•Type 1 learning gains are given a weight of 1

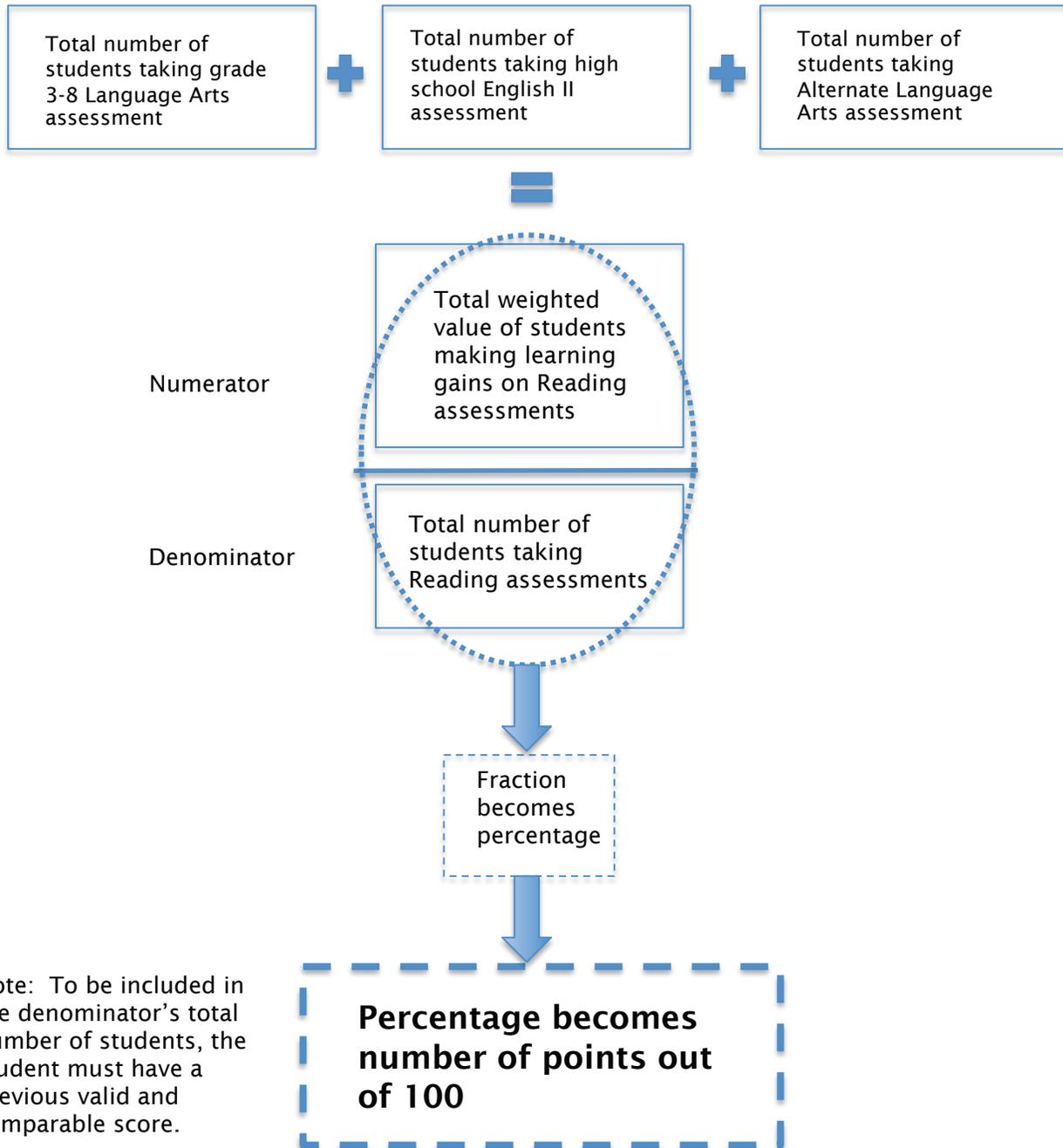
•Type 2 learning gains are given a weight of 1.2

•Type 3 learning gains are given a weight of 1.25

Note: If a student takes the high school assessment before 10th grade, that score is placed in an achievement category and banked until the student reaches 10th grade and then is compared to the achievement category of the assessment taken by the student in the year prior to taking the high school assessment

Exhibit 9 (cont.): Method of Determining Reading All-Growth Score with MDE's Accountability Standards, as of 2013-2014 Assessment Year

Calculation of Denominator



Note: To be included in the denominator's total number of students, the student must have a previous valid and comparable score.

SOURCE: MDE, PEER analysis.

Exhibit 10: Reading or Math All-Growth Calculation Example, as of 2013-2014 Assessment Year

Number of Students	Year 1 Achievement Category	Year 2 Achievement Category	Weight	Value (number of students x weight)
10	Bottom Minimal	Top Minimal	1	10
20	Bottom Minimal	Top Basic	1	20
10	Top Minimal	Top Minimal	0	0
10	Top Minimal	Proficient	1.2	12
20	Bottom Basic	Top Basic	1	20
10	Top Basic	Bottom Basic	0	0
10	Top Basic	Top Basic	0	0
2	Top Basic	Advanced	1.25	2.5
8	Proficient	Top Basic	0	0
40	Proficient	Proficient	1	40
20	Proficient	Advanced	1.25	25
20	Advanced	Proficient	0	0
20	Advanced	Advanced	1	20
Number of students who made learning gains →				149.5
Number of students who took the assessments →				200
Percent of students who made learning gains →				74.8%
Number of points for the All-Growth component (reading and math) →				74.8

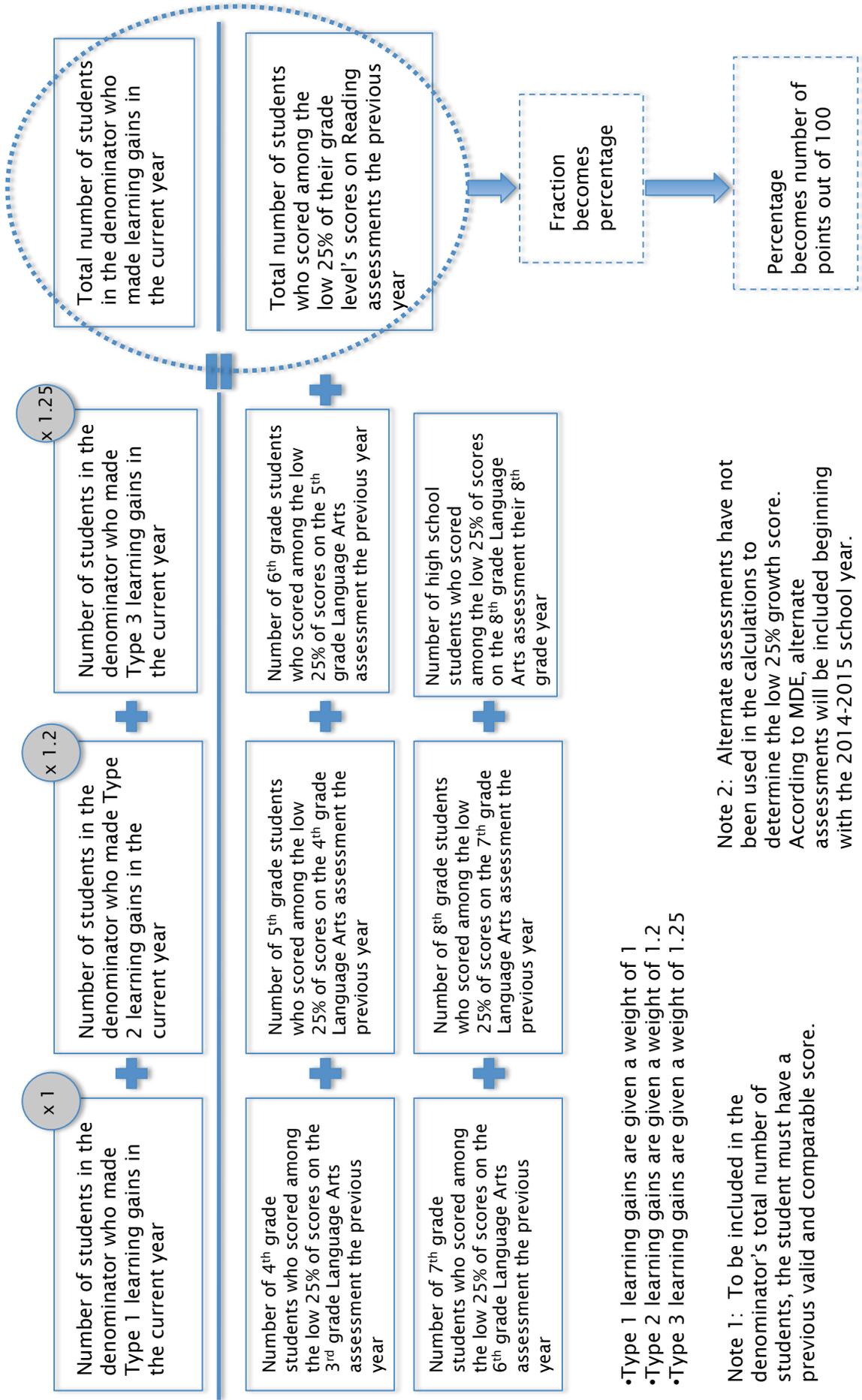
SOURCE: MDE and PEER analysis.

The second step in the process of determining the low 25% growth group score is to identify those students who are in the denominator who made learning gains between the previous year's test and the current year's test. As illustrated in Exhibit 11, for those students who made Type 1 learning gains, a weight of 1 is applied; for those students who made Type 2 learning gains, a weight of 1.2 is applied; and for those students who made Type 3

learning gains, a weight of 1.25 is applied. As illustrated in Exhibit 11, this process is conducted for each current year's grade level (fourth through eighth and high school) separately and then summed to determine the total number of students who were in the low 25% growth group in the previous year who also made learning gains on the current year's test. This number becomes the numerator of the low 25% growth group calculation.

This results in a percentage of low 25% growth students who made learning gains from one year's test to the subsequent year's test. This percent becomes the number of points for the low 25% growth (either reading or math) component. To illustrate, Exhibit 11 shows how the reading growth of the lowest 25% of students is calculated.

Exhibit 11: Method of Determining Reading Low 25% Growth Score with MDE's Accountability Standards, as of 2013-2014 Assessment Year



- Type 1 learning gains are given a weight of 1
- Type 2 learning gains are given a weight of 1.2
- Type 3 learning gains are given a weight of 1.25

Note 2: Alternate assessments have not been used in the calculations to determine the low 25% growth score. According to MDE, alternate assessments will be included beginning with the 2014-2015 school year.

Note 1: To be included in the denominator's total number of students, the student must have a previous valid and comparable score.

SOURCE: MDE, PEER analysis.

The Graduation Rate Calculation

In accordance with federal requirements, MDE calculates the graduation rate by determining the percentage of students who graduated in four years with a “regular high school diploma.”

When calculating the graduation rate, MDE uses the number of students who graduated in four years with a “regular high school diploma” as the numerator and the number of students who entered four years earlier as first-time ninth graders (with adjustments for deaths and transfers in and out) as the denominator.

MDE defines a “regular high school diploma” as the standard high school diploma that is fully aligned with the state’s academic content standards. The method of calculating the graduation rate is prescribed in *Code of Federal Regulations* Section 200.19. No exceptions are made for SCD students, students receiving an occupational diploma, students who receive a GED, or students who receive a certificate of achievement. The graduation rate is adjusted to account for deaths and transfers.

Exhibit 12, below, demonstrates a hypothetical calculation of the graduation rate.

Exhibit 12: Graduation Rate Calculation for a Hypothetical School or District

Numerator (<i>number who graduated in four years with regular high school diploma</i>)				138
Denominator = A-B+C				
<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	
<i>Number who entered as first-time ninth graders</i>	<i>Adjustments for death and emigration (transfers out)</i>	<i>Adjustments for immigration (transfers in)</i>	<i>Number who graduated in four years with a regular high school diploma</i>	
156	-8	+11	+138	159
Graduation Rate				86.8%

SOURCE: PEER analysis.

The Effectiveness of MDE's Accountability Standards in Measuring Schools' and Districts' Performance

Because of the way in which Mississippi's accountability standards are currently calculated, the standards do not provide stakeholders and the public with a clear picture of how Mississippi's schools and districts are performing. Not only does the calculation of the current standards make it impossible to compare one school or district to another, but also to compare a school or district to itself over time. Mississippi's standardized tests are carefully constructed to ensure that a student has mastered a certain level of competency; those tests alone provide the criterion/standard that should measure school performance.

PEER reached this conclusion based on the following:

- how achievement categories obscure student score data;
- issues presented by combining proficiency and growth into a single school grade;
- why MDE's emphasis on growth fails to demonstrate actual school or district performance; and,
- MDE's growth multipliers.

