

Washakie County School District #1



**“BETTER ASSESSMENT LEADS
TO BETTER LEARNING”**

Body of Evidence

A K-12 Assessment System

Approved thru the WDE Peer Review: Spring 2010

NONDISCRIMINATION STATEMENT

Washakie County School District #1 does not discriminate on the basis of race, color, national origin, sex, age, or disability in admission or access to, or treatment or employment in its educational programs or activities.

Inquiries concerning Title VI, Title IX, Section 504, and the Americans with Disabilities Act may be referred to the Wyoming Department of Education, Office for Civil Rights Coordinator, 2nd floor, Hathaway Building, Cheyenne, Wyoming 82002-0050 or (307) 777-5329, or the Office for Civil Rights, Region VIII, U. S. Department of Education, Federal Building, Suite 310, 1244 Speer Boulevard, Denver, CO 80204-3582, or (303) 844-5695 or TDD (303) 844-3417. This publication will be provided in an alternative format upon request

TABLE OF CONTENTS

Section 1: OVERVIEW	
Introduction	4
District Demographics	6
Student Demographics	7
District Assessment Plan	8
Graduation Requirements	10
System of Documentation	11
Summary of the District Body of Evidence	15
History of District Body of Evidence (with Changes and Rationale)	17
Section 2: ALIGNMENT	21
Section 3: CONSISTENCY	24
Section 4: FAIRNESS	28
Section 5: STANDARD-SETTING	31
Section 6: COMPARABILITY	35
APPENDIX A	38
APPENDIX B	44
APPENDIX C	46
APPENDIX D	58
APPENDIX E	68
APPENDIX F	73

Section 1: OVERVIEW

INTRODUCTION

Washakie County School District #1 firmly supports and sustains a balanced assessment program that...

1. Is systemically monitored for public accountability,
2. Provides information to enhance instructional practices,
3. Supports effectiveness of instructional strategies,
4. Measures student achievement and mastery of skills,
5. Provides appropriate data to improve student progress,
6. Engages parents in student learning, and
7. Prepares students to be effective citizens.

The Wyoming State Board of Education requires that each school district design, implement, and sustain an assessment system to provide each student with the educational opportunity to demonstrate mastery of knowledge and skills through a variety of assessments with multiple opportunities of the nine state content areas (to include language arts, mathematics, science, social studies, foreign language, fine and performing arts, health, physical education, and career/vocational) in compliance with the Wyoming Department of Education. This system is called the Body of Evidence (BOE).

The Washakie County School District #1 assessment plan includes a body of evidence document, which specifies K-12 assessment accountability and requirements for a student to graduate from Worland High School. Chapter 31 of the Wyoming Department of Education and State Board of Education Rules and Regulations indicate that a student must meet the graduation standards, requiring the successful completion of the following components, as evidenced by passing grades or by the successful performance on competency-based equivalency examinations:

- A. Four (4) school years of English;
- B. Three (3) school years of mathematics;
- C. Three (3) school years of science; and
- D. Three (3) school years of social studies, including history, American government and economic systems and institutions.

The district developed and implemented a course-based, common assessment body of evidence, compensatory in nature, to afford students and staff latitude to learning and teaching standards in nine content areas. This approach addresses the following:

- All courses involve standards-based instruction with each course clearly tied to all standards in each of the nine content areas.
- Evidence of student proficiency includes student performance on standards-based assessments and on district common assessments.
- The student standards' report card indicates student progress in meeting proficiency in each standard of each content areas with student performance on BOE assessments targeting specific standards/benchmarks, which impact student grades in a course.

- All Wyoming Content and Performance Standards are embedded in the district standards and are assessed.
- Each body of evidence assessment has an explicit rationale and is aligned to the Wyoming State Standards.
- All students are ensured equal educational opportunities district-wide, K-12.
- All body of evidence assessments reflect consistent decisions regarding student performance on the standards.
- Students are provided multiple opportunities in a variety of assessment modalities, using multiple formats to demonstrate knowledge and skills in each of the standards in all content areas. Assessments could include common projects or other types of measures that are standards-based.
- Common assessments are given in content areas when there are multiple teachers teaching the same course or grade, which reflects consistency in decisions regarding student performance on the standards and comparability to ensure equal educational opportunities for all students.
- Students have equal opportunity to learn the standards, while providing appropriate accommodations when necessary and allowing individual teacher choice and discretion.
- Students do not have to demonstrate proficiency on every benchmark to be proficient on the standard, allowing the majority of higher assessment scores to offset (compensate for) lower scores.
- Appropriate accommodations are available for qualifying students to have an equitable opportunity to demonstrate learning.

This document describes specifically how the district addresses the District K-12 Body of Evidence and the Washakie County School District #1 Graduation Policy.

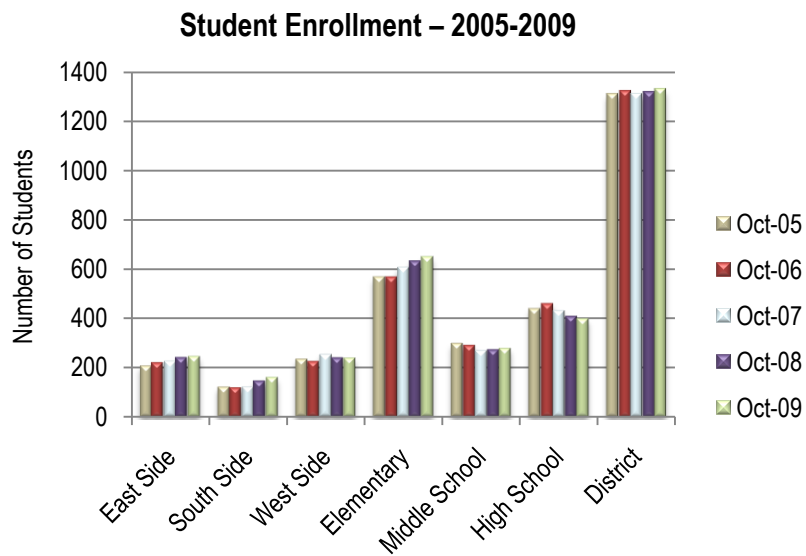
DISTRICT DEMOGRAPHICS

October 2009

Washakie County School District #1 is located in Worland, Wyoming and includes three elementary schools (K-5), one middle school (6-8), and one high school (9-12). The student population has sustained an enrollment of approximately 1300 students for the past five years.

STUDENT ENROLLMENT

Institution	Configuration	Enrollment
Washakie County School District #1	Grades K-12	1329
East Side Elementary School	Grades K-5	247
South Side Elementary School	Grades K-5	162
West Side Elementary School	Grades K-5	241
Total Elementary Schools	Grades K-5	650
Worland Middle School	Grades 6-8	280
Worland High School	Grades 9-12	399



STUDENT DEMOGRAPHICS

October 2009

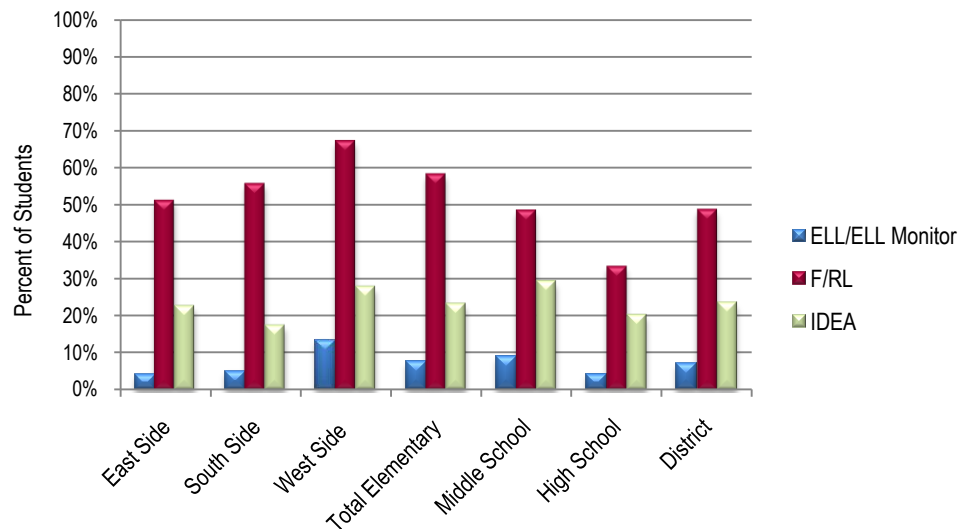
The Washakie County School District #1 student population is disaggregated according to...

- English Language Learners (ELL)
- Free/Reduced Lunch (F/RL)
- Special Needs (IDEA)

STUDENT DEMOGRAPHIC POPULATION

Institution	ELL	F/RL	IDEA	Enrollment
Washakie County School District #1	7%	49%	23%	1329
East Side Elementary School	4%	51%	22%	245
South Side Elementary School	5%	56%	17%	162
West Side Elementary School	13%	67%	28%	240
Elementary Schools (Total)	8%	58%	23%	647
Worland Middle School	9%	48%	29%	283
Worland High School	4%	33%	20%	399

STUDENT DEMOGRAPHIC POPULATION



DISTRICT ASSESSMENT PLAN

Balanced Assessment System

Assessment is an integral part of instruction, providing information about how well students are progressing toward state standards. Standards, curriculum, instruction, and assessment are interwoven in the learning process, and each informs the others.

- Standards are the *what* – defining what we expect students to know and be able to do.
- Curriculum is the *how* – defining a body of learning experiences that are designed to reach standards.
- Instruction is the *in what ways* – providing specific learning experiences and ways of differentiating experiences to scaffold student learning.
- Assessment is the *how well* – gauging the attainment of learning to inform instructional practices and curriculum.

A balanced approach to assessment includes a continuum of strategies within a range of frequency and purpose.

Formative assessment occurs continuously in the classroom, both within and between lessons. Information is used to adjust teaching strategies. Students receive frequent and meaningful feedback on their performances to include teacher observation, discussion, questioning, and non-graded class work. The focus of formative assessment is to determine what learning comes next for a student.

Benchmark assessment (BOE) occurs within, between, and among instructional units. Information is used to identify strengths and gaps in curriculum and instruction. Grade-level curriculum may be refined, and teachers may differentiate instruction for student groups based on student progress. These assessments could be teacher-developed or district-developed products or purchased commercially and linked to state standards. Benchmark assessments include midterm, end-of-unit assessments, semester assessments, and/or district-wide assessments. The focus of benchmark assessment is to determine how students are progressing and how well a program is working.

Large-scale assessment occurs annually. Information is used to develop strategic, long-term evaluation of curriculum and programming based on trends over time, and to monitor state, district, and school progress. Assessment instruments and procedures are standardized so that comparisons can be made across student groups. Examples of large-scale assessment include the PAWS®, ACT®, NAEP™, and AP® exams. The focus of large-scale assessment is to determine how schools, districts, and states are progressing, which is reflected in the district assessment matrix.

These assessment strategies provide information at differing intervals and for different purposes. Each one provides a different perspective, and one cannot take the place of another. Together, these assessments provide a balanced approach to assessment that informs decisions at the classroom, school, district, and state levels.

Balanced Approach to Assessments

Teacher Evaluation Strategies	BOE - Benchmark Assessments	Large Scale Assessments
Daily – Ongoing	Monthly/Quarterly/Semester	Annual
Student-Centered	Student/Classroom/School-- Centered	School/District/State-Centered
Formative	Interim and Summative	Summative

BALANCED ASSESSMENT SYSTEM

	Ongoing	BOE - Benchmark	Large Scale
Purpose	What learning comes next for this student?	How are students progressing? How well is this program working?	How are school/districts progressing? Is there student academic growth? Instructional Program Evaluation
Focus	Students Teachers	Students Teachers Parents Schools	Students Teachers Parents Schools District
Frequency	Occurs within and between lessons	Occurs within and between instructional units/midterm/semesters.	Occurs annually
Use of Results	<p>Within the instructional activity, information is used to change or to adjust teaching.</p> <p>Student receives frequent and meaningful feedback on performance.</p> <p>Teacher engages student in the monitoring of student learning.</p> <p>Teacher decides if the student needs remediation or enrichment.</p>	<p>Grade-level or content-based professional learning communities use data to identify strengths and gaps in curriculum and instruction</p> <p>Grade-level course curriculum may be changed, refined</p> <p>Teachers may modify instruction for student groups based on their progress.</p>	<p>School or district teams develop strategic, long-term evaluation of curriculum and programming based on trends over time</p> <p>States monitor school and district progress</p>
Professional Development	<p>Understanding how to gauge student progress toward a standard and how to scaffold instruction to reach that goal.</p> <p>Teacher reflection on what types of formative strategies are currently employed in the classroom.</p> <p>Implementation of new strategies.</p> <p>Opportunities to practice and to perfect usage.</p>	<p>If teacher-developed, school-team-developed, or district-developed products:</p> <ul style="list-style-type: none"> ♦ Identification of learning outcomes based on state content standards ♦ Development of assessment items & tests <p>If using a commercial product:</p> <ul style="list-style-type: none"> ♦ Learn features of the instrument and how the content links to district curricula and state standards <p>Using local products or commercial products:</p> <ul style="list-style-type: none"> ♦ Discussion of results ♦ Disaggregation & interpretation of data ♦ Sharing effective strategies across teams 	<p>Understanding and interpretation of large- scale assessment trends over time</p> <p>Data disaggregation and progress evaluation at the building level and at the grade level</p> <p>Opportunities for revising district or school instructional practices and curricula</p>
Examples	<p>Feedback</p> <p>Questioning and Interviewing</p> <p>Observations</p> <p>Discussions</p> <p>Ungraded Classwork</p>	<p>Midterm/Unit/Semester Assessments</p> <p>District assessments (BOE)</p> <p>BOE Consortia</p> <p>MAP</p> <p>Specific reading inventories</p>	<p>PAWS</p> <p>ACT</p> <p>PLAN</p> <p>ASVAB</p> <p>CLEP</p> <p>WorkKeys</p>

GRADUATION REQUIREMENTS

To earn a diploma from Worland High School...

1. A student is expected be in attendance for eight full semesters in grade 9 through 12 unless a petition for early graduation is granted and be registered for seven periods each day;
2. A student must successfully complete four school years of English, three of mathematics, three of science and three of social studies, to include history, American government, and economic systems and institutions, for a total of thirteen Carnegie units, as required by the State of Wyoming
3. A student must satisfactorily pass an examination on the principles of the Constitution of the United States and the State of Wyoming.
4. A student must earn an additional 12 credits during the four years of attendance of a possible 15 credits available;
5. A student must demonstrate evidence of proficient performance, at a minimum, on student content and performance standards for the common core of knowledge and skills in five of the nine content areas.

Carnegie Unit - Credit Requirements...

Grade	Required Courses	Semester(s)	Carnegie Unit
9*	English I (A and B)	2	1.0
	Math	2	1.0
	Physical Science (A and B) or Earth Science (A and B)	2	1.0
	Physical Education	1	0.5
	Health	1	0.5
	Electives	4	2.0
10*	English II (A and B)	2	1.0
	Math	2	1.0
	Biology (A and B)	2	1.0
	World History (A and B)	2	1.0
	Electives	6	3.0
11*	Introduction to Speech	1	0.5
	American Literature	1	0.5
	Math	2	1.0
	Science (A and B) (or Senior Year)	2	1.0
	U.S. History (A and B)	2	1.0
	Electives	6 or 8	3.0 or 4.0
12*	Senior English (A and B)	2	1.0
	Science (A and B) (or Junior Year)	2	1.0
	Government	1	0.5
	Electives	11	5.5
9-12	Computer Class	1	0.5
	Keyboarding I (or pass a keyboarding proficiency test)	1	0.5

*Flexible Grade Level Schematic

SYSTEM OF DOCUMENTATION

The Body of Evidence document of Washakie County School District #1 specifies the requirements for a student to graduate from Worland High School. Chapter 31 of the Wyoming Department of Education and State Board of Education Rules and Regulations indicate that in addition to earning the required number of Carnegie units (credits) to graduate a student must show proficiency on Wyoming's content and performance standards for the common core of knowledge and skills in the content areas approved by the Wyoming State Board of Education.

