



Oregon Department of Education

2012–2013

Technical Report

Oregon's Statewide Assessment System

Test Development

Volume 2

Last updated on August 1, 2012



Oregon's Statewide Assessment System
Technical Report: Volume 2, Test Development
Last updated on August 1, 2012

It is the policy of the State Board of Education and a priority of the Oregon Department of Education that there will be no discrimination or harassment on the grounds of race, color, religion, sex, sexual orientation, national origin, marital status, age or disability in any educational programs, activities, or employment. Persons having questions about equal opportunity and nondiscrimination should contact the state superintendent of public instruction at the Oregon Department of Education.

Oregon Department of Education
Office of Assessment and Information Services
255 Capitol Street NE
Salem, OR 97310
503-378-3600
<http://www.ode.state.or.us/>

Rob Saxton
Deputy State Superintendent of Public
Instruction

Doug Kosty
Assistant Superintendent

Rachel Aazzerah
Science and Social Sciences Assessment
Specialist

Saleem Ahmad
Research Analyst

Derek Brown
Manager, Essential Skills

Kathy Busby
Project Manager

Holly Carter
Assessment Operations and Policy Analyst

Ken Hermens
Language Arts Assessment Specialist

Beth LaDuca
NAEP State Coordinator

James Leigh
Mathematics Assessment Specialist

Brad Lenhardt
Extended Assessment Specialist

Michelle McCoy
ELPA and Assessment Implementation
Specialist

Laura Petschauer,
Accommodations Specialist

Steve Slater
Manager, Psychometrics & Validity

Sheila Somerville
Electronic Publishing Specialist

Tom Tinkler
Psychometric Specialist

Kathleen Vanderwall
Manager, Test Design and Implementation

Cari White
Lead, Assessment Production

This technical report is one of a series that describes the development of Oregon's Statewide Assessment System. The complete set of volumes provides comprehensive documentation of the development, procedures, technical adequacy, and results of the system:

Volume 1: Annual Technical Report

Volume 2: Test Development

Volume 3: Standard Setting

Volume 4: Reliability and Validity

Volume 5: Test Administration

Volume 6: Score Interpretation Guide

Volume 7: Alternate Assessment, Program Description and Statistical Summary

Volume 8: English Language Proficiency Assessment Test Development

Volume 9: English Language Proficiency Assessment Validity

All volumes can be found at <http://www.ode.state.or.us/search/page/?id=787>.

CONTENTS

1. OVERVIEW	1
2. GENERAL TEST DESIGN.....	2
2.1 ITEM TYPES	3
2.1.1 OAKS Assessments: Multiple-Choice Items	3
2.1.2 Writing Performance Assessment: Essay (Extended-Response) Items	3
2.1.3 Machine-Scored Graphic Response Items.....	Error! Bookmark not defined.
2.1.4 Nationally and Internationally Benchmarked Items.....	Error! Bookmark not defined.
3. ALIGNMENT TO ACADEMIC CONTENT STANDARDS	4
4. TEST SPECIFICATIONS	5
5. SCORE REPORTING CATEGORIES.....	6
5.1 MATH.....	6
5.2 READING/LITERATURE	7
5.3 SCIENCE	9
5.4 SOCIAL SCIENCES	10
5.5 WRITING.....	11
6. ITEM DEVELOPMENT	13
6.1 ITEM WRITING.....	13
6.2 COMMITTEE/PANEL REVIEW	17
6.3 EXPERT REVIEW	20
6.4 FIELD TESTING	21
6.5 TRANSLATION OF ITEMS TO SPANISH.....	21
6.6 ADDITIONAL EXPERT REVIEW OF ITEMS	22
7. ITEM USE AND RELEASE.....	23
8. TEST DEVELOPMENT.....	23
8.1 USING FIELD-TEST AND OPERATIONAL TEST DATA TO CHOOSE ITEMS	23
8.2 OAKS ONLINE ITEM POOLS	24
8.3 OAKS ONLINE ADAPTIVE ITEM SELECTION ALGORITHM	24
8.4 OAKS ONLINE TEST STOPPING RULES AND TARGET SCORE PRECISION	26
8.5 OAKS BRAILLETEST FORM DEVELOPMENT	ERROR! BOOKMARK NOT DEFINED.
8.6 COMPARABILITY OF OAKS ONLINE AND BRAILLE TESTS.....	ERROR! BOOKMARK NOT DEFINED.
9. EMBEDDED FIELD TESTS.....	26

10. ALIGNMENT OF STANDARDS AND ASSESSMENTS26
APPENDIX27

LIST OF TABLES

TABLE 1. NUMBER OF OPERATIONAL ITEMS AND GRADE LEVELS FOR ADMINISTRATION IN EACH SUBJECT4
TABLE 2. PERCENTAGE OF TEST ITEMS ASSESSING EACH SRC FOR MATH**ERROR! BOOKMARK NOT DEFINED.**
TABLE 3. PERCENTAGE OF TEST ITEMS ASSESSING EACH SRC FOR READING/LITERATURE9
TABLE 4. PERCENTAGE OF TEST ITEMS ASSESSING EACH SRC FOR SCIENCE.....10
TABLE 5. PERCENTAGE OF TEST ITEMS ASSESSING EACH SRC FOR SOCIAL SCIENCES11
TABLE 6. STRUCTURE OF ODE CONTENT AND ASSESSMENT PANELS/COMMITTEES17
TABLE 7. OAKS ONLINE TESTING ENGINE RULES25
TABLE 8. SUMMARY STATISTICS FOR PAPER/PENCIL FORMS, 2009-10**ERROR! BOOKMARK NOT DEFINED.**

