

2696



Commonwealth of Pennsylvania
STATE BOARD OF EDUCATION

February 27, 2009

Mr. Kim Kaufman
Executive Director
Independent Regulatory Review Commission
333 Market Street, 14th floor
Harrisburg, PA 17101

Re: Regulation # 006-312 (IRRC #2696)

Dear Mr. ^{Kim}~~Kaufman~~

In response to comments provided to the State Board of Education by the House Education Committee and Independent Regulatory Review Commission in reaction to proposed regulation #006-312 (IRRC #2696), the Department of Education commissioned Penn State University to perform a study of local assessment systems used by school district to determine whether students are proficient in reading, writing and mathematics in order to graduate from high school as currently required in 22 PA Code § 4.24(a). The Penn State research team found that only 5% of school districts have local high school graduation proficiency assessment systems in both math and reading that are both aligned with state academic standards and properly administered in practice to all students.

We believe that members of the Commission and your staff will find the study useful when considering this issue in the future.

Sincerely,

A handwritten signature in black ink that reads "Jim Buckheit".

Jim Buckheit
Executive Director

Local Assessment Validity Study Report
February 2009

Rayne A. Sperling, PhD

Jonna M. Kulikowich, PhD

The Penn State Research Team

Table of Contents

Introduction	4
Section I: District Characteristics for Reporting and Non-Reporting Districts	5
Table 1. Demographic Information for All Districts and for Reporting Districts.....	6
Table 2. Demographic Information for Non-Reporting Districts	7
Section II: A Description of the Materials Submitted to PDE	8
Database Development and Expert Panel Preparation.....	8
Characteristics of the Assessment Materials Submitted by Districts.....	9
Table 3. Descriptive Information for Materials Submitted by Districts	9
Table 4. Assessment Materials Submitted by Districts by Type.....	10
Multiple-Choice and Constructed-Response Assessments.....	10
Standardized Tests and Published Materials.....	11
Table 5. Standardized Tests and Published Materials Frequencies	12
Curricula	13
Portfolios.....	13
Courses and Tutoring.....	14
Section III: Local Assessment Measures of Materials Alignment and Reported Practices ..	14
Invitation to Expert Panelists	14
Purposes and Procedures for Expert Panelists	15
The Training Session	15
Coding Procedures for Alignment of Materials to the Proficiency Standards.....	16
Table 6. Example of Materials Coded by Expert Panels by Level	17
Results.....	19
Table 7. Materials Alignment Means and Standard Deviations of Scores for Reporting Districts	20
Table 8. Reading and Mathematics Materials Alignment Scores for Reporting Districts	21
Reliability of Rating Scale Codes	21
District Reported Practices as Measures of Proficiency	22
The Rating Scale for Reported Practices	22

Panel Discussions and Consensus Agreement on Practices	23
District Follow-up Procedures	23
Table 9. Agreement of Responses	24
Table 10. Proficiency Codes and Representative Practices	25
Table 11. Local Assessment Practices Means and Standard Deviations of Scores for Reporting Districts	26
Table 12. Reported Practices Scores for Reporting Districts	27
Table 13. Reported Practices by Alignment Ratings for All Reporting Districts	28
Table 14. Relations Between Alignment and Reported Practices in Mathematics and Reading	29
Conclusion	29
Appendices	
Appendix A. The Penn State Research Team	30
Appendix B. Fields that Correspond with Submitted Database.....	31
Appendix C. Proficiency Standards Experts Used as Reference	35
Appendix D. Example Proficiency Coding Sheet.....	38
Appendix E. Practice Rating Form Administered to Expert Panels	40
Appendix F. Expert Panel Evaluation Summaries	46
Appendix G. Local Assessment Survey.....	48

Introduction

In this report for the Local Assessment Validity study, we outline procedures conducted to store information in a research database about local assessments submitted to the Pennsylvania Department of Education (PDE) by school districts of the Commonwealth. Further, we provide initial descriptions of the assessment materials delivered by school districts in Pennsylvania based on the information stored in the research database.

This study is designed to examine the local assessments used by districts throughout the Commonwealth of Pennsylvania to measure proficiency of the academic standards in Mathematics and Reading at the secondary level. Pennsylvania Chapter 4 regulation states that *“Students shall demonstrate proficiency in reading, writing and mathematics on either the State assessments administered in grade 11 or 12 or local assessment aligned with academic standards and State assessments under § 4.52 (relating to local assessment system) at the proficient level or better to graduate.”*

This regulation guides our research. Our primary purpose is to describe the characteristics of the local assessments used by the individual districts. Appendix A provides the names and titles of the members of The Penn State Research Team.

To best describe the characteristics and nature of the local assessments, we formatted this report to include three major sections. Each section includes procedures, data sources, and findings, which we present descriptively. The first section describes characteristics of the districts that did and did not submit local assessment information. This section starts with the overall data collection procedure and documents team recording procedures and data sources. The second section describes the nature of the materials submitted by the districts based upon district characteristics. The third section describes the procedures employed and data collected to explore the degree to which materials and practices submitted by districts corresponded with the State proficiency standards.

In response to a request by the PDE sent to the districts via Penn Link Monday, July 28, 2008 and again Tuesday, August 12, 2008, districts provided the local assessments that they use to measure academic proficiency in Mathematics and Reading for those students who do not score proficient on the 11th grade administration of the PSSA or on the 12th grade retake of the exam. Under § 4.52, each district is to determine the characteristics of its individual local assessments. These local assessments might include a test, portfolio, curriculum materials, or other means. Regardless of the nature of the local assessment, each district is to have a local assessment practice that can measure proficiency on the academic standards for students whose performance is not deemed proficient through the PSSA testing. The purpose of the Local Assessment Validity Study was to describe the materials submitted as the local assessments by the districts. Additionally, we summarize information the districts provided about the types of practices they use to meet the proficiency requirements.

Section I: District Characteristics for Reporting and Non-Reporting Districts

In response to the Penn Link request, approximately 85% of districts submitted materials to PDE. Table 1 provides the basic demographic information of the school districts that submitted materials to PDE.

As noted in Table 1, the sample of districts that sent assessment materials is approximately 85% across Rural, Suburban, and Urban districts. Urban districts have a higher rate of non-proficiency than Suburban or Rural districts based upon PSSA performance. Across all district types, a substantial percentage of students in the Commonwealth do not score proficient in Mathematics and/or Reading given two opportunities on the PSSA.

Eighty-four percent of districts submitted local assessment practices or materials for review. The districts that submitted materials appeared representative of the Commonwealth. We further examined characteristics of the districts that did not submit materials in order to explore whether there were any trends among these non-reporting districts.

There were 79 school districts that did not submit materials after either the initial, or the follow-up, PDE requests for local assessment information. Table 2 presents descriptive information for these districts.

Table 1. Demographic Information for All Districts and for Reporting Districts.

	TOTAL	Rural	Suburban	Urban
All Districts	497*	226	231	40
Average District Enrollment 2006-2007	1163**	762	1300	2683**
Enrollment Range	85 – 47789	85 – 3760	188 - 9067	164 - 47789
Average Per-Pupil Expenditure (\$)	11925.45	11375.85	12474.21	11928.21
Expenditure Range (\$)	8174.25 – 23009.15	8174.25 – 20394.90	8484.56 – 23009.15	9058.45 – 15572.14
Average % failing both PSSA attempts	42.77	45.74	37.4	58.56
PSSA Range	10.2 – 93.1	19.3 – 77.3	10.2 – 75.4	26.7 – 93.1
Reporting Districts	418 (84.1%)	189 (83.6%)	195 (84.4%)	34 (85.0%)
Average District Enrollment 2006-2007	1185 (1072 w/o Districts of the First Class)	779	1295	2894 (1446 w/o Districts of the First Class)
Enrollment Range	85 – 47789	85 – 3760	188 - 9067	164 – 47789
Average Per-Pupil Expenditure (\$)	11988.41	11410.02	12556.58	11924.71
Expenditure Range (\$)	8174.25 – 23009.15	8174.25 – 20394.9	8634.41 – 23009.15	9058.45 – 15572.14
Average % failing both PSSA attempts	42.22	44.73	37.06	59.09
PSSA Range	10.2 – 92.2	19.3 – 77.3	10.2 – 75.3	33.6 – 92.2

*Although there are 501 school districts in PA, 4 do not have high schools, thus are not eligible for this investigation

** These numbers exclude Districts of the First Class

Table 2. Demographic Information for Non-Reporting Districts.

	TOTAL	Rural	Suburban	Urban
Non-reporting Districts(%)	79	37 (46.83)	36(45.57)	6 (7.59)
Average District Enrollment 2006-2007	1050	666	1295	1,646
Enrollment Range	317 – 3868	317 – 1,894	382 -3868	629 – 3,442
Average number of graduates 2006-2007	216	136	288	286
Graduation class Range	35 – 904	35-559	46 – 904	146 – 550
Average % failing both PSSA attempts	46.48%	51.30%	39.99%	55.72%
PSSA Range	17.8% - 93.1%	29.1%- 69.1%	17.8% - 75.4%	26.7% - 93.1%
Average Per-Pupil Expenditure (\$)	\$11,618.46	\$11,184.74	\$12,009.50	\$11,946.88
Expenditure Range (\$)	\$8,484.56 – \$16,121.68	\$8,921.52 - \$13,986.57	\$8,484.56 - \$16,121.68	\$10,787 - \$14,502.57
Average % Free & Reduced Enrollment	34.32%	37.35%	28.11%	52.93%
Free & Reduced Enrollment Range	2.51% – 82.08%	8.92% - 60.81%	2.51% - 82.08%	19.83% - 77.58%

*Excludes the 4 school districts that do not have high schools.