Diploma:

A high school diploma provides for one of the following endorsements, stated on the transcript of each student:

Advanced Endorsement

- Requires a student to demonstrate advanced performance in a majority of the areas of the common core of knowledge and skills and proficient performance in the remaining areas of the specified common core knowledge and skills, as defined by the uniform student content and performance standards
- Advanced in the majority (five) of the (nine) content areas and proficient in the remaining content areas

Comprehensive Endorsement

- Requires a student to demonstrate proficient performance in all areas of the common core of knowledge and skills as defined by the uniform student content and performance standards
- At least proficient in all (nine) content areas

General Endorsement

- Requires a student to demonstrate proficient performance in a majority (five) of the areas of the (nine) common core of knowledge and skills as defined by the uniform student content and performance standards
- At least proficient in the majority (five) of the (nine) content areas

Body of Evidence Plan Approach:

Washakie County School District #1 adopted a compensatory, course-based common assessment system for the District K-12 Body of Evidence. Common Assessments are administered through required and elective courses which provide standards-based instruction and assessments that afford students to meet the Body of Evidence (BOE) requirements in the following nine content areas:

- | | | |
|------------------|-----------------------|--------------------------------|
| 1. Language Arts | 4. Social Studies | 7. Career/Vocational Education |
| 2. Math | 5. Health | 8. Fine and Performing Arts |
| 3. Science | 6. Physical Education | 9. Foreign Language |

The approach for this system includes:

1. Use of course-based common assessments to provide opportunities to anchor grades from different teachers across the district;
2. Flexibility for common assessments to have the option to carry a certain percentage of weight with information from the course carrying the remainder of the weight, when necessary; and
3. Use of "on-demand" projects or other types of measures that are tied to the standards-based curriculum.

Advantages to the system include:

1. Helping to ensure equivalent opportunities for students to learn the standards;
2. Allowing individual teacher choice and discretion;
3. Encouraging teacher collaboration when designing and writing common assessments; and
4. Using common assessments of high quality to alleviate possible grading concerns.

Additional advantages involve:

1. Prioritizing common assessments and/or standards/benchmarks, through effective teacher collaboration;
2. Supporting allocation of time in examining student work and student expectations;
3. Encouraging staff to develop vertical alignment and transition between grades/school levels; and
4. Anchoring course grades.

Proficiency Levels:

Assessment items are developed at the benchmark level and meet the content and performance descriptor of one or more benchmarks within an assessment, and proficiency is determined through a compensatory approach. Levels for proficiency use the compensatory model and are subject to an increased requirement for proficiency, pending assessment development and standard/benchmark revisions.

Standards Level:

- The district deems a student proficient in a standard if at least 60% of benchmarks are met at the proficient or advanced level of the assessments.
- The district deems a student advanced in a standard if at least 60% of benchmarks are met at the advanced level of the assessments.

Content Level:

- The district deems a student proficient in a content area if at least 60% of standards are met at the proficient or advanced level of the assessments.
- The district deems a student advanced in a content area if at least 60% of standards are met at the advanced level of the assessments.

Graduation Requirement:

- **General Education Endorsement:** To fulfill the Body of Evidence graduation requirement for a General Education Endorsement, a student must be deemed proficient (or advanced) in the majority of the content areas (five of nine).
- **Comprehensive Endorsement:** To fulfill the Body of Evidence graduation requirement for a Comprehensive Endorsement, a student must be deemed proficient (or advanced) in all the content areas (nine of nine).
- **Advanced Endorsement:** To fulfill the Body of Evidence graduation requirement for a Advanced Endorsement, a student must be deemed advanced in the majority of the content areas and proficient in the remaining content areas.

Additional Opportunities to Demonstrate Proficiency:

- Additional opportunities that validate student proficiency to serve in lieu of performance standards are considered, as the ACT, PAWS, SAT, and any other type of district-approved assessment that is aligned to the state standards and meets approval of the administration.
- After consideration of additional proficiency opportunities, the student has the right to an appeal to the Board of Appeals. If the student is not satisfied with the results from the Board of Appeals, the student has the right to reconsideration the results of the appeal beyond the appeals board and must submit the request to the superintendent in writing.

Worland High School Board of Appeals

- If a student's high school transcript and proficiency report do not indicate that the student has met the requirements for graduation, the student has the right to appeal to the Board of Appeals. The Board of Appeals will consider any evidence presented by the student or on the student's behalf, and this board will decide if the student should be granted graduation status. The Board of Appeals consists of the principal, a counselor, and three classroom teachers from varying content areas.
- If the student has not met graduation requirements after a review of the student appeal, the Board of Appeals will have the authority to design an individualized student plan for the student to have an opportunity to demonstrate proficiency on the common core of knowledge and skills to meet graduation requirements. The plan can include items such as student interviews with staff members, an observation demonstrating knowledge and skills, a review of additional assessment information, records to student performance, completion of additional assessments, and/or other appropriate activity or performance that the Board considers appropriate to demonstrate knowledge and skills.
- A student has the right to appeal the decision of the Board beyond the Worland High School Board of Appeals, which must be done in writing to the school district administration.

Notification of Progress towards Body of Evidence and Graduation

- **Grade 8:** During the spring the high school counselor meets with students in the eighth grade and his/her parents prior to entry into high school to ease the transition to the next level schooling. This orientation is devoted to explaining grades, credits, and Body of Evidence requirements for graduation, as well as opportunities for college and career-ready planning.
- **Grade 9:** When a student enters the ninth grade, the student develops an Individual Graduation Plan, and during the following years of schooling, the counselor monitors student progress in meeting proficiency in all common core areas.
- **All High School Students:** Twice during the year, at the completion of each semester, high school students and their parents are informed of the student's progress toward meeting all graduation requirements. Within the District K-12 Body of Evidence, parents and students are notified regarding achievement of proficiency levels for standards in grade-level appropriate content areas.
- **All K-12 Students:** At the end of each school year, every parent is provided documentation regarding their child's progress toward meeting standards at the appropriate grade level.

- **Infinite Campus Access:** Parents and students have access to Infinite Campus regarding individual academic progress regarding coursework grades and standards' achievement.

Transfer Students

1. When a student transfers to Worland High School, the student's must transcripts are reviewed by the high school counselor to identify credits and evaluate grades to reflect the achievement of Wyoming content and performance standards/benchmarks.
2. When appropriate, the counselor will contact the previous school to verify coursework/curriculum and recommend or not recommend to the principal that the course grade reflects the level of proficiency in a content area.
3. If necessary a committee of content area teachers will review course documentation and determine the alignment with Wyoming content and performance standards/benchmarks.
4. Based on counselor and/or content committee recommendation, students could be required to complete one or more of the common assessments related to the content area to determine the level of student proficiency, or a students would be required to take coursework that allows them to demonstrate proficiency.
5. The school principal will make the final determination of proficiency on content area standards for transfer students.

SUMMARY OF THE DISTRICT BODY OF EVIDENCE

The Body of Evidence of Washakie County School District #1 is a K-12 systemic process that specifies the requirements and provides documentation for a student to graduate from Worland High School in compliance with Chapter 31 of the Wyoming Department of Education and State Board of Education Rules and Regulations. The district body of evidence assessment system is designed (1) to determine whether students have met graduation standards and (2) to provide a collection of evidence to espouse this decision, supporting inferences about students' mastery of the content areas and the district's capacity for ensuring equality of educational opportunity.

Washakie County School District #1 firmly supports and sustains a balanced assessment program that...

1. Is systematically monitored for public accountability,
2. Provides information to enhance instructional practices,
3. Supports effectiveness of instructional strategies,
4. Measures student achievement and mastery of skills,
5. Provides appropriate data to improve student progress,
6. Engages parents in student learning, and
7. Prepares students to be effective citizens.

The district assessment system uses a compensatory model, reflected in the Body of Evidence, which allows higher scores to offset lower scores in each standard in mathematics, language arts, science, social studies, foreign language, fine and performing arts, physical education, health, and career/vocational content areas.

The protocol for the district body of evidence is designed and implemented according to the following criteria:

Alignment: In a standards-based system, alignment involves all parts of a learning system (as assessments, teaching strategies, instructional materials, and professional development), working together to ensure student learning of the content standards. In terms of the relationship between standards and assessment, alignment refers to the match between items on the assessments and knowledge and skills represented by the standards. The district's assessment system ensures the process for assessment alignment of content standards and benchmarks, which is built directly from standards as a three-way process between curriculum, assessments, and standards. Assessment items/tasks are aligned to all content standards in the nine content areas as well as the curriculum; curriculum is aligned to standards and assessments; and all standards, in terms of content and cognition, are aligned to the curriculum and are assessed. Assessments reflect the cognitive depth of student performance described in the performance standards and measure the benchmark level to ensure an adequate sampling of the standards.

Consistency: The focus of the assessment system targets consistent decisions regarding student performance on the standards. The process to ensure consistency (reliability) in the district assessment system is based on data sources from (1) inter-rater agreement data of district open-ended assessments in each of the content areas and/or (2) item analysis using the Cronbach's Alpha for establishing reliability for closed ended assessments, with items having an exact answer per question. Teacher judgment is not used as a measurement construct in the Body of Evidence.

Fairness: The district assessment system is designed, implemented, and evaluated so that each student has an opportunity to demonstrate mastery of knowledge and skills through a variety of assessments with multiple opportunities in each of the content areas to maximize fairness. There is a formal process to monitor assessments to assure that there is no bias toward any group of students. Appropriate accommodations are used so that students with disabilities, English Language Learners, and identified 509 students have an opportunity to demonstrate learning. All assessments are at the benchmark level to ensure that multiple assessment opportunities and formats are used to maximize fairness. Student data from standards-based assessments are disaggregated to examine both the fairness of the system and opportunities for all students to learn the standards.

Standard Setting: The district uses the Bookmarking Method to establish cut scores at benchmark, standards, and content levels and has established a standard setting process that involves an item-based method for closed-ended items and student work-based methods for open-ended items. Determining cut scores at these levels supports a true compensatory model, for students do not need to demonstrate proficiency with every assessment, benchmark, or standard to be proficient in a content area. Open-ended assessments have performance descriptors of rubric design as a constant for identifying student performance levels. Student work is examined, arranging a sampling of student work from the highest scoring down the continuum to the lowest performing level. Proficiency levels, aligned to the performance standards and cognitive demand, are established according to the performance rubric. Anchor papers are selected to ensure identification of each proficiency level and consistency of transition from one proficiency level to another. Assessments with closed-ended items are matched with performance levels by item. A team of evaluators individually rate each set of work according to performance descriptors. Ratings are evaluated and performance levels are established.

Comparability: The district assessment system requirements are comparable for all students within a given class (or grade level) and across years. Common assessments are used throughout the district at all grade levels in all content areas to ensure comparability. If new assessment items/tasks are used in the body of evidence, district assessment procedures are implemented to maintain comparability annually and to ensure alignment, consistency, fairness, standard setting, and comparability. A rotation of assessments and content areas is used to guarantee consistency and comparability in each content area. Each criterion in the Body of Evidence is reinforced with professional training and is monitored continuously by staff through self-evaluation, instructional facilitators, and administrators.

The district body of evidence assessment system is an integral part of the balanced assessment system in the district assessment plan.

HISTORY OF THE DISTRICT BODY OF EVIDENCE

History of Protocol in Developing and Sustaining a Body of Evidence:

Washakie County School District #1 developed a Body of Evidence System at the beginning of the onset for the BOE requirement within the WDE guidelines involving the BOE Stakeholder Committee. The five criteria were addressed regarding method of protocol and implementation.

Upon completion, the BOE was submitted for peer review and was resubmitted within a two-year span for implementation. The BOE was submitted and filed by WDE at the Laramie office. When the WDE Superintendent's Memorandum (2004), regarding method of implementation, was received the following year, the district maintained compliance with the BOE guidelines through (1) continued monitoring of student performance in meeting the standards in all nine content areas at the high school level; (2) consistent reporting of student progress to parents and students in meeting proficiency of the standards at grades K-12; and (3) reporting standards' progress for students in the 11th and 12th grades to WDE during each school year via WDE Form 663.

Through the span of these school years, teachers examined student progress in meeting the standards and targeted individuals having difficulty meeting the graduation requirements in the district BOE. The BOE committee, with input from the entire staff, parents, and community members, selected a course-based and common assessment approach. Students were provided extensive opportunities to achieve proficiency levels of the standards through in-class activities, after-school tutoring, multiple opportunities and multiple varieties of district assessments, credit recovery through online resources (aligned with the standards), and summer school. Student data from the BOE was examined along with the district norm reference and state testing to identify areas of need to...

- Identify appropriate goals and interventions that target all students as well as identified subgroups;
- Improve instructional strategies and assessments; and
- Increase opportunities for students to demonstrate proficiency in each of the nine areas.

During 2007-2009 there was a conscious emphasis on district level assessments that target (1) student performance and (2) effective assessments to provide multiple opportunities to measure student performance. With the need to present the Body of Evidence for peer review during Spring, 2009, the district chose to empower a district-level committee to re-examine the present BOE for grades K-12, targeting the five BOE criteria in meeting student needs and WDE compliance. After reviewing the Guidelines in Developing an Assessment Plan as well as the *Wyoming Assessment Handbook*, the committee constructed district operational definitions and identified a comprehensive balanced assessment system, targeting standards-based, formative and summative assessments. The committee updated the district assessment matrix to identify summative assessments given throughout the district. It was apparent that there was a need to review each of the criteria to meet compliance with WS 21-3-110, 21-2-304, WS 21-9-101, and WS 21-9-102 and to revisit additional requirements with the mandated WDE BOE rubric. It was determined that the Body of Evidence should not be a "high school" document, since the body of evidence is a systemic process that impacts learning throughout the K-12 continuum.

Each criterion in the Body of Evidence was revisited and thoroughly examined using the protocol as recommended by the District Body of Evidence Committee and staff resulting in the following:

- ♦ **Alignment:** Alignment was revisited, which assured a three-way alignment process.
 - **Rationale:** Matrices had been fully developed at the middle and high school levels in all content areas, but this was not evident throughout the entire K-12 program. Blueprint matrices were examined – identifying multiple opportunities, multiple assessment types, and level of cognition. Assessments were revisited to examine, adjust, and reassure alignment with performance standards. All content areas were reviewed by the staff which supported a three-way alignment with (1) assessment items matched to standards, (2) the assessment of standards at the benchmark level represented an appropriate sampling of standards; and (#) curriculum supporting assessments and standards. Items were matched against the standards, the standards were checked to ensure they appropriately represent items; and curriculum aligned to the standards and assessments.
 - **Reference:** Supporting evidence of alignment procedures, compliance, and/or updates are referenced in the alignment matrix (Appendix B).

- ♦ **Consistency:** Staff members reviewed consistency of assessments and implemented district protocol regarding open- and closed-ended assessments.
 - **Rationale:** The staff had been involved in critiquing common assessments, but the inter-rater component and written protocol were lacking in assessment documentation. The present protocol involves teachers using Cronbach's Coefficient Alpha with closed-ended assessments and double scoring inter-rater reliability for open-ended assessments with appropriate rubrics based on performance descriptors.
 - **Reference:** Supporting evidence of consistency procedures, compliance, and/or updates are referenced in consistency protocol. (Appendix C)

- ♦ **Fairness:** Each K-12 assessment was reviewed for bias, identification and practice for use of accommodations, variety of assessments, disaggregation within subgroups, and participation. All district assessments have gone through a peer review for bias through an established district protocol. An accommodation statement to represent equitable opportunities for students to demonstrate proficiency was established district-wide. District data was reviewed regarding cognition, multiple formats, and multiple opportunities with a blueprint matrix, developed to record information for review and analyses.
 - **Rationale:** A district bias committee, established during the onset of the BOE, examined all assessments in the district. Additional assessments had been adjusted or added since that time, and there was a need to do a full blown peer review of bias on all body of evidence assessments. Documentation and analyses of fairness was recommended by the district committee.
 - **Reference:** Supporting evidence of fairness procedures, compliance, and/or updates are referenced in fairness procedures. (Appendix D)

- ♦ **Standard Setting:** Teachers identified cut scores and/or anchor papers identifying student performance levels, using the Bookmarking Method. Parents and community members were on the initial Body of Evidence committee, and the district secured input from parents and community members to establish performance at the standard and content levels. Using the bookmarking method, teachers identified student work (1) to establish cut scores and /or proficiency

descriptors according to rubric descriptors for extended response or performance/authentic assessments and (2) to determine performance levels of closed-ended assessments by establishing scaled scores matching each closed-ended item with each level of the performance standards. Each content area used an established district procedure to determine the scores that divide various performance levels on an assessment, which tied closely to proficiency descriptions that indicate what it means to be proficient.

