

Title 7: Mississippi Department of Education
Part 138: Excellence for All (E4A)
School and District Accountability Process 2013-2014

Overview

The Mississippi Department of Education (MDE) has a need to include schools and districts participating in the *Excellence for All* (E4A) program into its accountability system. *Excellence for All* is an innovative high school model developed by the National Center on Education and the Economy (NCEE). *Excellence for All* combines the tools provided by the world's best Board Examination systems, the high school structure used in many high performing countries, and American ideas of educational equality (Sibley & Jordan, February 2014, pg. 1).

Excellence for All is based on extensive research of international student assessments and functions on the premise that students who pass lower level examinations will be ready to enroll in community college without remediation and students who pass upper level examinations will be prepared for more selective four-year institutions. Mississippi currently has two programs (see Appendix B): *Cambridge International Examinations*, based on internationally benchmarked education programs, and *ACT Quality Core*, which fully aligns high school course standards in reading, writing, speaking and listening, language, and math to the Common Core State Standards.

A technical review of the alternate assessments being used by the grantees has concluded that the assessments are measuring different constructs and/or are unable to equate scores to the SATP2. A review of internal MDE documents—"Proposed Accountability Review Process for Districts and Schools Participating in *Excellence for All* 2013-2014" (Domaleski, 2014),

Cambridge IGCSE data (June 2013), ACT Technical Manual (2007), and external materials created by Research in Action, Inc. (RIA) for other state educational agencies—was used to create the review process and procedures articulated in this document. The final outcome of the process will be a recommendation to the Mississippi State Superintendent of Education for the performance classification to improve, remain the same, or decline with respect to the school’s or district’s prior year accountability outcome. That recommendation would be one of the following:

- Performance classification improves (e.g., from B to A)
- Performance classification is unchanged (e.g., remains B)
- Performance classification declines (e.g., from B to C)

The recommendations for performance classifications will be presented to the Commission of School Accreditation and the Mississippi State Board of Education (SBE) for final approval.

Background

Grant Purpose

The purpose of the grant was to award discretionary funds to assist eligible school districts implementing the *Excellence for All* program to provide students with the maximum opportunity for success after high school. The *Excellence for All* program will also reduce dropout rates and better prepare students for a global society. Evaluation of applications was based upon the following criteria:

- Program Goals
- Description of Proposed Project:
 - Course Offerings
 - Coursework Requirements

- Testing Dates
- Project Outcomes
- Method of Evaluating Progress
- Implementation Timelines
- Evidence of Broad-Based Support (attached letters of support from the school board, PTA, public organizations, and documentation of internal communication with district staff)
- Budget Narrative

Design: Review Process

Purpose

Mississippi has a process established through Mississippi Public School Accountability Standard 3.1.3 for schools and districts to review accountability decisions. Current policy provides schools and districts the right to present clear and convincing evidence that they were assigned an incorrect accountability rating and thus misclassified. Using preliminary data released to schools and districts, these schools and districts can file a request for review within a specified period of time before results are promulgated to the general public. These procedures are being used as the foundation needed to examine data about E4A grantees and include these schools and districts into Mississippi Statewide Accountability System.

Information and Data Requirements

E4A schools and districts will need to submit evidence and procedural information that demonstrate (a) the quality of the alternative assessments, (b) the integrity of the alternative performance indicators, and (c) program goals and objectives. Specifically, evidence must address the technical quality of the assessments used by the E4A schools and districts in lieu of

the statewide assessments. The procedures and associated metrics from the aforementioned assessments must detail how student achievement data was used to create accountability indicator proxies. Finally, each E4A school and district must articulate their initial (baseline) program goals and objectives as the reference point in determining growth/improvements in student achievements. The MDE will contact districts for student-level data.

In order to facilitate the organization and review of E4A grantee information and data, the MDE and RIA will assist grantees in organizing evidence/documentation within their accountability “portfolio.” The portfolio will contain three sections: (a) assessment quality, (b) performance indicators, and (c) goals and objectives. Information and data within each section will create “a body of evidence” used to support the performance classification (i.e., accountability rating). Table 1 below provides additional details regarding the types of information and data contained within the accountability portfolio.