PDE indicated the date materials were received and documented materials by districts as they were submitted. To assure accuracy in accounting for submitted materials, we compared the PDE checklist with the Penn State Database. There was 100% agreement between the PDE recorded districts and the Penn State Database. Eighty-three school districts did not send local assessment information or materials as requested from PDE. However, 4 of these school districts do not have high schools and as such are not included in further analysis. Of the remaining 79 school districts with high schools, approximately 46.83% were from rural areas, 45.57% were suburban school districts, and 7.59% were urban. Overall the percent of non-proficient students on the 11th grade PSSA and the 12th grade retake was 46.48%. The per-pupil spending for the 79 school districts averaged \$11,618.46 and the average percent of students enrolled in the free and reduced lunch program in these schools was 34.32%.

Descriptive comparisons between the reporting and non-reporting districts suggest that non-reporting districts had a smaller average enrollment and had lower per-pupil spending. The average percentage of students non-proficient after both attempts of the PSSA was higher for the

non-reporting schools overall. Further, the non-reporting districts had a higher percentage of free and reduced lunch overall when compared to the districts' average for the State at 34.32% with the State average of 30.87%. While these *averages* may be slightly higher for the non-reporting districts, the non-reporting districts' percentages fall within the *range* of districts overall on these critical variables.

Section II: A Description of the Materials Submitted to PDE

Database Development and Expert Panel Preparation

Schools responded to the PDE request by sending local assessment materials directly to PDE. As assessments were received, they were mailed to The Penn State Research Team. At the onset, it was unclear what types of materials districts might provide. As a research team it was therefore necessary to first explore what types of materials were submitted in response to the PDE request. Therefore, The Penn State Research Team first worked in pairs to record on large chart paper the nature of the materials sent by each district accompanied by information each provided regarding how the local assessments were used. After an initial examination of approximately 100 districts, The Penn State Research Team established a better understanding of the nature of the sent materials.

As indicated in the introduction, a primary purpose of the study was to develop a research database to store information about the assessments submitted to the PDE as well as any reported practices used by the districts to determine proficiency levels. Therefore, The Penn State Research Team next developed an initial database generated in part by clear themes found in the nature of the materials submitted and with data found in PDE Education Names and Addresses (EdNA: <http://edna.ed.state.pa.us>). We entered information for each district. Specifically, the initial database fields included demographic information for each district and for schools within district, in-take date, contact information for personnel at the district or school level, the nature of materials submitted, and whether there was a PSSA policy/practice relative to local assessments reported. We made one first review through all school district materials, created files for each district, and filed them alphabetically. This first review of the information submitted to the PDE allowed us to process all materials and to build the database based upon initial data codes. The Penn State Research Team then conducted a second, more comprehensive, review of all of the materials.

For the second examination of all district files, we again worked in pairs to assure accuracy in entering information. Our goals were to simultaneously complete an expanded database and also to prepare materials for Mathematics and Reading expert panels to code. The expanded database clearly explicated school-reported assessment practices designed to measure student proficiency of standards. It also included more descriptive fields about Mathematics and Reading assessments. For example, we coded materials on assessment type (e.g., district-created assessment, published test, curriculum) as well as item type (e.g., multiple-choice, constructed response). We also used a random number generator to assign each district a random number. We used this procedure so that we could remove all identifying information about school districts before review of their materials by the expert panels. Therefore, as we completed data entry for this second review of the materials, we removed all identifiers on materials that districts

submitted through redaction. Separate Mathematics and Reading folders were created for panelist coding and all were labeled with the districts' random numbers.

Characteristics of the Assessment Materials Submitted by Districts

There was variance in the nature of the materials submitted for review by the districts. This variance included, for example, that some districts submitted a letter and others submitted various materials, in some cases multiple binders or boxes of materials, to represent their local assessment. Some districts reported that they did not have a local assessment. Others noted that proficiency on their local assessment was not a requirement for graduation purposes. Many districts reported that their curriculum was aligned to standards, and, as such, if students passed their coursework their performance in Mathematics and Reading was deemed proficient. Some districts reported use of attendance or citizenship as measures of proficiency. In response to the PDE request, many districts submitted descriptions or examples of assessments, curriculum, or policy statements. Whatever each district sent was used to create a folder that represented the local assessment for that district. That is, the local assessment might not have been an individual test, or any test at all, but might include a collection of tests from several courses or tests and other materials, such as instructional materials, for multiple courses. In this section of the report, we note some of the characteristics of the materials submitted.

Table 3. Descriptive Information for Materials Initially Submitted by Districts.

Types of Materials Submitted by Districts Overall (n = 418, districts that sent materials)

Type of Material*	Mathematics Frequency (%)	Reading Frequency (%)
Multiple-choice Test	314 (75)	315 (75)
Constructed Response	269 (64)	273 (65)
Standardized/Published Test	140 (33)	150 (36)
Portfolio	33 (08)	40 (10)
Curriculum	36 (09)	35 (08)
Course	99 (24)	96 (23)
Tutoring	50 (12)	48 (11)

***Districts may have submitted more than one type of material.**

Table 4. Assessment Materials Initially Submitted by Districts by Type.

Type of Material	Mathematics			Reading		
	Rural	Suburban	Urban	Rural	Suburban	Urban
Reporting n	189	195	34	189	195	34
Multiple-choice Test (% within category)	133 (70)	161 (82)	20 (63)	133 (70)	161 (82)	21 (66)
Constructed Response (% within category)	110 (58)	141 (72)	18 (56)	112 (59)	142 (72)	19 (59)
Standardized/Published Test (% within category)	70 (37)	61 (31)	9 (28)	74 (39)	66 (34)	10 (31)
Portfolio (% within category)	16 (08)	16 (08)	1 (03)	15 (08)	24 (12)	1 (03)
Curricula (% within category)	14 (07)	20 (10)	2 (06)	16 (08)	18 (09)	1 (03)
Course (% within category)	39 (21)	56 (29)	4 (13)	36 (19)	55 (28)	5 (16)
Tutoring (% within category)	23 (12)	22 (11)	5 (16)	22 (11)	21 (11)	5 (16)

Multiple-Choice and Constructed-Response Assessments

Districts often submitted tests or sets of assessments that they employ to measure proficiency. These materials were coded based upon the nature of the items that they included. Some of these sent-materials included copies from standardized assessments, and these were included in the coding. However, the assessments often represented teacher- or district-constructed assessments. Some districts submitted one test or multiple versions of a single test. Other districts submitted copies of tests for various courses within their respective curriculum for coding. In some of these cases, the districts stated that the courses were aligned to standards and therefore tests in the courses represented their measure of proficiency. In other cases, the districts submitted a number of tests with no explanation. In either case, these materials were prepared for coding and were coded as districts' artifacts of their local proficiency measure. As illustrated in Table 3 and Table 4, many districts included multiple-choice items and constructed-response assessments within their measures of proficiency. Many submitted materials that employed both types of assessment formats. We coded constructed-response items to include both fill-in types of items as well as problems or essays. These types of materials were submitted by Rural, Suburban, and Urban districts for both Mathematics and Reading as noted in Table 4.

Standardized Tests and Published Materials

Numerous districts reported use of a variety of either packaged published materials or tests or standardized assessments. The number of districts that reported using these types of materials is larger than represented in the sent materials. Districts often reported using these materials but did not actually send artifacts for coding. That is, if a district said they used the 4Sight test, for example, but did not send a copy of that test, it was not coded as materials submitted. However, the reported use of any standardized or published materials for the purposes of local assessment was coded both with respect to alignment of the materials to the proficiency standards and the validity of the practices reported by the superintendent to determine proficiency levels.

Some districts reported use of student performance on, or completion of, college admissions tests such as the SAT, ACT, or PSAT as a measure of proficiency. Other districts reported use of national standardized tests such as the Terra Nova, Metropolitan 8, or the Stanford Achievement Test 9 or 10. Others used diagnostic assessments such as the Group Reading Assessment and Diagnostic Evaluation (GRADE) or the similar Mathematics assessment (GMADE), or the Degrees of Reading Assessment (DRP). Still others used general education tests such as the GED practice test or student success on the GED examination or the TABE. Other districts noted use of PASSKEY assessment or the ASVAB career exploration assessment as the measure used to establish students' proficiency on the Mathematics and Reading standards.

In addition to national standardized tests, some districts turned to items from other states' released achievement tests to use as measures of proficiency. Several districts also reported the use of the tests developed by Intermediate Unit 8 (IU8). The 4Sight tests for both Mathematics and Reading were commonly used local assessments. While many districts used the 11th grade 4Sight tests, others reported use of 4Sight tests below 11th grade level (e.g., 8th grade). Several districts reported the use of the practice items or the released PSSA items as measures of proficiency. In addition to stand-alone assessments, districts reported the use of PLATO and Study Island, as well as PSSA Coach, and other books and software designed for remediation of PSSA content.

Table 5 provides the names and frequency of use of standardized and published materials submitted by districts. It is important to note that some districts reported using several of these measures and others reported using these in conjunction with district-developed curricula and assessments. Some districts sent these materials or stated their use without additional explanation. In these cases, we observed what was sent but did not attempt to infer how these materials were used. Further, some districts reported that they used just the assessments from programs such as Study Island but did not use the program as a tutorial. In all, districts reported nearly 60 different measures and materials in this category. The most frequently reported measures in this category were the 4Sight assessments and Study Island followed by PSSA Coach, PDE released items, and PLATO.