- **Rationale:** Very few content areas had aligned performance descriptors with closed-ended items and rubrics with open-ended items and selected anchor papers for assessments. Teachers had saved work from student assessments and were able to identify exemplars that represented performance levels.
 - **Reference:** Supporting evidence of standard setting procedures, compliance, and/or updates are referenced standard setting data. (Appendix E)
- ♦ **Comparability:** The district committee reviewed protocol to ensure that all K-12 students receive and will continue to receive equitable learning opportunities that are fair, reasonable, impartial, just, and unbiased in every educational setting. Common assessments are used throughout the district at all grade levels in all content areas to ensure comparability. If new items/tasks are used in the BOE, assessment procedures are implemented to maintain year-to-year comparability. The district has a schematic that assures that assessments are administered, examined, and reviewed consistently each year, and comparability is monitored continuously..
- **Rationale:** The committee felt that the updates to the BOE should be supported with information that ensured that all K-12 students receive and will continue to receive equitable learning opportunities that are fair, reasonable, impartial just, and unbiased in every educational setting.
 - **Reference:** Supporting evidence of comparability procedures, compliance, and/or updates are referenced in comparability information. (Appendix F)

Changes/Adjustments to the Body of Evidence (with Rationale):

Spring 2009 the district Body of Evidence was submitted for peer review, and two areas of the district plan were considered unacceptable – consistency and standard setting. During this past year, each of those areas was addressed. Documents were reviewed, and the district committee involved all stakeholders - parents, staff, administration, and community members, providing clarification of and training in the process.

♦ **Consistency:**

- Concern: Team 1 - Strong procedures and rates, but open and closed ended item indices need to be completed.
Team 2 - Whole section on closed ended reliability makes it all highly suspect. Words, not a system.
- Action: Teachers submitted closed-ended student work, which was entered into a spreadsheet configuration developed by Jason Nicholas to determine the Cronbach Alpha Coefficient, an indicator of assessment reliability, and open-ended assessments were double-scored using assessment rubrics to evaluate student performance independently in order to establish consistency in scoring more than one correct response,

referred to as IRR (Inter-Rater Reliability). These practices were employed and documented with data throughout the district in each of the content areas at all grade levels.

♦ **Standard Setting:**

- Concern: Team 2 - No explanation of any detail of the cut score process. No link to proficiency. Stakeholders incomplete.
Team 3 – You've done great work setting cut scores at the assessment level. However, we can't identify the method/process you are using to make final decisions at the content level. Be clearer on your stakeholder participation.
- Action: The cut score process was revisited and clarified in greater detail. Stakeholders, including staff, parents, and community members, participated in the process, providing input and suggestions/concerns regarding cut scores, and cut scores were established at benchmark, standard, and content levels.

Section 2: ALIGNMENT

INTRODUCTION

Alignment represents a standards-based system with all components of a learning system (as assessments, teaching strategies, instructional materials, and professional development) working together to ensure student learning of content and performance standards. Alignment between standards and assessments represents a direct correlation between assessment items and core knowledge and skills in the curricula and standards.

When Wyoming Content and Performance Standards became available to all districts within the state, standards from nine content areas were adopted as a foundation for district standards. At this time, the district made a commitment to standards-based instruction with horizontal and vertical alignment of curriculum, instruction, and assessment. When information regarding a district body of evidence was received from the Wyoming State Board of Education, the district began developing a district assessment plan to align to the state content and performance standards.

The framework for the district assessment plan is designed to align to the district standards through a three-way process.

1. Curriculum/instruction is aligned (a) to content standards and benchmarks, congruent with levels of cognition in the performance standards and (b) to assessment items that measure standards/benchmarks at various levels of cognitive performance.
2. Standards and benchmarks are aligned (a) to assessments that consider cognition levels in the performance standards and (b) to curriculum/instruction which is standards-based delivery of content and performance standards.
3. Assessments are aligned (a) to standards and benchmarks, defining what we expect students to know and be able to do, and (b) to curriculum/instruction, measuring the attainment of learning to inform instructional practices and curriculum.

The district assessment plan was developed in this context, using Norm Webb's framework for expressing the cognitive complexity of performance alignment. To follow district protocol, assessments in the district plan were created and continue to be created directly from the standards to ensure concurrence between benchmarks/standards, assessments, and curriculum/instruction, interwoven in the learning process with each informing the other.

Assessment Alignment with Cognitive Demand

As assessments were developed, cognitive demand was aligned to performance standards to sustain the integrity of the depth-of-knowledge between standards and assessments, carefully analyzing if assessments effectively target and articulate cognitive levels in the performance standards. The BOE district committee reinforced the decision of the staff to align assessments to targeted benchmarks, using a finer grain level, to ensure an adequate sampling of standards in each of the nine content areas. Examination of assessments during bias review and student anchor paper selection provided opportunities for staff to evaluate content and performance (cognition) alignment.

Blueprinting and Matrices for Common Assessments

The blueprint matrix reflects the strength of the benchmark as well as the match between assessment type, cognition, and benchmark(s) and between assessment items and benchmarks assessed, ensuring three-way alignment.

The district assessment plan includes (1) a combination of assessments that are aligned with district content and performance standards; (2) standards that are assessed in terms of content and performance (cognitive complexity); (3) a standards-based curriculum that is aligned with standards and assessments; and (4) multiple assessment measures with various formats that maximize alignment between standards and assessments.

PROCESS OF ALIGNMENT

Alignment in the Assessment Plan:

The District Assessment Plan for the body of evidence is a course-based common assessment approach which consists of district assessments, Body of Evidence Consortium Activities, scientifically research-based aligned assessments, and alternate opportunities to demonstrate standards' proficiency. Each assessment that is used to measure student proficiency on the district standards report card is aligned with standards by benchmark. The district revisits the BOE annually referencing the guidelines for the district assessment plan, the District Assessment Matrix, and the District Assessment Matrix Blueprint by content.

A variety of assessment resources complete the blueprint for the district Body of Evidence.

- District-developed assessments were developed in the nine content areas. Each assessment was reviewed for content and cognition to ensure alignment to Wyoming content and performance standards and benchmarks. Blueprints for each assessment indicate the cognition level which marker standard/benchmarks assessed.
- The assessment plan includes a variety of research-based assessment resources that are aligned with Wyoming content standards and benchmarks, such as Study Island, Accelerated Reading, and My Access. Teachers pre-review all assessments that used for student evaluation in the Body of Evidence.
- Selected activities from the Wyoming Activities Consortium in math, language arts, science and social studies are embedded in the Body of Evidence, pending need and opportunity within the curriculum design.
- Health SCASS (CCSSO) assessments are used in the Health content area. Each assessment used in the Body of Evidence was reviewed for alignment to Wyoming content and performance standard/benchmarks. Blueprints for each assessment indicate the cognition level which marker the standards and benchmarks assessed.

The Alignment Process:

Each content area in all coursework is aligned to district content and performance standards and benchmarks, with state content and performance standards and benchmarks embedded within the district standards. The majority of assessments used throughout the district are common assessments, with course-based assessments providing a variety of authentic assessment opportunities for students to demonstrate proficiency as well as providing a way to anchor grades from different teachers across the district. Each course is a standards-based course with teachers teaching to the standards, and the

district has a standards' report card which is separate from the course grade. Assessments results are recorded on the standards' report card at the benchmark level. To validate that all state standards and benchmarks are taught to all students, a K-12 Benchmark Matrix is maintained in each content area by grade (Grades K-5) or by course (Grades 6-12). Additionally, the K-12 Benchmark Matrix in each content area verifies that assessments are aligned to standards and at a finer grain at the benchmark level. The frequency of assessments is recorded in the K-12 Cognition Matrix.

Standards-based instruction occurs at all grade levels in the district and is monitored and sustained in the K-12 District Assessment Matrix Blueprint by the curriculum director. Teachers are responsible for developing required standards based assessments, creating scoring rubrics, conducting bias review, and recommending cut scores which focus on standards at the benchmark level. The district uses a consistent process for alignment, and to ensure that assessments are measuring the standard, teachers conduct a back-check, reviewing student work on the assessments to reflect the intent of the assessment measure. This is accomplished by teachers examining student work while developing performance descriptors and identifying anchor papers during the standard setting process.

Steps for Alignment during Assessment Development:

The district ensures alignment between standard/benchmark(s) and/or assessments for each content area in the following steps:

1. Establish the grade level or course(s) the standards will be taught and assessed;
2. Examine the content area course matrix that indicates where the standards are taught (and aligns with the required Carnegie units at the high school level);
3. Identify the type and number of assessments that are necessary to establish student competency on the standards and benchmarks at various levels of cognition;
4. Reference the assessment blueprint matrix for each content assessment that describes the type of assessment item, the measure of cognition, and the timeline for administration, and examine blueprint matrices to design comparable assessments and identify the level of assessment coverage for all of the benchmarks in each standard;
5. Identify the cognition of the standard/benchmark(s) selected for assessment by the following:
 - Review the content standard/benchmark(s);
 - Identify verbs/adjectives in the performance standard/benchmark; and
 - Distinguish the levels of cognitive demand implied in the performance standard.

Section 3: CONSISTENCY

INTRODUCTION

Consistency targets coherent decisions about students' performance related to the standards. Within the Body of Evidence system the district ensures the following:

- Scoring consistency is assured for all assessments that measure student academic performance in meeting proficiency in the standards. The district implements scoring consistency through a protocol which assures that inferences drawn from the results of the assessments are consistent and not dependent on error due to raters or the quality of the assessments.
- The procedure for assessment data collection, analysis, and interpretation is used in the same manner for all students, yielding consistent decisions regarding student performance on the standards/benchmarks.
- All students are provided multiple opportunities with multiple formats to demonstrate proficiency on standards and are provided clear, explicit information regarding graduation requirements and transcript endorsements.
- An appeals process is available and accessible throughout the Body of Evidence system.

To support consistency in assessment scoring practices in the district K-12 body of evidence, instructional facilitators, grade level teachers, specialized teachers, and content-area teachers are trained in district procedures to effectively evaluate assessment items used to monitor student progress in meeting the standards. Training involved procedures used to ensure appropriate protocol for inter-rater reliability on open-ended assessments using performance rubrics (or descriptors) and on closed-ended assessments, using Cronbach's Coefficient Alpha. In-service time was available for all staff to work collaboratively to adjust/review all standards-based assessments.

PROCESS OF CONSISTENCY

The focus of the Body of Evidence system targets consistent decisions regarding student performance on the standards. The process to ensure consistency (reliability) in the district assessment system is based on data sources from (1) inter-rater agreement data of district open-ended assessments in each of the content areas and/or (2) item analysis using the Cronbach's Alpha for establishing reliability for closed-ended assessments, with items having one right answer per question. Teacher judgment is not used as a measurement construct in the Body of Evidence.

Open-ended Assessments

Open-ended assessments are an integral part of the Body of Evidence system to evaluate skills, abilities, and/or application of depth of knowledge (Norm Webb) and are scored with a two- or four-point, performance-based scoring rubric representing cognition levels, corresponding to the Wyoming performance standards. Each assessment is accompanied with explicit directions, requirements, scoring procedures and rubrics to ensure the reliability of the assessment task.

Protocol for Inter-rater Reliability for Open-ended Assessments:

- Training for scoring student work is provided to all teachers within each of the nine content areas, which is facilitated by the Curriculum Director, Instructional Facilitators, Content-Level and Grade-Level Teacher Leaders, and/or BOE Committee Members, who established the protocol prior to working with staff.

- Guidelines were established for submission of student work. The curriculum director identified the number of representative assessments to be collected for double scoring in each of the content areas and at all grade levels. Teachers selected a minimum of 25% of each assessment, preferably representing all levels according to proficiency descriptors congruent with the state performance standards. Student names were removed from all student work selected for scoring.
- Scoring is based on a rubric designed to evaluate each individual assessment, aligned to the identified benchmark(s) and performance standards. Teachers review the scoring rubric for clarification, if necessary. Each assessment is scored independently by two teachers, without conversation or collaboration. Scores from each assessment are compiled by a third party designee, who performs inter-rater reliability calculations on the Inter-rater Scoring Sheet. The results of assessment data is evaluated regarding exact and adjacent rating.
- Acceptable rates for inter-rater reliability scoring must meet or exceed 80% exact agreement and 98% exact and adjacent agreement. If the scorers are in agreement, scoring is consistent. It is acceptable if scorers score items in adjacent proficiency levels although exact agreement is preferable. Scores that differ by more than one level are discussed, and the designee consults with the curriculum director, who reviews all discrepancies (unacceptable inter-rater agreement) in the scoring. Consistent discrepancies in scoring may warrant a need to consider the following interventions:
 - (1) Conference with scorers,
 - (2) Provide an additional scorer to independently score the assessment,
 - (3) Evaluate the effectiveness of the scoring rubric,
 - (4) Revisit the prompt or question, and/or
 - (5) Provide training to improve scorer agreement.

Closed-ended Assessments:

Closed-ended assessments are administered to evaluate content knowledge as an item analysis using Cronbach's Alpha Coefficient to improve quality and to establish reliability of questions with items having an exact answer for each question. Questions such as multiple choice, specific response, and matching, are evaluated using this process.

Protocol for Closed-ended Assessments:

- Guidelines were established for submission of student work. The curriculum director identified the number or percent of representative assessments to be collected in each of the content areas and at all grade levels. Teachers obtained examples of all student work for each assessment selected.
- Each student response to every item in an assessment is recorded on a reliability spreadsheet (reliability calculator) as a correct or incorrect answer, provided by Jason Nicholas (Lander, Wyoming, November 2008) and designed to determine Cronbach's Alpha Coefficient, which is automatically calculated.
- A Cronbach's Alpha Coefficient rate to meet or exceed 70% was selected as an acceptable rate for all closed-ended assessment. If the coefficient falls below 70%, specific assessment items are revisited, and follow-up investigation into the assessment to determine potential issues with assessment items and/or instruction is necessary. This rate was deemed reasonable by teacher collaboration, as per recommendation of Dr. Alan Moore, WDE.

- If the coefficient were to fall below 70% (0.70), one of the following is addressed:
 - Number of assessment items (the greater the number of items, the greater the coefficient)
 - Appropriate assessment alignment
 - Collaboration among grade level or content teachers
 - Curriculum design
 - Instructional quality
- At a finer grain, individual items can be analyzed for student mastery, when success rate falls below the minimum acceptable rate of 70% (.70).

Teacher Judgment:

Teacher judgment is not used as a measurement construct in the district Body of Evidence plan.

5-Year Consistency Cycle – An On-going Process:

The district assessment plan addresses consistency as an on-going process. Reliability analysis of open-ended and closed-ended assessments is conducted on 20% of district benchmarks assessments each year resulting in a 5-year cycle for consistency. Rationale for this procedure allows for staff to...

- 1) Revisit assessments for revisions;
- 2) Adjust to adoptions/changes in standards and benchmarks;
- 3) Train new staff on procedures within the assessment system;
- 4) Adjust to textbook/program adoption; and
- 5) Afford staff to create more appropriate assessments.

Content areas, by grade level or content, determine assessments for review during each year. All completed data information is submitted to the Curriculum Director. Additional consideration is given to programs requiring changes in supplemental curricular needs, students at risk, RTI, differentiation of instruction, and students with special needs.

Each year specific content curricula (core and non-core) and instructional strategies are targeted for examination and review; new textbooks, curriculum/instruction, and technology resources are considered for adoption. Content areas rotate through the process within a five- to seven-year rotation, followed by one year of assessment pilot. Teachers address the "review" content area for examination and are afforded the latitude to extend assessment examination in all other content areas, especially for assessments that are need of change, replacement, and/or adjusted. Twenty percent (20%) of all assessments from content areas that are not considered for review in a specific year are revisited.

Note: When new assessments are developed, alignment, fairness, and preliminary standard setting will accompany assessment development, and fidelity to the process ensures consistency, comparability, and final standard setting during the following year.

Section 4: FAIRNESS

INTRODUCTION

Fairness in a body of evidence is designed so that it is not biased against any group of students and provides appropriate accommodations so that students with disabilities, 504 students, and Limited English Proficient students have fair access to the assessment system. Fairness assures that multiple assessment formats are included in the assessment system.

The Washakie County School District #1 Body of Evidence K-12 Assessment System was conceived to service all students, which includes all subgroups (as 504, special needs, gender, free/reduced lunch, ethnicity, and limited English proficient). The school improvement plans of all schools in the district include goals in reading, writing, and math, which provide interventions for all students and interventions for targeted subgroups not performing at the proficient level. This information is determined by PAWS data and is supported with data from assessments in the district assessment system.

The district assessment system is designed, implemented, and evaluated so that each student has an opportunity to demonstrate mastery of knowledge and skills through a variety of assessments with multiple opportunities in each of the content areas to maximize fairness. There is a formal district protocol to monitor assessments, assuring that there is no bias against any groups of students. Appropriate accommodations are used so that students with disabilities and English Language Learners (and other qualifying students, i.e. 504) have an opportunity to demonstrate learning. The results of the assessments are disaggregated to examine both the fairness of the system and opportunities for all students to learn the standards.