Table 1. Accountability Portfolio: Information and Data

Area	Data/Information	School Year (SY)
Assessment Quality	Science, Social Studies	2012 – 2013 2013 – 2014
Assessment Quality	Reading/ELA, Mathematics	2012 – 2013 2013 – 2014
Performance Indicators	Participation Rate	2012 – 2013 2013 – 2014
	Enrollment Rate	2012 – 2013 2013 – 2014
	FAY Exclusions	2012 – 2013 2013 – 2014
	Graduation Rates	2012 – 2013 2013 – 2014
Goals and Objectives (As defined during the application process)	Goal(s) Statements	2012 – 2013 2013 - 2014
	Program Objectives	2012 – 2013 2013 - 2014
	Other Program Information	

The information in Table 1 should also be augmented by actual assessment results on the respective assessments for both SY 2012-2013 and SY 2013-2014. These assessment data should be reported using the metric that allows comparability, meaning the set of standard scores and associated performance levels. Raw scores or actual response (unscored) data files are not needed for the portfolio. Other supplemental data, such as the performance of a comparative group (e.g., national population of test-takers for a given assessment) may be helpful in articulating how the performance results reflect “strong” status performance and/or “strong” improved performance. Participation rates on the respective assessments should be provided suggesting the eligible subpopulation of test-takers (denominator) and the actual test-takers (numerator) is consistent with the 95% requirement used in the Mississippi School Accountability System. The graduation data are calculated by the Mississippi Department of Education using with the business rules adopted by the Mississippi State Board of Education; thus, no additional data are needed for the portfolio.

Procedural Steps

Schools and districts participating in the *Excellence for All* (E4A) program are afforded an evidence submission period prior to public release to provide accountability data, including academic assessment, participation, and graduation data. This information will be organized into a three section portfolio. This organization will allow the MDE and an external review committee to examine the submitted information and data. The evidence examination will focus on three areas: (a) comparability, (b) representativeness, and (c) performance. These areas will then be assigned a rating based upon the “strength” of the evidence submitted.

All E4A schools and districts must provide or support the gathering of information necessary to meet the following requirements:

1. Submit all required information and data outlined in Table 1 as focused on the *E4A Accountability Rubric* to the Mississippi State Superintendent of Education no later than 30 calendar days after receiving the most current assessment results;
2. State with specificity any unique data or contextual factors the review committee should consider when applying the *E4A Accountability Rubric*; and,
3. Prepare a formal presentation on the accountability indicator proxies used by the school or district in meeting ESEA Flexibility/Section 1111(b) and State Board of Education's (SBE) policy.

Committee

The E4A Portfolio Review Committee is an *ad hoc* committee tasked with evaluating the information and data organized for each E4A school and district within the Accountability Portfolio. The committee members will examine the evidence against current accountability policies and business rules, including examining the validity of the alternative performance indicators. The committee will provide both technical and classification rating recommendations to the Mississippi State Superintendent of Education. All five committee members will have specialization in assessment, and at least two will have additional expertise in accountability systems. At least one member will have extensive knowledge of the Cambridge ACT programs and assessments.

Timeline

The E4A Accountability Portfolio Review is focused on minimizing information and data collection efforts by E4A schools and districts, while organizing a comprehensive “body of evidence” to support the performance classification assigned by the MDE’s accountability system. The proposed timeline integrates the tasks of the MDE, E4A schools and districts, committee members, and the SBE to assign accountability ratings that are supported by performance evidence documented within the E4A Accountability Portfolio. Table 2 below provides a timeline that meets the expectation of an October 2014, final accountability determination.

Table 2. E4A Timeline

Target Action Dates	Action
June 5, 2014	<i>Excellence for All</i> (E4A) grantees are notified of the details of the evaluation process, including a description of the process to be used for grade assignment, evidence to be submitted, and method for submitting evidence.
June 10, 2014	E4A grantees will provide feedback to MDE on the evaluation process.
June 24 and 27, 2014 July 10, 2014	E4A grantees participate in feedback meetings with MDE regarding the evaluation process.
July 23, 2014	State Board of Education Meeting. Note: The proposed evaluation plan was submitted to begin the Administrative Procedures Act (APA) process.
September 24, 2014	Deadline for RIA to complete the screening and analysis of the data and submit the analysis to the External Review Committee.
October 1, 2014	E4A grantee evaluation process.
October 3, 2014	E4A grantees notified of their 2014 preliminary performance classifications.