Table 5. Standardized Tests and Published Materials Frequencies.

Test	Frequency
4Sight	115
Study Island	72
PSSA Coach	35
PDE Released Items	28
PLATO	25
Intermediate Unit 08 Assessment	20
Measuring Up	13
Terra Nova	13
SAT	12
Princeton Review	9
NOCTI	7
PSAT	7
ACT	4
Cognitive Tutor	4
GED Practice Test	4
GMADE	4
GRADE	4
MAT 8	4
Aleks Software	3
Apangea	3
Jump Start	3
PassKey	3
PVASS	3
Maine Department of Education Sample Items	3
Stanford Achievement Assessments	3
ASVAB (Career Exporation)	2
Brain Child	2
Compass Learning	2
GATES Mac-Ginnitie	2
MAP	2
NOVA Net	2
PASA	2
PLAN	2
PSSA Finish Line	2
PSSA Test Prep Book	2
Scholastic Read 180	2
Oregon Department of Education Sample Items	2
Scantron Performance System	2
Skills Tutor	2
Success for All	2

Achieve 3000	1
Achievement Test of Basic Written English (Townsend Press)	1
AIMSWEB	1
AMSCO	1
APEX Learning	1
BlendedSchools.Net	1
Bridges	1
Capital Area Intermediate Unit Assessment	1
CTC	1
CTP 4	1
EduTest	1
E-Metrics	1
Homeroom.com	1
ICTC	1
Keys to Work	1
March 2 Success	1
MYAccess	1
Nevada Department of Education Sample Items	1
Smart Math	1
STAR Math	1
STAR Reading	1
TABE	1
WADE	1
Web Achiever	1
Woodcock Reading Battery	1

Curricula

Some districts submitted curriculum materials as their local assessment used to demonstrate proficiency on the Mathematics and Reading standards. The materials within this category represented considerable variance. Materials were coded as curricula if they included individual curricular materials or a full course curriculum with identified standards. Some districts submitted curricula with accompanying tests while others did not send course materials, but simply a course description.

Portfolios

Some districts reported that they utilize portfolios as their local assessment for students to demonstrate proficiency on the standards for Mathematics and Reading. We coded a portfolio as a local assessment for a district if either the district stated use of a portfolio or if they sent a portfolio description. As illustrated in Table 3, 33 school districts reported use of portfolios for Mathematics and 40 reported use of portfolios for Reading.

Through further examination of information districts provided about the assessment practices they employ, it appears that portfolios are typically used to measure proficiency in one of two ways. First, portfolios can be the last option if a student fails other local assessments, such as tests, or, second, they are used as a district's single local assessment. The contents of the portfolios reported as used can be divided into three broad categories: (1) portfolio content that is comprised of assignments/tests from published material such as Measuring Up, PSSA Coach, or Study Island; (2) portfolio content that is comprised of student-selected work from a course or a group of courses (these include tests and/or assignments); or, (3) portfolio content that is comprised of individualized assignments reflecting the areas in which an individual student scored non-proficient on the PSSA. Finally, some school districts indicated that the portfolio that students must create as part of their graduation requirement also is used as a means of determining proficiency on the academic standards.

Courses and Tutoring

Some districts reported the use of a stand-alone course or a tutoring program as their local assessment that students complete in order for the district to measure proficiency. Some of these courses and tutoring were reported in conjunction with commercially-available programs, such as Study Island or PLATO. Others were tutoring programs, summer school programs, or after-school programs conducted by the district. A local practice was coded as a course or a tutoring program based upon district description. Tables 3 and 4 present the frequencies and percentages of districts that reported courses and tutoring as a component of their local assessment practice.

In conclusion, there was wide variance in what districts submitted as their local assessments. Materials submitted included both published and locally-developed tests and other materials. All materials submitted were cataloged. If a district sent materials, those materials were included as local assessments and placed into a district folder for later coding. Materials were included whether or not the district submitted explanation or clarification regarding how their local assessments were used. As noted, all of the district folders were prepared for expert panel review. In the following section, we describe the procedures used to rate the degree to which local assessments aligned to proficiency standards and how the practices reported by districts were considered as potentially valid mechanisms to measure proficiency.

Section III: Local Assessment Measures of Materials Alignment and Reported Practices

Invitation to Expert Panelists

Expert practitioners from across the state were selected to review and rate the materials that districts submitted as their local assessment. PDE invited experts in Mathematics and in Reading to serve on the panels. These invitations were extended with assistance from Intermediate Units, superintendents, and through knowledge of previous work conducted by the practitioners in the Commonwealth. Experts were selected according to several criteria (e.g., the geographic locations within the Commonwealth of their respective affiliations; years of administrative/teaching experience). Experts were not paid for their work. Each expert received mileage, tolls, and meals in accord with State and Federal per-diem allotments. Reimbursement for substitute teachers' salaries was provided directly to the districts for those experts who were

teachers for whom classes had to be covered during scheduled dates of the coding. The expert panel was comprised of 24 members equally divided into two teams, a 12-member panel for Reading and a 12-member panel for Mathematics. Among the panelists, 18 are classroom teachers, six of whom are chairs in their respective departments; several teach at nearby colleges. Two panelists are administrators (a high school principal and a curriculum supervisor), two are PSSA coaches (a literacy coach and a mathematics coach), one is an instructional specialist, and one is an educational assessment specialist.

Panelists' experience with curriculum development and the PSSA is extensive and varied. As noted, two are PSSA coaches in their schools; while others have developed and taught PSSA refresher and remediation courses in their content areas. Twenty-two reported having experience with school-level curriculum development, and half reported direct experience with development of their districts' local assessments. Eight panelists participated in development of the PSSA on tasks such as item analysis. Panelists' overall years of experience in education ranged from 7 to 37 years. Together they brought nearly 450 years of experience to the coding task.

Purposes and Procedures for Expert Panelists

Members of the expert panels traveled to State College to evaluate assessment materials and practices. Content experts were first provided with a brief overview of the purposes of the local assessment study. After introductions by members of the panel and The Penn State Research Team, experts completed a pretest to assign codes to a sample of assessment practices used by school districts in their evaluation of the proficiency standards. The practices survey was constructed with 42 rating scale entries that emerged based on The Penn State Research Team's comprehensive review of information sent by districts as to how they implemented procedures using their assessment materials to determine proficiency in Mathematics and in Reading (see Appendix E). After completion of the pretest practices survey, experts separated to complete content-area specific coding training. This training, as next presented, was conducted to prepare experts to determine whether materials submitted by districts could result in data that could determine whether students met Mathematics or Reading proficiency standards. It is important to note that districts were not asked to submit actual responses of students. As such, panel members did not evaluate student performance relative to the standards. Panel members only examined assessment materials sent to the PDE.

The Training Session

Materials were selected for training to cover the variety of types of assessments sent by school districts (See Section II, Table 4). Some materials were included in training because they were frequently submitted by districts as a local assessment. Examples of these assessments included 4Sight, the Intermediate Unit 8 (IU8) test, and Study Island. Others, however, were selected due to their unique nature (e.g., curriculum, set of district-made midterm and final examinations). To prepare materials for coding by the expert panelists, we categorized all materials sent by school districts into four groups of assessments. These groups were labeled as T1 and T2 for types of tests sent and C1 and C2 for two categories of curriculum materials. Descriptions of each of these assessment categories are as follows:

- T1: Local assessments used by districts to measure proficiency on the standards (e.g., a district-created test used with high school seniors who failed to meet proficiency on the 11th grade PSSA and the 12th grade retake exam) with at least some description as to implementation of assessment to determine proficiency levels
- T2: Local assessments used by districts but with no general or detailed description of how they are used as a measure of proficiency (e.g., final exams for senior level English and Mathematics courses but with no accompanying explanation)
- C1: Curriculum materials used by districts to measure proficiency on the PA standards (e.g., PSSA remediation course description and related course and assessment materials used to measure proficiency) with at least some description as to implementation of assessment to determine proficiency levels
- C2: Curriculum materials used by districts but with no general or detailed description about how they are used to measure proficiency (e.g., course syllabi for senior level English or Math courses but with no accompanying explanation)

After all assessment materials were categorized in this manner, we created crates of materials for panelists organized by category type and ordered by randomly-generated numerical code. Specifically, for both Mathematics and for Reading, four organized systems of materials were prepared. For each content area, the system of materials was assigned to 1 of 4 expert panel groups (i.e., 8 teams total). Each expert panel group was represented by three members. The national advisers in Mathematics and Reading assessment, Kim Gattis and Julie Coiro, helped to prepare instructions for training so that all members in a group were directed to review each folder of material together and to arrive at a consensus code for rating the degree to which the materials in the folder could serve as a means to validly measure the proficiency standards.

We distributed each type of material (T1, T2, C1, C2) across the eight teams. The experts were blind to this coding scheme so they approached all materials with the same mindset. As mentioned previously, each district's *folder* was coded as the unit of analysis. Specifying the folder as the unit of analysis was important since it was possible that some districts could include just one test that would include all content areas listed in the standards. Comparatively, it was often the case that districts submitted sets of tests where each assessment represented one content area (e.g., Algebra, Geometry) referenced in the proficiency standards. Therefore, we deemed the set of materials (i.e., the folder), whether a single source or multiple lessons/tests, as the best representation of information to define each school district's local assessment.