The district assessment system ensures fairness through (1) a bias review of assessments, (2) accommodations for students in targeted subgroups when appropriate, (3) multiple opportunities to demonstrate proficiency in multiple formats, and (4) opportunities for all students to demonstrate proficiency. Student data are disaggregated by subgroup to ensure fairness for all students.

PROCESS OF FAIRNESS

Fairness - Bias

1. The Body of Evidence District Bias Committee, comprised of representatives from all school levels, reviewed K-12 assessments for bias. Each local assessment that was given to students in the district was reviewed. A separate report was recorded for each assessment, and suggestions/comments for staff members to address.
2. Staff members were instructed to heed the comments from the bias committee results and make necessary adjustments. If additional or replacement assessments were to be developed, teachers were instructed to carefully review for bias.
3. A district "Bias and Sensitivity" protocol was established to measure standards' assessments in the K-12 Body of Evidence when developing new, additional assessments. The district follows specific guidelines that addressed concerns as defined by age, disability, ethnicity, gender, national origin, race, religion, and sexual orientation, regarding bias which includes the following to:
 - Treat people with respect

- Minimize the effects of construct-irrelevant knowledge or skills.
- Avoid material that is unnecessarily controversial, inflammatory, offensive, or upsetting.
- Use appropriate terminology to refer to people.
- Avoid stereotypes
- Allow for balance in diversity
- Treat the sexes equally

Fairness - Accommodations for Students

1. All students have an equal opportunity to demonstrate proficiency. Appropriate accommodations are available to qualifying students. Accommodation policies apply to students with IEPs, students with 504 plans, ELL students with an ILP, students who qualify for Title I services, or to all students, and the district adheres to compliance with the 2008 Rules and Regulations: Chapter 31, Section 10 (f). The district recognizes that there could be additional circumstances for students who do not have a qualifying plan and could use accommodations during testing - permitting all students to use standard accommodations or permitting all students to use standard accommodations under certain circumstances (such as a student with a temporary disability). Documentation regarding student assessment accommodations is available to all teachers.
2. Accommodations are considered from the following criteria:
 - Use of instructional accommodations
 - Maintains the validity of the assessment
 - Individual student needs and characteristics
 - Purpose or nature of the assessment
 - Students' previous experiences with the recommended accommodations.
 - Confidentiality and training for administration of accommodations
 - Unfamiliar accommodations during test administration
3. Types of accommodations include presentation accommodations, equipment and materials accommodations, response accommodations, scheduling accommodations, and setting accommodations.

Fairness - Multiple Opportunities and Variety of Assessments

1. All students are provided multiple opportunities, using multiple formats, for demonstrating proficiency in all standards for all content areas. A variety of assessment types are available in all content areas with identifying indicators of various levels of cognition.

Fairness - Disaggregation of Assessment Results

1. Assessment results are disaggregated and viewed by the staff to identify areas of student proficiency, student needs/concerns, and assessment efficacy.
2. Assessment data is matched with student performance data on the state-wide PAWS test and the student MAP data.
3. Each school follows district protocol to identify strengths and weaknesses in student performance within the triangulation of data as well as other assessment resources available to specific schools (as STAR Reading,

DIBELS, STAR Math, STAR Early Literacy, etc.). Through the direction of the Building Leadership Team in each school, student data determines the need to identify specific interventions followed by specific action steps to target learning opportunities for all students as well as identified student subgroups. When necessary teachers are provided appropriate professional development opportunities to target these interventions and/or action steps.

4. Assessments are monitored regarding low performing students with specific needs, based on an IEP or ILP. Teachers are required to examine BOE assessments for bias and appropriate accommodations to meet individual student needs, where appropriate (as: potential bias in the assessment system and/or differences in opportunities to learn).

Fairness - Participation of All Students

1. All students have an equal opportunity to demonstrate proficiency by holding students accountable to participate in BOE assessments. Scheduling procedures allow for students to be assessed in each of the nine content areas. Appropriate accommodations are available to qualifying students to ensure all student participation.
2. All students are provided multiple opportunities, using multiple formats, to participate in demonstrating proficiency in all standards for all content areas.
3. The district respects the diversity of students by age, disability, ethnicity, gender, national origin, race, religion, and sexual orientation to provide opportunities for all students to participate. Procedures for ensuring fair participation of all students in the system are available in the WCSD #1 Administrative Procedures book.

Section 5: STANDARD SETTING

INTRODUCTION

The standard setting process determines scores that identify each performance level of proficiency on an assessment and/or collection of assessments. Washakie County School District #1 selected the Bookmarking Method to establish cut scores at benchmark, standard, and content levels and created a standard setting process that involved an item-based method for closed-ended assessment items and student work-based methods for open-ended assessment items. Determining cut scores at these levels supports a true compensatory model, for students do not need to demonstrate proficiency with every assessment, benchmark, or standard to be proficient in a content area.

The rationale for selecting the Bookmarking Method was based on the versatility of the process, useable with both multiple-choice (closed-ended) and constructed response (open-ended) items. Additionally, bookmarking is useful with rubrics and checklists, which are frequently used with authentic learning.

PROCESS OF STANDARD SETTING

When beginning the process of developing a district BOE, stakeholders, including parents, teachers, and community members, were assembled to create a description of proficiency and to determine supporting evidence to reflect that description. The stakeholders supported performance descriptors that the district adopted in each content area from the Wyoming Department of Education State Content and Performance Standards at established grade levels within the K-12 continuum and strongly emphasized district assessment congruency with the state performance standards.

Input from stakeholders supported the basis for content and standard level proficiency on a four-point scale, as...

1 = Below Basic Performance Level

2 = Basic Performance Level

3 = Proficient Performance Level

4 = Advanced Performance Level

Stakeholder support for establishing cut scores at the content and standard levels, as a compensatory application, was similar to that established at the state level for transcript endorsement. At the standard level, proficiency is identified with a score of three (3) or more on 60% of benchmarks, and advanced is identified with a score of four (4) or more on 60% of benchmarks. At the content level, proficiency is identified with a score of three (3) or more on 60% of standards and advanced is identified with a score of four (4) or more on 60% of standards.

K-12 assessments were developed by the staff to include closed-ended assessments and open-ended assessments with scoring rubrics. Throughout the process the staff was cognizant of alignment with the state performance standards, congruent with the appropriate depth of knowledge (Norm Webb's Depth of Knowledge). Samples of assessments were collected to represent levels of performance, and by consensus it was decided to have the staff collaboratively develop cut scores at the assessment level for benchmarks in each content area. Common assessments were created in all courses for each content area which determined student proficiency on a benchmark in a standard and provided multiple opportunities for students to achieve and demonstrate proficiency.

To establish benchmark cut scores, closed-ended assessments were examined to ensure congruency with performance standards, and performance descriptors were written for each assessment in each content area aligned to performance standards. Assessment items were examined, arranged from the most challenging down the cognitive continuum to the least challenging. Within each of the content areas, teachers individually ranked assessment items according to benchmark alignment and performance descriptors, independently identifying levels of proficiency by marking the transition point between below basic and basic, basic and proficient, and proficient and advanced for each assessment item written at the benchmark content level. After independent data was collected, teachers met as teams and evaluated assessment ranking, and performance levels were established with consensus from the team. Content and grade level teachers established the overall proficiency levels according to performance level descriptors for each assessment at the benchmark level.

Establishing cut scores for open-ended assessments involved performance descriptors of rubric design as a constant for identifying student performance levels. A collection of student work was examined, arranging a sampling of student work from the highest scoring down the continuum to the lowest performing level. Within each of the content areas, teachers individually ranked student work according to performance descriptors, independently identifying levels of proficiency by marking the transition point between below basic and basic, basic and proficient, and proficient and advanced for each assessment item written at the benchmark content level. Through team consensus proficiency levels, aligned to the performance standards and cognitive demand, were established according to the performance rubric, congruent with state performance standards. Anchor papers were selected to represent cut scores at each level of proficiency.

Protocol implemented in the standards setting process includes a minimum of three readers on a team (1) to verify student performance throughout an assessment; (2) to select anchor papers (or identify assessment items congruent with performance levels) and proficiency level transitions; and (3) to verify levels for anchor papers and sign off for each assessment. When assessment teams considered benchmark assessment scores, it was necessary to reflect and respect proficiency at the standard and content levels, which was established by all stakeholders. With the variety of assessment methods used in the district in all content areas, the benchmarking method of standard setting for the K-12 district affords the opportunity for all content areas to equitably evaluate student performance, maintaining consistency within the entire content arena and using the same scoring scale throughout all performance levels.

Established Cut Scores

Thus, the scoring protocol was implemented, and cut scores were established for each assessment at the benchmark level for standards in each of the content areas. According to the assessment blueprint matrix, students are provided multiple opportunities to achieve proficiency. All assessment scores are reported on the student Standards' Report Card at the benchmark level during a student's tenure in school, and scores are aggregated to determine proficiency.

Proficiency Levels:

Assessment items are developed at the benchmark level and meet the content and performance descriptors of one or more benchmarks within an assessment, and proficiency is determined through a compensatory approach.

Standards Level:

- The district deems a student proficient in a standard if at least 60% of benchmarks are met at the proficient or advanced level of the assessments.
- The district deems a student advanced in a standard if at least 60% of benchmarks are met at the advanced level of the assessments.

Content Level:

- The district deems a student proficient in a content area if at least 60% of standards are met at the proficient or advanced level of the assessments.
- The district deems a student advanced in a content area if at least 60% of standards are met at the advanced level of the assessments.

Graduation Requirement:

- General Education Endorsement: To fulfill the Body of Evidence graduation requirement for a General Education Endorsement, a student must be deemed proficient (or advanced) in the majority of the content areas (five of nine).
- Comprehensive Endorsement: To fulfill the Body of Evidence graduation requirement for a Comprehensive Endorsement, a student must be deemed proficient (or advanced) in all the content areas (nine of nine).
- Advanced Endorsement: To fulfill the Body of Evidence graduation requirement for an Advanced Endorsement, a student must be deemed advanced in the majority of the content areas and proficient in the remaining content areas.

Rationale for Establishing Cut Scores by Stakeholders:

Proficiency level majorities were established at the standard and content level to support fairness and validity of the assessment system. Majority consideration is supported with the compensatory approach by Wyoming statute with a general endorsement of proficiency in five of nine content areas and an advanced endorsement of advanced proficient in five of nine proficient content areas. Concerns of stakeholders under consideration during the standard setting process were (1) how prepared will the student be at graduation, and (2) how good is good enough. After careful examination of multiple opportunities and rigor of assessments at the benchmark level, stakeholders expressed fairness to students and discussed the outcome of the results. Will students be adequately prepared when they graduate? Is this a measure of a complete education for college or career-ready? Through consensus of the stakeholders, averaging benchmark scores did not seem an option, especially with data accumulated over a four year span affording students the opportunity to improve proficiency during the students' four years of schooling. At the onset of the development of the Body of Evidence numerous stakeholder meetings resulted in examining a variety of options for students to achieve proficiency at the high school level. During 2008-2009 and 2009-2010, the district BOE Committee and parent-community voices supported the district standard setting process and adjusted the outcomes for graduation supporting cut scores described.

PROGRESS IN MEETING THE STANDARDS FOR GRADUATION

...Protocol for High School Graduation

Communication with Students and Parents...

- ♦ Pre-High School: Students are made aware of the requirements to graduate from high school and review opportunities to qualify for the Hathaway scholarship. An orientation day for parents and students, presented by high school personnel (including guidance counselors, principals, and/or teachers), provides a timeline with requirements for high school graduation for the incoming class. Students are given a copy of the high school enrollment guide and are directed toward the requirements for graduation, indicated in the schematic in the guidebook. To reinforce the importance of this information, students are referenced to the graduation requirements upon registration. Additionally, this information is conveyed (a) to parents of special needs students during student IEP meetings, (b) to parents of ELL students, and (c) to other parents when requested.
- ♦ At the end of each semester during a student's tenure at the high school, students are notified regarding the status of completion of standards, Carnegie units, and required course compliance in each of the content areas in meeting graduation requirements. A hard copy of this information is sent to parents, and parents are welcome to visit with school personnel regarding student progress. Additionally, parents and student have the option to access the student information system, Infinite Campus, to monitor student academic progress.
- ♦ During high school class meetings at the beginning of each school year, students at each grade/class level are informed about graduation requirements in addition to other opportunities, as the Hathaway Scholarship and credit recovery. During junior and senior class meetings, students are informed about opportunities for scholarships.
- ♦ If a student falls behind during any year, a counselor reviews graduation requirements for the student and provides counseling for meeting content standards' requirements and for credit recovery opportunities, such as (1) approved distant education classes/courses and (2) summer school or after-school tutoring.
- ♦ Special needs students, English language learners, 504 students, and other students on individual learning plans are carefully monitored during the student's schooling. The process of progress monitoring is done on an individual basis by counselors, ELL staff, and/or special education staff. This information is communicated frequently during the year and is shared during the spring of each year at scheduled IEP or ILP meetings.

Continuous Communication in Meeting Standards for Graduation...

- ♦ Parents have access to Infinite Campus (electronic), updated and aggregated daily, to view their child's progress in meeting standards (Standards Report) and grades (Carnegie Units) in addition to many other items of interest.
- ♦ High school parents are informed twice during the year and K-8 parents are informed annually regarding student progress in meeting standards.
- ♦ Students meet with the guidance counselor after developing an Individual Graduation Plan during Grade 9.

Meeting the Standards for Graduation...

- ♦ Students are provided multiple opportunities to attain proficiency of the standards in all nine content areas during their four years of schooling. Assessments are written at the benchmark level, overlapping in various courses and are administered during students' schooling. The district uses a compensatory method to determine proficiency.

Section 6: COMPARABILITY

INTRODUCTION

Comparability ensures that all students receive equal educational opportunities. The district balanced assessment system requirements are comparable for all students in each school within a given class (or grade level) and across years. Common assessments are used throughout the district at all grade levels in all content areas. If new items/tasks are used in the body of evidence, assessment procedures are implemented to maintain year-to-year comparability. Comparability is monitored continuously by staff through self-evaluation, grade level or content departments, instructional facilitators, and administrators. Staff members are involved in the process of developing the assessments as well as in implementing them, following the protocol for alignment, consistency, fairness, and standard setting and are expected to sustain the process.

PROCESS

The effectiveness of the assessments, the resulting data, and assessment protocol are monitored at the district level to sustain comparability at the K-12 district level.

Assessment/Item Replacement:

- ◆ Not all assessments or assessment items maintain credibility across the years; thus, the following procedure, referenced in the district assessment plan, is established for assessment replacement. As standards are revisited at the state and district levels, there is a need for assessment revision and/or replacement. Teachers and/or administrators may request improvements in assessments to target cognition and/or item congruency with curricular program alterations or revisions. Changes in any assessments are submitted with a rationale for change to central office before implementation.
- ◆ The district consistently monitors assessment revisions and assessment item replacement for all benchmark assessments. If there is a need to replace an assessment or assessment items, a panel of teachers, familiar with the content or grade, reviews such revisions and/or replacements. Alignment, consistency, fairness, standard setting, and comparability are to be addressed to ensure appropriate replacement.
- ◆ Accommodations and/or alternate assessments are encouraged to meet individual student needs, especially students within identified subgroups.
- ◆ If a course is added to a program of study in the district, any assessment that is given in the course that is used to measure student achievement in any of the district content and performance standards must follow the district protocol for compliance in the district body of evidence.
- ◆ For assessment completion, revision, or replacement, refer to the district guidelines which cover the following:
 - Assure alignment with content and performance standards.
 - Review for comparable agreement on performance alignment and scoring agreement
 - Review for bias, using the district protocol for bias review.
 - Identify necessary accommodations that may be necessary to meet the needs of various subgroups, as ELL, SPED, F/RL, ethnicity, gender, etc.
 - Develop appropriate performance descriptors to evaluate student performance when first administration of items or assessment

- Select anchor papers after administration

Staff Training:

- ♦ All K-12 staff is aware of the district assessment plan and is trained in the body of evidence criteria and district assessment protocol.
- ♦ New staff members are mentored in assessment protocol and are monitored through the process as active participants in assessment practices. Instructional facilitators, cognizant of the district body of evidence protocol, provide guidance and support during the mentoring period.
- ♦ Assessment training for all K-12 instructional personnel, involving performance descriptors, examination of student work, and scoring practices, is provided annually or on demand and supported one-on-one by content or grade level teams. The district staff development plan reflects support in understanding the district assessment plan and district BOE protocol.