Target Action Dates	Action
October 10, 2014	Deadline for E4A grantees to submit a Request for Internal Review (appeal) of their preliminary performance classification.
October 13, 2014	MDE Internal Review Committee meets to review E4A grantee(s) appeal(s) to the results of the evaluation process.
October 13, 2014	E4A grantees notified of the Internal Review Committee's decision.
October 14, 2014	Performance classifications for E4A grantees will be submitted to the Commission on School Accreditation for approval.
October 16-17, 2014	State Board of Education Meeting Note: Comments received during the APA process will be presented to the Board.
October 16-17, 2014	Performance classifications for E4A grantees will be submitted to the State Board of Education for approval.

*Note: All dates are subject to change, pending SBE approval.

Evaluation Criteria

The proposed evaluation process is based on three dimensions: (a) comparability, (b), representativeness, and (c) performance.

Dimension I: Comparability

Comparability addresses the extent to which data are available in order for a credible comparison to be made with traditional state accountability standards and outcomes. This could occur due to one or more of the following three factors:

- a. The data are the same as that used by schools and districts not participating in E4A (e.g., graduation rate and/or SATP2 scores).
- b. The data are in one of the areas in which the Mississippi's Technical Advisory Committee (TAC) had a greater degree of confidence in outcomes from the linking

studies. [Specifically, the TAC concluded that the comparability results from the Cambridge IGSE English First Language test and the Cambridge IGCSE American History were more defensible than the outcomes from all other studies.]

- c. The data with respect to performance are so obviously far above or below Mississippi state standards for proficiency that knowing the exact relationship with SATP 2 is moot. For example, Percentage Uniform Mark (PUM) scores associated with Cambridge IGCSE grades of A or B are internationally regarded as high benchmarks for academic achievement. Or, the ACT benchmarks for College and Career Readiness are nationally regarded as commendable levels of achievement. A claim of comparably high performance with respect to the state standards for proficiency could be supported for a school or district in which the majority of all students and all subgroups scored at such laudable levels.

The E4A Portfolio Review Committee will make an overall judgment considering these three factors holistically to arrive at one of the following conclusions:

1. **Strong support for comparability.** There are multiple sources of data (at minimum, a majority of the indicators) in which there is strong evidence of comparability with the data used in the state accountability system. For example, most of the indicators evaluated are the same indicators used in the state system and where the indicators are different, the results from the comparability studies are more trustworthy (i.e., Cambridge English or American History) or obviously above/below state standards based on established benchmarks. Further, the standard for assessment quality for all of the “alternate assessments used by the grantees” are of at least equal technical quality and rigor as the statewide assessments.

2. **Moderate support for comparability.** There are some but less than a majority of the indicators evaluated in which there is strong evidence of comparability with the data used in the state accountability system. Further, the standard for assessment quality for **most** of the “alternate assessments used by the grantees” are of at least equal technical quality and rigor as the statewide assessments.
3. **Weak support for comparability.** There are no indicators or very few indicators evaluated in which there is strong evidence of comparability with the data used in the state accountability system. Further, the standard for assessment quality for **very few or none** of the “alternate assessments used by the grantees” are of at least equal technical quality and rigor as the statewide assessments.

Dimension II: Representativeness

Representativeness addresses the extent to which the data evaluated reflect the performance of all students served by the school or district. For example, if credible and comparable data exist for an indicator, but only for a small and non-representative group of students, that data are regarded as far less useful than an indicator available for the full population of students. It is recommended that the review committee make an overall judgment to arrive at one of the following conclusions:

1. **Strong support for representativeness.** There is strong evidence that the overwhelming majority of indicators (no more than 1 or 2 exceptions) are based on performance of all students or a sample that is highly representative of all students served by the school or district. Further, the assessment results represent no less than 95% of all eligible test-takers and no systematic exclusion of students (e.g., ELL) is evident. The sample design

ensures every high school student participates in an assessment that measures rigorous content standards in reading, mathematics, and science, although some students may participate at different grade levels at different time during their high school experience.

2. **Moderate support for representativeness:** There is evidence available but it is less than strong in support of the representativeness of the indicator(s) or strong evidence is not available on 2 or more indicators. Further, the assessment results represent no less than 95% of all eligible test-takers and no systematic exclusion of students (e.g., ELL) is evident. The sample design does not ensure all high school students participate in an assessment that measures rigorous content standards in reading, mathematics, and science, although some students may participate at different grade levels at different time during their high school experience.
3. **Weak support for representativeness.** There is little to no evidence available to evaluate representativeness. Further, the assessment results represent less than 95% of all eligible test-takers and systematical exclude students (e.g., ELL) is evident. The sample design does not ensure all high school students participate in an assessment that measures rigorous content standards in reading, mathematics, and science, although some students may participate at different grade levels at different time during their high school experience.