Coding Procedures for Alignment of Materials to the Proficiency Standards

In two separate rooms, the panel experts in Mathematics and in Reading assigned consensus codes to folder materials. Experts in Mathematics education worked with the national adviser in Mathematics assessment, Kim Gattis, in one conference room. Likewise, experts in Reading

education worked in another conference room with the director of the project, Rayne Sperling. In each location, the four expert teams assigned codes independently. As such, their dialogues were independent of one another. Kim Gattis and Rayne Sperling intervened during the coding sessions only as panelists had any questions about procedures. Therefore, neither provided any instruction or recommendations as to code assignment, nor did they confirm any sense of accuracy given a consensus code recorded by an expert team.

Rating scale materials. Each team of experts was provided two forms to facilitate coding. First, they were given the list of the proficiency standards to which they could refer as they examined each folder of materials (See Appendix C for references of the proficiency standards in Mathematics and in Reading, respectively).

Each expert rating group had one code-assignment sheet. As displayed in Appendix D, this rating scale sheet presented the folder code in the first column. The second column presented material code(s). The third column presented letters (i.e., A, B, C, D, E) that coincided with the specific list of standards against which the materials submitted by districts could be evaluated as a means of determining proficient performance. The fourth column presented space for ranking proficiency. In the final column, expert panel group members could record comments. In many instances, group members documented reasons why they assigned a particular rating scale code. Additionally, for several sets of folder entries, the experts conveyed that they viewed the assessment materials as excellent ways to measure proficiency. The scale used by experts to judge the degree to which the assessment materials aligned to proficiency standards is provided in the left hand column of Table 6.

Table 6. Examples of Materials Coded by Expert Panels by Level.

	Mathematics	Reading
0 = No content areas represented; no alignment of outcomes to standards.	<ul style="list-style-type: none"> • No assessment information provided • Calculus Final Exam • Content area assessment used as Mathematics proficiency (e.g., Science test or English exam) • Assessment with errors in items or answers • Portfolio without Mathematics content • Curriculum materials with no observable outcome measure 	<ul style="list-style-type: none"> • No assessment information provided • Below grade level assessments (e.g., 8th grade Reading test) • Isolated vocabulary test • Isolated grammar test • Content area assessment used as reading proficiency (e.g., American Government assessment or Science test) • Curriculum materials with no observable outcome measure • Writing assessments

		used as measure of reading proficiency
1= Some content areas represented; some outcomes are aligned.	<ul style="list-style-type: none"> • Course overview or table of contents, but aligned to proficiency standards • Assessment that does not match standards • Assessment above eligible content that may include some standards • Assessment far below grade level that may include some standards • Assessment that includes low and high level items in a single content area (e.g., Trigonometry) • A single midterm or final exam for one particular content area (e.g., Algebra or Geometry). 	<ul style="list-style-type: none"> • Standardized test that does not align to PA Standards • Above or below grade level assessment with or without text that matches some standards • Good example that only partially addresses limited standards • No variety of texts, for example, all narrative without attempt to measure standards • Collection of tests not aligned to standards • Assessment with emphasis on recall not reading
2=Many to most content areas represented; most outcomes are aligned.	<ul style="list-style-type: none"> • Assessment or collection of assessment that may be missing only one area (e.g., Geometry or Probability) • Collection of chapter tests and quizzes • Assessment that includes content beyond the standards • Assessment that has all standards but students are given option to complete sections 	<ul style="list-style-type: none"> • Many but not all standards addressed • Inappropriate text materials but attempt to meet standards • Assessment that addresses outdated standards • Lacks systematic coverage of standards (e.g., may not have propaganda/bias, or poetry as part of assessment) • Released PSSA items
3=All content areas represented; all outcomes are aligned.	<ul style="list-style-type: none"> • An assessment or a collection of assessments that include at least one item for each standard (Even though some students may not take all assessments) 	<ul style="list-style-type: none"> • Grade level text(s) present with items/tasks that meet all proficiency standards. • Performance

	<ul style="list-style-type: none"> • 4-Sight assessment • IU 8 assessment • Study Island assessments. • An assessment that covers all proficiency standards as either one test or a collection of tests with emphasis on use of knowledge in problem solving. 	<p>assessment representative of standards</p> <ul style="list-style-type: none"> • Study Island assessments
--	---	--

When data were available, each district was given a materials rating for Mathematics and for Reading. When interpreting the ratings of district materials, there are a few important considerations. First, in cases where a district used more than one assessment, the materials rating recorded represented the *highest* rating given to any assessment the district reported using. For example, if a district used a locally-developed test that experts rated as a ‘1’ and another standardized assessment, given a ‘2’, a ‘2’ was recorded as the district materials rating. Second, the codes for materials were assigned based upon the contents of a district *folder* as the unit of analysis. Therefore, if a district provided multiple assessments and across the assessments all proficiency standards were addressed, then the district would receive a ‘3’. It may be, however, that not every student would take every assessment represented in the district folder. Third, no consideration was made for number of items or depth of understanding required of students in answering the items. If any item(s) or task(s) that represented a proficiency standard were present within the materials then they were coded as such.

Results

As shown in Table 7, experts in the Mathematics rating groups assigned higher average materials codes than experts in the Reading rating groups overall. The mean of 2.77 approximates the highest code that could be assigned given the scale (i.e., 3 = All content areas represented; all outcomes are aligned to proficiency standards). Table 8 provides the frequencies of materials overall and for rural, suburban, and urban districts. Because rating categories define an ordinal scale of measurement, we also present the frequencies per category in Table 13. This cross tabulation is a summary that combines information for both ratings of alignment to materials standards as well as ratings used to evaluate reported use of practices. The rating scale categories for practices appear as rows and the ratings for alignment appear in the columns.

As indicated in Table 7, 321 of the 418 (i.e., approx. 77%) folders evaluated by the Mathematics expert panel groups received a code of 3. Comparatively, Reading expert panel groups only assigned this highest rating code for 99 of 418 (i.e., approx. 24%) of the folders they reviewed. This lower percentage is reflected in their overall average (1.79), as indicated in Table 7. For both Mathematics and for Reading, averages reported for type of district looked similar to the overall averages presented in Table 7 for each discipline. Some districts did not provide

materials with their practice or practice with their materials. These districts therefore, although they responded to the PDE requests, will not have data for one of these category codes. These districts' information is represented as 'missing' in subsequent tables.

Table 7. Materials Alignment Means and Standard Deviations of Scores for Reporting Districts.

Group (n)	Mathematics		Reading	
	Mean (S.D.)	Missing	Mean (S.D.)	Missing
All (418)	2.77 (0.57)	31	1.79 (0.95)	36
Rural (190)	2.78 (0.58)	18	1.75 (1.01)	23
Suburban (196)	2.79 (0.53)	9	1.81 (0.90)	9
Urban (32)	2.61 (0.79)	4	1.82 (0.86)	4

Score range for all groups is 0 – 3.

Table 8. Reading and Mathematics Materials Alignment Scores for Reporting Districts.

Mathematics

Group (n)	0	1	2	3	Missing
All (418)	6 (.01)	12 (.03)	48 (.11)	321 (.77)	32 (.08)
Rural (190)	4 (.02)	2 (.01)	21 (.11)	145 (.76)	18 (.09)
Suburban (196)	1 (.01)	7 (.04)	23 (.12)	156 (.80)	9 (.05)
Urban (32)	1 (.03)	2 (.06)	4 (.13)	20 (.63)	5 (.16)

Reading

Group (n)	0	1	2	3	Missing
All (418)	39 (.09)	105 (.25)	138 (.33)	99 (.24)	37 (.09)
Rural (190)	21 (.11)	47 (.25)	51 (.27)	48 (.25)	23 (.12)
Suburban (196)	15 (.08)	54 (.28)	71 (.36)	47 (.24)	9 (.05)
Urban (32)	3 (.09)	4 (.13)	16 (.50)	4 (.13)	5 (.16)

Reliability of Rating Scale Codes

For many commonly used assessments, consensus was established among the expert panels. For example, Study Island was aligned as a '3' for both Mathematics and Reading experts. For approximately 20 percent of the remaining school district folders that included less-frequently-used assessments, two panelist groups provided ratings in an effort to establish rating agreement. A procedure was implemented such that every 5th folder was randomly assigned to two expert groups. The two expert groups varied such that all combinations of groups were represented.

Kappa coefficients were then computed to determine the consistency between rating pairs given the folders assigned to one group which were also evaluated by 1 of the 3 other expert panel groups. Specifically, eight Kappa coefficients were computed, one per expert panel group across

both disciplines (i.e., Mathematics and Reading). The eight coefficients ranged from .53 to .82 for these less frequently-occurring materials. In summary, therefore, while members within panel groups assigned ratings with 100% (i.e., consensus) agreement, across-group ratings for materials were not always the same. These variations in scoring assignment by groups who rated the same sets of materials may be related to both the amount and variety of information submitted by districts and assembled into their folders as units of analysis. Further, the between-group variance of the randomly assigned expert panels may also have been a factor in coding.

District-Reported Practices as Measures of Proficiency

As mentioned previously, experts also completed a practices survey. This survey was administered prior to rating the degree to which materials aligned to the proficiency standards as well as directly after coding by expert panel groups was completed. The 42 survey items were constructed by Project Directors, Sperling and Kulikowich, based on an exhaustive review of the practices reported by school districts regarding use of their assessment materials to determine proficiency status (see Appendix E). The entries were designed to cover the range of materials submitted, from curricula to district-developed tests. Additionally, the project directors read each reported practice saved in the research database twice. This was done to present actual language on the survey used by the districts as their reported practices. For example, several districts reported that students could retake an examination multiple times until proficiency level was reached. Other districts, by comparison, reported that students had a limited number of opportunities to retake an examination. After so many attempts, their diplomas would be withheld if proficiency status was not reached. Every effort was made to capture these types of distinctions in the construction of the practice survey stems.