How Achievement Categories Obscure Student Score Data

MDE's use of achievement categories obscures actual student test score data because all scores in an achievement category are basically considered to be equal, despite the wide range of scores within a category. Determining proficiency by calculating the percentage of students whose scores are in the top two achievement categories, as described in MISS. CODE ANN. Section 37-17-6 (5) (c) (i) (1972), compounds the problem because the range of scores deemed "proficient" is even wider, indicating an insensitive measurement instrument.

In accordance with MISS. CODE ANN. Section 37-17-6 (5) (c) (i) (1972), MDE uses the percentage of students who scored proficient and advanced on the current state assessments to determine proficiency. As described previously, in determining a school's or district's accountability grade, MDE places student assessment scores into one of four achievement categories (minimal, basic, proficient, or advanced). At this point, all scores in each achievement category are basically considered to be equal, despite the wide range of scores within a category. For example, the score of a student at the very bottom of the proficient category (i. e., one point above the basic category) is considered the same as a student's score at the top of the advanced category.

The problem of using achievement categories is compounded because MDE adds the “proficient” and “advanced” categories to determine the percentage of students in calculating proficiency (see page 16). As shown in Exhibit 13, page 29, a scale score of 189 on the third grade MCT2 Language Arts assessment would be included in the calculation to determine the school’s or district’s proficiency. Another student’s scale score of 150 on the same assessment would also be considered equally “proficient” because MDE’s calculation of the proficiency component includes students in both achievement categories.

MDE already necessarily converts students’ raw test scores to scale scores for comparability purposes. The scaling of student scores should be the extent of the score adjustment process; any further adjustments obscure actual student performance. In order for a school’s or district’s true student performance to be represented by its accountability grade, MDE should stay as close as possible to students’ score data.

Further, a school or district could have objectively better or worse scale scores than another school or district but receive an identical number of points for its proficiency component. For example, the hypothetical school or district shown in Exhibit 6 on page 18 received 47.5 total component points. As illustrated in Exhibit 14, page 30, another school or district could have a different distribution of scale scores yet also receive 47.5 total component points.

In the example in Exhibit 14, the hypothetical school or district with 95 advanced students and 105 basic students scored objectively better than the hypothetical school or district in Exhibit 6 on page 18, that had 35 advanced, 60 proficient, 55 basic, and 50 minimal students. Further, the hypothetical school or district in Exhibit 6 scored objectively better than the hypothetical school or district in Exhibit 14 with 95 proficient and 105 minimal students. All three hypothetical schools or districts could have the same number of students, have objectively different distributions of scores, yet receive the same number of proficiency component points. This is an indicator of an insensitive measurement instrument.

Exhibit 13: May 2014 MCT2 Test Score Summary Showing Raw Score and Scale Score* Ranges for Each Achievement Category

Assessment	Minimal		Basic		Proficient		Advanced	
	Scale Score Range	Raw Score Range						
Grade 3 Language Arts	112-137	0-18	138-149	19-26	150-161	27-35	162-189	36-50
Grade 4 Language Arts	107-137	0-15	138-149	16-24	150-161	25-35	162-188	36-50
Grade 5 Language Arts	105-137	0-19	138-149	20-30	150-163	31-46	164-188	47-60
Grade 6 Language Arts	109-137	0-18	138-149	19-28	150-165	29-45	166-189	46-60
Grade 7 Language Arts	109-138	0-19	139-149	20-28	150-167	29-51	168-189	52-70
Grade 8 Language Arts	104-137	0-21	138-149	22-33	150-166	34-55	167-188	56-70
Grade 3 Math	107-137	0-15	138-149	16-24	150-165	25-37	166-189	38-45
Grade 4 Math	112-140	0-16	141-149	17-23	150-164	24-37	165-187	38-45
Grade 5 Math	114-141	0-17	142-149	18-25	150-163	26-41	164-187	42-50
Grade 6 Math	114-141	0-18	142-149	19-25	150-163	26-41	164-186	42-50
Grade 7 Math	107-142	0-20	143-149	21-26	150-163	27-39	164-189	40-50
Grade 8 Math	117-142	0-17	143-149	18-23	150-163	24-39	164-184	40-50

MDE's use of achievement categories obscures actual student test score data because all scores in an achievement category are basically considered to be equal, despite the wide range of scores within a category. Then MDE adds the "proficient" and "advanced" categories to determine the percentage of students in calculating proficiency. For example, a scale score of 189 on the third grade MCT2 Language Arts assessment would be included in the calculation to determine the school's or district's proficiency. Another student's scale score of 150 on the same assessment would also be considered equally "proficient" because MDE's calculation of the proficiency component includes students in both achievement categories.

*MDE defines a raw score as the number of questions a student answered correctly. MDE defines scale scores as statistical conversions of raw scores that adjust for variations in the difficulty of items on different tests and permit valid comparison across all test administrations within a particular subject area or grade.

SOURCE: MDE.

Exhibit 14: How Schools or Districts with Objectively Different Scale Scores Could Receive the Same Number of Proficiency Component Points

Hypothetical School/District 1

Number of Students	Achievement Category	Value
105 ←	Minimal	
0	Basic	
95 ←	Proficient	
0	Advanced	
Number of students scoring proficient or advanced →		95
Total number of students taking the assessment →		200
Percent of students scoring proficient or advanced →		47.5%
Number of points school or district would get for proficiency component →		47.5

Hypothetical School/District 2

Number of Students	Achievement Category	Value
0	Minimal	
105 ←	Basic	
0	Proficient	
95 ←	Advanced	
Number of students scoring proficient or advanced →		95
Total number of students taking the assessment →		200
Percent of students scoring proficient or advanced →		47.5%
Number of points school or district would get for proficiency component →		47.5

SOURCE: PEER analysis of issues with using achievement categories.

Problems with Combining Proficiency and Growth into a Single School Grade

Due to the way MDE's accountability grade components are structured, combining proficiency and growth to determine a school's or district's accountability grade may not present the most accurate picture of actual student performance.

As described previously, the components that determine a school's or district's grade are:

- Reading Proficiency: 100 points;
- Reading Growth (All Students): 100 points;
- Reading Growth (Low 25% of Students): 100 points;
- Math Proficiency: 100 points;
- Math Growth (All Students): 100 points;
- Math Growth (Low 25% of Students): 100 points;
- Science Proficiency: 100 points in 700-point system and 50 points in 900-point system;

- U. S. History Proficiency (schools and districts with a 12th grade): 50 points; and,
- Graduation Rate-All Students (school and districts with at 12th grade): 200 points.

As this list demonstrates, growth (whether for all students or the low 25%) accounts for a large portion of a school's or district's accountability grade--potentially, 400 of the 700 or 900 possible points. Growth of the lowest 25% of students is worth just as many points as all growth in the reading and math proficiency components. Thus, the way the components are currently structured makes it possible for growth to affect a school's or district's grade greatly.

In reference to accountability system changes, MISS. CODE ANN. Section 37-17-6 (4) (g) (1972) states that MDE is "authorized and directed to change the school and school district accreditation rating system to a simple 'A,' 'B,' 'C,' 'D,' and 'F' designation based on a combination of student achievement scores and student growth." However, PEER believes that to ensure an accurate picture of actual student performance, the Legislature should amend this CODE section to allow for separate proficiency and growth indicators.

PEER believes that growth is a very important factor in school performance, but if the way growth is calculated affects a school's or district's grade in such a way that it no longer demonstrates true student performance at that school or district, MDE's overall purpose of the accountability standards is not being fulfilled. If the purpose of the accountability standards is to improve student achievement and increase the level of accountability of schools and districts, then more emphasis should be placed on proficiency--how a student actually performs on the assessments.

Why MDE's Emphasis on Growth Fails to Demonstrate Actual School or District Performance

Because of the way MDE has structured its accountability standards, in certain situations (such as a student whose score places him or her in the low 25% of scores), a student's growth from one achievement category to another could be counted up to three times in the determination of the school's or district's accountability grade for a given year. Additionally, a school or district could appear to have made substantial growth gains, which might actually be inaccurate.

As noted previously, MDE's accountability standards have an all-growth component that measures the percentage of students who made learning gains from one year to the next. In addition to including students who made learning gains in the proficient or advanced category, this component also

includes students who made learning gains in the minimal or basic categories.

Therefore, the all-growth component counts the growth of students in all achievement categories. This is, in a way, duplicative for the students in the proficient or advanced categories, because these students' achievement scores reflect their growth. In essence, for proficient or advanced students, the standards add growth upon growth.

Further, the standards measure the growth of the low 25% of students' scores in the same way that they measure all-growth. Given that the low 25% has already been included in the all-growth component, the low 25% of scores component also counts growth upon growth. See Exhibit 15, page 33, for a visual representation of counting growth upon growth.

Because of the way MDE has structured its accountability standards, in certain situations (such as a student whose year one score places him or her in the low 25% percent of scores), a student's growth from one achievement category to another could be counted up to three times in the determination of a school's or district's accountability rating for a given year.

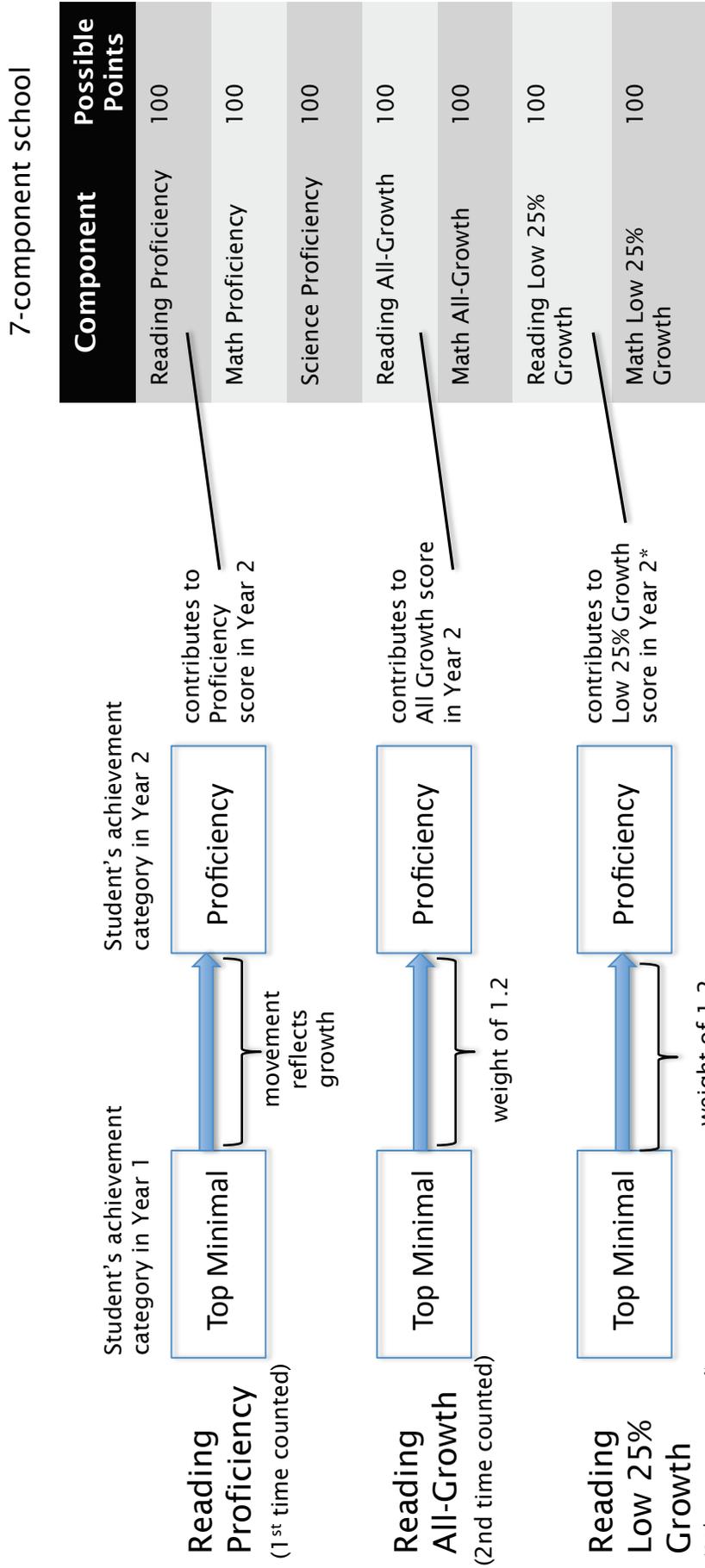
Also, as illustrated in Exhibit 16, page 34, a hypothetical school or district could appear to have achieved 75% growth despite the fact that there were not any students that actually improved by an achievement category.