Year-to-Year Maintenance and Implementation:

- ♦ Instructional facilitators, grade level teacher leaders, and content teacher leaders are trained to guide staff and maintain consistency in implementation of assessments administration and scoring procedures (including common assessment rubrics) through peer-to-peer training and mentoring staff new to the district. Anchor papers, as exemplars, are collected continuously by the grade level teams and content specific groups to use for updated, effective exemplars.
- ♦ All staff involved in assessments is encouraged to systemically and systematically evaluate assessments through student performance data and student feedback.

5-Year Consistency Cycle – An On-going Process:

The district assessment plan addresses consistency as an on-going process. Reliability analysis is conducted on 20% of district standards' assessments each year resulting in a 5-year cycle for consistency. Rationale for this procedure allows for staff to...

- 1) Revisit assessments for revisions;
- 2) Adjust to adoptions/changes in standards and benchmarks;
- 3) Train new staff on procedures within the assessment system;
- 4) Adjust to textbook/program adoption; and
- 5) Afford staff to create more appropriate assessments.

Content areas, by grade level or content specialty, can determine assessments for review during each year. All completed data information is submitted to the Curriculum Director.

Each year specific content curriculum, (core and non-core) and instructional strategies are targeted for careful examination; new textbooks, instructional practices, and technology resources are considered for adoption. Each content area rotates through the process within a five- to seven-year rotation, followed by one year of assessment pilot. Teachers address the content area for examination and are afforded the latitude to extend assessment examination in all other content areas for consideration, especially for those that are need of change, replacement, and/or adjustment. Content specific programs will examine 20% of all assessments. When new assessments are developed, alignment, fairness, and standard setting will accompany assessment development, and fidelity to the process ensures comparability within the Body of Evidence.

***Note:** Additional consideration will be given to programs requiring changes in supplemental curricular needs, students at risk, RTI, differentiation of instruction, and students with special needs. The district realizes that compliance with WDE requirements for the body of evidence criteria must be dynamic, constantly changing to meet WDE Content and Performance Standards that will begin to undergo changes during the 2009-2010 school year. This will affect the cognition and benchmark blueprint matrices annually and impact each criterion, especially the comparability criterion, which will impact the protocol for continuous assessment development as content areas undergo revisions of the WDE Content and Performance Standards on a continuous rotational cycle. The established BOE district protocol is designed to adjust to the dynamic nature of the process.

Adding, Deleting, and/or Changing a District (Body of Evidence) Assessment:

The district affords teachers the latitude to change assessments, when the need arises, to address situations as frequent assessment use, curricular changes, changes in the district/state standards/benchmarks, increasing multiple opportunities, or assessment variation. Whatever situation might be prevalent, assessment development must go through the established district body of evidence criteria, following district protocol for alignment, consistency, fairness, standard setting, and comparability adopted by the Washakie County School District #1 Board of Trustees. Minor changes (less than 25%) to an established and approved assessment does not warrant the major criteria protocol, but changes do require the endorsement of the content area department in grades 6-12 or the grade level team for grades K-5 to ensure district body of evidence compliance.

APPENDIX A – BOE OVERVIEW

...WDE Body of Evidence Glossary

Advanced Endorsement Diploma: Advanced in the majority of 9 content areas – Proficient in the remaining content areas

Alignment: is the core idea that in a standards-based system, all parts of the system work together to ensure student learning of the content standards. In terms of the relationship between standards and assessment, alignment refers to the match between the items on the assessments and the knowledge and skills represented by the curriculum and the standards.

Assessment: designed to measure students' progress toward meeting the content standards at specific benchmarks, providing information about the achievement and performance levels of the students.

Assessment System: a well-articulated set of assessments, designed to provide a cohesive array of information on student performance, which contributes toward supporting inferences related to the identified purposes and has the following characteristics:

- a. Comprehensively addresses the content and performance standards or other well-defined learning targets.
- b. Provides students with multiple opportunities, using multiple formats, to demonstrate their knowledge and skills related to the Wyoming Content and Performance Standards.
- c. A clear and explicit rationale for each individual assessment.

It should be able to support inferences about students' mastery of the common core content areas and the districts' capacity for ensuring equality of educational opportunity.

Balance of Representation: "The balance-of-representation criterion is used to indicate the extent to which items are appropriately distributed across standards" (Webb). Assessment items/tasks should match the standards in proportion/weight.

Body of Evidence Assessment System: Determines whether students have met graduation standards and to provide a collection of evidence to support this decision.

Categorical Concurrence: "The categorical concurrence between the standards and assessment is met if the same or consistent categories of content appear in both documents" (Webb). The categorical concurrence between the assessment and the standards is assured if the assessments that comprise the assessment system are faithfully built directly from the standards.

Comparability: Ensure that all students receive equal educational opportunities.

Compensatory Model: Students DO NOT have to demonstrate proficiency on every benchmark to be proficient on the standard, allowing higher scores on some measures to offset (compensate for) lower scores on other measures - the simple average.

Comprehensive Endorsement Diploma: At least proficient in all content areas

Conjunctive Model: Scores on **all** measures used must be above the criterion point (cutscore) for the student to have met the overall standard - student would have to be above the cut score on EACH measure to be considered proficient

Consistency: Indicates the focus is on the assessment system - the set of assessments should yield consistent decisions about students' performance related to the standards.

Content Standards: Clear and rigorous statements that target what students should know and be able to do as a result of instruction in the common core of knowledge and skill areas. These should be clear and rigorous.

Course-Based Approach: Grades in courses are based on achievement of the standards

Course-Based and Common Assessment Approach: Course-based approach that includes the use of common assessments

Credibility: Confidence in the school system (best by involving various stakeholder groups, such as key business people, local legislative representatives, parents, and other key community members).

Depth-of-Knowledge Consistency: “Depth-of-knowledge consistency between standard and assessments indicates alignment if what is elicited from students on the assessment is as demanding cognitively as what students are expected to know and do as stated in the standards” (Webb). Alignment should be judged in terms of matches between test items and standards and performance levels established for the assessment, targeting the cognitive levels in the standards.

Disjunctive Model: Students only need to demonstrate proficiency on any one of the multiple measures used to be considered proficient

District-Based Approach: Relies on stand-alone assessments at key checkpoints during high school career

Fairness: Need to be concerned about whether the set of assessments leads to unfair inferences about students’ performance on the standards and the issue is based on the analyses of the assessment issues.

General Endorsement Diploma: At least proficient in the majority of content areas

Mixed-Model: combines features of conjunctive, disjunctive, and/or compensatory, as using both compensatory and conjunctive approaches might require students to score above a threshold score on every assessment (conjunctive), but above an average score on the full set of assessments (compensatory).

Mixed Model Approach: Represent a combination of multiple approaches

Multi-District Approach: Relies on collaboration among districts to reduce workload and to produce higher-quality assessments than can be produced by an individual district

Multiple Measures: Used to increase the validity, alignment, consistency, and fairness of the assessment system and represent different formats and strategies to enhance the fairness and alignment of the system. Enough measures should be employed so that the alignment and fairness criteria are improved over a single assessment.

Performance Standards: describe the characteristics of students at various levels of performance. They describe “how good is good enough?” These should clearly differentiate what students “look like” at various stages of learning the content, and the performance standards should be clearly measurable.

Range-of-Knowledge Correspondence: *“The range-of-knowledge criterion is used to judge whether a comparable span of knowledge expected of students by a standard is the same as, or corresponds to, the span of knowledge that students need in order to correctly answer the assessment items/activities. The criterion for correspondence between span of knowledge for a standard and the assessment considers the number of objectives within the standard measured with at least one related assessment item/activity”* (Webb, 1999, p. 8). Alignment should occur at a finer grain (benchmark level) than simply matching tests items to broad standards, meaning that assessments should be built from a blueprint that reflects the appropriate weight of each benchmark.

Sampling: Using a random number of selections/assessments or at critical points to reduce the number and amount of time devoted to district-level assessments, yet still yield adequate data for school improvement and accountability. A district can get a representative sample of student knowledge by collecting data at critical points and intervals throughout the K-12 continuum.

Standard Setting: The process of determining the scores that divide various performance levels (i.e., cutscores) on an assessment with cutscores tied closely to proficiency descriptions – should know what it means to pass or be proficient and there should be evidence that the cutscore reflects this description.

Standardized Assessment Approach: Results of standardized assessments such as norm-referenced tests (NRT), district achievement-level tests (DALT), and PAWS

Standards-Based System: include assessments that are closely linked to standards and assessment results that are reported in terms of the performance standards.

Two-Way Alignment: All items on assessments can be matched to one or more standards and all the standards (content and performance) are adequately assessed, occurring at the benchmark level or finer grain to ensure that the standards are appropriately sampled. This does not mean that each benchmark should be assessed. Items are matched against the standards and the standards are checked to make sure they appropriately represent the items.

Validity: The degree to which a test measures what it is intended to measure - it is the inferences in the context of a particular use that are valid or not.

Wyoming Body of Evidence Assessment System: Provides students with multiple opportunities, using multiple formats, to demonstrate their knowledge and skills related to the Wyoming Content and Performance Standards.

Approaches Considered toward BOE Design...

Approach	Descriptor	Advantages	Disadvantages
Course-Based	<ul style="list-style-type: none"> – clearly tied to standards – grades in courses are based on achievement of the standards – Body of Evidence consists of course grades from these "standards" courses. 	<ul style="list-style-type: none"> – increases teacher buy-in and ownership – could be the most instructionally useful – results are timely need to meaningfully adjust instruction – could receive valuable feedback for modifying or targeting instruction 	<ul style="list-style-type: none"> – impacts accuracy and consistency of grading practices which lead to credibility – compromises comparability with more than one teacher teaching the standards in a particular content area – requires more work/responsibility for individual teachers.
Course-Based & Common Assessment	<ul style="list-style-type: none"> – course-based approach that includes the use of common assessments – provides a way to anchor grades from different teachers across the district – common assessments carry a certain percentage of weight, and information from the course carry the remainder of the weight – common assessments carry enough weight to realistically anchor student grades – can be "on-demand" for projects or other types of measures tied to standards (i.e. BOE Consortium) 	<ul style="list-style-type: none"> – helps to ensure equivalent opportunities for students to learn the standards – allows individual teacher choice and discretion – encourages teacher collaboration with teachers designing and writing common assessments – BOE Consortium Activities allow for opportunity to collaborate (without having to do all of the work) – using common assessments of high quality can help alleviate some of the grading problems 	<ul style="list-style-type: none"> – weighting common assessments may lead to potential problems as a result of teacher grading practices – requires time for teachers to collaborate
District-Based	<ul style="list-style-type: none"> – relies on stand-alone assessments at key checkpoints during high school career – assessments could be at the end of specific courses or grades – results from assessments comprise the Body of Evidence system if set of assessments used measure all of the standards appropriately as documented through alignment evidence 	<ul style="list-style-type: none"> – fewer records to keep – fewer decisions to make – easier to ensure technical quality as long as exams are updated regularly and there are enough assessments in each content area to ensure reliability – may have more credibility with the public since easy to explain 	<ul style="list-style-type: none"> – could decrease teacher buy-in if they felt that students' performance in their class did not matter since only students' scores on the district assessments count – doubtful that on-demand tests could cover the full range of standards – district would need other types of assessments to measure all the standards (e.g., speaking/listening). – need to represent a range of formats to satisfy fairness and alignment – using a single type of assessment would be a serious disadvantage – if the same assessments are used over time, they must be updated (and equated to some extent) to ensure year-to-year comparability
Mixed Models	<ul style="list-style-type: none"> – represent a combination of multiple approaches 	<ul style="list-style-type: none"> – a district's BOE system does not have to look the same for each content area because some content areas might lend themselves to one approach over another – allows districts to rely on the best features of these multiple approaches – minimizes disadvantages and maximizes advantages 	<ul style="list-style-type: none"> – represent a combination of disadvantages included in course-based, common assessment, and district-based
Washakie #1	<p>Selected... Course-Based & Common Assessment</p>	<p>Additional Advantages involves...</p> <ul style="list-style-type: none"> – opportunity to collaboratively weight common assessments and/or standards/benchmarks, through effective teacher collaboration – supports allocating time for teachers to collaborate in examining student work and establish student expectations – encourages staff to develop vertical alignment and transition between grades/school levels – anchors course grades 	

...Right to Appeal

Worland High School Board of Appeals

If a student's high school transcript and proficiency report do not indicate that the student has met the requirements for graduation, the student has the right to appeal to the Board of Appeals. The Board of Appeals will consider any evidence presented by the student, or on the student's behalf, and decide if the student should be granted graduation status.

The Board of Appeals consists of the following members:

- High School Principal
- High School Counselor
- Classroom Teacher #1 (from varying content areas)
- Classroom Teacher #2 (from varying content areas)
- Classroom Teacher #3 (from varying content areas)

If the student has not met graduation requirements after a review of the student appeal, the Board of Appeals will have the authority to design an individualized student plan for the student to have an opportunity to demonstrate proficiency on the common core of knowledge and skills to meet graduation requirements.

Individualized Student Plan could include the following:

- Student interviews with staff members
- An observation demonstrating knowledge and skills
- A review of additional assessment information, as
 - Records relating to student performance,
 - Completion of additional assessments, and/or
 - Other appropriate activity or performance that the Board considers appropriate.

A student has the right to appeal the decision of the Board beyond the Worland High School Board of Appeals, which must be done in writing to the school district administration.

...Balanced Assessment System

Assessment is an integral part of instruction, providing information about how well students are progressing toward state standards. Standards, curriculum, instruction, and assessment are interwoven in the learning process, and each informs the others.

- Standards are the *what* – defining what we expect students to know and be able to do.
- Curriculum is the *how* – defining a body of learning experiences that are designed to reach standards.
- Instruction is the *in what ways* – providing specific learning experiences and ways of differentiating experiences to scaffold student learning.
- Assessment is the *how well* – gauging the attainment of learning to inform instructional practices and curriculum.

A balanced approach to assessment includes a continuum of strategies within a range of frequency and purpose.

Formative assessment occurs continuously in the classroom, both within and between lessons. Information is used to adjust teaching strategies. Students receive frequent and meaningful feedback on their performances to include teacher observation, discussion, questioning, and non-graded class work. The focus of formative assessment is to determine what learning comes next for a student.

Benchmark assessment (BOE) occurs within, between, and among instructional units. Information is used to identify strengths and gaps in curriculum and instruction. Grade-level curriculum may be refined, and teachers may modify instruction for student groups based on their progress. These assessments might be teacher-developed or district-developed products, or they are purchased commercially and linked to state standards. Benchmark assessments include midterm, end-of-unit assessments, semester assessments, and/or district-wide assessments. The focus of benchmark assessment is to determine how student groups are progressing or how well a program is working.

Large-scale assessment occurs annually. Information is used to develop strategic, long-term evaluation of curriculum and programming based on trends over time, and to monitor state, district, and school progress. Assessment instruments and procedures are standardized so that comparisons can be made across student groups. Examples of large-scale assessment include the PAWS®, ACT®, NAEP™, and AP® exams. The focus of large-scale assessment is to determine how schools, districts, and states are progressing, which is reflected in the district assessment matrix. (Appendix A: 25-29)

These assessment strategies provide information at differing intervals and for different purposes. Each one provides a different perspective, and one cannot take the place of another. Together, these assessments provide a balanced approach to assessment that informs decisions at the classroom, school, district, and state levels.

Balanced Approach to Assessments

Teacher Evaluation Strategies	BOE - Benchmark Assessments	Large Scale Assessments
Daily – Ongoing	Monthly/Quarterly/Semester	Annual
Student-Centered	Student/Classroom/School--Centered	School/District/State-Centered
Formative	Interim and Summative	Summative

APPENDIX B – ALIGNMENT

Useful information for completing the...

District Assessment Matrix

Assessment Types	
Performance/Authentic Assessment	Writing Prompt
Selected Response	Extended Project
Constructed Response	Observation
Portfolio	Journal
Participation/Interaction	Checklist

Levels of Cognition	
AKA: Depth of Knowledge - Cognitive Complexity - Cognitive Demand	
Recall	Recall fact, information, or procedure.
Application	Application of Skill/Concept: includes use of information, knowledge, or procedure.
Strategic Thinking	Strategic Thinking requires reasoning, developing a plan or sequence of steps; has complexity or more than one possible answer.
Extended Thinking	Extended Thinking requires an investigation, time to think and process multiple conditions of the problem/task.