The Rating Scale for Reported Practices

As with the codes for alignment of materials to the proficiency standards, each stem on the survey was evaluated using a four-point categorical scale. A score of 0 indicated that “the practice as reported **cannot** ensure proficiency level in Mathematics and/or Reading is met.” A score of 1 indicated that the information provided “is **insufficient** to determine whether proficiency in Mathematics and/or Reading is met.” A score of 2 indicated that there “are **some good** practices reported to determine proficiency level; however, more information is needed.” A score of 3 represented reported practices that were deemed “a **valid** system of procedures to determine proficiency in mathematics and/or reading.”

Again, each panelist rated each practice statement ($n = 42$) twice, once before evaluating the materials used by the districts (i.e. tests and curricula), and once after the materials were evaluated. Panelists were also invited to record comments about the practices. Three measures of central tendency were calculated: the mean, median, and mode. Standard deviations were between .5 and 1, making the mean a poor measure. Mode was used to decide the classification of the practice, with the median as an additional source of information when needed. The median and mode were identical for most practices. Mathematics and Reading experts agreed on the viability of many reported practices, especially those practices that received either lower or higher scores.

Panel Discussions and Consensus Agreement on Practices

After completion of the posttest survey for practices, the national adviser in Mathematics assessment, Kim Gattis, and Co-Project Director Kulikowich, discussed the survey ratings with the expert panelists in Mathematics. Co-Project Director, Sperling, had a similar discussion with the Reading specialists. These debriefing sessions lasted approximately 30 minutes. The purpose of the discussions was to ask panelists to review the practice statements and to come to consensus as to those reported practices that absolutely could not be implemented by a district to determine proficiency as well as those procedures that could be considered “best practices” in evaluating proficiency in reliable and valid ways.

District Follow-up Procedures

As noted, when districts initially responded to the PDE requests for information regarding Mathematics and Reading local assessments, there was great variance in the types of materials submitted by districts. There was also variance in the amount of detail that districts provided regarding their local assessment practices. Some districts provided detailed, step-by-step procedures for determining proficiency for students who did not reach proficiency on the senior retake PSSA. Other districts reported simply that students ‘received tutoring’ or ‘took a class’ or ‘took an alternative assessment.’ In an effort to obtain additional information to assure the accuracy of ratings of local practice given to districts, PDE staff members phoned many districts and asked district personnel for additional details regarding their local assessments. As not all districts were able to be reached, to assure all districts were represented consistently and accurately, a follow-up survey was sent to each of the 418 districts for which information was submitted in compliance with the initial PDE Penn Link requests.

The survey form was sent electronically by PDE to the superintendents of all reporting districts. The primary purpose of the survey was to provide superintendents an opportunity to verify and/or modify the description of assessment practices associated with the materials submitted for their respective school districts. Superintendents were asked to examine the materials and practices recorded for their districts and to either confirm the practice as accurate or to make any necessary corrections to the information provided about their school district’s local assessment materials and practices. In addition, superintendents were requested to specifically address aspects of their district practice. First, superintendents reported whether students must pass the local assessment to graduate or if alternative provisions remain for students to graduate without passing the local assessment. Second, they were asked to report whether students are permitted to retake or resubmit the assessment in part or whole in order to obtain a passing score. Third, they responded to whether the local assessment used is part of course or cumulative grades that are used for graduation purposes. Superintendents returned this information to The Penn State Research Team through a filtered email address specifically designated for the return of the surveys. Each survey response was printed. All responses and any alterations were documented and entered into corresponding codes in the database. All of the new data were included with existing information and were used in the coding practices. The survey sent to superintendents is displayed in Appendix H.

Of the 418 responding districts, 327 responded to the survey and most provided additional information. Table 9 provides the coded responses of the districts overall and by type of district. A majority of districts took the opportunity to amend the information provided regarding their local assessment practices. Even when superintendents responded that the summary of their local assessment practice was accurate, they often included notes or comments with their survey responses.

Table 9. Agreement of Responses (n = 418, districts that sent materials).

Type of District	Yes (1)	No (2)	Yes-No (3)	No Response	Missing
All	161 (.39)	150 (.36)	2 (.05)	14 (.03)	91 (.22)
Rural	71 (.37)	63 (.33)	2 (.01)	9 (.05)	45 (.24)
Suburban	78 (.40)	76 (.39)	0 (0)	5 (.03)	37 (.19)
Urban	12 (.38)	11 (.34)	0 (0)	0 (0)	9 (.28)

The new data provided by the superintendents were incorporated into the existing database. These data, the sets of consensus codes, and a review of the pretest and posttest survey responses grounded the development of a coding rubric for local assessment practices. Sperling and Kulikowich used this coding rubric as they assigned a code for every practice reported by districts as found in the research database. Table 10 presents the scoring rubric and descriptive anchors used for reported practices coding. As indicated in Table 10, districts that used attendance as a measure of proficiency or did not have a proficiency requirement for graduation were coded a '0'. Other practices and anchors illustrate some of the common practices found across districts.

Table 10. Proficiency Codes and Representative Practices.

To code practices for Mathematics and Reading for each of the districts, we relied on the following rubric grounded in the expert panelists' consensus ratings. A practice was not given a rating higher than the alignment rating. That is, if a local assessment was given an alignment rating of '2', it could not receive a practice rating greater than '2'.

Proficiency Code	Representative Practices
<p>0 = The practice as reported <i>cannot</i> ensure proficiency level in Mathematics and/or Reading is met.</p>	<ul style="list-style-type: none"> • No proficiency requirement for graduation • Graduation project serves as proficiency standard • Summer course attendance after graduation • Remedial course attendance without any explanation of how grades are assigned • Courses without an objective measure or portfolios with no standard rubric system • Tests not aligned to standards • Cumulative grades for which aligned objective measure is not weighted more than 50% of grade • Unlimited retakes of unsecured assessments in part or full
<p>1 = The information about practices provided is <i>insufficient</i> to determine whether proficiency in Mathematics and/or Reading is met.</p>	<ul style="list-style-type: none"> • Retakes of aligned assessments that exceed the number of forms of the assessment (if perfectly aligned) • When information was not adequate to fully evaluate the practice and the alignment was at least a '2' • Ambiguous security or retake opportunities of an aligned assessment
<p>2 = There are <i>some good</i> practices reported to determine proficiency level; however, more information is needed.</p>	<ul style="list-style-type: none"> • When information was not adequate to fully evaluate the practice and the alignment was a '3' • Valid practice with a alignment of '2' • Fully aligned assessment with ambiguous security or retake opportunities • Fully aligned and secure assessment with unclear information about how the scores or ratings determine that proficiency levels have been met • Aligned assessment without clear criteria for proficiency

<p>3 = The practices reported represent a <i>valid</i> system of procedures to determine proficiency in Mathematics and/or Reading.</p>	<ul style="list-style-type: none"> • Fully aligned secure assessment with retakes that do not exceed number of assessment forms • Fully aligned assessment with randomly-generated (e.g., computer-adaptive, computer-based) items • Fully aligned and secure assessment that has scores or cut-off points to measure proficiency
---	--

Therefore, when complete data were available, school districts that submitted materials and reported practices received codes both for alignment of their materials to proficiency standards as well as ratings of the degree to which the reported practices could determine whether proficiency levels of performance had been attained. Table 11 complements these results with average ratings for the practices overall for both Mathematics and Reading. Table 12 provides the summary of local assessment practices overall and by district type.

A few observations can be made. First, for both Mathematics and Reading, averages for the practice codes are lower than those for the material alignment codes. Second, these averages are consistent across type of district. Third, as with the materials ratings, the practice ratings were higher for Mathematics than for Reading. As noted in Table 12, 31 districts' Mathematics practices were coded a '3' while 19 Reading practices reported by districts received the highest rating.

Table 11. Local Assessment Practices Means and Standard Deviations of Scores for Reporting Districts.

Practice Scores

Group (n)	Mathematics		Reading	
	Mean (S.D.)	Missing	Mean (S.D.)	Missing
All (418)	0.90 (0.97)	8	0.69 (0.91)	10
Rural (190)	0.98 (0.98)	5	0.75 (0.95)	7
Suburban (196)	0.84 (0.95)	3	0.65 (0.86)	3
Urban (32)	0.75 (1.02)	0	0.56 (0.95)	0

Score range for all groups is 0 – 3.

Table 12. Reported Practices Scores for Reporting Districts.

Mathematics

Group (n)	0	1	2	3	Missing
All (418)	184 (.44)	117 (.28)	77 (.18)	31 (.07)	9 (.02)
Rural (190)	74 (.39)	58 (.31)	36 (.19)	17 (.09)	5 (.03)
Suburban (196)	92 (.47)	52 (.27)	37 (.19)	12 (.06)	3 (.02)
Urban (32)	18 (.56)	7 (.22)	4 (.13)	2 (.06)	1 (.03)

Reading

Group (n)	0	1	2	3	Missing
All (418)	231 (.55)	96 (.23)	61 (.15)	19 (.05)	11 (.03)
Rural (190)	98 (.52)	47 (.25)	26 (.14)	12 (.06)	7 (.04)
Suburban (196)	111 (.57)	45 (.23)	31 (.16)	6 (.03)	3 (.02)
Urban (32)	22 (.69)	4 (.13)	4 (.13)	1 (.03)	1 (.03)

In the next paragraphs we examine the relationships between alignment and practices codes for Mathematics and Reading and also explore the relationships between Mathematics and Reading materials and practices. As presented in Table 13, 31 of 418 districts had highest codes for both their alignment of assessment materials to proficiency standards as well as reported practices that could determine a valid evaluation as to whether proficiency status could be attained. Many of these districts employ the use of Study Island or 4Sight, both materials that received a ‘3’ code for Mathematics coupled with a secure assessment procedure. For Reading, this frequency was lower; 19 districts. This is in part due to experts’ consensus that students must engage with a text passage as part of the local assessment in order for the assessment to attain a score of ‘3’ for alignment. It is important to reiterate that all codes were assigned solely based on information submitted to the PDE for review as well as information reported by school district personnel

about uses of their local assessments to determine proficiency levels. As illustrated in Table 10, stringent practices were required by districts to attain a ‘3’ score on local assessment practices.