LIST OF FIGURES

FIGURE 1. SAMPLE OREGON ITEM WRITING FORM.....16
FIGURE 2. SAMPLE OREGON CONTENT AND ASSESSMENT PANEL ITEM REVIEW FORM19

1. OVERVIEW

A series of technical reports was commissioned in 2006 to provide information about the technical and procedural characteristics of Oregon’s Statewide Assessment System (OSAS). The OSAS was created by the Office of Assessment and Information Services in the Oregon Department of Education (ODE), with considerable participation and involvement from Oregon educators.

The technical reports are intended to summarize and inform audiences by compiling existing documentation from a variety of sources into a single, easily accessible document. Consisting of nine volumes, the reports describe the development, operational procedures, and technical features of the assessment system.

- Volume 1: Annual Technical Report
- Volume 2: Test Development
- Volume 3: Standard Setting
- Volume 4: Reliability and Validity
- Volume 5: Test Administration
- Volume 6: Score Interpretation Guide
- Volume 7: Alternate Assessment, Program Description and Statistical Summary
- Volume 8: English Language Proficiency Assessment Test Development
- Volume 9: English Language Proficiency Assessment Validity

The annual report (Volume 1) describes student performance and documents changes to the system and assessment-related activities undertaken during the previous year. ODE updates volumes 2 through 6 as new information becomes available or as new procedures are implemented. Volume 7, which is updated annually, describes the results of the alternate assessment administered to students with disabilities. Volumes 8 and 9, also updated annually, describe the results of the English Language Proficiency Assessment.

Together, the reports describe Oregon public school students’ progress toward meeting the academic achievement standards and the process and technical adequacy through which this progress is measured.

Volume 2: Test Development describes the development and maintenance of the Oregon Assessment of Knowledge and Skills and the Writing Performance Assessment and is updated as test development procedures change.

2. GENERAL TEST DESIGN

Oregon’s Statewide Assessment System (OSAS) consists of multiple choice Oregon Assessments of Knowledge and Skills (OAKS) in Mathematics and Reading/Literature aligned to grade-level content standards, Science and Social Sciences aligned to benchmark level content standards, a Writing Performance Assessment, and the English Language Proficiency Assessment (ELPA)¹. The OAKS are used for NCLB accountability and for accountability on the State Report Card. In addition, OAKS Reading/Literature and OAKS Mathematics are assessment options available for high school students to demonstrate proficiency in the Essential Skills of Reading and Mathematics under Oregon’s high school diploma requirements. Students first enrolled in grade 9 in 2008-2009 are the first group of students responsible for proficiency in Reading for the high school diploma. Students first enrolled in grade 9 in 2010-2011 are the first group of students responsible for proficiency in Math for the high school diploma.

The Writing Performance Assessment is used for school and district accountability on the State Report Card. In addition, starting with students first enrolled in grade 9 in 2009-2010, high school students may use the Writing Performance Assessment as one assessment option to demonstrate proficiency in the Essential Skill of Writing under Oregon’s diploma requirements.

The OAKS are summative assessments, which are generally designed to assess student learning at the end of an instructional period. Summative assessments are typically used for program accountability and to assign achievement level scores to students. Summative assessments are not designed as diagnostic tools for student placement or as formative assessments. Given the specific focus and purpose of summative assessments, the OAKS can only be used as part of a collection of evidence regarding the academic needs of individual students. The primary purpose of the OAKS is to ascertain the achievement level of individual students and compare that achievement with the Achievement Standards established by the State Board of Education. Local schools and districts can compare student performance by grade, school, and district to results throughout the state.

OAKS is administered using an online format (OAKS Online) and is an adaptive assessment, which means that the items presented to the student vary in difficulty based on the student’s performance on the previous item. Therefore, the state creates a grade-level item pool rather than a single pre-made test for each grade level. The computer selects questions based on the answer a student gives to a test item, which in turn determines the difficulty of the next item that the computer will select. Because the computer “pushes” students to find out their highest ability, OAKS Online tests will appear to be moderately difficult to virtually every student. All items are presented at the student’s appropriate grade level via OAKS Online and are available in both English and English/Spanish format for mathematics, science, and social sciences. ODE also offers a Grade 3 Spanish reading/literature assessment which is available as a native-language assessment.

For students with visual impairments who benefit from access to Braille, ODE provides access to the adaptive online format of OAKS Mathematics, Reading/Literature, Science, and Social Sciences assessments through a Braille interface of OAKS Online

¹ Volume 8: English Language Proficiency Assessment Test Development provides information on the test development of the ELPA.