Dimension III: Performance

Performance addresses the extent to which student performance, with respect to graduation rate and academic achievement, is at an exemplary high level or improving at an extraordinary pace for all students and all subgroups.

1. **Strong Performance.** There is strong evidence for the overwhelming majority of indicators that graduation and achievement are at an exemplary high level or improving at an extraordinary pace for all students and all subgroups. Further, all assessment results have achievement at or above the 50th percentile of all test-takers for a given year and/or improvement rates are significantly outpacing all other test-takers. Graduation rates are in the top 75th percentile for the State of Mississippi.
2. **Moderate Performance.** Evidence of performance does not support a conclusion that graduation and achievement are at an exemplary high or low level or improving at an extraordinary pace for all students and all subgroups. Further, most assessment results have achievement at or above the 50th percentile of all test-takers for a given year and/or improvement rates are significantly outpacing all other test-takers. Graduation rates are in the top 50th percentile for the State of Mississippi.
3. **Weak Performance.** There is strong evidence for the overwhelming majority of indicators that graduation and achievement are at an extraordinary low level or declining at an extraordinary pace for many students and subgroups. Further, most assessment results have achievement below the 50th percentile of all test-takers for a given year and/or improvement rates are significantly lower than all other test-takers. Graduation rates are in the bottom 50th percentile for the State of Mississippi.

Rubric Design

The E4A Accountability Portfolio evidence will be organized to address three dimensions within a holistic scoring rubric, the *E4A Accountability Rubric*. The *E4A Accountability Rubric* is comprised of three dimensions: (a) comparability, (b) representativeness, and (c) performance. These dimensions are evaluated in terms of different sections within each school's or district's portfolio: (a) assessment quality, (b) performance indicators, and (c) goals and objectives. The horizontal axis provides the "rating" continuum ranging from "Strong" evidence to "Weak" evidence.

Build: Implementation Sequence

Workflow

The macro-level workflow shown in Figure 1 provides the major steps involved in the *E4A Accountability Portfolio Review* process. This workflow is provided as an overview of the procedures from the initial stage of orientation training to the final submission of the reviewed assessments, performance indicators, and goals and objectives.