Why MDE Emphasizes Growth

MDE emphasizes growth in order to ensure that lower performing schools or districts that are improving positively contribute to their school's or district's accountability grades and, as required by state statute, to emphasize the progress of the lowest twenty-five percent of students in the school or district.

During discussions at the task force meetings, task force members expressed their desire to focus on growth in order to ensure that when lower performing schools and districts improve, although they may not reach the proficient or advanced categories, that their growth is recognized in some way and contributes to the school's or district's accountability grade. Further, MISS. CODE ANN. Section 37-17-6 (1972) states that MDE should place an "emphasis on the progress of the lowest twenty-five percent (25%) of students in the school or district." However, the statute does not specify how this should or could be done.

Exhibit 15: Illustration of Counting Growth upon Growth Utilizing MDE's Accountability Standards, as of 2013-2014 Assessment Year



Scores are added, allowing for one action by the student (moving from one achievement category to another) to count 3 times.

* Assuming that the student's test score was in the low 25% of test scores for his grade level in Year 1

SOURCE: PEER analysis.

Exhibit 16: How a School or District Could Appear to Make Substantial Growth Gains

Number of Students	Year 1 Achievement Category	Year 2 Achievement Category	Weight	Value (number of students x Weight)
50	Bottom Minimal	Bottom Minimal	0	0
55	Bottom Minimal	Top Minimal	1	55
95	Proficient	Proficient	1	95
Number of students who made learning gains →				150
Number of students who took the assessments →				200
Percent of students who made learning gains →				75%
Number of points for the All-Growth component (reading and math) →				75

SOURCE: PEER analysis.

Options for Presenting Growth

If proficiency scores are accurate, comparing proficiency scores from one year to the next or reporting scale scores divided by total possible scale score points are other ways of showing whether a school or district improved from one year to the next.

If proficiency scores are accurate, comparing proficiency scores from one year to the next would be another way of showing whether a student improved from one year to the next. For example, if a school or district had a higher proficiency score in Year Two than in Year One, then that school or district showed improvement, or growth. The difference between the two proficiency scores would represent the growth from Year One to Year Two.

Another option to present growth would be to no longer use proficiency and report scale scores divided by total possible scale score points (in the form of a percentage). Reporting this percentage would keep results close to original scores and allow for a better representation of actual growth. This computation could also be done for the student population in the low 25%.

MDE's Growth Multipliers

MDE's assignment of weights for learning gains appears to be arbitrary and results in the obfuscation of data, which impedes MDE from reaching its goal of improving student achievement and increasing school and district accountability.

In addition to MDE's calculation of growth upon growth, the added multipliers that MDE allows for certain types of growth appear to be arbitrary.

As noted previously, the all-growth and the low 25% growth components of the accountability grade incorporate three different weights (1.0, 1.2, and 1.25) for the three different types of growth. (See page 18 for an explanation of when each of these weights is applied.) According to MDE, these weights were developed based on a similar weighting system implemented in Florida. Beyond this, MDE could provide PEER with no rationale for the values of the weights.

Because the accountability standards place a student's score in an achievement category at the very beginning of the process (therefore establishing an ordinal ranking), the precision of true growth is lost. As shown in Exhibit 8, page 21, a student's score might show little growth, but because that score crossed one or more points that divide one achievement category from another, that growth has more weight than another student's growth that was greater numerically but whose score crossed fewer achievement category division points.

Assigning a numerical multiplier to ordinal data is always arbitrary because by the nature of the measurement, ordinal data establishes a rank order. For example, the advanced achievement category ranks higher than the proficient achievement category, which ranks higher than the basic achievement category, which ranks higher than the minimal achievement category. In essence, the achievement categories have been ranked; however, assigning each student to an achievement category has masked the true performance of the student. Assigning a multiplier to the ordinal ranks to reflect the advancement of ranks further clouds true performance.

MDE must be able to depict accurately a school's or district's performance if effective changes are to occur within Mississippi's education system. While it might seem beneficial to provide incentives for schools and districts to encourage them to reach a higher level of achievement, if those incentives obfuscate data regarding actual student performance, the ultimate goal of improving student achievement and increasing school and district accountability has not been reached.

Clarity and Accuracy of the Accountability Standards' Presentation of Schools' and Districts' Performance

This chapter addresses:

- the “better of two years” and “pausing” waivers;
- how MDE determines accountability grades for six-component schools;
- MDE’s creation of assessment benchmarks and cut-points; and,
- changes in graduation requirements.

“Better of Two Years” and “Pausing” of Schools’ and Districts’ Grades

Although MDE developed its current accountability standards for use in the 2013-2014 assessment year, because of the implementation of college-and career-readiness standards that year, MDE has used “better of two years” or “pausing” adjustments to schools’ and districts’ accountability grades. This has allowed schools and districts to be represented by grades that do not accurately reflect current performance and makes it impossible to compare a school or district to itself or to other schools or districts over time.

As described previously in this report, MDE developed its current accountability standards to take the place of the standards that were used in the 2012-2013 assessment year. The current standards were to be implemented during the 2013-2014 assessment year.

However, according to MDE, the 2013-2014 assessment year was considered a transitional year for accountability grades because it was the first academic year that schools were expected to implement fully Mississippi’s college- and career-ready standards. Because this was a transition year with respect to assessments, the U. S. Department of Education approved a request by MDE for a one-year waiver from applying the new accountability grades to schools and districts. This waiver allowed MDE to use a “better of two years” adjustment in assigning grades to districts.

How MDE has Assigned Accountability Ratings to Schools and Districts

MDE has applied a “better of two years” adjustment to its 2013-2014 assessment data, meaning it has possibly used, for some schools and districts, 2012-2013 assessment data for the 2013-2014 assessment year. Further, MDE has requested a federal waiver, or a “pause,” from calculating 2014-2015 assessment data in order to apply the official 2013-2014 grade to the 2014-2015 assessment year.

As illustrated in Exhibit 17, page 38, the “better of two years” adjustment meant that after having calculated the actual accountability grades for each school and district, MDE could decide, for each school and district, to apply the calculated grade based on the 2013-2014 assessment results or to retain the previous year’s grade. For example, if a district’s 2012-2013 assessments yielded a grade of an A and its 2013-2014 assessments yielded a grade of a B, MDE would apply the A grade as the official grade for the 2013-2014 assessments.

Also, MDE has requested a “pause” from the U. S. Department of Education for the 2014-2015 assessment year. This “pause” means that rather than calculating actual accountability grades for each school and district for that assessment year and choosing the “better of two years,” if approved by the U. S. Department of Education, MDE may automatically apply the previous year’s accountability grade. For example, rather than calculating a grade for the 2014-2015 assessment results, MDE may be allowed to retain automatically the grade applied for 2013-2014. In the example described previously, the 2013-2014 grade could have actually been calculated as a B but under the “better of two years” scenario, an A grade could have been applied to the 2013-2014 assessments. Under the “pause,” if granted, the 2014-2015 assessment data could have actually yielded an F, but the district could be permitted to apply an A under a combination of the “better of two years” and “pausing.”

As of April 28, 2015, MDE’s request for a “pause” for the 2014-2015 assessment year was still pending with the U. S. Department of Education (USDE). According to MDE staff, the 2014-2015 assessment year is the last year that the USDE will entertain a request for a “pause.” According to the USDE, annual accountability grade calculations are to resume for the 2015-2016 assessment year.

Exhibit 17: Hypothetical Example of “Better of Two Years” and “Pausing” Adjustments

Assessment Year	Calculated Rating	“Better of Two Years” Rating*	“Pause” Rating*
2012-2013	A	--	--
2013-2014	B	A	--
2014-2015	Not calculated	--	A

* MDE publishes both the calculated grade (where applicable) and the “Better of Two Years” or “Pause” grade to its website; however, the official grade used by the districts is the “Better of Two” or “Pause” grade.

SOURCE: PEER analysis.

The Effect of “Better of Two Years” and “Pausing” on Accountability Ratings

The use of “better of two years” and “pausing” on accountability grades obscures the performance of students on the assessments, therefore preventing MDE from making accurate comparisons among schools or districts to each other or to themselves over time. Further, accountability grades could reflect the accountability standards as they were calculated in a previous year rather than as they should be calculated in the current year.

The use of “better of two years” and/or “pausing” results in the over-arching problem that the accountability grades do not reflect the performance of the students on the assessments. This prevents MDE from making accurate comparisons among schools and districts in a given year and from accurately making comparisons of a school or district against itself over time. As it stands, by using the “better of two years” and “pausing,” the only option is for schools and districts to (appear to) improve.

As illustrated in Exhibit 17, the school and district grades that MDE has reported for the 2013-2014 assessment year and the grades it could potentially report for the 2014-2015 assessment year (pending approval) could actually reflect the grades from the 2012-2013 assessment results. Calculating accountability grades in this way obscures the fact that a school or district could have performed poorly in the current year.

These grades could also reflect the accountability standards as they were calculated in the 2012-2013 assessment year rather than the accountability standards as they are currently

calculated. These scenarios are possible because MDE could use the “better of two years” for the 2013-2014 assessment year, could potentially use a “pause” for the 2014-2015 assessment year (pending approval), and therefore could possibly apply the 2012-2013 assessment results for up to three consecutive years.

Because of this, in some cases the public might not have actually seen the accountability results as MDE has purported they are calculated. It is possible that for some schools and districts, the public has actually seen the accountability results as they were calculated for the 2012-2013 school year. The purpose of accountability standards, described on page 3, outlines the type of information and capabilities the assessment system will allow Mississippi to obtain and do. However, the “better of two years” and “pausing” practices undermine the purposes described in state statute.

How MDE Determines Accountability Grades for Six-Component Schools

Rather than determining cut-points for the accountability grades for schools without a twelfth grade and without a science assessment (i. e., six-component schools), MDE determines these accountability grades based on the actual distribution of grades for seven-component schools. This results in six-component schools reflecting the performance and growth of seven-component schools rather than their own performance and growth.

The staff at MDE, the Technical Review Committee, the Accountability Task Force, and RIA established A-F cut-points for the seven-component schools and districts and the nine-component schools and districts. As previously mentioned, seven-component schools are schools without a twelfth grade but with science assessments (for example, a school composed of only 5th and 6th grades) and seven-component districts are districts without a high school but with science assessments. Nine-component schools are schools with a twelfth grade and with science assessments and nine-component districts are districts with a high school and with science assessments.

In addition to the seven-component schools and districts and the nine-component schools and districts, there are also schools that have only six components, which are schools without a twelfth grade and without science assessments. Examples of six-component schools would be schools composed of only 3rd and 4th grades or only 6th and 7th grades. Six-component districts do not exist because all districts have science assessments.

As illustrated in Exhibit 18 on page 40, rather than determine the A-F cut-points for the six-component schools in the same manner as the seven-component and nine-component cut-points were determined, MDE takes the A-F distribution of the actual grades of schools that have seven components and applies that distribution to the six component schools. MDE calls this “linking” or “equating.” For example, as shown in

Exhibit 18: Illustration of How MDE Determines Grades for Six-Component Schools Using its Accountability Standards, as of 2013-2014 Assessment Year

Grade	Year 1		Year 2		Year 3	
	Percent of 7 component schools - ACTUAL	Percent of 6 component schools - LINKED	Percent of 7 component schools - ACTUAL	Percent of 6 component schools - LINKED	Percent of 7 component schools - ACTUAL	Percent of 6 component schools - LINKED
A	15%	15%	10%	10%	5%	5%
B	30%	30%	25%	25%	20%	20%
C	45%	45%	40%	40%	35%	35%
D	5%	5%	15%	15%	20%	20%
F	5%	5%	10%	10%	20%	20%

Step 1: determine the actual letter grade for each 7-component school based on the 7-component school cut-points

Step 2: determine the percentage of 7-component schools with each letter grade

Step 3: apply, or “link,” the 7-component distribution (percentages) of letter grades to the 6-component schools

Step 4: do this for each year

SOURCE: MDE, PEER analysis.

Exhibit 18, suppose that in Year 1 the actual distribution of A-F grades for seven-component schools was such that 15% of the schools made an A, 30% made a B, 45% made a C, 5% made a D, and 5% made an F. MDE then applies, or “links,” that distribution (i. e., the percentages for each A-F grade) to the six-component schools. This method forces the six-component schools into the seven-component distribution.

Because the six-component schools’ accountability grades rely on the distribution of the seven-component schools’ grades each year, as the seven-component distribution changes, so does the six-component distribution. In other words, using this method results in the six-component schools reflecting the performance and growth of the seven-component schools rather than their own performance and growth.

Further, if a seven-component school appeals its grade, it may change the A-F grade distribution of the seven-component schools. As a result of the “linking” or “equating” process, the new seven-component school distribution will be applied to the six-component schools, potentially changing one or more six-component schools’ grade(s) even though nothing warranted a change in grade at the six-component school(s). The staff at MDE is aware of this problem and according to MDE, in May 2015 the Board of Education approved a rule that would fix this problem.