While the Writing Performance Assessment is primarily administered using a paper-based format, students have the option to test online instead. ODE provides the Writing Performance Assessment in both English and Spanish, as well as in both Braille and Large Print formats. Finally, ODE offers an alternative assessment (OAKS Extended) for all required content areas (the optional Social Sciences assessment is only available online). Additional information on specific assessment options and testing formats is included in Volume 5: Test Administration (available at <http://www.ode.state.or.us/search/page/?id=787>) and in Part VI – Students and Assessment Options of the Test Administration Manual (available at <http://www.ode.state.or.us/go/tam>). Additional information on ELPA test administration is included in Volumes 8 and 9: English Language Proficiency Assessment (available at <http://www.ode.state.or.us/search/page/?id=787>) and in Appendix H of the Test Administration Manual.

All tests are developed to be representative and valid measures of the grade-specific content knowledge required by Oregon’s academic content standards. The tests are designed to serve the following goals:

- Provide instructionally useful evaluation of individual student progress toward mastery of the academic content standards
- Guide instructional program improvement
- Ensure that the state is progressing toward the state and federal goals for high standards for all
- Determine proficiency in the Essential Skills for the purpose of awarding students a high school diploma
- Inform the public of school success

To facilitate accessibility for all students, the tests are designed according to the principles of universal design, including the use of plain language.

2.1 Item Types

2.1.1 Multiple-Choice Items

The OAKS operational tests consist primarily of multiple choice items. Each item generally measures a single score reporting category (SRC) or strand, and the test pool provides items from a range of difficulty levels within each SRC. Items typically have four response options, although some items have either 3 or 5 response options. OAKS multiple choice items do not use “none of the above” or “all of the above” as response options. Reading/literature, Science, and Social Sciences items may be grouped together into modules or related questions that are linked to a single passage or stimulus (e.g., data table, diagram). Tests generally contain 40 to 50 operational items.

2.1.2 Machine-Scored Graphic Response Items

In addition to multiple choice items, the OAKS operational tests for Mathematics and Science contain machine-scored graphic response items. ODE first field tested this new item type for both Mathematics and Science in the spring of 2010. ODE will continue field testing new machine-scored graphic response items for Mathematics and Science. Machine-scored graphic response items ask students to plot their answers on a grid and allow students to demonstrate their knowledge and skills

in a more complex fashion than permitted by multiple choice items. These items are scored using a rubric and may be weighted to be worth multiple points.

2.1.3 Writing Performance Assessment: Essay (Extended-Response) Items

The Writing Performance Assessment consists of multiple prompts from which the student selects a single prompt to compose a response. Both the paper-based and online format of the test consist of four prompts. Each prompt assesses one of four modes of writing: narrative, imaginative, expository, or persuasive. The prompts are designed to address experiences and interests common to the student's age level, allowing students to write from experience. Note: due to budgetary restrictions, the Writing Performance Assessment will only be offered to students in Grade 11 in 2012-13.

Table 1.
Number of Operational Items and Grade Levels for Administration in Each Subject

Subjects	Grade levels	Length (number of Items)
OAKS		
Reading/Literature	3–8, 11	40-50
Mathematics	3–8, 11	40
Science	5, 8, 11	45
Social sciences	5, 8, 11	50
Writing Performance Assessment (all extended response)		
Writing	11	1 essay

3. ALIGNMENT TO ACADEMIC CONTENT STANDARDS

Academic content standards identify what students are expected to know and be able to do in the content areas. All Knowledge and Skills tests are written to, or aligned with, Oregon's academic content standards. Under Oregon law, ODE develops the academic content standards using a process of broad-based consensus on the important academic goals for K-12 education in the state. This process relies on extensive input from Oregon educators and other members of the educational community, including members of professional organizations, district administrators, post-secondary institutions, and other community members. The current content standards for Reading/Literature and Social Sciences were originally approved by the State Board of Education in 1996 after an extensive public review process and again in 2000 – 2004 after they were revised. The State Board of Education approved new Mathematics content standards for grades 3 – 8 in 2007 and for high school in 2009; OAKS Mathematics has assessed students under the new mathematics content standards since 2010-2011. The State Board of Education approved new science content standards for grades 3 – 8 and high school in 2009; OAKS Science has assessed students under the new science content standards since 2011-2012.

Items on the OAKS and the Writing Performance Assessment are written to represent Oregon’s content standards, and tests are composed of items such that the emphasis of the tests matches the emphasis of the content standards. Achievement standards define what students must do to meet or exceed Oregon’s academic content standards. Volume 3: Standard Setting describes the process that ODE followed in setting the new achievement standards that went into effect for Mathematics in 2010-2011 and for Reading and Science in 2011-2012. The academic content and achievement standards and their year of adoption by the State Board of Education are available on ODE’s Web site, <http://www.ode.state.or.us/teachlearn/real/>:

Math: [newspaper/Newspaper_Section.aspx?subjectcd=MA](http://www.ode.state.or.us/teachlearn/real/newspaper/Newspaper_Section.aspx?subjectcd=MA)

English/language arts: [newspaper/Newspaper_Section.aspx?subjectcd=ELA](http://www.ode.state.or.us/teachlearn/real/newspaper/Newspaper_Section.aspx?subjectcd=ELA)

Science: [newspaper/Newspaper_Section.aspx?subjectcd=SC](http://www.ode.state.or.us/teachlearn/real/newspaper/Newspaper_Section.aspx?subjectcd=SC)