We next explored the nature of the relationship between Reading and Mathematics assessments. As presented in Table 14, 98 districts had materials that were rated a ‘3’ for both Mathematics and Reading. Regarding practices, 18 districts had both a Reading and a Mathematics practice that was coded a ‘3’. Overall 18 districts had a rating of ‘3’ for materials in both Reading and Mathematics and a rating of ‘3’ for both Mathematics and Reading practices.

Table 13. Reported Practices by Alignment Ratings for All Reporting Districts.

Mathematics Practices by Alignment Overall (n = 383, 91.6% of reporting districts)

Alignment→ Practices ↓	0	1	2	3
0	6 (.016)	10 (.026)	35 (.091)	108 (.281)
1	0 (.000)	1 (.003)	11 (.029)	105 (.273)
2	0 (.000)	0 (.000)	2 (.005)	74 (.193)
3	0 (.000)	0 (.000)	0 (.000)	31 (.081)

Reading Practices by Alignment Overall (n = 376, 90.0% of reporting districts)

Alignment→ Practices ↓	0	1	2	3
0	38 (.101)	73 (.194)	74 (.197)	17 (.045)
1	0 (.000)	30 (.080)	39 (.104)	26 (.069)
2	0 (.000)	0 (.000)	23 (.061)	37 (.098)
3	0 (.000)	0 (.000)	0 (.000)	19 (.051)

Note: Cell entries are frequency counts, with proportion of n for chart, not overall n (e.g., proportions calculated on the first table are out of the 383, for which we have complete Mathematics alignment by practices information, not on the 418 total of reporting districts).

Table 14. Relations Between Alignment and Reported Practices in Mathematics and Reading.

Alignment: Reading by Mathematics Alignment Overall (n = 378, 90.4% of reporting districts)

Reading→ Mathematics ↓	0	1	2	3
0	6 (.016)	0 (.000)	0 (.000)	0 (.000)
1	2 (.005)	6 (.016)	3 (.008)	0 (.000)
2	6 (.016)	22 (.058)	17 (.045)	1 (.002)
3	24 (.063)	74 (.195)	118 (.312)	98 (.259)

Practices: Reading by Mathematics Alignment Overall (n = 407, 97.1% of reporting districts)

Reading→ Mathematics ↓	0	1	2	3
0	180 (.442)	3 (.007)	1 (.002)	0 (.000)
1	41 (.101)	73 (.179)	1 (.002)	0 (.000)
2	5 (.012)	17 (.042)	54 (.133)	1 (.002)
3	4 (.010)	3 (.007)	5 (.012)	18 (.044)

Conclusion

This study represents the most comprehensive examination to date of Pennsylvania district-level local assessments that serve as an alternative high school graduation requirement. In its review of the hundreds of assessments provided by 418 school districts statewide, the research team noted diversity in the type, format, and quality of materials that districts submitted and reported that they use to determine students' proficiency levels in Mathematics and Reading to meet graduation requirements.

Review of the materials by expert panels found variance in the degree of connection between the assessments and 11th grade proficiency standards. While Mathematics assessments were more often rated as aligned, ratings were considerably lower for Reading assessments.

Because alignment is a necessary but insufficient criterion for determining validity of an assessment, the research team also evaluated the administration and use of local assessment practices based on information provided by district staff. The researchers noted considerable variance among districts.

Finally, the research team examined ratings for alignment and practice and reported that for Mathematics, thirty-one school districts (8.1%) had ratings of '3' for both, while nineteen (5.1%) districts received ratings of '3' for both alignment and practice in Reading. Based on criteria established by the panels, evidence of alignment to standards and practices that could result in valid measures of proficiency was present from 5 percent of school districts statewide.

Appendix A: The Penn State Research Team

Rayne Sperling, Co-Principal Investigator

Jonna Kulikowich, Co-Principal Investigator

Crystal Ramsay, Project Coordinator

Toni Betaudier, Research Associate

Whitney Zimmerman, Research Associate

Kelli Higley, Research Associate

Appendix B: Fields that Correspond with Submitted Database

- A. Materials (from Summer 2008 PDE request) received:
 - a. 1= Yes, we did receive material.
 - b. Blank= We did not receive any materials.
- B. Date: Date stamped by PDE. If district sent something but there was no visible date stamp, field was left blank. If materials were submitted based on follow-up phone calls from PDE, a date was not indicated for those materials.
- C. School ID: Randomly-generated and assigned school ID number
- D. See ID#: For school districts with multiple schools that only sent one set of materials; This field tells which School ID number to refer to for the information for the school. Schools with same numbers are from same districts.
- E. School AU number: PDE-generated school number
- F. School Name: Name of the high school; linked to EDNA.
- G. Enrollment, 2006-07
- H. District enrollment, 2006-07
- I. District Count: Number of schools in the school district
- J. District quartile
- K. Fail quartile
- L. Expenditure quartile
- M. School District: Linked to EdNA
- N. IU Affiliation
- O. City
- P. County
- Q. School Type: Information taken from the Ensuring Success for All High School Graduates document provided by PDE
- R. % of graduates who did not score proficient/advanced in Mathematics AND Reading on the 2005-06 11th grade PSSA AND the 12th grade retest: Information taken from the Ensuring Success for All High School Graduates document provided by PDE (District level data)
- S. Number of graduates who did not score proficient/advanced in Mathematics AND Reading on the 2005-06 11th grade PSSA AND the 12th grade retest: Information taken from the Ensuring Success for All High School Graduates document provided by PDE (District level data)
- T. Number of graduates 2006-2007: Information taken from the Ensuring Success for All High School Graduates document provided by PDE (District level data)
- U. Per-pupil spending: Information taken from the “2006-07 Total Exp per ADM” document provided by PDE. (District level data)
- V. Free and reduced lunch: % free and reduced enrollment from “2007-08 Building Data” document provided by PDE (District level data)
- W. Administrator: Name of administrator; Some of these names are linked to EDNA.
- X. Title of administrator
- Y. Contact Person: Based on the cover letters sent by schools, some indicated who the contact person is regarding the material sent and who to contact if further information is needed. In cases where no contact person is listed, defer to Column W.
- Z. Contact person (Column Y) phone number

- AA. District PSSA policy included:
- a. 0 = No policy sent, or no policy can be inferred from the materials that were sent.
 - b. 1 = Yes, District stated X as their policy; cannot not infer that a proficiency standard is required. This is what the school says they used.
- AB. District graduation policy included
- a. 0 = No policy sent
 - b. 1 = Sent policy
- AC. Tests provided for Mathematics?
- a. 0 = No, test(s) not provided
 - b. 1 = Yes, test(s) provided
- AD. Scoring system provided for Mathematics test(s) (e.g., rubric, scoring key)?
- a. 0 = No, scoring system not provided
 - b. 1 = Yes, scoring system provided
- AE. Multiple-choice items included on Mathematics assessments?
- a. 0 = No
 - b. 1 = Yes
- AF. Open-ended, or constructed response, items included on Mathematics assessments?
- a. 0 = No
 - b. 1 = Yes
- AG. Standardized/published tests in Mathematics?
- a. 0 = No
 - b. 1 = Yes
- AH. Portfolio/project in Mathematics?
- a. 0 = No
 - b. 1 = Yes
- AI. Scoring system provided for Mathematics portfolio/project (e.g., rubric, scoring key)?
- a. 0 = No
 - b. 1 = Yes
- AJ. Mathematics curricula?
- a. 0 = No
 - b. 1 = Yes
- AK. Mathematics course?
- a. 0 = No
 - b. 1 = Yes
- AL. Mathematics tutoring?
- a. 0 = No
 - b. 1 = Yes
- AM. Tests provided for Reading?
- a. 0 = No, test(s) not provided
 - b. 1 = Yes, test(s) provided
- AN. Scoring system provided for Reading test(s) (e.g., rubric, scoring key)?
- a. 0 = No, scoring system not provided
 - b. 1 = Yes, scoring system provided
- AO. Multiple-choice items included on Reading assessments?
- a. 0 = No
 - b. 1 = Yes

- AP. Open-ended, or constructed response, items included on Reading assessments?
- 0 = No
 - 1 = Yes
- AQ. Standardized/published tests in Reading?
- 0 = No
 - 1 = Yes
- AR. Portfolio/project in Reading?
- 0 = No
 - 1 = Yes
- AS. Scoring system provided for Reading portfolio/project (e.g., rubric, scoring key)?
- 0 = No
 - 1 = Yes
- AT. Reading curricula?
- 0 = No
 - 1 = Yes
- AU. Reading course?
- 0 = No
 - 1 = Yes
- AV. Reading tutoring?
- 0 = No
 - 1 = Yes
- AW. School District: Linked to EdNA
- AX. District PSSA proficiency policy: Description of what districts do, given materials they sent.
- AY. Policy calls information (based on calls made by PDE personnel)
- AZ. Proficiency information for superintendents (text included in December follow-up letters from Secretary Zahorchak to superintendents)
- BA. Materials sent: List of school district materials received by Penn State Research Team (list included in December follow-up letter to superintendents)
- BB. Return response
- 0 = No, superintendent response was not returned.
 - 1 = Yes, superintendent response was returned to follow-up requested by **December 23, 2008.**
 - 2 = Yes, superintendent response was returned to follow-up requested by **December 22, 2008.**
- BC. Agreement
- 1 = Yes, District agreed with proficiency information (Column AY), as sent.
 - 2 = No, District did not agree with proficiency information (Column AY), as sent.
 - 3 = District responded both Yes and No to question of agreement.
 - NR (No Response) = District returned letter as requested, but provided NO information to any questions.
- BD. Summary District response: District's amended response to Column AY. (Blank field indicates no amendment.)
- BE. Summary of Part C revisions: District's summary response to questions of whether local assessment is required to graduate, whether retakes are allowed, and the

degree to which local assessment scores are factored into graduation decisions.
(Blank field indicates no additional response.)