The Creation of Assessment Benchmarks and Cut-Points

MDE’s current process for determining accountability grades is not being driven by student performance; rather, a Mississippi teachers’ group determines the benchmarks for student performance. MDE, the task force, and the Technical Review Committee, with the help of a consultant from Research in Action, determine the cut-points for establishing the accountability grades each year, maintaining significant control over the outcome of accountability grades.

According to the *Mississippi Curriculum Test, Second Edition, Technical Manual for 2013-2014*, MDE and Pearson each have a role in determining the questions that become part of the core questions for the MCT2. The process is as follows:

1. Pearson identifies item writers, who are then trained by Pearson and develop potential items (i. e., test questions) for field testing.
2. Pearson and the item writers determine the initial field questions for potential use on the MCT2.
3. MDE and Mississippi educators--a group of MDE personnel and teachers--approve or disapprove the potential questions.
4. Pearson implements MDE’s approved edits and the field test questions are included on the next MCT2 test.

5. Pearson provides results of all statistics and analysis on field-tested questions, although the field test questions do not count toward a student's score.
6. Pearson and MDE then meet to review actual student performance on the recently field-tested questions. At this point, MDE and Pearson determine which questions are accepted and which questions are rejected as future core questions. The questions that are accepted are available for use during the immediately upcoming test administration.

Determining the Benchmarks and Cut-Points

Regarding assessment benchmarks, a group consisting of teachers from throughout the state of Mississippi has determined the assessment benchmarks, while MDE facilitates the process. Regarding A-F cut-points, MDE, the task force, and the Technical Review Committee, with the help of a consultant from Research in Action, Inc., established the current cut-points.

According to MDE staff, a Mississippi teachers' group determines the benchmarks on the assessment tests that illustrate "what proficiency looks like." According to the *Mississippi Subject Area Testing Program, Second Edition (SATP2) Technical Manual for 2013-2014*, MDE selects sixteen Mississippi teachers from a list of nominees made by the superintendents. Once this group of teachers has determined the benchmarks for the achievement categories for particular assessments, a policy review team of fifteen Mississippi teachers reviews the standards and reaches a conclusion to submit to the State Board of Education for approval.

As discussed on page 15, MDE determines the A-F cut-points assigned to schools/districts.

Effect of MDE's Processes to Determine Benchmarks and Cut-Points on Accountability Grades

The processes used to determine achievement category benchmarks, A-F cut-points, and the number of possible points for each accountability component are subjective rather than criterion-based. Moreover, the placement of benchmarks and cut-points can affect the magnitude of trends, possibly giving some schools and districts an advantage in their accountability grades.

The process that the Mississippi teachers' group uses to create the benchmarks (i. e., what scores make a student "proficient") and the process MDE uses to create the A-F cut-points (determining schools' and districts' accountability grades) appear to be subjective rather than criterion-based. MDE should ensure that when a student, school, or district is identified as "proficient" that there is a defensible criterion for being proficient or a clear explanation as to why a student has developed a mastery of the material.

Also, regarding the components of the accountability standards, MDE and the Accountability Task Force determine the number of possible points for each of the components, extending further the subjectivity of the accountability grades.

As noted previously, when MDE determined the A-F cut-points, it utilized two years' worth of student assessment data as its starting point. The effect of this is that the A-F cut-points applied to that data set are normative, meaning they are determined based on the rankings of that specific population. Those normative cut-points are then applied to future data sets obtained from different populations.

The placement of achievement category benchmarks and A-F cut-points can affect the magnitude of trends. For example, as explained previously on pages 27-28, two students' assessment scores can technically make identical growth gains, but depending on where the achievement category benchmark lies, one student's growth may cross over the mark while the other's does not. In this scenario, one school or district could appear to have greater growth than another even though actual growth gains may be identical. This, in turn, would give that school or district an advantage in its overall composite score and therefore an advantage in its A-F accountability grade.

A system that adjusts scores for one growth gain and not another with identical gains cannot provide an accurate picture of actual school or district performance. As a result, as the standards are currently written, MDE will not be able to satisfy the overall purpose of the accountability system or to communicate true school and district performance to education stakeholders.

Accountability Task Force Power

The Accountability Task Force has the ability to recommend changes to the weights of the components that determine a school's or district's grade based on how those changes affect impact data.

As noted previously, the Accountability Task Force, composed of superintendents, principals, and teachers from across the state, has the ability to make recommendations to the Commission on School Accreditation. (See Appendix A, pages 49-50, for the composition of the task force and the commission). If the commission approves these recommendations, they will be presented to the Board of Education. Once approved by the board, the changes will be enacted across schools and districts.

PEER observed that certain recommendations that were considered for commission approval were recommended based on impact data that was provided to the task force. For instance, as described on page 13, the components that determine a school's or district's grade each carry a certain amount of weight. For the 900-point system, reading and math, subjects that are tested more frequently, receive more weight

(100 points each) than science or history (subjects that are not tested as often). To determine how much weight each component should receive, MDE provided the task force with data that showed how different component weights would affect Mississippi schools' and districts' grades. For example, the task force wanted to see how many schools would receive a letter grade of A, B, C, D, or F if the U. S. history component counted 75 points instead of 50 points. Based on PEER's observation, such a discussion regarding point weights seemed to be based on concerns about how the changes might affect the accountability grade that a school or district might receive rather than concerns about actual performance.

This type of decisionmaking does not align with the overall purpose of the accountability standards. The goal is not to make schools and districts "look good," but to provide an accurate picture of how schools and districts are performing. MDE should ensure that task force recommendations support the purpose of the accountability standards so that appropriate changes, when necessary, can be made.

Changes in Graduation Requirements

In March 2015, the State Board of Education voted to approve additional options that will allow more students to graduate. This will allow schools and districts to have better graduation rates, which count 20% of the overall accountability grade.

In 2013, the federal government required that the graduation component had to account for twenty percent of a school's or district's accountability grade. As noted previously, a school's or district's graduation rate is the number of students who graduate in four years with a regular high school diploma divided by the number of students who entered four years earlier as first-time ninth graders.

According to MDE, prior to January 2014, a student could not graduate high school in Mississippi unless he or she passed each subject area test (English II, Algebra I, U. S. History, and Biology I). However, the Mississippi Board of Education voted in March 2015 to allow students to graduate if they fail one or more of their subject area tests but meet certain other requirements. The perception is that MDE has made graduation more easily attainable, thus allowing schools and districts to have better graduation rates (which count 20% of the accountability grade). See Appendix D, page 63, for graduation options approved in March 2015.

PEER believes that changes such as this should be criterion-driven and based on logic and rationale. In changing the graduation requirements, MDE relied on the fact that other states do not require students to pass subject area tests as the reason for changing its graduation requirements. PEER questions whether this decision was made in the best academic interest of the students or if that decision reflects the desire to improve the graduation rate of schools.

MDE should ensure that all changes made to the accountability standards, or the pieces that affect the standards, are not made for any other purpose except to “improve student achievement and increase the level of accountability for both school districts and individual schools.” Lowering graduation standards with little defense to demonstrate how the additional options for graduation are equivalent and comparable to a standard high school diploma is in direct contrast to the purpose of MDE’s accountability standards.

In order for MDE to have defensible accountability standards, those standards should accurately reflect a school’s or district’s performance for the most recent academic year. If accountability grades are obfuscated or if they do not reflect current performance, the standards fail to fulfill their purpose and results cannot be used effectively to make any statewide policy, school, or district level decisions.

Recommendations

1. In order for a school's or district's student proficiency to be represented accurately by its accountability grade, MDE should report performance grades that reflect student assessment score data as closely as possible. This could be done by:

- eliminating the use of the four achievement categories (minimal, basic, proficient, and advanced); or,
- reporting scale scores divided by total possible scale score points (in the form of a percentage).

To accomplish this, the Legislature should amend MISS. CODE ANN. Section 37-17-6 (5) (c) (i) (1972).

(Note: When *proficiency* is referenced in other recommendations in this report, it is with the assumption that an accurate proficiency measure will be utilized.)

2. In order to communicate and report student proficiency and student growth accurately and to prevent either proficiency or growth from greatly affecting a school's or district's accountability grade, MDE should separate proficiency and growth into two separate grades.

MDE could do so by assigning a letter grade (A thru F) for proficiency, followed by another indicator to represent growth. The department could use a letter grade to demonstrate proficiency and an arrow that indicates direction to reflect whether a school has made adequate learning gains. For example, a school that made learning gains and earned a B in proficiency would have a grade of B↑. However, a school that earned a B in proficiency, but did not make adequate learning gains, would have a grade of B↓.⁵

For the separation of scores to take place, the Legislature would need to amend MISS. CODE ANN. Section 37-17-6 (4) (g) (1972) to allow for separate proficiency and growth indicators.

3. To ensure that a school's or district's growth is represented accurately in its accountability grade, MDE should indicate growth by a student's improvement from one year to the next in the accurate proficiency grade.

⁵The growth component is not a measure of performance; it seeks to communicate where a school or district stands relative to current performance. An A school or district that earns an A in proficiency would not have much (if any) room for growth and would not necessarily have an arrow indicator following the school's or district's grade.

MDE uses growth multipliers of 1, 1.2, or 1.25 to indicate greater growth, but any multiplier or incentive that alters an original score takes a rating farther away from accurately demonstrating true performance.

4. To ensure that a school's or district's grade for a given year is a direct representation of that school's or district's performance for that year, MDE should instruct schools and districts to report and publicize not only their official grade, but also their "paused" or "waived" grades in any school year that is considered a transitional year. Allowing schools and districts the opportunity to publicize the better grade of two years, or an outdated school grade, does not provide a clear picture of current performance.

Further, to ensure that schools' and districts' grades can be reliably compared to those of other schools or districts for that year and that a single school or district can analyze its performance over a period of time, MDE should report schools' and districts' grades using the same accountability standards (as opposed to a previous year's standards or a previous year's grades).

5. To ensure that the A through F cut-points and assessment benchmarks are directly related to student mastery over material, MDE should develop a defensible criterion for being "proficient."
6. To ensure that the accountability standards accomplish what they are designed to accomplish, MDE should ensure that task force recommendations support the purpose of the accountability standards so that appropriate changes, where necessary, can be made.
7. In the best interest of the students and to acknowledge the distinct honor of successfully completing high school, MDE should develop a method to ensure that the changes made to the graduation options are equivalent and comparable to a standard/regular high school diploma.
8. The Legislature should enact legislation requiring that the Mississippi Department of Education submit any proposed changes to the school accountability standards to the appropriations and education committees of the House and Senate and to the Executive Director of the Legislative Budget Office one year before those standards would become effective. Such submission should also include a statement of estimated economic impact detailing how the proposed changes could impact the development of recommendations for the funding of the adequate education program. This is important because school districts' accountability grades are figured into the MAEP formula⁶ and any changes in the

⁶Components of the MAEP funding formula process are defined in MISS. CODE ANN. Section 37-151-5

way that a “successful” district (currently, a district receiving a C accountability grade) is defined will affect the calculation of the MAEP funding formula and thereby affect the amount of funding requested by MDE and ultimately the amount of funding received by school districts.

(1972). Currently, districts receiving a grade of C are classified as “successful” and if other statistical requirements are met, their expenditures form the base of the MAEP funding formula. Expenditures from districts receiving higher grades (A or B) or lower grades (D or F) impact the statistical calculations used in the MAEP formula, but expenditures from these districts do not otherwise impact the funding formula. The MAEP funding formula is calculated every four years, with adjustments for inflation during the intervening years. The most recent recalculation was for FY 2015. (A full recalculation of the MAEP funding formula will be completed for FY 2019.)

As noted above, MDE uses the MAEP formula to determine the amount of funding necessary to fund all schools at the funding level of the schools used in the formula that met the “successful” level of student performance. However, if the classification of student performance is flawed, as is illustrated in this report, the assumptions underlying the selection of schools to be used in the computation of funding are also flawed from a performance perspective.

Appendix A: Composition of Commission on School Accreditation, Accountability Task Force, Technical Review Committee, State Board of Education, and Technical Advisory Committee

Commission on School Accreditation

Congressional District	Name	Position	City	Term Expires
1 st	Lee Childress	Superintendent	Corinth	June 2017
1 st	John Paul Mistilis	Teacher	Oxford	June 2017
1 st	Eddie Prather	School Board	Tupelo	June 2016
2 nd	Mary Jean Gates	Non-Educator	Cleveland	June 2016
2 nd	Sean Brewer	Principal	Madison	June 2016
2 nd	Robert Tyner, Jr.	School Board	Clarksdale	June 2016
3 rd	Kenny Bush	Non-Educator	Philadelphia	June 2017
3 rd	Michael Miller	Non-Educator	Brandon	June 2015
3 rd	Heather Westerfield	Teacher	Pearl	June 2016
4 th	Chuck Blackwell	Non-Educator	Ellisville	June 2017
4 th	Ann Jones	Non-Educator	Jackson	June 2015
4 th	L. C. Tennin, Jr.	Non-Educator	Jackson	June 2016
5 th	Henry Arledge	Superintendent	Gulfport	June 2017
5 th	Michael Lindsey	Principal	Gulfport	June 2015
5 th	Anthony Montgomery	Non-Educator	Gulfport	June 2017

SOURCE: MDE.