Social Sciences: [newspaper/Newspaper_Section.aspx?subjectcd=SS](http://www.ode.state.or.us/teachlearn/real/newspaper/Newspaper_Section.aspx?subjectcd=SS)

The content and achievement standards are also described in greater detail in Volume 3: Standard Setting. Existing evidence of the alignment between the assessments and the content standards is described in Volume 4: Reliability and Validity. Both of these volumes are available online at <http://www.ode.state.or.us/search/page/?=1305>. ODE contracted with West Ed to collect independent evidence of alignment between the test items and the content standards for Reading/Literature and Science, as well as a content standard review for Social Sciences. ODE contracted with Educational Policy Improvement Center (EPIC) to collect independent evidence of alignment between OAKS Mathematics test items and the new mathematics content standards. EPIC also did a follow-up alignment review of Reading/Literature in 2010 and of Science in 2011 and verified that items written subsequent to the West Ed study continued the strong alignment found in the initial study.

4. TEST SPECIFICATIONS

ODE’s test specifications reflect skill expectations that are outlined in Oregon’s content standards. These specifications, individually created for each content area and grade level, establish guidelines for selecting test content and writing test items. The specifications lead to a “test blueprint” that lays out for the test item writers the item format and number of questions to be written and tested in each Score Reporting Category (SRC) for each grade and content area. They include the percent of items measuring each SRC, the item format (e.g., MC), psychometric properties used in item selection and forms development, and the arrangement or selection of items on operational tests. The specifications determine both the composition of the item pool as well as the rules for item selection. Development of both the general education (including English/Spanish side-by-side) and Braille formats of OAKS Online are governed by the same test specifications.

Test specifications explain the overall design of a test and describe the specific content that appears on a test, conveying to teachers what their students can expect on state assessments and what they are responsible for assessing through classroom assignments and local performance assessments. By serving as the foundation for test development, the specifications clearly define the domains for which score inferences are desired.

For each subject, test specifications reflect the skills outlined in the content standards. In turn, the content standards are benchmarked to well-defined, rigorous, nationally accepted standards. The content of the Reading test specifications reflects the skill expectations outlined in the content standards developed, in part, to represent the content identified by the National Standards for the Language Arts² and A Compendium of Standards and Benchmarks for K–12 Education³ and to correlate to the skills assessed on the reading portion of the National Assessment of Educational Progress (NAEP).⁴ Oregon’s Mathematics Content Standards for Grades K-8 and High School Mathematics Content Standards were written to reflect the approach in the National Council of Teachers of Mathematics (NCTM) publication “Curriculum Focal Points for Prekindergarten through Grade 8 Mathematics: A Quest for Coherence.”⁵ They are organized around three core standards for grades K-8 and eight core standards for high school. Science standards were developed, in part, to correlate with the Knowledge and Skills assessed on science standards on the National Assessment of Educational Progress and to align with the National Science Standards⁶ and the Benchmarks for Science Literacy.⁷ The social sciences standards were developed using “match/gap analysis” with NAEP Frameworks in U.S. History, Geography, and Civics, and the national standards in Economics and World History.

Since 2006, Oregon has identified SRCs/content standards for which more items of certain difficulty levels and cognitive complexity were written. These items introduce flexibility in test specifications by specifying a small range rather than a set percent of items measuring each SRC (increasing the likelihood that tests will not deviate from test specifications), and recognizing the balance required between test information, or statistical precision, and content representation in item selection.

Test specifications for the Oregon Statewide Assessment are available at <http://www.ode.state.or.us/search/page/?=496>.

5. SCORE REPORTING CATEGORIES

Below are descriptions of the Score Reporting Categories (SRCs) for each content area and tables showing the percentage of questions on a given grade-level test that assess each SRC in that content area.

5.1 Mathematics

The State of Oregon Mathematics Content Standards for Kindergarten through Grade 8, 2007, were written to align closely to the *NCTM Focal Points*; and the *Revised High School Mathematics Content Standards, 2009*, were written to build on that structure.⁸ The math tests are designed to assess the skills identified by

² National Council of Teachers in English, www.ncte.org

³ McREL and ASCD, www.mcrel.org/standards-benchmarks

⁴ <http://nces.ed.gov/nationsreportcard/>

⁵ <http://www.nctm.org/standards/content.aspx?id=270>

⁶ National Research Council, www.nas.edu

⁷ American Association for the Advancement of Science, www.project2016.org

⁸ For more information, visit <http://www.nctm.org>.

these standards overall and at the three Core Standards for each grade 3 – 8 and three topic areas for high school: Algebra, Geometry, and Statistics.

Table 2 shows the Core Standards for each of the grade levels and the percentage of test questions on the test at each grade level that assess that category. For more detailed descriptions by core standard, please see the Mathematics test specifications and blueprints posted online at <http://www.ode.state.or.us/search/page/?id=496>.