Performance Matrix: Webb’s Depth of Knowledge

Webb	Recall	Application (of Skill/Concept)	Strategic Thinking	Extended Thinking
Description	Recall of a fact, information, or procedure	Use of information, conceptual knowledge, procedures (e.g. two or more steps), etc.	Requires reasoning, developing a plan or sequence of steps; has some complexity with cognitive demands as somewhat complex and abstract	Requires an investigation; time to think and process multiple conditions of the problem or task
Verbs as Descriptors	Recall of a fact Definition Term Identify Recognize Use Measure Describe Explain	Classify Organize Estimate Construct Make observations Collect Display Compare Explain Describe Interpret Summarize Infer Determine fact or opinion	Reasoning Planning Using evidence Draw conclusions Cite evidence Explain Use concepts Analyze and describe Support Edit	Connect Relate Consider alternatives Design Conduct Make connections Combine and synthesize Critique Analyze Examine and explain, Describe and illustrate Create Identify and generate

Coordinating Webb's Depth of Knowledge to Bloom's Taxonomy...

Performance Matrix: Bloom's Taxonomy (Old)

Old Bloom's	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
Description	Usually rote memory of factual material	Involves translating from one level of abstraction to another (e.g., give an example of a concept or principle)	The ability to use learned material in novel situations	The ability to break down material into its component parts so that its organizational structure may be understood (e.g., distinguishing fact from opinion)	The ability to put parts together to form a new whole – usually involved in producing a product	The ability to judge the value of material (statement, novel, poem, research report) for a given purpose. (this is like analysis but with logical and defensible value judgment)
Verbs as descriptors	Tell, list, define, label, recite, memorize, find, repeat, name, record, fill in, recall, relate	Locate, summarize, explain, identify, describe, report, discuss, review, paraphrase, restate, retell, show, outline, rewrite	Demonstrate, use, construct, record, diagram, revise, reformat, illustrate, interpret, dramatize, practice, organize, translate, manipulate, convert, adapt, research, calculate, operate model, order, display, implement, sequence, integrate, incorporate	Compare, contrast, classify, critique, categorize, solve, deduce, examine, differentiate, appraise, distinguish, experiment, question, investigate, categorize, infer	Compose, design, hypothesize, create, formulate, invent, develop, refine, produce, transform	Judge, predict verify, assess, justify, rate, prioritize, determine, select, decide, value, choose, forecast, estimate

Performance Matrix: Bloom's Taxonomy (New)

New Bloom's	Remembering	Understanding	Applying	Analyzing	Creating	Evaluating
Description	Exhibit memory of previously learned material by recalling facts, terms, basic concepts, and answers.	Demonstrate understanding of facts and ideas by organizing, comparing, translating, interpreting, giving descriptions, and stating main ideas.	Solve problems to new situations by applying acquired knowledge, facts, techniques, and rules in a different way.	Examine and break information into parts by identifying motives or causes. Make inferences and find evidence to support generalizations.	Compile information together in a different way by combining elements in a new pattern or proposing alternative solutions.	Present and defend opinions by making judgments about information, validity of ideas, or quality of work based on a set of criteria.
Verbs as descriptors	Choose, define, how, label, find, list, match, name, omit, recall, relate, select, show, spell, tell, what, when, where, which, who, why	Classify, compare, contrast, demonstrate, explain, extend, illustrate, infer, interpret, outline, relate, rephrase, show, summarize, translate	Apply, build, choose, construct, develop, experiment with, identify, interview, make use of, model, organize, plan, select, solve, utilize	Analyze, assume, categorize, classify, compare, conclude, contrast, discover, dissect, distinguish, divide, examine, function, inference, inspect, list, motive, relationships, simplify, survey, take part in, test for, theme	Adapt, build, change, choose, combine, compile, compose, construct, create, delete, design, develop, discuss, elaborate, estimate, formulate, happen, imagine, improve, invent, make up, maximize, minimize, modify, original, originate, plan, predict, propose, solution, solve, suppose, test, theory	Agree, appraise, assess, choose, compare, conclude, criteria, criticize, decide, deduct, defend, determine, disprove, dispute, estimate, evaluate, explain, importance, influence, interpret, judge, justify, mark, measure, opinion, perceive, prioritize, prove, rate, recommend, rule on, select, support, value

APPENDIX C – CONSISTENCY

...Consistency: Open-ended Assessments

Examining Student Work...

Open-ended Assessments: Open-ended assessments are an integral part of the Body of Evidence system, to evaluate skills, abilities, and/or application of depth of knowledge (Norm Webb) and are scored with a two- or four-point, performance-based scoring rubric representing cognition levels, corresponding to the Wyoming performance standards. Each assessment is accompanied with explicit directions, requirements, scoring procedures and rubrics to ensure the reliability of the assessment task.

Open-ended Assessment Scoring Procedures to Ensure Inter-rater Reliability...

- Teacher:
 1. Select a minimum of 25% student papers for each assessment for review, preferably representing all performance levels according to proficiency descriptors, congruent with the state performance standards.
 2. Remove student names from all student work selected for scoring.
 3. Select and identify two scorers and a scoring designee.
- Scorers:
 1. Review the assessment (or assessment item) scoring rubric for clarification, if necessary.
 2. Score each assessment independently without conversation or collaboration.
 3. On completion of scoring submit scores to the scoring designee.
- Designee:
 1. Compile data and perform inter-rater reliability calculations on the Inter-rater scoring sheet.
 2. Evaluate the results of assessment data regarding exact and adjacent rating.
 3. Determine acceptable inter-rater reliability scoring, which must meet or exceed 80% exact agreement and 98% exact and adjacent agreement.
 4. If the scorers are in agreement, deem that the scoring is consistent.
 5. It is acceptable if scorers score items in adjacent proficiency levels as long as the percentage of a combination of adjacent and exact agreement is equal to or greater than 98%.
 6. Consult and discuss scores that do not meet percentage range and/or scores that differ by more than one level with the curriculum director.
- Curriculum Director: Review all discrepancies (unacceptable inter-rater agreement) in the scoring. Consistent discrepancies in scoring may warrant a need to consider the following interventions with follow through:
 - (6) Conference with scorers,
 - (7) Provide an additional scorer to independently score the assessment,
 - (8) Evaluate the effectiveness of the scoring rubric,
 - (9) Revisit the prompt or question, and/or
 - (10) Provide training to improve scorer agreement.

Note: The identical process is used for open-ended assessment items as constructed response items.

Inter-rater Scoring Sheet (Open-ended Assessment or Item)

Content Area: _____

Course: _____ Standard/Benchmark(s) Assessed: _____

Assessment (Item): _____ Date: _____

Student Paper	Scorer #1 Rubric Score	Scorer #2 Rubric Score	Inter-rater Agreement	
			Exact	Adjacent
1				
2				
3				
4				
5				
6				
7				
8				

Record rubric score data from the Individual Scoring Form for each Scorer

Indicate if the scores were either exact or adjacent

- Enter number of exact, adjacent, and exact-adjacent.
- Determine % age of exact, adjacent, and exact-adjacent
- Is the data acceptable or not?
- If not, provide a recommendation for consideration

Designee:				Comments:
Inter-Rater Agreement	Exact	Adjacent	Exact and Adjacent	
	# _____	# _____	# _____	
	_____ %	_____ %	_____ %	

The forms to record information are located on the district I-Drive at
I-Drive > Central Office > Curriculum > BOE

Further Action Designee and Curriculum Director	Accepted	Recommendation: <input type="checkbox"/> Conference with scorers <input type="checkbox"/> Provide an additional scorer to independently score the assessment <input type="checkbox"/> Evaluate the effectiveness of the scoring rubric <input type="checkbox"/> Revisit the prompt or question <input type="checkbox"/> Provide training to improve scorer agreement
	Yes <input type="checkbox"/>	
	No <input type="checkbox"/>	

Scoring Form (Open-ended Assessment or Open-ended Item) Content Area: _____

Course: _____ Standard/Benchmark(s) Assessed: _____

Assessment (Item): _____ Date: _____

Scorer: _____ Designee: _____

Student Paper	Rubric Score	Comments	Anchor Recommendation
1			
2			
3			
4			
5			
6			
7			
8			

Using the rubric accompanying the assessment, and record the score that you have deemed correct.

Check for Anchor Paper consideration.

Optional: Include any comments.

Additional Comments:

Inter-rater Scoring Sheet (Open-ended Assessment or Item)

Content Area: _____

Course: _____ Standard/Benchmark(s) Assessed: _____

Assessment (Item): _____ Date: _____

Student Paper	Scorer #1 Rubric Score	Scorer #2 Rubric Score	Inter-rater Agreement	
			Exact	Adjacent
1				
2				
3				
4				
5				
6				
7				
8				

Designee:	Wise			Comments:
Inter-Rater Agreement	Exact	Adjacent	Exact and Adjacent	

Further Action Designee and Curriculum Director	Accepted	If no, recommendation: <input type="checkbox"/> Conference with scorers <input type="checkbox"/> Provide an additional scorer to independently score the assessment <input type="checkbox"/> Evaluate the effectiveness of the scoring rubric <input type="checkbox"/> Revisit the prompt or question <input type="checkbox"/> Provide training to improve scorer agreement
	Yes <input type="checkbox"/>	
	No <input type="checkbox"/>	

Scoring Form (Open-ended Assessment or Open-ended Item) Content Area: _____

Course: _____ Standard/Benchmark(s) Assessed: _____

Assessment (Item): _____ Date: _____

Scorer: _____ Designee: _____

Student Paper	Rubric Score	Comments	Anchor Recommendation
1			
2			
3			
4			
5			
6			
7			
8			

Additional Comments:

Scoring Form (Open-ended Assessment or Open-ended Item) Content Area: _____

Course: _____ Standard/Benchmark(s) Assessed: _____

Assessment (Item): _____ Date: _____

Scorer: _____ Designee: _____

Student Paper	Rubric Score	Comments	Anchor Recommendation
1			
2			
3			
4			
5			
6			
7			
8			

Additional Comments:

Consistency: Closed-ended Assessments...

...Examining Student Work

Closed-ended Assessments: As part of the Body of Evidence system closed-ended assessments are given to evaluate content knowledge as an item analysis using the Cronbach's Alpha Coefficient to improve the quality and to establish reliability of questions, with items having an exact answer for each question. Questions such as multiple choice, specific response, and matching, etc. are evaluated using this process.

Closed-ended Assessment Scoring Procedure...

- Teacher:
 - Obtain all examples of student work for the assessment.
 - Record a score of "1" for correct or "0" for incorrect for each student response to every assessment item in an assessment on a reliability spreadsheet provided by Jason Nicholas (Lander, Wyoming, November 2008), which is automatically calculated.
 - After all item scores are entered, identify the Cronbach's Alpha Coefficient.
 - Submit the Cronbach's Alpha Coefficient rate to the content department head or grade level leader.
- Department Head or Grade Level Leader (or Designee):
 - Record all coefficient rates for each assessment.
 - Determine if the rate is acceptable or not.
 - If the coefficient rate meets or exceeds 70% (.70), the assessment has met an acceptable rate, meaning that if additional students were to take the assessment, scores would have that identified rate of the same results. If the coefficient falls below 70%, identify and revisit specific assessment items are with appropriate considerations.
 - Submit a tally of all scores to the Curriculum Director.
- Curriculum Director:
 - Review scores for all assessments.
 - If the coefficient falls below 70%, investigate factors within the assessment to determine potential issues with assessment items and/or instruction, such as:
 - Number of assessment items (the greater the number of items, the greater the coefficient)
 - Examine individual items within the assessment
 - Appropriate assessment alignment
 - Collaboration among grade level or content teachers
 - Curriculum design
 - Instructional quality

Close-ended Assessment Worksheet Using a Reliability Calculator

Cronbach's Alpha is calculated automatically as data is entered.
 Example: 89% (.8948), meets the district acceptable rate of 70% (.70).

Cronbach's Alpha	0.8948
Split-Half Correlation	0.7226
Spearman-Brown Prophecy	0.8389
KR-21	1.0199
KR-20	0.8948
Mean for Test	49.9231
Standard Deviation for Test	12.5113

Reliability Calculator

Grade High School
 Content Area Math
 Course Algebra I
 Teacher Froyd
 School Worland High School
 Assessment Spring Semester Exam
 Standard 1, 3, 4, 5

Questions 50	Students 23
-----------------	----------------

	Question 1	Question 2	Question 3	Question 4	Question 5	Question 6	Question 7	Question 8	Question 9
Student 01	1	0	0	1	1	0	0	1	1
Student 02	1	1	1	1	1	1	1	1	1
Student 03		1	1	1				1	1
Number of Questions and Students Populates Automatically		0	1	1				0	1
		1	1	1	1	1	1	1	0
Student 06	1	1	0	1	1	1	1	0	1
Student 07	1	0	0	1	1	0	1	0	1
Student 08	1	1	1	1	1	1	1	1	1
Student 09	1	0	1	0	1	0	0	1	1
Student 10	0		0	1	1	1	1	0	1
Student 11		1	0	0	1	0	0	1	0
Record a "1" for a correct response		1	0	1	1	1	1	0	1
Student 13	0	1	0	0	1	1	1	0	1
Student 14	1	1	1	1		1	0	0	1
Student 15	0	0	1	1	0		0	0	1
Student 16	1	0	0	0	1				1
Record a "0" for an incorrect response									1
Student 17	0	1	1	1	1				1
Student 18	1	0	0	0	1	0	0	0	0

Steps:

1. The format of the spreadsheet is located on the district I-Drive at I-Drive > Central Office > Curriculum > BOE
2. Enter information about the assessment.
3. Record a "1" for each correct item and a "0" for each incorrect item.

Note: If using this format with another assessment, remove data using the "Clear Contents" command.
 If pasting data directly in the spreadsheet use the "Paste Special" command.
 The Item Statistics tab provides additional information that reflects individual item information.
 Save entered data using a separate file, to keep the original file clean.

Cronbach Alpha Coefficient Data: Tally Sheet

Content Area: _____

Date: _____

Department Head or Grade Level Teacher:

- Record information collected on each assessment.
- Complete all columns. If there no item concerns, please indicate "None."
- Include additional sheets if necessary.
- Date and return to Curriculum Director when complete.

Name the assessment used to measure the benchmark.

Course	Assessment	Quarter	Cronbach α	Concern Item Number	Teacher

Identify the course.

- Indicate when the assessment was given.
- Record the coefficient as a decimal.
- Indicate specific items of concern.
- Name the teacher administering the assessment

APPENDIX D – FAIRNESS

...Bias and Sensitivity Review Procedure

Reader #1:

1. Examine the assessment that is used to evaluate student academic performance to meet standards in the content areas, targeting design, directions, content, and performance.
2. Using the checklist, if present identify concerns in the following:
 - a. Bias
 - b. Stereotyping
 - c. Verbiage/Material
3. At the end of each section, indicated recommendation as USE, REVISE, OR DO NOT USE.
4. Record the information on the “tally” worksheet. Add comments at the bottom of this page for concerns/issues regarding bias and sensitivity.
5. Give the assessment to the Reader #2 for review.

Reader #2:

1. Follow the procedure as for Reader #1 (Steps 1-4).
2. Give the assessment to the Reader #3 for review.

Reader #3

1. Follow the procedure as for Reader #1 (Steps 1-4).
2. Tally the scores.
3. Determine overall recommendation from the bias team.

When all assessments are reviewed, return all bias data to the administration office.

Thank you for all that you do.

Bias and Sensitivity Review

Date: _____

Content Area: _____

Grade Level: _____

Assessment Title: _____

Bias Review Summary...

- No gender bias
- No racial bias
- No ethnic bias
- No religious bias

- No age bias
- No other bias
- Accessibility for LEP
- Accessibility for students with disabilities

Recommendation...

Use

Revise

Do Not Use

Stereotyping Review...

- No gender stereotypes
- No racial stereotypes
- No religious stereotypes
- No socio-economic stereotypes
- No stereotyping of persons with disabilities

- No insensitivity to cultural representation/groups
- No age stereotypes
- No ethnic stereotypes
- No occupation stereotypes
- No geographic/regional stereotypes

Other (specify): _____

Recommendation...

Use

Revise

Do Not Use

Verbiage/Material Review...

Significantly more or less familiar to some groups based on:

- Language
- Socio-economics
- Geographical/regional areas
- Prior knowledge
- Prior experience

Other (specify): _____

Recommendation...