BF. Materials revisions: District's revisions to Column AZ. (Blank field indicates no amendment.)

BG. Other notes: Notes included by The Penn State Research Team or additional notes supplied by school districts.

BH. Mathematics alignment of materials to proficiency standards as coded by Mathematics panelists

- a. 0 = No content areas represented; no alignment of outcomes to standards
- b. 1 = Some content areas represented; some outcomes are aligned
- c. 2 = Many to most content areas represented; most outcomes are aligned
- d. 3 = All content areas represented; all outcomes are aligned

BI. Reading alignment of materials to proficiency standards as coded by Reading panelists

- a. 0 = No content areas represented; no alignment of outcomes to standards
- b. 1 = Some content areas represented; some outcomes are aligned
- c. 2 = Many to most content areas represented; most outcomes are aligned
- d. 3 = All content areas represented; all outcomes are aligned

BJ. Mathematics reported practices as coded by Mathematics panelists

- a. 0 = The practice cannot ensure proficiency level in Mathematics is met.
- b. 1 = The information about practices provided is insufficient to determine proficiency level.
- c. 2 = There are some good practices reported to determine proficiency level; however, more information is needed.
- d. 3 = The practices reported represent a valid system of procedures to determine proficiency in Mathematics.

BK. Reading reported practices as coded by Reading panelists

- a. 0 = The practice cannot ensure proficiency level in Reading is met.
- b. 1 = The information about practices provided is insufficient to determine proficiency level.
- c. 2 = There are some good practices reported to determine proficiency level; however, more information is needed.
- d. 3 = The practices reported represent a valid system of procedures to determine proficiency in Reading.

BL. Notes

Appendix C: Proficiency Standards Experts Used As Reference

MATHEMATICS
Proficiency Standards
Performance Level Descriptors*

Practical and Real World Problems

- A. Demonstrates understanding of and ability to use different forms of real numbers; uses estimation and operations on real numbers to solve multi-step problems, including problems involving proportional relationships.
- B. Uses formulas to solve problems involving two and three-dimensional measurements of standard and composite geometric shapes; manipulates multi-step formulas; demonstrates the relationships of a change in length and changes in perimeter, circumference, area and volume.
- C. Uses properties and relationships of parts of circles, triangles and quadrilaterals to solve problems; applies the concepts of congruence and similarity in problem-solving settings; describes measures and relationships (perpendicular/parallel with respect to slope) of segments in a coordinate plane.
- D. Writes algebraic expressions and linear and nonlinear equations to describe graphs or patterns; solves problems represented as systems or compound inequalities or quadratic equations; simplifies algebraic expressions in problem-solving situations.
- E. Reads and constructs graphical representations of data; uses box-and-whisker plots to represent data; draws conclusions based on measures of central tendency; uses counting techniques to determine probability; makes predictions based on data sets, probability, graphs and scatter plots.

Material Codes	Proficiency Standards
1=Published test	0 = No content areas represented; no alignment of outcomes to standards.
2=Individual local test	1 = Some content areas represented; some outcomes are aligned.
3=Collection of local tests	2 = Many to most content areas represented; most outcomes are aligned.
4=Curriculum or course materials	3 =All content areas represented; all outcomes are aligned.
5=On-line or published tutorial or self-paced study	
6=Project or portfolio standards or rubric	
7=Other: Note in comments	

*<http://www.pde.state.pa.us/a and t/lib/a and t/Grade 11 Math PLDs.pdf>

READING
Proficiency Standards
Performance Level Descriptors*

Grade-level fiction and nonfiction

- A. Applies a variety of strategies to determine meanings of words, including synonyms and antonyms, using context clues and word parts.
- B. Makes inferences, draws conclusions, and generalizes, using textual support; makes within and among text-to-text connections.
- C. Identifies and explains main ideas; summarizes text.
- D. Interprets and analyzes:
 - purpose of text (e.g., narrative, informational)
 - organizational patterns (e.g., sequencing, compare/contrast)
 - relationships among literary elements (character, setting, plot, theme, tone, style, mood, symbolism)
 - use of figurative language (e.g., simile, metaphor), author’s style, and point of view.
- E. Interprets and analyzes the use of facts and opinions and analyzes the effectiveness of bias and propaganda in nonfiction.
- F. Describes and analyzes the sequence of steps in a list of directions; analyzes and evaluates graphics and charts.

Material Codes	Proficiency Standards
1=Published test	0 = No content areas represented; no alignment of outcomes to standards.
2=Individual local test	1 = Some content areas represented; some outcomes are aligned.
3=Collection of local tests	2 = Many to most content areas represented; most outcomes are aligned.
4=Curriculum or course materials	3 = All content areas represented; all outcomes are aligned.
5=On-line or published tutorial or self-paced study	
6=Project or portfolio standards or rubric	
7=Other: Note in comments	

*[http://www.pde.state.pa.us/a and t/lib/a and t/Grade 11 Reading PLDs.pdf](http://www.pde.state.pa.us/a_and_t/lib/a_and_t/Grade_11_Reading_PLDs.pdf)

Appendix D: Example Proficiency Coding Sheet

Material Codes	Proficiency Standards
1=Published test 2=Individual local test 3=Collection of local tests 4=Curriculum or course materials 5=On-line or published tutorial or self-paced study 6=Project or portfolio standards or rubric 7=Other: Note in comments	0 = No content areas represented; no alignment of outcomes to standards. 1 = Some content areas represented; some outcomes are aligned. 2 = Many to most content areas represented; most outcomes are aligned. 3 =All content areas represented; all outcomes are aligned.

ID	Material Code(s)	Proficiency Standards	Proficiency	Comments
		A B C D E		
		A B C D E		
		A B C D E		
		A B C D E		

Mathematics Proficiency Codes

A=Numerical properties and number relationships

B=Measurement of geometric shapes

C=Use of geometric concepts in problem-solving

D=Algebraic concepts and problem solving

E=Data analysis, probability, and statistics

Material Codes	Proficiency Standards
1=Published test 2=Individual local test 3=Collection of local tests 4=Curriculum or course materials 5=On-line or published tutorial or self-paced study 6=Project or portfolio standards or rubric 7=Other: Note in comments	0 = No content areas represented; no alignment of outcomes to standards. 1 = Some content areas represented; some outcomes are aligned. 2 = Many to most content areas represented; most outcomes are aligned. 3 =All content areas represented; all outcomes are aligned.

ID	Material Code(s)	A B C D E F	Proficiency	Comments
		A B C D E F		
		A B C D E F		
		A B C D E F		
		A B C D E F		

Reading Proficiency Codes

- A. Vocabulary in Context
- B. Inferences, Conclusions, Connections & Support
- C. Main Idea and Summarize

- D. Interpret & Analyze Style/Structure
- E. Interpret & Analyze Fact/Opinion/Bias
- F. Analyze sequential directions & graphs/charts

Appendix E: Practice Rating Form Administered to Expert Panels

Expert Panel ID: _____

Practices to Meet Proficiency Level

In addition to curriculum and test materials, many school districts provided information about the practices they implement to ensure students' performance level in mathematics and reading is proficient. Below are a series of practices used by the school districts. Practices are the means by which school districts gather evidence to determine proficiency.

Directions: Examine each practice and evaluate it using the following scale:

- 0 = The practice *cannot* ensure proficiency level in Mathematics and/or Reading is met.
- 1 = The information about practices provided is *insufficient* to determine whether proficiency in Mathematics and/or Reading is met.
- 2 = The are *some good* practices reported to determine proficiency level; however, more information is needed.
- 3 = The practices reported represent a *valid* system of procedures to determine proficiency in Mathematics and/or Reading.