Table 2.
Percentage of Test Items Assessing Each SRC for Math

Math	Score Reporting Category (Core Standard)		
	First Core Standard	Second Core Standard	Third Core Standard
Grade 3	35	35	30
Grade 4	35	35	30
Grade 5	35	35	30
Grade 6	35	35	30
Grade 7	35	35	30
Grade 8	40	30	30
High School	50	30	20

5.2 Reading/Literature

The reading/literature specifications reflect the skill expectations that were outlined in the content standards adopted by the State Board of Education for implementation during the 2005–2006 school year. These standards were developed, in part, to correlate with the skills assessed on the reading portion of the National Assessment of Educational Progress. As a result, Oregon’s SRCs for reading use similar terminology in their definitions as those used for the 2005 NAEP Reading Framework. On the basis of the standards, the Reading/Literature tests are designed to assess literacy skills overall in the following six SRCs:

- **Vocabulary.** In this skill area, students use appropriate strategies to determine the meaning of unknown words. For the items on the state assessment, students are asked to focus primarily on context clues. Passages providing context clues include well-known, high frequency words that explain the meaning of the target word. The clues may be stated directly in a phrase or in sentences before or after use of the target word, or they may be found through careful reading of the entire text. At some grade levels, students may also be asked to use context clues to determine the meanings of words with multiple meanings or of phrases, such as idioms and figurative expressions.
- **Read to perform a task.** When reading to perform a task, students use skimming and scanning techniques to search for information in what is termed “practical” text. Depending on the grade level, practical text may include charts, schedules, directions, recipes, forms, maps, graphs, or job and consumer-related materials. The reader’s purpose is to look for information in order to do something. At grade 8 and at the high school level, questions ask

students to synthesize information and reach logical conclusions, not simply to understand the selection's content.

- **Demonstrate general understanding.** Students show a general understanding by accurately responding to questions about material that is explicitly stated in the text. After reading informational text, students might be asked to identify an article's topic statement, recall the correct sequence of events, or identify important details that were stated in the reading passage. Similarly, after reading literary text, students might be asked questions about the sequence of events in the plot or asked to identify details or events that were critical to the development of the plot.
- **Develop an interpretation.** To develop an interpretation, students must look beyond what is explicitly stated in a selection and show a more complete understanding of what was read. For informational text, questions include drawing inferences about the author's meaning, making predictions about forthcoming information in the text or events that are likely to occur in the future, and drawing conclusions about reasons for actions when those reasons are not explicitly stated. For literary text, students make predictions about events likely to happen later in the story, interpret the story to uncover its themes, and draw conclusions about traits present in the character and motivations for his or her actions.
- **Examine content and structure: informational text.** Examining content and structure requires students to critically analyze and evaluate text. Students stand apart from the text, consider it objectively, and evaluate its quality and effectiveness. For informational text, questions ask students to consider the author's purpose and style. Depending on the grade level, students may be asked about instances in which the author has relied on facts or opinion; which arguments or statements have support; whether the passage has evidence of bias; and what structural elements are present in the work. At the upper grades, students may be asked to compare information and make connections across parts of a text or between texts. This reporting category is not assessed at grade 3.
- **Examine content and structure: literary text.** Examining content and structure requires students to critically analyze and evaluate text. Students stand apart from the text, consider it objectively, and evaluate its quality and effectiveness. For literary text, students evaluate the use of literary elements and devices and the impact and purpose of their use within a selection. Questions may ask students to examine selections to determine their mood or tone and to determine how authors achieved that mood or tone. Students may be asked literary genre questions at specific grades (poetry at grade 6 and drama at the high school level, for example). At the upper grades, students may be asked to compare the treatment of themes and make connections between two literary selections. This reporting category is not assessed at grades 3 and 4.

Table 3 shows the Score Reporting Categories for each of the grade levels and the percentage of test questions on the test at each grade level that assess that category.

Table 3.
Percentage of Test Items Assessing Each SRC for Reading/Literature

Reading/ Literature	Score Reporting Category (strand)					
	Vocabulary	Read to perform a task	Demonstrate general understanding	Develop an interpretation	Examine content and structure: informational text	Examine content and structure: literary text
	SRC 1	SRC 2	SRC 3	SRC 4	SRC 5	SRC 6
Grade 3	28	16	28	28	N/A	N/A
Grade 4	25	13	25	25	12	N/A
Grade 5	21	13	21	21	12	12
Grade 6	20	12	20	20	14	14
Grade 7	20	12	20	20	14	14
Grade 8	20	12	18	20	15	15
High School	20	12	16	20	16	16

5.3 Science

The science specifications reflect the skill expectations outlined in the science content standards, which were adopted February 26, 2009, by the State Board of Education. These standards were developed, in part, to correlate with the knowledge and skills assessed on science standards on the National Assessment of Educational Progress and align with the National Science Standards. Oregon’s science tests are designed to assess literacy skills overall and in the following SRCs:

- **Structure and Function (SRC 1):** Understand living and non-living things have characteristics, form and function, and are composed of components that function together to form systems.
- **Interaction and Change (SRC 2):** Understand components in a system can interact in dynamic ways, within or without that system, and may result in change.
- **Physical Science (SRC 5):** Understand structures and properties of matter, forms of energy, and changes that occur in the physical world.
- **Life Science (SRC 6):** Understand structures, functions, and interactions of living organisms and the environment.
- **Earth and Space Science (SRC 7):** Understand physical properties of the Earth and how those properties change. Understand Earth’s relationship to other objects in the Universe.
- **Scientific Inquiry and Engineering Design (SRC 8):** Understand science process concepts and skills that characterize the nature and practice of science. Scientific Inquiry (SRC 3) is a systematic process that includes proposing testable hypotheses, collecting, analyzing, and interpreting data to produce evidence-based explanations and new explorations. Engineering Design (SRC 4) is a process of formulating problem statements, identifying criteria and constraints, testing solutions, and incorporating modifications based on test data and communicating the recommendations.