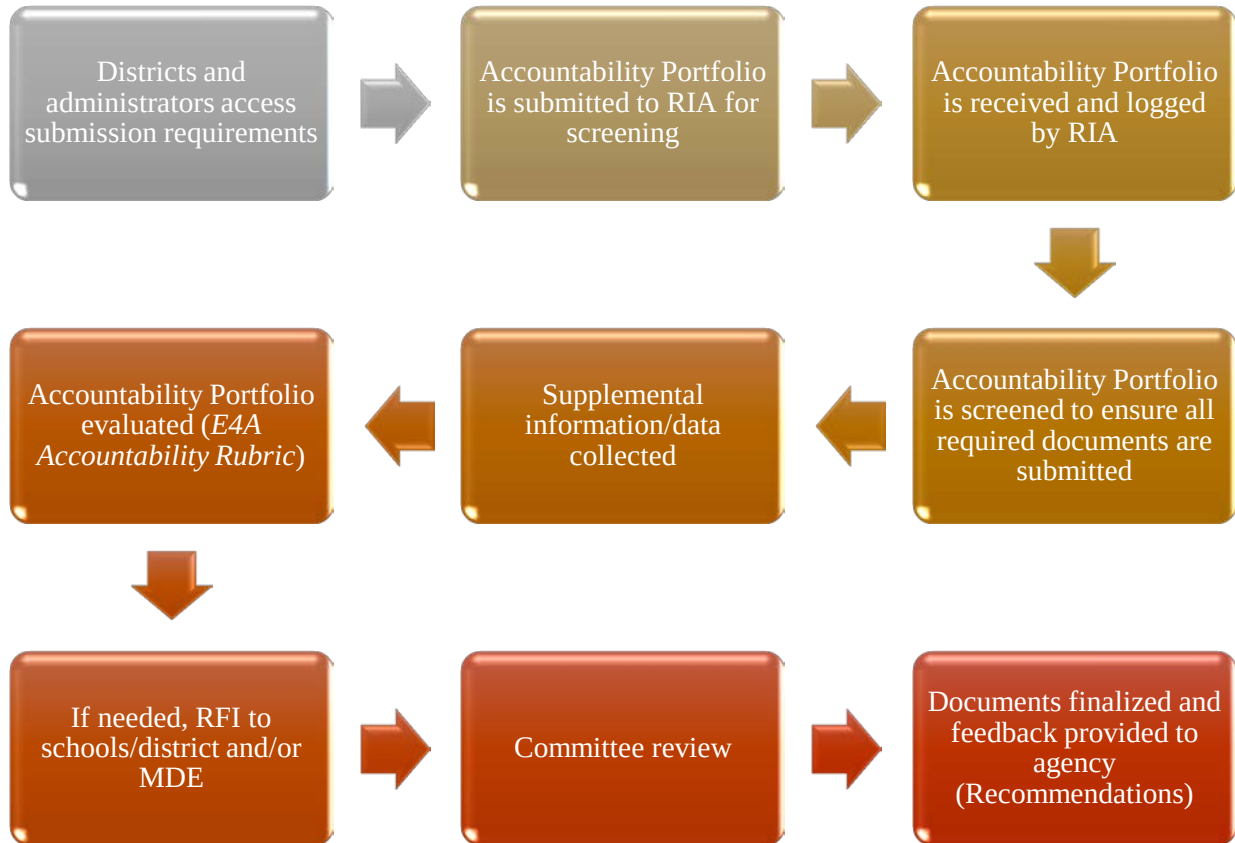


Figure 1. Macro-Level Workflow

Procedures

Based on the macro-level workflow, a detailed sequence of steps and decision logic has been documented to provide a standardized sequence of activities and process paths for decision-making.

Procedural Phases

Screening

This phase involves conducting a preliminary review of the information and data within the district's accountability portfolio. Any missing documents will be identified prior to the Reviewing phase. A screening checklist will be developed and applied for each of the submitted

portfolios to ensure comparability of information and data. Supplemental information on the assessment quality will be collected by the respective vendors and the MDE.

Evaluating

Each portfolio will be evaluated using the required *E4A Accountability Rubric*. For external assessments, the respective Buross Report, housed at the University of Nebraska, as well as any additional technical information provided by the vendor/publisher will be extracted and used as evidence aligned to the applicable descriptor with the *E4A Accountability Rubric*. The Buross Report will be used as the “gold standard” for basing responses and recommendations to the agency. Additional information and data (e.g., enrollment) collected by the MDE and validated by districts will be included within the accountability portfolio.

Reviewing

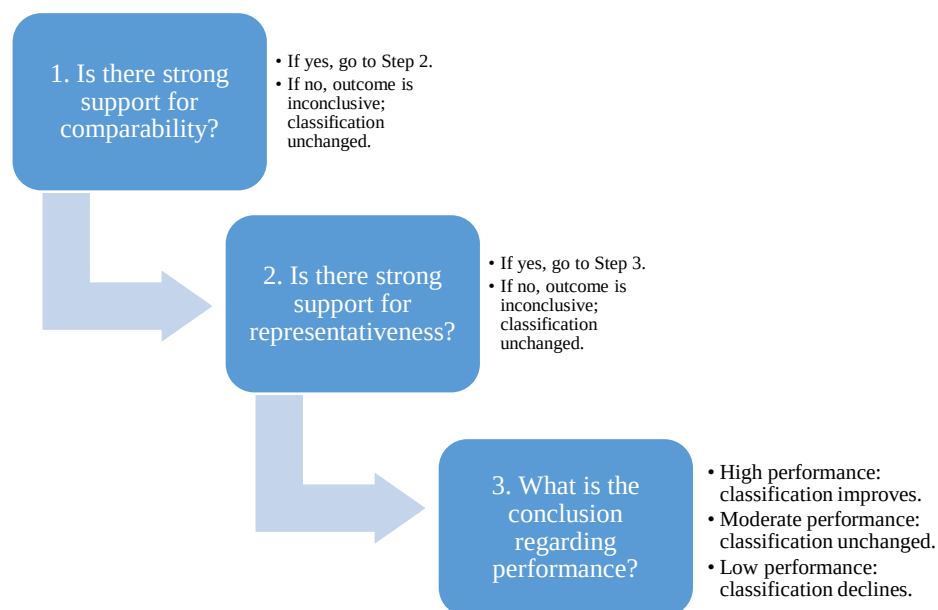
Committee members will be provided with each school’s and district’s portfolio prior to the on-site meeting and will evaluate the information and data using the criteria found within the accountability portfolio. These “preliminary” results, notes, and comments will be brought to the on-site meeting and discussed among the five committee members.