Original Accountability Task Force Members (2012-2013)

Name	Position
Lee Childress	Superintendent/Commission on School Accreditation Representative
Dennis Dupree	Superintendent
Roy Gill	Director of Curriculum and Instruction/Accountability
Kim Hubbard	Teacher, Grade 4
Rebecca Ladner	Superintendent
Richard Morrison	Assistant Superintendent/State Board of Education Representative
Therese Palmertree	Superintendent
Eddie Peasant	Principal
Adam Pugh	Superintendent

SOURCE: MDE.

2014-2015 Accountability Task Force

Name	Role (District)
Amy Carter	Assistant Superintendent (Meridian)
Brian Marshall	Principal (Rankin)
Christopher Williams	Assistant Superintendent (Ocean Springs)
Darren Edwards	Superintendent (West Tallahatchie)
Jason Sargent	Accountability Director (JPS)
Jay Foster	Superintendent (Senatobia)
JoAnn Malone	Superintendent (MSB)
Manika Kemp	Principal (Clarksdale)
Richard Morrison	Assistant Superintendent (Rankin) [State Board of Education Representative]
Robert Lamkin	Principal (McComb High)
Ryan Kuykendall	Assessment Director (DeSoto)
Sean Brewer	Principal (Madison County) [Commission Rep]
Todd English	Superintendent (Booneville)
Tracee Thompson	Teacher (JPS, 2013 Milken Award Winner)
Vikki Landry	Federal Programs Director (Bay-Waveland)
Warren Nance	Accountability Specialist (Harrison)
Legislative Chairs	
Grey Tollison	Chair, Senate Education Committee
John Moore	Chair, House Education Committee

SOURCE: MDE.

Technical Review Committee

Name	District/Agency/Company
J.P. Beaudoin	Research in Action
William Buchanan	Harvard Strategic Data Fellow/MDE
Lee Childress	Corinth (Commission on School Accreditation/Task Force Representative)
Irene Dearman	University of Southern Mississippi
Roy Gill	Harrison County (Task Force Representative)
Christy Hovanetz	Foundation for Excellence in Education
Richard Morrison	Rankin County (State Board of Education/Task Force Representative)
Sharon Schattgen	University of Missouri
Jennifer Weeks	DeSoto County

SOURCE: MDE.

State Board of Education

Name	Position	City	Term Expires
John R. Kelly	Chair	Gulfport	July 2020
Richard Morrison	Vice-Chair	Brandon	July 2023
O. Wayne Gann	At-Large Representative	Corinth	July 2015
Kami Bumgarner	Teacher Representative	Madison	July 2018
Charles McClelland	At-Large Representative	Jackson	July 2019
Rosemary G. Aultman	At-Large Representative	Clinton	July 2022
Johnny Franklin	1 st Supreme Court District Representative	Bolton	July 2016
Carey M. Wright	State Superintendent of Education	Jackson	-----
Karen Elam	3 rd Supreme Court District Representative	Oxford	July 2021
William Harold Jones	At-Large Representative	Petal	July 2017

SOURCE: MDE

Technical Advisory Committee

Name	Agency/Company/University
Chad Buckendahl	Alpine Testing Solutions
Chris Domaleski	The National Center for the Improvement of Educational Assessment, Inc.
George Englehard	The University of Georgia and Emory University
Gerunda Hughes	Howard University Office of Institutional Assessment
Robert Lee	Department of Elementary and Secondary Education, Malden, MA

SOURCE: MDE.

Appendix B: Glossary of Terms and Concepts Related to MDE’s Accountability Standards

Acceleration

According to MDE’s technical manual, the acceleration component refers to the percentage of students taking and passing the assessment associated with accelerated courses such as Advanced Placement, International Baccalaureate, Advanced International Certificate of Education, or industry certification courses approved by the State Board of Education.

The acceleration component will consist of “participation” and “performance” components.

Alternate Assessments

The DLM Math, Language Arts, and the MAAESF Science high school tests are taken once in high school. The growth component is not counted for DLM and MAAESF scores for twelfth graders if it has been more than three years since the previous test. For example, if a student takes the high school test in twelfth grade and the previous test was in eighth grade, that student’s score is not counted in the growth component.

According to MDE, the DLM assessment was pilot tested in the 2013-2014 school year. Scores for that year’s DLM were not included in the calculation of a school’s grade, but were included in the 2014-2015 school year. A new alternate assessment will be administered in the 2015-2016 school year.

Back Mapping

For any elementary or middle school that does not have reading or math scores because the school does not have the required grade level, the scores from the students in the next higher grade in the tested subject within the same district will be applied to the student’s lower elementary school of origin. In order for the scores to be applied, the student must meet full academic year requirements at the lower grade, the current school, and if there is a gap in years, anywhere in the district for the years in between.

Banking

Scores of students taking Algebra I, Biology I, English II, or U. S. History end-of-course assessments in a grade below tenth grade will be “banked” for proficiency/achievement and growth calculations until the student is in the tenth grade and then

applied to (a) the student's school of origin where he or she took the assessment and (b) the student's tenth-grade school (if the student met full academic year requirements the year he or she was assessed and during the student's tenth-grade year).

College and Career Readiness

An independent, nonpartisan, nonprofit education reform organization called Achieve describes college and career readiness in this way:

From an academic perspective, college and career readiness means that a high school graduate has the knowledge and skills in English and mathematics necessary to qualify for and succeed in entry-level, credit-bearing postsecondary coursework without the need for remediation -- or put another way, a high school graduate has the English and math knowledge and skills needed to qualify for and succeed in the postsecondary job training and/or education necessary for their chosen career (i.e. community college, university, technical/vocational program, apprenticeship, or significant on-the-job training).

To be college and career ready, high school graduates must have studied a rigorous and broad curriculum, grounded in the core academic disciplines, but also consisting of other subjects that are part of a well-rounded education. Academic preparation alone is not enough to ensure postsecondary readiness but it is clear that it is an essential part of readiness for college, careers, and life in the 21st century.

Simply put, "college and career readiness" is the umbrella under which many education and workforce policies, programs and initiatives thrive. From high-quality early education and strong, foundational standards in elementary school to rigorous career and technical education programs and college completion goals, college and career readiness is the unifying agenda across the P-20 education pipeline.⁷

⁷According to the Education Commission of the States, an organization that tracks state policy trends, translates academic research, and provides unbiased advice for state leaders to learn from each other, local communities and states are trying to create a "seamless system of education" in which all levels of education--preschool through college--educate as one system instead of several. These efforts most commonly are named K-16, P-16, or P-20 systems.

Exclusion of Students with Banked Scores from the Low 25%

The scores of students who have taken the SATP2 tests before tenth grade, and whose scores are banked until tenth grade, are excluded from the process of determining the low 25% of students in a given year. According to the Mississippi Department of Education, students who make higher grades tend to take the SATP2 tests before tenth grade, and therefore those students who make higher grades have banked SATP2 scores. By excluding such banked SATP2 scores from the process of determining the low 25% of students, higher-performing students will not be included in the low 25% and therefore the low 25% has a better chance of more accurately reflecting the lower performing students.

Full Academic Year Requirements

In order for a student to meet Full Academic Year (FAY) requirements and be included in the proficiency and growth calculations, the student must have been enrolled (regardless of attendance) for at least 75% of the days from September 1 (of school year) to the first day of testing. If a student meets FAY requirements at a school other than the school where he or she is enrolled at the time of testing, the student's scores will count at the school where he or she met FAY requirements. FAY requirements will be calculated at the school level as well as at the district level. Therefore, it is possible for a student who transfers within a district to meet FAY requirements for a district and be included in the calculations for the grade assignment for the district but not be included in the calculations for a school. Scores of all students will be included in the state-level calculations regardless of FAY status.

No Child Left Behind Act of 2001

The NCLB Act of 2001 was meant to close the achievement gap with accountability, flexibility, and choice, so that no child is left behind. The act aims to bring all students up to the proficient level on state tests and to hold states and schools more accountable for results. NCLB requires all districts and schools receiving Title I funds to meet the state's adequate yearly progress goals for their total student populations and for specified demographic subgroups, including major ethnic/racial groups, economically disadvantaged students, limited English proficient students, and students with disabilities.

What are Title I Schools?

Title I schools are schools where at least 35 percent of the children in the school attendance area are from low-income

families or at least 35 percent of the enrolled students are from low-income families and are eligible to receive federal Title I funds. The proportion of low-income families is most frequently measured by the percent of students eligible to receive free and reduced-price lunch. Title I funds are to be used for programs designed to improve the academic achievement of children from low-income homes. Over half of all public schools receive funding under Title I. Schools that do not receive federal Title I funds are considered non-Title I schools. The federal government does not require states to give non-Title I schools an improvement status, although some states may choose to do so.

What is Adequate Yearly Progress?

No Child Left Behind requires states to measure “adequate yearly progress” for schools receiving Title I funds, with the goal of all students reaching the proficient level on reading/language arts and mathematics tests by the 2013-2014 school year. States must define minimum levels of improvement as measured by standardized tests chosen by the state. Adequate yearly progress targets must be set for overall achievement and for subgroups of students, including major ethnic/racial groups, economically disadvantaged students, limited English proficient students, and students with disabilities.

If a school receiving Title I funding fails to meet adequate yearly progress goals for two or more years, it is classified as a school “in need of improvement” and faces the following consequences:

School transfer options: When a Title I school fails to meet adequate yearly progress goals for **two or more** consecutive years, parents of children in a school “in need of improvement” have the choice to transfer their children to schools that are not identified as “in need of improvement” and not identified by the state as persistently dangerous schools.

Supplemental services: When a Title I school fails to meet adequate yearly progress goals for **three or more** consecutive years, students in schools “in need of improvement” are eligible for state-approved supplemental educational services, which include tutoring or other extra education services that provide academic aid to students. Parents can choose from a list of supplemental service providers generally available on state Department of Education websites.

Corrective action: When a Title I school fails to meet adequate yearly progress goals for **four consecutive years**, the district must implement at least one of the following corrective actions: replace school staff; implement new curriculum; decrease the authority of school-level administration; appoint outside experts to advise the school; extend the school year or school day; and/or restructure the internal organization of the

school. Also, the district must continue to provide school transfer options and supplemental services.

Restructuring (planning): When a Title I school fails to meet adequate yearly progress goals for **five consecutive years**, the district must prepare a plan to restructure the school. The restructuring plan must include one of the following alternative governance arrangements: reopen the school as a public charter school; replace all or most of the school staff, including the principal; enter into a contract to have an outside entity operate the school; arrange for the state to take over operation of the school; or any other major restructuring of the school's governance arrangement. In addition, the district must continue to provide school transfer options, supplemental services, and corrective actions.

Restructuring (implementation): When a Title I school fails to meet adequate yearly progress goals for **six consecutive years**, the district must implement the plan developed in the previous year to restructure the school. In addition, the school transfer options, supplemental services, and other corrective actions from the previous years continue.

NCLB requires states to align tests with state academic standards and test students on an annual basis in reading and math in grades 3 through 8 and at least once during grades 10 through 12. In addition, it requires the National Assessment of Educational Progress reading and mathematics tests to be administered to a sample of fourth and eighth graders in each state every other year in order to make cross-state comparisons. NCLB also mandates school districts to hire teachers designated as "highly qualified" to teach core academic subjects in Title I programs. Finally, states must issue annual report cards for schools and districts.

SATP2 High School Tests

The SATP2 Algebra I, Biology I, English II, and U. S. History end-of-course assessments are traditionally taken in tenth grade but could be taken earlier or later (see explanation of "banking"). Each subject area test is only taken once unless the student fails the test, at which point the student may re-take the test.

Significant Cognitive Disabilities

Students with Significant Cognitive Disabilities (SCD) have separate assessments that they must be deemed eligible to take.

In order to participate in SCD assessments, students must meet all three of the following criteria for having SCD:

- The student demonstrates significant cognitive deficits and poor adaptive skill levels (as determined by that student's

comprehensive assessment) that prevent participation in the standard academic curriculum or achievement of the academic content standards, even with accommodations and modifications.

- The student requires extensive direct instruction in both academic and functional skills in multiple settings to accomplish the application and transfer of those skills.
- The student's inability to complete the standard academic curriculum is neither the result of excessive or extended absences nor primarily the result of visual, auditory, or physical disabilities; emotional behavioral disabilities; specific learning disabilities; or social, cultural, or economic differences.