Student information from 2011-12 OAKS Online Science will be reported through six Score Reporting Categories (SRC) including the four Science Core Standards and three Science subjects as sub categories. Scientific Inquiry (SRC 3) and Engineering Design (SRC 4) scores will be combined and reported together as SRC 8 (Science Processes).

Table 4 shows the score reporting categories (SRC) for each grade tested and the percentage of questions on a given test administration that would assess that category. Test items aligned to SRC 1 and SRC 2 will also be reported to a subject SRC of Physical Science, Life Science, or Earth and Space Science. But, each test item in SRC 1 or SRC 2 will only be counted once toward a student's summary science score. In 2011-12, test items aligned to SRC 3 and SRC 4 will be reported only once as SRC 8.

Table 4.
Percentage of Test Items Assessing Each SRC for Science

Science	Core Score Reporting Category				Subject Score Reporting Category			
	SRC 1	SRC 2	SRC 3*	SRC 4*	SRC 5	SRC 6	SRC 7	SRC 8*
	Structure and Function	Interaction and Change	Scientific Inquiry	Engineering Design	Physical Science	Life Science	Earth and Space Science	Science Processes
Grade 5	25	50	13	12	25	25	25	25
Grade 8	30	45	13	12	25	25	25	25
High School	30	45	13	12	24	27	24	25

* Scores from SRC 3 and SRC 4 are combined and reported as SRC 8 for 2011-12

5.4 Social Sciences

Social Science tests are not required by the state but are available for districts, schools, or individual students for the purpose of program or individual student evaluation. They measure knowledge across 6 SRCs:

- **History**
 - **Historical skills.** Design and implement strategies to analyze issues, explain perspectives, and resolve issues using the social sciences.
 - **U.S. History.** Understand the importance and lasting influence of issues, events, people, and developments in U. S. history. Relate significant events and eras in United States history to past and present issues and developments.
 - **World history.** Understand the importance and lasting influence of significant eras, cultures, issues, events, and developments in world history. Relate significant events

and eras in world history to past and present issues and developments.(Not assessed at grade 5)

- **Civics and government.** Understand and apply knowledge about governmental and political systems and the rights and responsibilities of citizens.
- **Economics.** Understand economic concepts and principles and how available resources are allocated in a market economy.
- **Geography.** Understand and use geographic skills and concepts to interpret contemporary and historical issues.

Table 5 shows the SRCs for each of the grade level groups (grades 4-5 tested at grade 5; grades 6, 7, and 8 tested at grade 8; and grades 9-11 tested at high school) and the percentage of test questions on the test at each grade level group that assesses that category.

Table 5.
Percentage of Test Items Assessing Each SRC for Social Sciences

Social Sciences	Score Reporting Category (strand)					
	Civics and government	Economics	Geography	Historical skills	United States history	World history
	SRC 1	SRC 2	SRC 3	SRC 4	SRC 5	SRC 6
Grade 5	25	25	25	10	15	N/A
Grade 8	20	20	20	10	15	15
High School	20	20	20	10	15	15

5.5 Writing

The Writing Performance Assessment is open-ended, requiring students to write an essay in response to one of multiple provided prompts. ODE provides the Writing Performance Assessment in two formats: paper-based and online. Whether administered in the paper-based or the online format, the Writing Performance Assessment is not timed, but students generally complete the assessment within 3 class periods (approximately 120 minutes).

For both the paper-based and the online format, students are presented with four prompts from which to choose a single topic to complete their writing sample. The prompts reflect the four modes Oregon assesses for writing: Narrative, Expository, Persuasive, and Imaginative. Each prompt covers one of four writing modes:

Narrative (also called *personal narrative*) writing recounts a personal experience based on something that really happened. The paper has a clear, identifiable storyline that is easy to recognize, follow, and paraphrase. All details work together in an integrated way to create a complete story with a beginning, middle, and an end. There is a focus with a controlling idea, central impression, sense of change, or something learned or gained by the writer. Events move along, staying within that focus,

with some sort of a narrative structure, often chronological. (For purposes of statewide assessment, narrative is distinguished from imaginative writing in that narrative prompts focus on real-life experiences, whereas imaginative prompts are meant to be fiction.)

Imaginative (also called *fictional narrative*) writing describes a situation or story that did not happen and is based on the writer's imagination. The writer might create a scene, situation, and character(s); might predict what could happen under hypothetical circumstances; or might solve a hypothetical problem using a creative approach. Imaginative writing often, but not always, takes the form of a short story. In some of the most effective imaginative writing, the writer uses his or her knowledge of the world, people, or situations to make the situation or story seem realistic, but, as in all fictional writing, the writer is not bound by the constraints of reality. Strong imaginative writing may contain, as appropriate, elements of fantasy, drama, humor, the unusual, the unexpected, or suspense. Reader reactions often range from a sense of being challenged or intrigued to a sense of feeling delighted or amused.