Use

Revise

Do Not Use

Content Area: _____ Grade Level: _____

Assessment Title: _____

Reader #1: _____

Reader #2: _____

Reader #3: _____

Evaluation		Reader #1			Reader #2			Reader #3		
		Use	Revise	Do Not Use	Use	Revise	Not Use	Use	Revise	Do Not Use
Bias	Review Recommendation									
Sensitivity	Stereotyping Recommendation									
	Verbiage Recommendation									
Recommendation (Sensitivity and Bias)										

Recommendation (Circle)	Use	Revise	Do Not Use
-----------------------------------	-----	--------	------------

Comments:

...Equal Treatment of the Sexes: Roles of Men and Women

Educational and Career Opportunities

- Assessment items do not stereotype gender
- Males and females are treated as individuals
- Men or women do not target primary or secondary roles
- There is equal participation of males and females with the same frequency.
- Career choices demonstrate opportunities for both men and women.

Job Stereotypes

- There are no job stereotypes implied for men and women.
- There is no job choice inferior to another.
- Job choice is open to all people.
- Women's full participation in the work force is assumed.
- Men and women are shown equally at all professional levels.
- Men's and women's worth is not predicated on income level or status.
- Men and women have same motivation for work.

Lifestyles

- There is no stereotyping in regard to lifestyle.
- Male and female activities are equitable
- Interests are based on individuality, not on gender.
- There are no implications of effects on family if a woman works.
- Both males and females are engaged in home maintenance and/or activities.

Historical and Cultural

- There are no gender limitations on rights, opportunities, or customs.
- Language or events of a specific time period reflect bias of that time or culture.
- There are diversified choices of case histories (examples) to represent the broadest spectrum of historical/cultural reality.

Traits

- There is no exaggeration of physical differences between males and females.
- There are no implications of male or female aspects of personality, experience, or modes of thought.

Parallel Language, Titles, and Names

- There is use of parallel language for men and women, avoiding subtle stereotyping of women by roles.
- Men's and women's names should be treated comparably.
- There is unnecessary reference on a woman's marital status.
- The same nomenclature is used for the same job or position by a man or woman.

AVOID BIAS: GUIDELINES TO CONSIDER...

The groups of primary concern for these guidelines are defined by age, disability, ethnicity, gender, national origin, race, religion, and sexual orientation.

♦ **Treat People with Respect**

Language and Illustrations: Use language and illustrations that show respect for people in different groups. Avoid language and images that are belittling, condescending, contemptuous, derogatory, demeaning, disrespectful, exclusionary, insulting, or patronizing, unless such content is required for valid measurement.

Societal Roles: Assessments should demonstrate that people in different groups are found in a wide range of societal roles and contexts. Do not over-represent members of any group in examples of lower-status occupations or in examples of inappropriate, foolish, unethical, or criminal behavior.

Problems and Beliefs: Do not treat the problems of any of the groups of primary concern. An impartial tone should be maintained unless a different tone is necessary for valid measurement.

Ethnocentrism: Avoid any indication that any particular group is superior to other groups, or is the standard against which all other groups are measured. Do not use language that assumes all test takers are citizens of the United States.

Underlying Assumptions: Avoid content or images based on underlying assumptions about groups of people. As...“The pioneers and their wives crossed the plains,” assumes that only men were pioneers.

♦ **Minimize the effects of construct-irrelevant knowledge or skills.**

Construct-irrelevant Difficulty: If construct-irrelevant knowledge or skill is required to answer an item, and the knowledge or skill is not equally distributed, then the item is not fair. As...

- **Charts, maps, graphs, and other visual stimuli:** Avoid the use of charts, maps, graphs, and other visual stimuli if they are arbitrarily chosen as one of many possible means of testing some particular point. Charts, maps, graphs, and other visual stimuli are acceptable if the ability to read and understand them is part of the construct.
- **Unnecessary difficult words, figures of speech, idioms, or syntactic structures:** Avoid unnecessary difficult language unless the purpose of the item is to measure the ability to deal with such language. Use the simplest and most straightforward language consistent with valid measurement.
- **Elitism.** Avoid the use of words or topics generally associated with wealthier social classes, such as *penthouse*, *polo*, *regatta*, *yacht*, and the like, unless required for valid measurement. Avoid unnecessary depictions of people spending large amounts of money unless required for valid measurement or related to the purpose of the test.
- **Specialized farm-related words:** Avoid requiring knowledge of uncommon farm-related words, unless required for valid measurement.
- **Specialized legal words:** Avoid requiring knowledge of uncommon legal terms when the goal is to measure some other skill or knowledge, unless the goal is to measure knowledge of legal terminology.
- **Military topics:** Avoid items with a primary focus on military topics, such as armed forces, battles, military strategy, wars, weapons, etc., unless required for valid measurement. Avoid unnecessary use of specialized words related to weapons, unless that is the intended point of measurement.
- **Specialized political words.** Avoid requiring knowledge of uncommon words related to politics unless required for valid measurement.
- **Regionalisms:** Unless required for valid measurement, avoid words, phrases, and concepts more likely to be known by people in some regions than in others.
- **Religious knowledge:** Do not require specific knowledge about any religion to answer an item unless such knowledge is part of the construct. If the knowledge is part of the construct, take care to use only the information about religion required for valid measurement.
- **Scientific and technical words:** Even though the specifications of many verbal skills tests call for items in the content category *science*, the goal is to measure verbal ability rather than knowledge of scientific subject matter. Avoid requiring knowledge of specialized words unless required for valid measurement.
- **Spatial skills:** Avoid the measurement of spatial skills (visualizing how objects or parts of objects relate to each other in space) unless the measurement of such skills is the purpose of the item.
- **Sports.** If knowledge of a sport is needed to answer an item, the item should not appear in a general skills test. Even if knowledge of the sport is not needed to answer an item, avoid unnecessary references to particular sports in the setting of an item.
- **Specialized words associated with tools and machinery:** Avoid requiring knowledge of specialized words associated with tools and machinery, unless required for valid measurement.

- Specialized words associated with transportation: Avoid requiring knowledge of specialized vehicles, such as ketch or biplane, or of parts of vehicles unless it is required for valid measurement.
- United States culture: Assessments are often taken by ELL students who may not be familiar with United States culture. Do not require specific knowledge of United States culture to obtain the answer to an item, unless the item is supposed to measure such knowledge.

♦ **Avoid material that is unnecessarily controversial, inflammatory, offensive, or upsetting.**

Problematic topics: Material that is unnecessarily controversial, inflammatory, offensive, or upsetting should be avoided whenever possible. Some reasonably controversial material may be necessary for valid measurement, however, even in skills tests. For example, if the ability to compare and contrast two points of view about a topic is required, the topic must be controversial enough to allow at least two defensible points of view. If such material must be used, it should be handled in a conscientious, balanced, sensitive, and objective manner. Preferably, the least controversial, inflammatory, offensive, or upsetting material that will meet the requirements should be used.

Note: Material that is neutral or pleasing to some groups may be offensive to other groups.

- Topics to avoid: Some topics are so sensitive that the best strategy is to avoid them if it is possible to do so - as, abortion, abuse of people/animals, euthanasia, experimentation on human beings/animals that is painful or harmful, genocide, human sexual behavior, including contraception, hunting or trapping for sport, rape, Satanism, torture, witchcraft. Avoid language or images that deal with topics as controversial, inflammatory, offensive, or upsetting as the following unless required for valid measurement or the purpose of the communication.
- Topics requiring extreme care: There are topics that need not be avoided entirely or must be handled in a very sensitive manner. Treat topics that are as controversial, inflammatory, offensive, or upsetting as the following in ways that limit their problematic aspects unless a different treatment is required for valid measurement or the purpose of the communication.
- Accidents, illnesses, or natural disasters: Avoid dwelling on gruesome, horrible, or shocking aspects of accidents, illnesses, or natural disasters unless required for valid measurement. Other aspects of those topics are acceptable - the prevention of accidents, the causes of illness, or the occurrences of natural disasters.
- Death and dying: Do not focus on gruesome details associated with death and dying, unless required for valid measurement. A statement that someone died in a particular year or that a disease was responsible for a certain number of deaths is acceptable.
- Evolution: It is undeniable that evolution is controversial, so it is preferable not to focus on the topic. However, evolution is a core concept in biological science, and topics associated with evolution are important in other disciplines. The most sensitive aspect of evolution appears to be the evolution of human beings.
- Religion: It is preferable not to focus on religion, unless required for valid measurement. If required, materials about religion should be as objective as possible. Do not support or oppose religion in general or any specific religion, and do not praise or ridicule the practices of any religion.
- Slavery: Except as required for valid measurement in assessments, slavery should not be the main focus of any material. Mention of the topic in skills tests is possible – as, a passage on the abolitionist movement would be acceptable if the passage touched on slavery. A discussion of the economic value of slaves would not be acceptable.
- Suicide or other self-destructive behavior: It is acceptable to mention that a person committed suicide, but it is not acceptable to focus unnecessarily on various means of suicide or to glorify suicidal behavior. Similarly, details of substance abuse should be avoided unless required for valid measurement.
- Violence and suffering: Do not focus on violent actions, on the detailed effects of violence, or on suffering unless required for valid measurement. Violence and suffering are too pervasive in art, biology, history, literature, and in most aspects of human and animal life to exclude them completely. But, do not dwell unnecessarily on the gruesome, shocking aspects of violence and suffering.
- Materials for international populations: Some images or descriptions of people and their interactions that are acceptable in the United States may be offensive to people in certain other countries - avoid culture-specific taboos. As a general rule, if some aspect of any content or image is believed to be offensive or confusing to some students, it should be avoided if the aspect is not required for valid measurement and the elimination of the aspect will not offend or confuse students. Avoid images or descriptions of people dressed in tight/revealing clothing; people who are posed immodestly; men and women touching each other; and men and women in intimate settings.

- **Advocacy:** Do not use assessments to advocate any particular cause, lifestyle, or ideology, unless required for valid measurement. Items should be neutral and balanced. Do not “take sides” on any controversial issue unless required for valid measurement. Sensitive items should be no more controversial than is required for valid measurement.
- **Requirements for research reports:** If it is necessary to focus on controversial and/or inflammatory issues, use language that is emotionally neutral to discuss the issues.

♦ **Use appropriate terminology to refer to people.**

- **Appropriate terminology:** Do not attach unnecessary labels to people. If group identification is necessary, it is generally appropriate to use the terminology that group members prefer. Explicitly derogatory names for groups should not be used. Reduce the use of *minority* and *majority* to refer to groups of people.
- **Appropriate terminology for people with disabilities.** If someone with a disability is referenced, focus on the person rather than the disability. Avoid terms that have negative connotations or reinforce negative judgments. Do not use the term *handicap* to refer to a disability. Avoid euphemistic or patronizing terms. Avoid implying that someone with a disability is sick. Refer to “Avoid/Use” Chart
- **Appropriate terminology for African American people.** The terms *Black* and *African American* are both acceptable. Note that *Black* should begin with an uppercase letter.
- The terms *Negro* and *Colored* are not acceptable except when embedded in clearly historical contexts as required for valid measurement, or in the names of organizations. (Because *Black* is used as a group identifier, avoid the use of *black* as a negative adjective, as in *black magic*, *black day*, or *black hearted*.)
- **Appropriate terminology for Asian American people.** The terms *Asian American*, *Pacific Island American*, and *Asian/Pacific Island American* should be used as appropriate. If possible, use specific terminology such as *Bangladeshi American*, *Chinese American*, *Filipino American*, *Japanese American*, and so forth. Do not use the word *Oriental* to describe people.
- **Appropriate terminology for Hispanic American people.** The terms *Hispanic American* and *Latino American* are both acceptable and may be used as appropriate. For women, use the term *Latina* in place of *Latino*. Though *Chicano* and *Chicana* as terms for Mexican Americans are accepted by some groups, they are rejected by others. It is therefore best to avoid using the words. Where possible, use a specific group name such as *Cuban American*, *Dominican American*, *Mexican American*, and so forth as appropriate.
- **Appropriate terminology for Native American people.** The terms *American Indian* and *Native American* are both acceptable. Avoid use of the term *Eskimo* for people who are more acceptably called *Alaskan Natives* or *Inuit*. Indigenous people in Canada are often referred to as *members of the First Nations*. Whenever possible, it is best to refer to a people by the more specific group names they use for themselves. However, this name may not be commonly known, and it may be necessary to clarify the term the first time it is used, as in the following example: “The *Diné* are still known to many other peoples as the *Navajo*.” Many Native Americans prefer the words *nation* or *people to tribe*.
- **Appropriate terminology for White American people.** The terms *White* and *Caucasian* are both acceptable, but *White* is becoming the preferred term. (Note that *White* should begin with an uppercase letter.) The term *European American* may also be used. The term is gaining some currency because of its parallelism with *African American*, *Asian American*, etc.
- **Appropriate terminology for women and men.** Women and men must be referred to in parallel terms. When women and men are mentioned together, both should be indicated by their full names, by first or last name only, or by title. Do not, for example, indicate men by title and women by first name. *Ladies* should be used for women only when men are being referred to as *gentlemen*. Similarly, when women and men are mentioned together, women should be called *wives*, *mothers*, *sisters*, or *daughters* only when men are referred to as *husbands*, *fathers*, *brothers*, or *sons*. People of one sex must not be described by physical attributes when people of the other sex are described by mental attributes or professional position. Irrelevant references to a person’s appearance or attractiveness are not acceptable.

Women should generally be referred to as *women*. If they are 18 or older, they should not be referred to as *girls*. Men should generally be referred to as *men*. If they are 18 or older, they should not be referred to as *boys*. Language that assumes that all members of a profession are members of one sex is unacceptable. Generic terms, such as *poet*, *doctor*, and *nurse*, include both men and women, and modified titles such as *poetess*, *woman doctor*, or *male nurse* are not acceptable. Role labels such as *scientist*, *executive*, and so forth include both men and women. Do not use expressions such as *the soldiers and their wives* that assume only men fill those roles unless such is the case in some particular instance. Do not couple generic role words

with gender-specific pronouns or actions unless a particular person is being referenced. Do not, for example, use terminology that assumes all kindergarten teachers or food shoppers are women, or that all college professors or car shoppers are men.

Using *he* or *man* to refer to all people is not acceptable. Avoid the use of generic *he* or *man* unless it is included in historical material. Alternating generic *he* and generic *she* is unacceptable because neither word should be used to refer to all people. Avoid *he/she* and *(s)he*.

- Appropriate terminology for people who are bisexual, gay, lesbian, or transgendered. Issues of human sexuality should be avoided unless required for valid measurement, as indicated in Guideline 3. Therefore, identify people by sexual orientation only when it is construct-relevant to do so. Do not use the labels gratuitously. The words *bisexual*, *gay*, *lesbian*, and *transgendered* are all acceptable. Because some people assume that *gay* refers to men only, use *gay* or *gay people* only when prior reference has specified the gender composition of this term. Avoid using the term *homosexual* outside of a scientific or historical context. Do not use the term *queer* to refer to sexual orientation, even though the term has some currency among gay and lesbian political activists and scholars.
- Appropriate terminology for older people. It is best to refer to older people by specific ages, for example, *people age 65 to 75*. It is also acceptable to refer to *older people*. Avoid *elderly* as a noun. Minimize the use of euphemisms such as *senior citizens* or *seniors*. Tests in certain content areas such as medicine may use terms such as *old-old* or *oldest-old* that are not appropriate in general usage.

♦ Avoid stereotypes

- ♦ A stereotype is a conventional, overgeneralized, and oversimplified conception of the characteristics of a group of people. Stereotypes attribute characteristics to groups on the basis of age, disability, ethnicity, gender, national origin, race, religion, or sexual orientation. Stereotypes ignore differences among members of the group. Generally, any statement or image that ascribes the same characteristic to all, or even most, members of a group (unless the group was composed on the basis of that characteristic) is a stereotype.
- Do not imply that all members of a group share the same culture, as opposed to recognizing substantial cultural variations within groups. Avoid stereotypes in language and images unless required for valid measurement, or necessary for research. If stereotypes are required, use the least offensive stereotypes that will result in valid measurement.
- The terms *stereotypical* and *traditional* overlap in meaning but are not synonymous. Depicting an individual engaged in a traditional activity (such as a woman cooking) does not necessarily constitute stereotyping. As long as the material as a whole does not depict members of a group engaged exclusively in traditional activities, such references are acceptable.
- Avoid using phrases that encapsulate stereotypes such as *Dutch uncle*, *Indian giver*, *run like a girl*, *women's work*, *man-sized job*, and so forth. In addition, no group defined by age, disability, ethnicity, gender, national origin, race, religion, or sexual orientation should be stereotyped as superior or inferior to any other such group with regard to any of the following, or similar, characteristics.