	Coursework/Curriculum with or without Testing Practices	Rating Scale Evaluation	Comments
1	Students must achieve proficiency levels in their required courses <i>AND</i> demonstrate proficiency on district-made tests.	0 1 2 3	
2	Enrollment in a senior-level math or reading class. Students must earn a grade of C or higher in that class based on assignments, quizzes, midterm, final exams, and/or projects.	0 1 2 3	
3	Enrollment in a remediation course using Study Island. Students take the local assessment comprised of past PSSA items. If students do not reach proficiency levels, then tutoring continues. Students retake examination until proficiency levels are met.	0 1 2 3	

4	Completion of laboratory courses using 4Sight and Terra Nova assessments as diagnostic tools. Assessment scores determine proficiency level.	0	1	2	3	
5	One of the following can be used to determine proficiency: a) passing a math and/or reading proficiency PSSA Coach class; OR , b) Score at least proficient on the 4Sight assessment.	0	1	2	3	
6	Enrollment in a Study Island remediation course. No testing required.	0	1	2	3	
7	Enrollment in a remediation course where students complete released items from the PDE website, PLATO, 4Sight, and the PSSA Coach. No exit test is administered.	0	1	2	3	
8	The PLATO learning system is used to help students in areas where proficiency levels have not been obtained. The mastery tests are then used to measure proficiency levels. Students must score 80% or higher on EACH mastery test.	0	1	2	3	
9	Completion of a remediation course followed by a final exam in the Senior Year in Algebra 2 or English 12.	0	1	2	3	
10	Attendance at a Saturday weekend class for 15 hours without testing.	0	1	2	3	
11	Enrollment in a <i>Princeton Review</i> Course. After class, administration of an alternative assessment. Students can enroll in class and take the assessment up to two times. If not meeting levels of proficiency on second time, diploma is withheld.	0	1	2	3	

12	Tutoring and remediation are provided to students using <i>Measuring Up</i> until teachers believe proficiency levels are met. After tutoring, students take a local, district-made assessment to determine levels of proficiency.	0	1	2	3	
13	Passing a course in summer school. No testing is required.	0	1	2	3	

14	Enrollment in either a PSSA remediation course or use of 4Sight as an alternative assessment if student does not pass the remediation course. The 4Sight alternative assessment can be taken up to 3 times. To meet the proficiency requirement, students must pass this assessment 2 out of 3 times.	0	1	2	3	
15	For reading, proficiency levels can be met by passing an AP English or English 12 course. For math, proficiency levels can be met by passing an AP Calculus course.	0	1	2	3	
16	Non-proficient students will be provided a plan to assist them in achieving the standards. Students' plans may be adjusted to allow additional instruction time in the areas of deficiency. The Intermediate Unit 8 Test is then administered after the instructional plan is completed to determine proficiency.	0	1	2	3	

	Other Practices	Rating Scale Evaluation				Comments
17	The graduation project.	0	1	2	3	
18	Report card marks.	0	1	2	3	
19	Completion of an activity to show good	0	1	2	3	

	citizenship.		
20	Career inventory measure.	0	1 2 3
21	WebQuest search task activities for reading or math with a final presentation.	0	1 2 3
22	Administration of tests for grades 10 or lower.	0	1 2 3
23	Total credits earned in mathematics and/or reading that meet school district graduation requirements.	0	1 2 3

	Specific Testing Practices	Rating Scale Evaluation	Comments
24	Passing midterm and final exams in all Senior-level math and/or English classes.	0	1 2 3
25	Scores on the PSAT/SAT Quantitative and/or Verbal Subscales	0	1 2 3
26	Use of a math or reading lab proficiency portfolio. Students must include entries for all proficiency standards. A district-developed rubric is used to determine proficiency levels.	0	1 2 3
27	Demonstration of proficiency on the Metropolitan Achievement Tests in math and/or reading without test preparation.	0	1 2 3
28	Demonstration of proficiency using a district-developed local assessment (aligned to standards) that does not include released items of the PSSA.	0	1 2 3
29	Use of a computerized supervised test built to align with PSSA.	0	1 2 3
30	Administration of the Stanford Achievement 10 Test. Students can take test as many times as needed.	0	1 2 3
31	Use of dictionaries on reading local assessments.	0	1 2 3
32	Administration of the Intermediate Unit 8 test. Students have two opportunities to demonstrate	0	1 2 3

	proficiency. If they do not pass the second time, then the diploma is withheld.		
33	Inclusion of portfolio entries that contain artifacts displaying proficiency. Completion of portfolio is followed by a senior exit interview. Students present their collection of work during the interview. Scripted questions are provided to the interviewer. Selected items correspond to the standards. The interviewer asks the questions to each student on an individual basis. The interview has rubrics and the responses are analyzed by the administrative testing team.	0	1 2 3
34	A conditional diploma system. If students do demonstrate proficiency level is met in math and/or reading, then they can graduate. They must continue to take district local assessment until proficiency level is met. Students then receive official diploma.	0	1 2 3
35	Use of one of three practice tests from <i>Measuring Up</i> .	0	1 2 3
36	Use of standards-based assessments as blocks or selected subsets of items where proficiency levels are not met in reading and/or math for specific performance standards. The blocks of items are selected and scored by a district assessment team or center.	0	1 2 3
37	Administration of Study Island posttests. Students can take test as needed until proficiency level is attained.	0	1 2 3
38	Passing a GED examination.	0	1 2 3
39	Use of tests developed by states other than Pennsylvania. Test items are not released by the school district for practice.	0	1 2 3

40	Use of calculators on mathematics tests (district-made and/or standardized assessments).	0 1 2 3	
41	Completion of a district-developed local multiple-choice assessment in reading and/or math where all items scores have been examined for reliability and validity by the district. The items are secured. The school district administers the test two times each year. Students have two opportunities to demonstrate proficiency level is met.	0 1 2 3	
42	Use of a weighted point rubric for different assessment system batteries where multiple indicators are used. District-developed local assessments have highest point value followed by standardized tests (e.g., Metropolitan Achievement Test) followed by 4Sight Grade 11 Benchmark Tests).	0 1 2 3	

Appendix F: Expert Panel Evaluation Summaries

Expert Panel Evaluations of the Coding Task

Panelists were asked to rate the extent to which they felt that their professional expertise was respected; that coding materials were well-organized and the process expedient; and that the experience contributed in a meaningful way to their own professional development. Generally, panelists reported a mutual respect among fellow panelists and between panelists and The Penn State Research Team members. While some noted that the variety of materials made the task inherently challenging, they agreed that the boxes of assessments they were required to code were well-organized. Many commented on the benefits to their professional development. One panelist noted that “Learning occurred at an exponential rate for me.” Another said, “From the sharing among other professionals to learning about the education policy in the making, I learned a great deal.”

Panelists also commented on three aspects of their expert panel experience:

- The process of coding the local assessments (i.e., team process, materials)
- Reactions to materials sent by school districts (i.e., What surprised you?, In positive ways? In ways that concerned you?)
- Overall experience as a panelist

Comments relative to the process were very positive. Panelists found the team process to be especially helpful. They appreciated the collaboration, the compromise, the consensus-building, and the efficiency that such an approach afforded. Many agreed that occasional larger-group interaction complemented the team structure. They commented on the value of training at the outset, periodic regrouping and debriefing, and opportunities to discuss and clarify both the process and their thinking. In spite of this generally effective process, panelists reported that, at times, it was challenging to code consistently in the face of such variety among the local assessment materials.

Participants had strong reactions to the materials they encountered during coding. One reading panelist remarked that, “The scope of materials was staggering.” Math panelists described the materials using phrases such as “wide range of assessment,” “variance of materials,” and “extremes.” Reactions to these varied assessments were both positive and negative. Panelists noted that they found “some good assessments” and that some were “thorough and professional.” In contrast, they were troubled by “the plethora of recall materials,” “the lack of specific district assessments,” “a lack of concern for quality and true purpose of assessment,” “the basic level of many assessments,” and “that my school district does not have an alternative assessment in place.”

While panelists reported a positive experience with the coding process and mixed reactions to the materials, there was broad consensus that overall their participation in the expert panel was a valuable one. Reactions included:

- *I enjoyed this experience very much. I feel empowered to go back to my school and lead others.*

- *Very interesting experience! I learned so much and I have come to a new appreciation of the purpose of the PSSA. I know I will view my own assessments very differently from now on.*
- *Overall, this has been extremely helpful looking at all the different “proficiency” measures used by the state and collaborating with a strong group of educators.*
- *The overall experience was exceptional. I learned a vast amount of information that I will carry home with me to help not only my district, but my own classroom as well.*
- *The experience was very valuable and allowed me to develop future ideas based on my analysis. There were many great ideas for addressing performance level descriptors.*
- *Enjoyed the experience, especially the lively exchange of opinions which on some occasions caused me to change my own original opinion.*

Appendix G: Local Assessment Survey

SUMMARY OF LOCAL ASSESSMENT PRACTICES
[MM] SCHOOL DISTRICT

The information below reflects your district's submission of local assessments to the Department of Education as well as responses to any follow-up phone calls from Department staff. Please review the summary for accuracy and make any changes or corrections in the space provided. **Save your form (using district name) and return to localassess@psu.edu by December 23, 2008.** Thank you.

A. MATERIALS SUBMITTED: [MM]

B. For last year's (2008) graduating students who did not score proficient or advanced on the 11th grade PSSA (or 12th grade retest), how are the results or outcomes from these assessment materials used to determine proficiency and the awarding of a diploma?

DISTRICT RESPONSE: [MM]

- Yes, this summary is accurate
- I wish to amend and/or correct the summary (please type any changes or corrections in the space below)

C. If not addressed by the summary above, please answer the following:

- 1. Must students pass the local assessment to graduate?**
 - Yes
 - No
 - an alternative assessment or provision remains for non-proficient, regular education (non-IEP) students

- 2. Are students permitted to retake (or resubmit) the assessment in whole or in part to achieve a passing score?**
 - Yes (please identify the number of retakes permitted and describe your district's efforts to protect the security of assessment material)
 - No

- 3. Are local assessment scores factored in to course or cumulative grades that are used for graduation purposes?**
 - Yes
 - No