Expository writing informs, explains, clarifies, or defines. The writing informs or amplifies the reader's understanding through a carefully crafted presentation of key points, explanations, and supportive detail. The writing contains clear ideas that are focused and fully explained. When appropriate, the writer uses a variety of credible resources or personal knowledge to gather accurate, relevant information that provides a strong base of support in the form of facts, examples, illustrations, incidents, or explanations. Strong writers show a concern for audience and purpose by carefully selecting words, elaborative detail, and stylistic devices; they also recognize that greater stylistic distance may be required in a formal, academic paper than in an informal, personal paper, but that expository writing can be lively, engaging, and indicative of the writer's commitment to the topic.

Persuasive writing attempts to convince the reader to agree with a particular point of view or to persuade the reader to take specific action. The topic must be debatable: there are clearly reasons for more than one point of view. Persuasive writing differs from expository in that it does more than explain; the writer also takes a stand and endeavors to persuade the reader to take that same stand. Strong persuasive writers support their clearly stated position with reasoned arguments supported by credible evidence, facts, anecdotes, and statistics; if used, emotional appeals are well balanced by these and other objective forms of documentation. Strong writers also address other points of view but acknowledge or counter points without seeming to shift positions. When a specific audience has been identified, the nature of the arguments and the style of presentation are designed to appeal to that audience.

Writing essays measure skills across six SRCs: ideas and content, organization, sentence fluency, conventions, voice, and word choice. The traits of voice and word choice are scored even though they are not tied to the state performance standard. However, because teachers and other writing experts realize the instructional value of these traits, the statewide assessment continues to include them in scoring. As a result, students will receive feedback and consider these traits as valuable components of writing along with the required traits.

6. ITEM DEVELOPMENT

Oregon’s item development process is consistent with industry practice and takes approximately two years, including writing, reviewing, and field-testing new items. Just as the development of Oregon’s content and achievement standards is an open, consensus-driven process, the development of test items and prompts to measure those constructs is grounded in a similar philosophy.

6.1 Item Writing

For both OAKS and the Writing Performance Assessment, item writing takes place during item writing workshops, in which Oregon teachers across the five main content areas write and review items. In addition, Oregon has begun conducting remote item writing through its secure electronic item tracking system.

Item writers are typically Oregon teachers who have received training in item construction, are familiar with test specifications, and have demonstrated skill in writing items that pass content and sensitivity panel review. Item writers receive compensation for their time and travel expenses. Among other security precautions, ODE requires item writers to sign confidentiality forms assuring that they will work with the items in a secure manner.

All items are written to measure specific sub-domains of the content standards at a variety of specified levels of cognitive complexity and estimated difficulty. Cognitive complexity is represented by the following classification, developed from Bloom’s educational taxonomy:⁹

- **Recall:** Recall, label, or locate information; define or describe facts or processes.
- **Skill/Concept (Basic Application):** Use information or conceptual knowledge, often requiring two or more steps; summarize, classify, or explain information or processes; make predictions or generalizations; solve problems.
- **Strategic thinking:** Analyze, critique, compare or contrast; create new information; or organize presented information.
- **Extended thinking:** Make connections and extensions (exclusively assessed in the Writing Performance Assessment).

Through both item writing workshops and remote item writing, writers draft items, document rationale of distracters, and conduct peer reviews of each other’s items. Examples of items are provided, and facilitators provide process guidance and additional review. Writers and reviewers evaluate the strength and clarity of the match between the drafted item and the standard it measures, intended cognitive complexity, and intended difficulty. All issues are worked out or solved multiple times by multiple reviewers who verify that distracters are plausible, that answers are correct, and that each item has only a single correct answer.

⁹ Bloom, B. S. (ed.), Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: Handbook I: Cognitive domain*. New York: David McKay.

Item writing and passage selection are guided by the following principles for each of the item types.

Multiple-choice items:

- have one correct response option.
- contain plausible distractors that represent feasible misunderstandings of the content.
- provide options that are grammatically parallel in structure and length.
- represent the range of cognitive complexity and include challenging items for students performing at all levels.
- are appropriate for students in the assigned grade in terms of reading level, vocabulary, interest, and experience.
- arrange numerical answers in either ascending or descending order, except in items involving the ordering of numbers.
- are embedded in a real-world context.
- do not provide answers or hints to other items in the set or test.
- are in the form of questions or sentences that require completion.
- use clear language and are not worded in the negative unless doing so provides substantial advantages in item construction.
- are free of absolute wording, such as “always” and “never,” and have qualifying words (e.g., least, most, except) printed in small caps for emphasis.
- reflect the diversity of Oregon students and avoid emotionally-charged issues such as death, violence, drug and alcohol abuse, criminal activities, or the occult.
- are free of ethnic, gender, political, and religious bias.

Machine-scored graphic response items:

- represent the range of cognitive complexity and include challenging items for students performing at all levels.
- assess the content standards more directly and in different ways than is possible through multiple choice items
- are appropriate for students in the assigned grade in terms of reading level, vocabulary, interest, and experience.
- are embedded in a real-world context.
- do not provide answers or hints to other items in the set or test.
- use clear language and are not worded in the negative unless doing so provides substantial advantages in item construction.
- are free of absolute wording, such as “always” and “never,” and have qualifying words (e.g., least, most, except) printed in small caps for emphasis.
- reflect the diversity of Oregon students and avoid emotionally-charged issues such as death, violence, drug and alcohol abuse, criminal activities, or the occult.