• Contribution to society	• Industriousness
• Dependence on welfare	• Leadership ability
• Dialect or language usage	• Morality
• Generosity	• Physical appearance
• Honesty	• Quality of culture
• Impulsiveness	
- Special requirements for tests. Unless the testing of stereotypes is specifically required for valid measurement, avoid stereotypes in tests even as sources of wrong answer choices. Test takers who select a wrong answer believe it is correct, so their belief in the legitimacy of a stereotype may be reinforced. Passages that explore the process of stereotyping, without including negative stereotypes, are acceptable in skills tests. Content tests may include stereotypes to the extent required for valid measurement.
- Special requirements for research. To avoid the reinforcement of stereotypes in discussions of group differences in research and statistical documents, do not imply that there is a biological or social cause not demonstrated by the research.

♦ Allow for Balance in Diversity

- Gender Balance. In skills tests, women and men should be reasonably equally represented. In addition to roughly balancing numbers of people of each gender, the status of the men and women shown should be reasonably equivalent.

A mention of Albert Einstein in one item is not balanced by a mention of a placeholder female name such as Jane in another item. The gender balance of content tests should be appropriate to the content area. In occupational tests, do not depart greatly from the gender distribution of the members of the occupation. For example, showing half of the nurses in a test as male or half of the automobile mechanics as female would be unrealistic and possibly disconcerting to test takers. At least some of the men and women shown should be equivalent in status. For example, do not show all the doctors as male and all the nurses as female, or vice versa.

- **Racial and Ethnic Balance.** The ideal racial and ethnic balance in a test would reflect the diversity of the test-taking population. It is not feasible, however, to show members of every group in the test-taking population in every test. As a reasonable compromise, in skills tests made for use primarily in the United States, strive to have about 20 percent of the items that mention people represent African American people, Asian American people, Latino American people, and/or Native American people. Try to depict several different groups in a test.
- If there is insufficient context in an item to indicate group membership in other ways, representation may be accomplished by using the names of reasonably well-known real people in various groups, or by using placeholder names commonly associated with various groups (e.g., Juan, Kimani, Latisha, Matsuko, Ram). Do not add unnecessarily to the linguistic loading of the item by using names that are inordinately difficult for test takers to decode.
- For skills tests with mixed domestic and international populations and for skills tests made for general international use, the representation of diversity among 20 percent of the items that mention people may include African people, Asian people (including Indian people), Latino people, and such indigenous peoples outside of the United States as the Guarani Indians of Paraguay.
- For skills tests with mixed domestic and international populations, at least one of the groups represented should be from the United States. For skills tests made for general international use, it is acceptable, but not required, to include any of the United States groups listed above. Tests made for specific countries should reflect the test-taking populations in those countries.
- In some content tests, such as history tests or literature tests, the proportion of items dealing with diverse groups may be fixed by the test specifications. If the proportions are not fixed by the test specifications, try to meet the representational goals given for skills tests to the extent allowed by the subject matter. If the names of people appearing in content tests are part of the subject matter (e.g., Avogadro's number, Heimlich maneuver, Jay's treaty), the items are not counted as including people for the purpose of calculating the number of items in which diversity should be represented.
- For all tests, if it is compatible with valid measurement, try to achieve at least a rough parity in status of the people depicted in different racial and ethnic groups. Do not, for example, depict all social workers as White and all clients as Black or Hispanic.

Comparison Examples – Interesting Options

AVOID	USE
men and ladies	men and women or ladies and gentlemen
man and wife	husband and wife
girl's and men's	women's and men's or girl's and boy's
Albert Einstein and Mrs. Mead	Dr. Einstein and Dr. Mead, Einstein and Mead, or Albert Einstein and Margaret Mead
chairman	chair, chairperson , leader, presiding officer, or moderator
<p>Note: <i>Chairman</i> and <i>chairwoman</i> are acceptable when referring to specific men and women. Do not use <i>chairman</i> for a man and <i>chairperson</i> for a woman because the terms are not parallel.</p>	
the consumer...she	consumers...they or consumer...he or she
the assistant...she	the assistants...they or the assistant...he or she
the breadwinner...his earnings	breadwinners...they or breadwinner...he or she
confined to a wheelchair or wheelchair bound	uses a wheelchair
Negative connotations afflicted, crippled, confined, inflicted, invalid, sickly, pitiful, victim, or unfortunate, handicap	
Euphemistic terms as special, physically challenged, inconvenienced	
Patronizing terms as inspirational, courageous, achieving success in spite of a disability, overcoming a disability	
patient	client
the blind	a blind person or a person who is blind
deaf and dumb or deaf mute	deaf (adjective) or hard of hearing
Down's syndrome or Mongoloid	Down syndrome
Learning disabled	Person with a learning disability.
Mentally ill	person with a psychological disability, person with a emotional disability
Mentally retarded, Retarded, Retard	a person with mental retardation, developmentally disabled , or developmentally delayed
Paraplegic or quadriplegic as nouns	Paraplegic or quadriplegic as adjectives
Spastic	Physical disability
manmade	synthetic, artificial
fireman, mailman, salesman, insurance man, foreman	firefighter, mail carrier, sales representative, Insurance agent, supervisor
Mankind or man	humanity , people, or human beings
sexual preference	sexual orientation
If a student studies, he will learn. If a student studies, she or he will learn. If a student studies, he or she will learn.	If students study, they will learn. Students who study will learn. A student who studies will learn.
manpower	workers, labor, personnel, or work force

APPENDIX E - STANDARD SETTING

...Establishing Cut Scores

For all Stakeholders...

Basic Standard Setting Information:

1. What is Standard Setting?
 - Performance standards describe how good is “good enough.”
 - Standard Setting is the process to convert these performance descriptions into operational definitions.
2. “How good?” is a rather political.
 - There is no “right answer;” it is judgment - not everyone will agree
 - One must balance the idealistic view of high school graduates versus the practical realities of failure
 - Supports the importance for an identified process with documentation
3. BOE issues to consider...
 - Assessments will change over time
 - There are multiple formats
 - Assessments are collected over multiple years
 - Aggregation issues can influence the decision about a student.
 - A system approach needs to be implemented to make the system effective.
4. Scoring is not Standard Setting
 - It is difficult to design a single task to perfectly represent the performance standard.
 - Does not resolve the question of “good enough” to graduate.
5. General Steps to follow...
 - Identify performance descriptors
 - Establish standard-setting approach(es) – can use a variety of methods
 - Select team and comfortable working conditions
 - Consider alignment, consistency, cognition levels, and fairness
6. Performance Descriptors
 - Begin with a general descriptor that becomes more specific descriptor after linking student work or test items to the descriptions.
 - “Clear and explicit” are critical to clarify “good enough”
7. Documentation and Process are key
 - Work may be challenged legally by those who either feel the assessment is too rigorous or too easy
 - Use a defensible process: who is involved and protocol is followed
 - Keep copies of how you communicated results
 - Document everything

Cut scores used for determining proficiency at...

The Content Level...

Proficiency Cut scores at the Content Level in all Core and Non-Core Subjects		
CUT SCORES AT THE CONTENT LEVEL		
	All Standards in all Core Subjects	All Standards in all Non-Core Subjects
PROFICIENCY	A SCORE OF 3 OR MORE ON 60% OF STANDARDS	A SCORE OF 3 OR MORE ON 60% OF STANDARDS
ADVANCED	A SCORE OF 4 OR MORE ON 60% OF STANDARDS	A SCORE OF 4 OR MORE ON 60% OF STANDARDS

The Standard Level...

Proficiency Cut scores at the Standard Level in all Core and Non-Core Subjects		
CUT SCORES AT THE STANDARD LEVEL		
	All Standards in all Core Subjects	All Standards in all Non-Core Subjects
PROFICIENCY	A SCORE OF 3 OR MORE ON 60% OF BENCHMARKS	A SCORE OF 3 OR MORE ON 60% OF BENCHMARKS
ADVANCED	A SCORE OF 4 OR MORE ON 60% OF BENCHMARKS	A SCORE OF 4 OR MORE ON 60% OF BENCHMARKS

The Benchmark Level...

The standard setting process determines scores that identify each performance level of proficiency on an assessment and/or collection of assessments. Washakie County School District #1 selected the **Bookmarking Method** to establish cut scores at benchmark, standards, and content levels and created a standard setting process that involved an item-based method for closed-ended items and student work-based method for open-ended items. Determining cut scores at these levels supports a true compensatory model, for students do not need to demonstrate proficiency with every assessment, benchmark, or standard to be proficient in a content area.

The rationale for selecting the Bookmarking Method was based on the versatility of the process, useable with both multiple-choice (closed-ended) and constructed response (open-ended) items. Additionally, bookmarking is useful with checklists, which are frequently used with authentic assessments and a two-point rubric.

Cut Scores at the Benchmark (Assessment) Level...

Directions to Establish Cut Scores for Closed-ended Assessment (Items) for Each Assessment...

1. Examine each assessment item, and align to the performance level for each benchmark. Use performance descriptors established for the assessment. You may choose to use the state performance standards for the descriptors to assure congruency.
2. Indicate the level of performance for the item.
3. Record the item number on the performance descriptor sheet by performance level, and set aside for team collaboration.
4. Using the bookmark method, arrange the items from the most challenging level to the least level of cognition. (Note: Placing the items on sheets of paper beginning at the least level of cognition to the most challenging level would be acceptable in lieu of "stacking" the items.)
5. Select two readers from the staff to examine the items.
6. Provide a performance rubric for each reader, and instruct the readers to (independently) examine the "stack" of items, and place a marker at the transition point between below basic and basic, basic and proficient, and proficient and advanced.
7. When all markers have been placed, compare your identified level marks with the marks indicated by the two readers.
8. Discuss the difference in the location of the three marks at each level. Develop a consensus regarding the cut points between the levels, regarding what is or is not acceptable.
9. Once the cut points are established by consensus, record the items in the performance descriptor as an "anchor" for each performance level.
10. Within the pilot year, collect trend data and record student data by item to note consistency (Cronbach's Alpha Coefficient) of scoring; revisit items when necessary.
11. Turn in a clean copy of the assessment and the performance descriptor for each assessment to the Curriculum Director.

Directions to Establish Cut Scores for Open-ended Assessments (Items) for Each Assessment...

1. Examine the assessment task, and create either a two- or four-point rubric that is aligned to the performance level for each benchmark. Use performance descriptors established for the assessment. You may choose to use the state performance standards for the descriptors to assure congruency.
2. Be explicit with the rubric descriptors, and if necessary itemize expectations within each rubric descriptor. Using the performance standards for your rubric descriptors is very acceptable, but make certain that if clarification is necessary, you include the clarification in the performance level rubric.
3. Select a sampling of student work and using the Bookmarking Method, arrange the sampling of student work from the highest scoring down the continuum to the lowest performing level.
4. Select two readers from the staff to examine the assessment task.
5. Provide the performance rubric for each reader, and instruct the readers to (independently) examine the "stack" of items, and place a marker at the transition point between below basic and basic, basic and proficient, and proficient and advanced. If anchor papers have been established through the consistency process, those papers can be used as indicators for marker placement.
6. When all markers have been placed, compare your identified level marks with the marks indicated by the two readers.
7. Discuss the difference in the location of the three marks at each level. Develop a consensus regarding the cut points between the levels, regarding what is or is not acceptable.
8. Once the cut points are established by consensus, select "anchor" papers for each performance level, and possibly include anchor papers that transition from one performance level to another. If anchor papers have been established through the consistency process, review anchor papers for consistency in scoring according to cut points.
9. Within the pilot year, collect trend data and record student data by performance level to note consistency of scoring; revisit when necessary.
10. Turn in a clean copy of the assessment, copies of anchor papers, and the performance descriptor (rubric) for each assessment to the Curriculum Director.

Procedures for Establishing Performance Descriptors

(Reference District Standards)

- 1 - Examine performance standards descriptors, noting the four performance levels.
- 2 - Develop a brief description (expectations) for each proficiency level, congruent with the district performance standards to include Below Basic, Basic, Proficiency, and Advanced for:
 - Closed-ended items in an assessment
 - Open-ended items in an assessment
 - Closed-ended and open-ended items in an assessment (combination)
- 3 - Write performance descriptors in the space below. (Document will scroll as you write)
 - If necessary (but not required), separate descriptors may be necessary for specific items.
 - If writing performance levels for both closed-ended and open-ended items in one assessment, include expected levels for both types of Items at each level.

NOTE: A performance descriptor rubric is necessary for both closed-ended and open-ended assessments.

If you have a rubric to score student work...

- The rubric should be congruent with the state/district performance standards
- If congruent, simply attach a copy of the rubric to the worksheet – no need to copy over.
- If not congruent, refer to the district performance standards and adjust appropriately.

If the assessment overlaps standards and/or benchmarks, make certain to include additional proficiency descriptors or develop an additional rubric.

Content: _____

Performance Descriptor Rubric

Course: _____

Assessment: _____

Level	Description
Advanced (Benchmarks)	
Proficient (Benchmarks)	
Basic (Benchmarks)	
Below Basic (Benchmarks)	

APPENDIX F – COMPARABILITY

...5-Year Consistency Cycle – An On-going Process

Comparability: The district assessment system requirements are comparable for all students within each district within a given class (or grade level) and across years. Common assessments are used throughout the district at all grade levels in all content areas to ensure comparability. If new items/tasks are used in the BOE, assessment procedures are implemented to maintain year-to-year comparability. Comparability is monitored continuously by staff through self-evaluation, instructional facilitators, and administrators.

20% of district benchmarks assessments each year resulting in a 5-year cycle for consistency. Rationale for this procedure allows for staff to...

- 1) Revisit assessments for revisions;
- 2) Adjust to adoptions/changes in standards and benchmarks;
- 3) Train new staff on procedures within the assessment system;
- 4) Adjust to textbook/program adoption; and
- 5) Afford staff to create more appropriate assessments.

Content areas, by grade level or content, determine assessments for review during each year. All completed data information is submitted to the Curriculum Director. Additional consideration is given to programs requiring changes in supplemental curricular needs, students at risk, RTI, differentiation of instruction, and students with special needs. All students receive equal educational opportunities, to include ELL, Special Needs, 504, free/reduced lunch, and Title I students.

The district continues to maintain a K-12 Body of Evidence assessment system by addressing the principles of BOE design, which includes identified district procedures for alignment, consistency, fairness, comparability, and standard setting. If there is more than a 25% change in a BOE assessment, the assessment must be reviewed, following the district procedures each of the assessment design criteria.

Each year specific content curricula (core and non-core) and instructional strategies are targeted for examination and review; new textbooks, curriculum/instruction, and technology resources are considered for adoption. Content areas rotate through the process within a five- to seven-year rotation, followed by one year of assessment pilot. Teachers address the “review” content area for examination and are afforded the latitude to extend assessment examination in all other content areas, especially for assessments that are need of change, replacement, and/or adjusted. Twenty percent (20%) of all assessments from content areas that are not considered for review in a specific year are revisited.

Note: When new assessments are developed, alignment, fairness, and preliminary standard setting will accompany assessment development, and fidelity to the process ensures consistency, comparability, and final standard setting during the following year.

Rotation Schedule for Assessment Review

Instruction, Textbook, Assessment, and Technology Review Year

Content	Instruction, Textbook, Assessment, and Technology Review Year										
	Social Studies	Science	Physical Education	Math	Language Arts	Health	Foreign Language	Fine and Performing Arts	Career Vocational		
2009-2010	K-5	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%
	6-8	20%	20%	20%	20%	T	20%	20%	20%	20%	20%
2010-2011	9-12	20%	20%	20%	T	T	20%	20%	20%	20%	20%
	K-5	T	20%	20%	20%	T	20%	20%	20%	20%	20%
2011-2012	6-8	20%	20%	20%	20%	P	20%	20%	20%	20%	20%
	9-12	20%	20%	20%	P	P	20%	20%	20%	20%	20%
2012-2013	K-5	P	20%	T	20%	P	20%	20%	20%	20%	T
	6-8	20%	T	T	20%	20%	20%	20%	20%	20%	T
2013-2014	9-12	20%	T	T	20%	20%	20%	20%	20%	20%	T
	K-5	20%	T	P	20%	20%	20%	20%	20%	20%	P
2014-2015	6-8	20%	P	P	20%	20%	20%	20%	20%	20%	P
	9-12	20%	P	P	20%	20%	20%	20%	20%	20%	P
2015-2016	K-5	20%	20%	20%	T	20%	20%	20%	20%	20%	T
	6-8	T	20%	20%	P	20%	20%	20%	20%	20%	T
	9-12	T	20%	20%	20%	20%	20%	20%	20%	20%	T

- T = Textbook/Curriculum Adoption
- 20% = Assessment Review (Annual) – Reliability Analysis
- P = New Assessment Pilot (Textbook/Curriculum Adoption Following Year) – Reliability Analysis