The outcome of the review process will be a recommendation to the Mississippi State Superintendent of Education for the performance classification to improve, remain the same, or decline with respect to the school or district’s prior year accountability outcome. That recommendation would be one of the following:

1. Performance classification improves (e.g., from B to A).
2. Performance classification is unchanged (e.g., remains B).
3. Performance classification declines (e.g., from B to C).

To support a recommendation of assigning a more or less favorable classification, there should be strong support for both comparability and representativeness. Under such circumstances, the outcome of the third criterion should influence the decision. For example, if there is strong evidence that the overwhelming majority of indicators are comparable and representative **AND** performance is at an exemplary high level or improving at an extraordinary level, then these conditions might support assigning the school a more favorable outcome.

Conversely, if there is strong evidence of both comparability and representativeness **AND** performance is very low or declining at an extraordinary pace, then these conditions might support assigning the school a less favorable classification. However, when circumstances reveal that comparability and/or representativeness is moderate or low, the outcome is likely inconclusive and no adjustment in classification can be supported. The committee may support an adjustment in classification if the ratings for comparability or representativeness are moderate **IF** accompanied by particularly compelling performance data. The following diagram illustrates the proposed evaluation process for decisions:



Finalizing

This phase is the last phase of the process and involves a final read thru of all documents prior to submission to the agency. Any omissions, flags, or issues observed will be identified at this phase and corrections made. A final recommendation for each E4A grantee will be documented in the E4A portfolio and within an executive summary. Table 3 below outlines the proposed recommendation summary.

Table 3. Committee Recommendation Summary

School	District	Rubric Performance	Recommendation
School A	District A	High Performance	Classification Improves
School B	District B	Moderate Performance	Classification Unchanged
School C	District C	Low Performance	Classification Declines
School D	District D	High Performance	Classification Improves
School E	District E	Moderate Performance	Classification Unchanged

Note: This evaluation process was approved as a temporary rule based on a finding by the State Board of Education that the rule only confers a benefit or removes a restriction on the public or some segment thereof. The business rules of the Mississippi Statewide Accountability System were revised to include the process of evaluation and reporting for districts and schools participating in the *Excellence for All* program for school year 2013-2014. On February 21, 2014, the United States Department of Education (ED) granted the request to permit students in five (5) districts participating in the *Excellence for All* program to take the Cambridge

International Exams and/or the ACT End of Course/End of Domain exams in place of the state assessments, and to use the results from those assessments in accountability decisions. The temporary rule is necessary for the State Board of Education's approval of the final performance classifications for these five (5) districts on October 16-17, 2014.

Appendix A: E4A Accountability Rubric

General Directions

The enclosed accountability rubric is designed to examine the quality characteristics of alternate accountability indicators found within the *Excellence for All* (E4A) portfolio. The rubric is comprised of technical requirements organized into three dimensions: (a) comparability, (b) representativeness, and (c) performance. Each dimension is comprised of three areas of focus, similar to the three-part portfolios developed for each E4A district. Each area of focus within a given dimension is rated holistically using categorical assignments based upon the reviewed evidence.

Reviewer's Task

- Step 1.** Review each part of the portfolio (i.e., information, data, and documents) in terms of comparability, representativeness, and performance.
- Step 2.** Assign a category to each focus area within a particular dimension using the following rating scheme:
 - a. (S) = strong evidence
 - b. (M) = moderate evidence
 - c. (W) = weak evidence
- Step 3.** List information, data, and document references supporting each assigned rating.
- Step 4.** Add notes and/or comments articulating nuances of the evidence reviewed.
- Step 5.** Compile results for each dimension into the overall Summary Matrix.

Summary Matrix: _____

<school name><district name>

Preliminary Recommendation (Circle One)

- Accountability classification improves (e.g., from B to A)
- Accountability classification is unchanged (e.g., remains B)
- Accountability classification declines (e.g., from B to C)

Dimension	Portfolio Section	Evidence Rating (S), (M), (W)	Comments/Notations
Comparability	Assessment Quality		
	Performance Indicators		
	Goals and Objectives		
Representativeness	Assessment Quality		
	Performance Indicators		
	Goals and Objectives		
Performance	Assessment Quality		
	Performance Indicators		
	Goals and Objectives		

Dimension 1: Comparability

Comparability addresses the extent to which data are available in order for a credible comparison to be made with traditional state accountability standards and outcomes.