Quality of Distribution Index

Prior to the current standards, which began in the 2013-2014 school year, schools and districts were graded based on their Quality of Distribution Index (QDI) score. The QDI measures the distribution of student performance on state assessments around the cut-points for Basic, Proficient, and Advanced performance. The formula for QDI was as follows:

$$\text{QDI} = \% \text{Basic} + (2 * \% \text{Proficient}) + (3 * \% \text{Advanced})$$

SOURCE: MDE, Achieve website, USDE No Child Left Behind website.

Appendix C: Process to Determine Total Number of Points for Schools and Districts

Summary of Steps

- Break out the enrollment by grade level (broken out by traditional and SCD students) for all grade levels 3 through 12 in the school or district
- Determine the test(s) taken in each grade level for traditional and SCD students
- Determine the number of test-takers in each grade level
- Determine the proficiency scores⁸ for Years 1 and 2, the growth scores (from Year 1 to Year 2), and the graduation rate for Year 2. These total to either 7 or 9 component scores (explained below) that are then summed to get one composite score. Seven component scores are used for schools and districts without a twelfth grade. Nine component scores are used for schools and districts with a twelfth grade.
- Determine letter grade based on composite score.
- Calculate participation rate to determine whether letter grade stays or drops.

Steps to Determine Reading Proficiency Score (Possible 100 Points)

- For Year 1, for each 3rd grader in the school or district, assign him/her an ID number and determine his/her scale score on the MCT2 Reading and the reading alternate assessment (for SCD students). Repeat this process for all grade levels 4th through 8th.
- For each high school student (or student who took the SATP2 English II) in the school or district, assign him/her an ID number and determine his/her scale score on the SATP2 English II.
- For each high school SCD student, assign him/her an ID number and determine his/her scale score on the reading alternate assessment.
- Determine the benchmarks for the achievement categories (minimal, basic, proficient, advanced) for each grade level test. Each grade level and each test have different benchmarks that have already been determined by Mississippi teacher groups.
- Determine in which achievement category each student in each grade level (3rd through 8th, traditional high school, and SCD high school) falls.
- Count the total number of students (grade levels 3-8, traditional high school, and SCD high school) who took the MCT2 Reading test, the SATP2 English II test, or the reading alternate assessment who scored proficient or advanced. This is the numerator.

⁸The proficiency scores calculated based on this process must be adjusted so that no more than 1% of SCD students taking the alternate assessments score proficient or above, according to a December 2003 U. S. Department of Education regulation. To do this, replace the original alternate assessment proficiency numerator with 1% of total alternate assessment test takers (assuming that the original numerator is greater than the 1%; if the numerator is less than 1%, keep the original numerator). Then use the adjusted numerator with the other proficiency scores to determine an adjusted composite numerator. Use the original denominator.

- Determine the total number of students (grade levels 3-8, traditional high school, and SCD high school) who took the MCT2 Reading test, the SATP2 English II test, or the reading alternate assessment. This is the denominator.
- This fraction is the percentage of students who scored proficient or above in Reading. Multiply this percentage by 100 and round to the nearest one decimal place. Then multiply that number by a coefficient of one (since it is on a 100-point scale) to determine the points awarded for that component. This is the Reading Proficiency Score.

Steps to Determine Math Proficiency Score (Possible 100 Points)

- The Math Proficiency Score process is the same as the Reading Proficiency Score process except replace MCT2 Reading with MCT2 Math, English II with Algebra I, and the reading alternate assessment with the math alternate assessment.

Steps to Determine Science Proficiency Score (Possible 100 Points for Schools and Districts Without a 12th grade; Possible 50 Points for Schools and Districts With a 12th grade)

The Science Proficiency Score process is the same as the Reading and the Math Proficiency Score processes, except for the following differences:

- Replace MCT2 with MST2, use the science alternate assessment for SCD, and use the SATP2 Biology I test as the traditional high school test.
- Only grade levels 5 and 8 take the MST2 test and the science alternate assessment, so out of grade levels 3-8, there will only be grades 5 and 8.
- After multiplying the percentage by 100 and rounding to the nearest one decimal place, then multiply it by a coefficient of 0.5 for schools and districts with a 12th grade since it is on a 50-point scale. Multiply it by a coefficient of 1 for schools and districts without a 12th grade since it is on a 100-point scale. This is the Science Proficiency Score.

Steps to Determine History Proficiency Score (Possible 50 Points)

The History Proficiency Score process is the same as the Reading, Math, and Science Proficiency Score processes, except for the following differences:

- There is only one test, the SATP2 U. S. History test taken in high school. There are no tests for grade levels 3-8 nor for SCD.
- After multiplying the percentage by 100 and rounding to the nearest one decimal place, then multiply it by a coefficient of 0.5 since it is on a 50-point scale. This is the History Proficiency Score.

Steps to Determine Reading All-Growth Score (Possible 100 Points)

- For Year 2, repeat the “Steps to Determine Reading Proficiency Score” up through determining in which achievement category each student in each grade level falls (Year 1 has already been calculated; see above).
- Only use Year 1 students that have valid scores in the Reading tests for both Year 1 and Year 2 (so they can be compared) and who are FAY in Year 2 in the school or district.

- Year 1 students in grade level 3 are going to be compared to Year 2 students in grade level 4, etc.
- Determine the conditions under which a student would make a “learning gain” and that condition’s corresponding coefficient (1, 1.2, or 1.25).
- For the minimal and basic achievement categories for both Year 1 and Year 2, determine what the mid-point⁹ is on those scale scores and then further divide those two achievement categories (minimal and basic) into four sub-categories (bottom half minimal, top half minimal, bottom half basic, top half basic) and assign each minimal or basic student to a sub-category for that same year.
- Comparing Year 1 grade level 3 students (MCT2 Language Arts, alternate assessment Language Arts) to Year 2 grade level 4 students (MCT2 Language Arts, alternate assessment Language Arts), determine which students made learning gains and which coefficient should be applied to each student. Do this for each grade level, comparing the following:
 - Year 1 grade level 4 (MCT2, alternate assessment) to Year 2 grade level 5 (MCT2, alternate assessment)
 - Year 1 grade level 5 (MCT2, alternate assessment) to Year 2 grade level 6 (MCT2, alternate assessment)
 - Year 1 grade level 6 (MCT2, alternate assessment) to Year 2 grade level 7 (MCT2, alternate assessment)
 - Year 1 grade level 7 (MCT2, alternate assessment) to Year 2 grade level 8 (MCT2, alternate assessment)
 - Year 0 grade level 8 (MCT2) to Year 2 high school (SATP2 English II)
 - Year (whenever grade level 8 alternate assessment was taken) to Year 2 high school alternate assessment (unless it has been more than 3 years between the two tests)
- Apply the appropriate coefficient to each student. In cases in which a student qualifies in more than one “learning gain” category, apply only the highest coefficient.
- Add the number of total Year 2 students who made learning gains (with their weighted coefficients already applied) in grade levels 4-8, traditional high school, and SCD high school. This is the numerator.
- Add the total number of students in Year 2 in grade levels 4-8, traditional high school, and SCD high school who took the test. This is the denominator.
- This fraction is the percentage of students who made “learning gains in reading.” Multiply this percentage by 100 and round to the nearest one decimal place. Then multiply that number by a coefficient of one (since it is on a 100-point scale) to determine the points awarded for that component. This is the Reading All-Growth Score.

⁹The lowest two performance/proficiency levels will be split into half at the mid-point of the range. In the event that the range is an odd number and cannot be split into two equal halves, the lower half of the performance/proficiency level will be one point larger than the upper half. (Example: If the range of the performance/proficiency level is thirteen scale score points, the bottom half of the range will be seven scale score points and the upper half of the range will be six scale score points.)

Steps to Determine Math All-Growth Score (Possible 100 Points)

The Math All-Growth Score process is the same as the Reading All-Growth Score process except replace MCT2 Reading with MCT2 Math, English II with Algebra I, and the reading alternate assessment with the math alternate assessment.

Steps to Determine Reading Low 25% Growth Score (Possible 100 Points)

- Do not include SCD students (alternate assessment) in these calculations.¹⁰
- For each grade level (3rd through 8th) for Year 1, rank (from highest score to lowest score) the Reading scale scores. These students need to have full academic year in the district for Year 1 and for Year 2 and to have taken the Reading test(s) for both years so that they can be compared.
- For each grade level (3rd through 8th) take the total number of test takers and divide that number by 4. If the number is not a whole number, round up to the nearest whole number.
- Using that whole number, count from the bottom (worst scores) up the list of students. Determine the “boundary score” for the resulting student. Include any student with that boundary score or worse in the “Low 25%” category.
- Using the calculations that were already made for Reading All-Growth for Year 2, determine if Year 1’s Low 25% made learning gains in Year 2.
- Add the total number of (Year 1 Low 25% cohort) Year 2 students who made learning gains (with their weighted coefficients already applied) in grade levels 4-8 and high school. This is the numerator.
- Add the total number of (Year 1 Low 25% cohort) students in Year 2 in grade levels 4-8 and high school who took the test. This is the denominator.
- This fraction is the percentage of the low 25% students who made “learning gains in reading” from Year 1 to Year 2. Multiply this percentage by 100 and round to the nearest one decimal place. Then multiply that number by a coefficient of one (since it is on a 100-point scale) to determine the points awarded for that component. This is the Reading Low 25%-Growth Score.

Steps to Determine Math Low 25%-Growth Score (Possible 100 Points)

The Math Low 25%-Growth Score process is the same as the Reading Low 25%-Growth Score process except replace Reading data with Math data.

Steps to Determine Graduation Rate (Possible 200 Points)

- The number of students who graduate in four years with a regular high-school diploma is the numerator.

¹⁰According to MDE, the DLM assessment was pilot tested in the 2013-2014 school year. Scores for that year’s DLM were not included in the calculation of a school’s grade, but were included in the 2014-2015 school year. A new alternate assessment will be administered in the 2015-2016 school year.

- The number of students who entered four years earlier as first-time 9th graders, with adjustments for deaths, emigration, and transfers in and out, is the denominator.
- This fraction is the graduation rate.

SOURCE: PEER analysis.

Appendix D: MDE's Graduation Options

In the event that a student fails one or more subject area tests, that student can still be eligible for graduation if he or she meets certain requirements outlined by MDE. An excerpt from MDE's other graduation options is quoted below:

*While it is possible that a student will meet one of the options below before taking the subject area test, this policy states that a student is eligible to use any of these options once he or she has failed to pass any required end-of-course subject area test **one time**. Specifically, students may meet the graduation requirement outlined in State Board Policy 3803 by attaining any one of the measures outlined below for each of the subject area tests listed.*

Algebra I:

*Obtain a score of **17** or higher on the Math subject subscore of the ACT.*

Earn a C or higher in an entry level, credit-bearing dual enrollment/dual credit/college credit course with a MAT prefix.

*Obtain an ASVAB AFQT score of 36 **plus one** of the following:*

Earn a CPAS (Career Planning and Assessment System) score that meets the attainment level assigned by Federal Perkins requirements.

*Earn an approved Industry Certification as specified in the Career Pathway's Assessment Blueprint and outlined in **Appendix A-5** in the current edition of the Mississippi Public School Accountability Standards.*

*Obtain the Silver Level on the ACT WorkKeys **plus one** of the following:*

Earn a CPAS (Career Planning and Assessment System) score that meets the attainment level assigned by Federal Perkins requirements.

*Earn an approved Industry Certification as specified in the Career Pathway's Assessment Blueprint and in **Appendix A-5** in the current edition of the Mississippi Public School Accountability Standards.*

Biology I

Obtain a score of **17** or higher on the Science subject subscore of the ACT.

Earn a C or higher in an entry level, credit-bearing dual enrollment/dual credit/college credit course with a BIO prefix.

Obtain an ASVAB AFQT score of 36 **plus one** of the following:

Earn a CPAS (Career Planning and Assessment System) score that meets the attainment level assigned by Federal Perkins requirements.

Earn an approved Industry Certification as specified in the Career Pathway's Assessment Blueprint and in **Appendix A-5** in the current edition of the Mississippi Public School Accountability Standards.

Obtain the Silver Level on the ACT WorkKeys **plus one** of the following:

Earn a CPAS (Career Planning and Assessment System) score that meets the attainment level assigned by Federal Perkins requirements.

Earn an approved Industry Certification as specified in the Career Pathway's Assessment Blueprint and in **Appendix A-5** in the current edition of the Mississippi Public School Accountability Standards.

English II

Obtain a score of **17** or higher on the English subject subscore of the ACT.

Earn a C or higher in an entry level, credit-bearing dual enrollment/dual credit/college credit course with an ENG prefix.

Obtain an ASVAB AFQT score of 36 **plus one** of the following:

Earn a CPAS (Career Planning and Assessment System) score that meets the attainment level assigned by Federal Perkins requirements.