Portfolio Section	Evidence Reviewed	STRONG (S) There are multiple sources of data (at minimum, a majority of the indicators) in which there is strong evidence of comparability with the data used in the state accountability system.	MODERATE (M) There are some but less than a majority of the indicators evaluated in which there is strong evidence of comparability with the data used in the state accountability system.	WEAK (W) There are no indicators or very few indicators evaluated in which there is strong evidence of comparability with the data used in the state accountability system.
Assessment Quality				
Performance Indicators				
Goals and Objectives				
Comments/Notes				

Dimension 2: Representativeness

Representativeness addresses the extent to which the data evaluated reflect the performance of all students served by the school or district.

Portfolio Section	Evidence Reviewed	STRONG (S) There is strong evidence that the overwhelming majority of indicators (no more than 1 or 2 exceptions) are based on performance of all students or a sample that is highly representative of all students served by the school or district.	MODERATE (M) There is evidence available but it is less than strong in support of the representativeness of the indicator(s) or strong evidence is not available on more than 1 or 2 indicators.	WEAK (W) There is little to no evidence available to evaluate representativeness.
Assessment Quality				
Performance Indicators				
Goals and Objectives				
Comments/Notes				

Dimension 3: Performance

Performance addresses the extent to which student performance with respect to graduation rate and academic achievement is at an exemplary high level or improving at an extraordinary pace for all students and all subgroups.

Portfolio Section	Evidence Reviewed	STRONG (S) There is strong evidence for the overwhelming majority of indicators that graduation and achievement are at an exemplary high level or improving at an extraordinary pace for all students and all subgroups	MODERATE (M) Evidence of performance does not support a conclusion that graduation and achievement are at an exemplary high or low level or improving at an extraordinary pace for all students and all subgroups	WEAK (W) There is strong evidence for the overwhelming majority of indicators that graduation and achievement are at an extraordinary low level or declining at an extraordinary pace for many students and subgroups
Assessment Quality				
Performance Indicators				
Goals and Objectives				
Comments/Notes				

Appendix B: PROGRAMS FACT SHEET

Summary of Programs from Vendor Web Pages

Cambridge International Examinations

University of Cambridge International Examinations is the world's largest provider of international education programs and qualifications for 5–19 year olds. Its qualifications are taken in over 160 countries and recognized by universities, education providers and employers across the world. The Cambridge International Examinations are part of the Cambridge Assessment Group, a non-profit organization and a department of the University of Cambridge. Their mission is to deliver world-class international education through the provision of curricula, assessment and services. Cambridge International Examinations is committed to extending access to the benefits of high-quality education around the globe.

Web Page: <http://www.cie.org.uk>

ACT Quality Core[®]

QualityCore[®] high school course standards in Reading, Writing, Speaking and Listening, Language, and Math are fully aligned to the Common Core State Standards. QualityCore[®] allows educators, administrators, and policymakers to:

- Focus on fewer, clearer, and higher course standards that are the essentials for college and career readiness.
- Use formative assessments to guide instruction and make targeted interventions.
- Measure student progress through benchmark and end-of-course assessments.
- Provide professional development resources and training opportunities for teachers and leaders.
- Customize instruction to the needs of specific classrooms, schools, and districts.

Web Page: <http://www.act.org/qualitycore>

Appendix C: PRESENTATION TEMPLATE

Executive Summary (three key points)

Key Point 1. <Statement regarding the overwhelming information/data supporting a “strong” rating on each of the three dimensions>

Key Point 2. <Statement regarding the additional benefits leveraged by the programs inclusion in the schools/district>

Key Point 3. <Statement regarding anticipated further results given the relatively short implementation history>

Persuasive Evidence

(recommended 1 page limit per Key Point)

Key Point 1.

- a. <supplemental empirical data>
- b. <additional qualitative information>
- c. <anecdotal facts and observations>
- d. <case study and testimonials >

Key Point 2.

- a. <supplemental empirical data>
- b. <additional qualitative information>
- c. <anecdotal facts and observations>
- d. <case study and testimonials >

Key Point 3.

- a. <supplemental empirical data>
- b. <additional qualitative information>
- c. <anecdotal facts and observations>
- d. <case study and testimonials >