Earn an approved Industry Certification as specified in the Career Pathway's Assessment Blueprint and in **Appendix A-5** in the current edition of the Mississippi Public School Accountability Standards.

Obtain the Silver Level on the ACT WorkKeys **plus one** of the following:

Earn a CPAS (Career Planning and Assessment System) score that meets the attainment level assigned by Federal Perkins requirements.

Earn an approved Industry Certification as specified in the Career Pathway's Assessment Blueprint and in **Appendix A-**

5 in the current edition of the Mississippi Public School Accountability Standards.

U. S. History

*Obtain a score of **17** or higher on the reading subject subscore of the ACT.*

Earn a C or higher in an entry level credit-bearing dual enrollment/dual credit/college credit course with a HIS prefix

*Obtain an ASVAB AFQT score of 36 **plus one** of the following:*

Earn a CPAS (Career Planning and Assessment System) score that meets the attainment level assigned by Federal Perkins requirements.

*Earn an approved Industry Certification as specified in the Career Pathway's Assessment Blueprint and in **Appendix A-5** in the current edition of the Mississippi Public School Accountability Standards.*

*Obtain the Silver Level on the ACT WorkKeys **plus one** of the following:*

Earn a CPAS (Career Planning and Assessment System) score that meets the attainment level assigned by Federal Perkins requirements.

*Earn an approved Industry Certification as specified in the Career Pathway's Assessment Blueprint and in **Appendix A-5** in the current edition of the Mississippi Public School Accountability Standards.*

SOURCE: MDE website. State Board Policy 3804, as approved January 2014 and amended March 2015.

PEER's Response to the Department of Education's Response

The Mississippi Department of Education submitted a response package to the PEER Committee's report entitled *A Review of Accountability Standards of the Mississippi Department of Education*. The response package consisted of three parts: a letter to the Executive Director; a response to key conclusions and recommendations of PEER's report; and the resumé, vitae, and publication lists of individuals involved in the design and review of MDE's accountability standards. Although PEER has developed a detailed response to MDE's response package, the Committee chose to include in this report only this general statement, which prefaces the first two portions of the response package (included in pages 69 through 77).

The PEER Committee has a rich forty-two-year history as Mississippi's non-partisan legislative oversight entity. PEER utilizes a staff with diverse educational backgrounds and experience to make disciplined inquiry into a variety of service and policy areas to provide useful, objective information for legislative decisionmaking. The project and review teams for this project represents well over a century of experience in legislative auditing and assistance. Persons familiar with the profession of legislative auditing know that such staff need not be "experts" in the area they are reviewing; rather, they need only be proficient in the skills required to conduct an objective evaluation. As noted at the beginning of the report, PEER's task was to determine, through the eyes of an independent reviewer, whether MDE's accountability standards adequately measure school performance and present a clear, accurate picture of how schools and districts are actually performing. PEER sought to assess the structural suitability of the accountability system in light of its statutory purposes. PEER conducted this review in the same manner that it conducts all such reviews--examining records, interviewing personnel, analyzing information, and working from a set of predetermined objectives to make its assessment.

In its response, MDE made multiple assertions that some of PEER's conclusions were unsupported by evidence or were "technical misinterpretations." In summary, PEER notes the following regarding the following major areas in which PEER's professional opinion differs:

- MDE expressed no qualms with how PEER presented the actual process of calculating a school's or district's accountability grade, but with PEER's analysis of what that grade actually communicates to the public, a judgment that rightly should be made by an independent party. For example, the report pointed out that MDE's practice of applying the "better of two years" or "pausing" waiver adjustments to schools' and districts' accountability grades does not accurately reflect current performance. MDE responded that it provides schools' and district's results both with and without the waiver on its website. Although true, PEER believes that reporting waiver-based scores goes against the statutory aim of the accountability system and that using them invites confusion. Also, PEER pointed out problems with combining proficiency and growth into a single accountability grade and recommended that MDE separate proficiency and growth into two separate grades. The department responded that placing more emphasis on proficiency "disadvantages schools, especially those that serve traditionally low achieving students" and did not explain why reporting two separate grades for proficiency and growth would not be workable. PEER stands by its conclusion that the way in which MDE calculates schools' and districts' grades

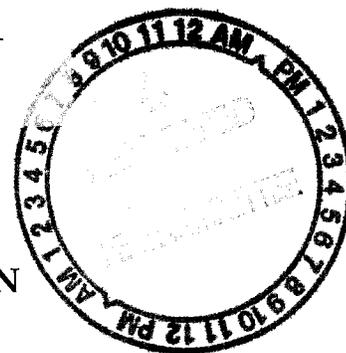
does not provide a clear, accurate picture of how schools and districts are actually performing.

- Some of the conclusions with which MDE takes issue were based directly on information provided by MDE during fieldwork, but were argued against in the department's response, or were topics on which PEER had asked for additional information during fieldwork, but such was not provided. For example, when PEER questioned MDE regarding the reasons and process behind converting a raw score to a scaled score, MDE responded that "using raw scores alone contains too much random measurement of error" and directed PEER to two technical manuals for the statewide assessments, both of which noted that "scale scores are statistical conversions of raw scores that adjust for variations in the difficulty of items on different tests and permit valid comparison across all test administrations within a particular subject area [or grade]." However, in its response package, MDE states that "scale score based indicators are prone to false-precision." Also, when PEER requested data from MDE to show that the additional options for graduation were equivalent to a regular high school diploma, MDE provided a list of other states' graduation requirements, rather than the results of actual research showing evidence that the changes in the options would render equivalent results.
- MDE continues to combine measurement with "incentivizing," a combination that PEER finds inappropriate. In its response package, the department states that certain growth metrics are weighted in order to incentivize improvement in performance. PEER believes that the accountability system is intended as a measurement, not as an incentive. (See MISS. CODE ANN. Section 37-16-1 [1972].)
- Although MDE insists to the contrary, PEER believes that the nature of the Accountability Task Force's involvement in development of the accountability system was at times subjective and inappropriate when designing a system of measurement and reporting. PEER did not assert either that the task force lacked expertise or that its approach was not supported in the literature, but that the approach was not appropriate in furthering the statutory aims of the accountability system. Regardless of how qualified MDE's subject matter experts might be, if they are making decisions regarding the accountability system based on how such changes might affect a school's or district's grade, these decisionmakers have moved away from designing a system that reports actual performance.

PEER notes that the Department of Education's staff responded in a manner typical of the staffs of agencies, institutions, and programs that the Committee reviews. They took issue with some of the report's conclusions and recommendations and believe that they have good reason to do so. During the exit process, PEER worked diligently to perfect the report, making revisions and clarifications in the draft based on a point sheet MDE developed. In conclusion, PEER stands by this report and believes that it does not contain unsupported conclusions or misinterpretations, only differences of opinion regarding the educational accountability system.



MISSISSIPPI DEPARTMENT OF EDUCATION



Carey M. Wright, Ed.D.
State Superintendent of Education

July 10, 2015

James Barber, Executive Director
Joint Legislative PEER Committee
Woolfolk Building
501 North West Street, Suite 301A
Jackson, Mississippi 39201

Dear Mr. Barber,

The Mississippi Department of Education (MDE) has reviewed the PEER Committee report "A Review of Accountability Standards of the Mississippi Department of Education." While the evaluation is well-written and organized, the evaluators' chain-of-logic and evidence does not support many of the conclusions. The report does not provide either empirical or research-based evidence to support many conclusions. Further, technical misinterpretations suggest subject-matter experts in the fields of psychometrics, accountability systems, and statistics may not have fully reviewed the findings or recommendations.

Specifically, the report outlined the following concerns and public perceptions related to the performance standards, which are part of the Mississippi Statewide Accountability System. MDE's rationale regarding the conclusions follows each of these concerns.

PEER Concern: MDE's calculation of students' scores on assessments, upon which the accountability standards are based, does not offer a clear picture of how schools and districts are actually performing, nor does it show whether schools and districts have achieved what they are supposed to have achieved.

MDE Response: Many state education accountability systems use performance categories precisely because it clearly incorporates the assessment standard into the system. Indeed, Mississippi's criterion referenced tests are designed for the explicit purpose of classifying with the highest degree of precision student performance with respect to these standards. Any assertion that scale scores should be the primary driver of accountability determinations is incomplete without consideration of the error associated with interpreting unit level changes in scores which may be within the standard error of measurement. In other words, scale score based indicators are prone to false-precision that is not recognized in the recommendations.

PEER Concern: The U. S. Department of Education's voluntary waiver that allows schools to choose the higher of two years' grades (for the 2012-2013 and 2013-2014 school years) does not make it possible to compare a school's accountability grade from one year to another or to compare the grades of several schools or districts over time.

MDE Response: MDE has provided data on its public reporting website with all the resulting calculations, with and without the "waived" result, since the State Board of Education (SBE) approved the results on October 17, 2014. The file is available on MDE's public reporting webpage, as well as through MDE's official [news release](#) regarding 2014 Accountability Results.

PEER Concern: Certain multipliers included in a school's or district's accountability grade calculation result in inflation of schools' and districts' grades.

MDE Response: The growth metric includes weighting achievement levels in order to incentivize improvement in student performance from one year to the next. The growth indicators are combined with status indicators to create an overall score. This compensatory design reduces the misclassification of schools and districts. The evaluators' recommendation would increase the misclassification of the accountability determinations, thus resulting in spurious results. The evaluators' provided no empirical data supporting their recommendation. Further, they provided neither actual nor simulated analytics demonstrating that additional conjunctive decisions would have misclassification (i.e., Kappa) indices similar the SBE's current design.

PEER Concern: Some of MDE's Accountability Task Force's decision-making practices do not align with the overall purpose of the accountability standards.

MDE Response: The agency selected and applied a standard-setting approach that has extensive research supporting the validity of the approach. The agency used panelists who were subject matter experts in the area of educational leadership and management (e.g., principals, supervisors, superintendents) to recommend cut-scores for the State Board of Education's consideration.

PEER Concern: MDE's ability to create assessment benchmarks and cut-points or the accountability grades gives the department an inappropriate amount of control over these grades.

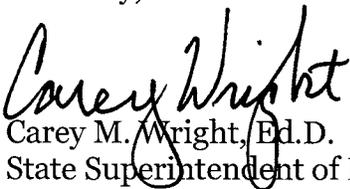
MDE Response: The literature has demonstrated the volatility of growth-based metrics. This volatility is in part ameliorated by including status indicators, using multiple data waves, and ensuring only reliable assessment results (low standard error of measure) are included. The inclusion of growth-based metrics with status-

based metrics increases the robustness of the accountability design. The addition of a growth-metric improves the goodness-of-fit of the accountability model to predict accountability ratings.

The documents that follow support the existing accountability system. Each recommendation from the PEER Committee report is addressed. The resumes for MDE's Technical Advisory Committee (TAC) and key members of the Technical Review Team provide evidence of the quality of the team vetting the system. It should be noted that these individuals have extensive experience in designing accountability and assessment systems and are recognized as national experts in their fields. I requested the resumes of those who conducted the review and only received the degrees they obtained, as shown in the attachment. I question the technical sufficiency of the PEER Review Team, which is needed to evaluate, draw conclusions, and formulate design recommendations for state assessment and accountability systems.

The performance standards of the Mississippi Statewide Accountability System are a high-quality set of indicators of school effectiveness, intended to improve student outcomes. The system, developed through a multi-year deliberate process of stakeholder feedback and research-based best practice, reflects a very different philosophical approach to school accountability than the report illustrates. The MDE has worked diligently to ensure reliability and validity in these indicators of school success, incorporating both state and federal statute into one coherent model to measure school performance.

Sincerely,


Carey M. Wright, Ed.D.
State Superintendent of Education

Attachments: MDE Response to PEER Accountability Report
Vitas/Resumes of:

- Joint Legislative PEER Committee
- MDE Technical Advisory Committee for Assessment and Accountability
- MDE Technical Review Team, External Assessment and Accountability Members

Mississippi Department of Education
Response to Recommendations from the PEER Committee report “A
Review of Accountability Standards of the Mississippi Department of
Education”

Recommendation 1. In order for a school's or district's student proficiency to be represented accurately by its accountability grade, MDE should report performance grades that reflect student assessment score data as closely as possible. This could be done by:

- eliminating the use of the four achievement categories (minimal, basic, proficient, and advanced); or,
- reporting scale scores divided by total possible scale score points (in the form of a percentage).

To accomplish this, the Legislature should amend MISS. CODE ANN. 37-17-6(5)(c)(i) (Note: When *proficiency* is referenced in other recommendations in this report, it is with the assumption that an accurate proficiency measure will be utilized.)

MDE Response: The report raises objections about the use of performance categories to signal student performance, indicating that this is less preferable to an indicator based solely on scale scores (pp. 29-30). However, it is important to note that many state education accountability systems throughout the nation use performance categories precisely because it clearly incorporates the assessment standard into the system. Indeed, Mississippi's criterion referenced tests are designed for the explicit purpose of classifying with the highest degree of precision student performance with respect to these standards. Furthermore, any assertion that scale scores should be the primary driver of accountability determinations is incomplete without consideration of the error associated with interpreting unit level changes in scores which may be within the standard error of measurement. In other words, **scale score based indicators are prone to false-precision that is not recognized in the recommendations.**

Recommendation 2. In order to communicate and report student proficiency and student growth accurately and to prevent either proficiency or growth from greatly affecting a school's or district's accountability grade, MDE should separate proficiency and growth into two separate grades.

MDE could do so by assigning a letter grade (A thru F) for proficiency, followed by another indicator to represent growth. The department could use a letter grade to demonstrate proficiency and an arrow that indicates direction to reflect whether a school has made adequate learning gains. For example, a school that made learning gains and earned a B in proficiency would have a grade of B+. However, a school that earned a B in proficiency, but did not make adequate learning gains, would have a grade of B-.

For the separation of scores to take place, the Legislature would need to amend MISS. CODE ANN. Section 37-17-6 (g) (1972) to allow for separate proficiency and growth indicators.

MDE Response: A central assertion of the report authors is that the influence of growth corrupts or clouds the accountability outcomes. For example, on page 29 the authors state, “Mississippi’s standardized tests are carefully constructed to ensure that a student has mastered a certain level of competency; those tests alone provide the criterion/ standard that should measure school performance.” Again on page 33 the authors reason, “If the purpose of the accountability standards is to improve student achievement and increase the level of accountability of schools and districts, then more emphasis should be placed on proficiency.” Taken to its natural conclusion, the application of this statement creates a system that systematically disadvantages schools, especially those that serve traditionally low achieving students. A well-designed and fair system must include multiple indicators including, especially, elements that reward student growth. Furthermore, to the extent that equity (or improving performance of at-risk students) is a priority outcome, indicators must be included that signal such progress and reward ongoing efforts. To this point, the Council of Chief State School Officers (CCSSO) document entitled *CCSSO Roadmap for Next-Generation State Accountability Principles*, which was developed by a representative group of state education leaders and experts in the field of assessment and accountability, explicitly promotes as principle 3, “Focus on student outcomes on a variety of indicators including those of both status and growth” (see:

[http://www.ccsso.org/Documents/2011/Roadmap_for_Next-](http://www.ccsso.org/Documents/2011/Roadmap_for_Next-Generation_Accountability_2011.pdf)

[Generation Accountability 2011.pdf](http://www.ccsso.org/Documents/2011/Roadmap_for_Next-Generation_Accountability_2011.pdf)). From the same seminal report, the authors identify the following option for states, “Focus particularly on and weigh more heavily the achievement of the lowest-performing students.” **In short, there is a strong rationale for and extensive practice nationally in support of design decisions represented in the Mississippi system. Conversely, there is a growing recognition nationally, that systems that rely exclusively or heavily on status test performance do not represent best practice.**

Recommendation 3. To ensure that a school's or district's growth is represented accurately in its accountability grade, MDE should indicate growth by a student's improvement from one year to the next in the accurate proficiency grade. MDE uses growth multipliers of 1, 1.2, or 1.25 to indicate greater growth, but any multiplier or incentive that alters an original score takes a rating farther away from accurately demonstrating true performance.

MDE Response: On page 34 of the report, there are multiple references to growth being counted three times in the model. Growth is only counted twice, and then ONLY if the student is in the Low subgroup. The support for the three-time argument is that a student who performs at proficient or advanced, by nature, is growing. However, that is not always the case. In fact, statewide results for school year 2013-14, of the students meeting proficiency, approximately 80% of students met both proficiency and growth, 11% met proficiency but not growth, and 9% met proficiency and exceeded growth.

On page 37, in the last sentence of the last paragraph, the following statement is made: “if those incentives obfuscate data regarding actual student performance, the ultimate goal of improving student achievement and increasing school and district accountability

has not been reached.” Reviewing statewide results from school year 2013-14 reveals that, if the weights are removed from the growth calculations, approximately one point makes the difference in weighted and unweighted values. One point of increase is more incentive than obfuscation. The report also asserts that “MDE could provide no rationale for the values of the weights”; however, incentivizing growth is the rationale, as noted in the report.

On page 36 the authors suggest two approaches for “presenting growth.” The first is to report proficiency rates annually and focus on changes in rates. Simply put, this is not growth. This refers to a trend analysis which is not based on matched cohorts of the same students. Achievement differences over time using such a method cannot be disentangled from annual changes in school enrollment. The authors further suggest an approach whereby the scale score is reported as a percentage of the total possible scale score. This is not technically appropriate or defensible. Such a practice completely departs from the measurement model used for the tests—Item Response Theory (IRT)—and would substantially distort precision and meaningful score interpretation.

Recommendation 4. To ensure that a school's or district's grade for a given year is a direct representation of that school's or district's performance for that year, MDE should instruct schools and districts to report and publicize not only their official grade, but also their “paused” or “waived” grades in any school year that is considered a transitional year. Allowing schools and district the opportunity to publicize the better grade of two years, or an outdated school grade, does not provide a clear picture of current performance.

Further, to ensure that schools' and districts' grades can be reliably compared to those of other schools or districts for that year and that a single school or district can analyze its performance over a period of time, MDE should report schools' and districts' grades using the same accountability standards (as opposed to a previous year's standards or a previous year's grades).

MDE Response: The report raises concerns about being able to compare scores from year to year (pp. 39-40). However, MDE has released to the public all the SY 13-14 calculations, with and without the “pause” or “waived” grade. Further, MDE intends to run the SY14-15 model, as the values are required by the U.S. Department of Education for federal accountability purposes. The first sentence of the last full paragraph on page 40 indicates that “the public might not have actually seen the accountability results as MDE has purported they are calculated.” MDE has provided data on its public reporting website with all the resulting calculations, with and without the “waived” result, since the State Board of Education approved the results on October 17, 2014. The file is available on MDE public reporting webpage, as well as through MDE’s official [news release](#) regarding 2014 Accountability Results.

Recommendation 5. To ensure that the A through F cut-points and assessment benchmarks are directly related to student mastery over material, MDE should develop a defensible criterion for being “proficient.”

MDE Response: The method for setting standards for both the state tests and the state accountability system seem to be questioned as too ‘subjective’ rather than ‘criterion-based’ (p.45). The process for setting standards for the state tests follows established, research-based protocols used broadly for state achievement tests. The Bookmarking (Lewis, Mitzel, & Green, 1996) and the Contrasting Group (Livingston & Zieky, 1982) methods used to establish the academic achievement standards (AAS) are internationally recognized approaches. These methods were used in a majority of statewide assessment systems reviewed by the United States Department of Education’s Peer Review process in meeting the requirements of 34 C.F.R. 200. “The performance standards provide qualitative descriptions of the intended distinctions between adjacent levels of performance. The cut scores are points on the score scale, with one cut score associated with each performance standard (Cizek, 2001, pg. 55). This process relies on both judgments and data. It is precisely because of this process, that a criterion-based claim can be supported for the assessments. Moreover, utilizing a broad-based task force to make judgments based in part on data, as is the practice in Mississippi, to reach decisions about the accountability system standards is a common and acceptable practice nationally to identify performance criteria. If the authors have another proposal for establishing the key performance criteria for either the tests or the accountability system, it was not apparent to the reader. The process evidence demonstrates the technical quality of the standard-setting methodology and the nationally recognized experts associated with implementing the procedures. Claims asserting the “defendable criterion” are neither supported by psychometric experts or empirical evidence.

Recommendation 6. To ensure that the accountability standards accomplish what they are designed to accomplish, MDE should ensure that task force recommendations support the purpose of the accountability standards so that appropriate changes, where necessary, can be made.

MDE Response: The practice of engaging stakeholders in the decision-making process is one that is widely used across various disciplines. Allowing a diverse group of current practitioners to have a voice in the development of the accountability system helps to strengthen its efficiency, reduce risks, and increases transparency and understanding around its design. The role of the Accountability Task Force (ATF) is to advise MDE, State Commission on School Accreditation, and State Board of Education (SBE) on the business rules and practices that improve the implementation of the accountability system and its measure of district and school performance. The ATF reviews and discusses thoroughly the purpose of the accountability standards as well as data, where necessary, to ensure the decisions are equitable for all districts and schools across the state. As noted on page 46 of the report, all of the recommendations generated through the work of the ATF are vetted through the Technical Review Team and Technical Advisory Committee to determine the soundness of the decisions being brought forward as well as to ensure the decisions are statistically/technically defensible. In evidence of the quality of the vetting process, the resumes for the Technical Advisory Committee, as well as for the members of the Technical Review Team with expertise in national educational accountability research, are provided at the end of this response. Prior to the approval of these recommendations by the SBE, the general public are afforded an opportunity to provide feedback to the SBE on concerns of the proposed

recommendations through the Administrative Procedures Act (APA) process. This feedback are considered and changes are made as necessary prior to the final ruling of the SBE.

Recommendation 7. In the best interest of the students and to acknowledge the distinct honor of successfully completing high school, MDE should develop a method to ensure that the changes made to the graduation options are equivalent and comparable to a standard/regular high school diploma.

MDE Response: The agency applied weighting to both academic achievement results and graduation rates consistent with SBE's policy. The emphasis on students graduating on time with a regular diploma is enumerated in the weighting. The emphasis on students graduating from high school and either entering the workforce or attending college is supported by the research. An extensive body of research exists showing that lifetime earnings of high school graduates are significantly higher than for those that dropout of schools (Card, 1999; Cheeseman, & Newburger, E., 2014). Further, the research also shows strong evidence that high school dropouts are incarcerated (Anderson, 1999), unemployed, and use government support services (Grossman, 2006) at rates much higher than graduates. The graduation options that result in the issuance of a standard, high school diploma use multiple indicators (i.e., course credit earned, standardized assessment results). Other states' policies allow for additional performance evidence to be considered, such as remediation opportunities, attendance, grade-point-average, portfolios, electives, or occupational readiness examinations, as well as other nationally recognized assessments (see Education Commission of the States, Indiana Department of Education). These nationwide changes reflect the divergence of the K-12 systems in preparing students to be ready to enter the workforce or to move to post-secondary education. The high-stakes "gate keeper" test policies of the late 1990s as the only indicator of post-secondary readiness have been determined to be fraught with negative consequences (e.g., disadvantaging students with disabilities and English language learners). Further, the MDE ensures that changes made to graduation options are equivalent and comparable to a standard/regular high school diploma through the input of our Technical Advisory Committee (resumes attached) and the use of national research in the field, as well as the thorough vetting of decisions through the Commission on School Accreditation, the State Board of Education, and public comments garnered through Administrative Procedures Act (APA) process.

Recommendation 8. The Legislature should enact legislation requiring that the Mississippi Department of Education submit any proposed changes to the school accountability standards one year before those standards may become effective. Such submission should also include a statement of estimated economic impact detailing how the proposed changes could impact the development of recommendations for the funding of the adequate education program. This is important because school districts' accountability grades are figured into the MAEP formula and any changes in the way that a "successful" district (currently, a district receiving a C accountability grade) is defined will affect the calculation of the MAEP funding formula and thereby affect the

amount of funding requested by MDE and ultimately the amount of funding received by school districts.

MDE Response: As authorized through section 203(2) of the Mississippi Constitution, the State Board of Education formulates “policies according to the law for implementation by the State Department of Education.” Following statewide procedure, the SBE finalizes policy through the Administrative Procedures Act (APA) process, as codified in MISS. CODE ANN. § 25-43-1 through 4. The legislation governing this process provides for public comment and review of any changes and potential impacts.

Overall, MDE is disappointed in the quality of the report. Based on the extensive and research-based approach used to develop the Mississippi Statewide Accountability System, MDE has tremendous confidence it accurately reflects the progress of districts and schools in a most transparent manner.

PEER Committee Staff

James A. Barber, Executive Director

Legal and Reapportionment

Ted Booth, General Counsel
Ben Collins
Barton Norfleet

Administration

Barbara Hamilton
Larry Landrum
Rosana Slawson
Gale Taylor

Quality Assurance and Reporting

Ava Welborn, Director
Tracy Bobo

Performance Evaluation

Lonnie Edgar, Team Coordinator
David Pray, Team Coordinator
Jennifer Sebren, Team Coordinator
Kim Cummins
Matthew Dry
Matthew Holmes
Angela Norwood
Michael Surratt
Jenell Ward
Jade Watters
Sarah Williamson
Julie Winkeljohn
Ray Wright

Performance Accountability

Linda Triplett, Director
Kirby Arinder
Alicia Davis
Anna Johnson
Jessica Kelly
MeriClare Steelman