- are free of ethnic, gender, political, and religious bias.

Selected reading passages:

- represent literary (fiction), informative (nonfiction), and practical selections (e.g., nontraditional text pieces, including tables, charts, glossaries, indexes).
- have a definite beginning, middle, and end and a sense of completeness.
- are of high interest and appropriate readability for the grade level.
- are of appropriate length for the grade level—
 - Grade 3, 300 words or less (150–250 on average)
 - Grade 4, 400 words or less (250–350 on average)
 - Grade 5, 500 words or less (350–450 on average)
 - Grade 6, 600 words or less (500–600 on average)
 - Grade 7, 700 words or less (600–700 on average)
 - Grade 8, 800 words or less (700–800 on average)
 - High School, 1,000 words or less (800–900 on average)
- do not represent material that is so widely anthologized or taught that students may have already studied the selection.
- include real-world texts (from consumer, workplace, or such public documents as letters to the editor, newspaper and magazine articles, thesaurus entries) to the extent possible.
- include material by writers from Oregon or the Pacific Northwest.
- reflect the diversity of Oregon students and avoid emotionally-charged issues such as death, violence, drug and alcohol abuse, criminal activities, or the occult.
- are free of ethnic, gender, political, and religious bias.

Writing prompts:

- are of high interest and readability appropriate to the grade level.
- represent the range of cognitive complexities and include challenging items for students performing at all levels.
- are appropriate for students in the assigned grade, in terms of reading level, vocabulary, interest, and experience.
- are embedded in a real-world context in which students may have experience.
- reflect the diversity of Oregon students and avoid emotionally-charged issues such as death, violence, drug and alcohol abuse, criminal activities, or the occult.
- are free of ethnic, gender, political, and religious bias.

6.2 Committee/Panel Review

ODE convenes a series of advisory groups to advise ODE both on assessment-related policy and on item development. ODE seeks to ensure that membership on these advisory groups reflects the demographics of Oregon’s student population. Each advisory group has approximately 15–35 members who serve three-year terms with one-third of the members rotating out each year and being replaced by new representatives.

Table 6 describes the structure of these panels.

Table 6.
Structure of ODE Content and Assessment Panels/Committees

Committee/Panel	Number of Members	Meeting Frequency	Who Nominates Members?
Assessment Policy Advisory Committee	20–25	2-3 times a year	School districts, Confederation of Oregon School Administrators (COSA), Oregon School Board Association (OSBA), Oregon Education Association (OEA), Educational Service Districts (ESDs), and Oregon Parent and Teacher Association (OPTA)
Sensitivity Panel ¹⁰	12–15	2–4 times a year	School districts, OEA, ESDs, and self-nominate (application process)
English/Language Arts Content and Assessment Panel	25	1-2 times a year	School districts, OEA, ESDs, and self-nominate (application process)
Mathematics Content and Assessment Panel	35	4 - 6 times a year	School districts, OEA, ESDs, and self-nominate (application process)
Science Content and Assessment Panel	35	4- 6 times a year	School districts, OEA, ESDs, and self-nominate (application process)
Social Sciences Content and Assessment Panel	25	1 - 2 times a year	School districts, OEA, ESDs, and self-nominate (application process)
English Language Proficiency Content and Assessment Panel	15	2-3 times a year	In Spring 2011, ODE recruited membership through applications sent to Title III directors and District Test Coordinators. In the future, a broader application process will be used.
Spanish Translation Review and Resolution Panel	8	2-3 times a year	New in 2011-12. In Spring 2011, ODE recruited membership through applications sent to Title III directors in districts with above average Spanish ELL populations. In the future, a broader application process will be used.

Note. Oregon’s Accommodations and Modifications Review Panel is not described here.

Source: <http://www.ode.state.or.us/teachlearn/testing/dev/panels/structurecapanel.doc>

¹⁰ Suspended for 2012-2013 school year due to restricted budget and reduced item development. Sensitivity review in 2012-2013 will be conducted by a subset of the full panel.

Panel members commit to up to 6 school days of service with an additional 3 or 4 days during the summer. Panels are often convened remotely rather than in person as secure technology improvements allow distributed work. Although committee members on district contracts are not compensated for their service, they do receive travel reimbursement for committee travel of more than 70 miles, and substitute teachers are provided for service during the school year. When classroom teacher members and other educators work for ODE during non-contract time, they are compensated at an hourly wage as temporary employees of ODE.

The Assessment Policy Advisory Committee consists of representatives from Oregon school districts, schools, and ESDs who are knowledgeable about assessment-related issues. The purpose of the Committee is to advise ODE on both the procedural and policy implications of Oregon's assessment system, as well as the feasibility of proposed improvements to Oregon's assessment system. Committee members provide input regarding the various elements of the state assessment system such as educational technology, electronic reporting, operational assessment issues, and test administration.

In addition to seeking advice on assessment-related policy, ODE requires that all items generated for use on Oregon statewide assessments must pass a series of rigorous reviews before they can be used in field and operational tests. All items go through both a content and a sensitivity review as part of the item development process; only those items that measure the grade-level expectations and meet both overall quality and sensitivity criteria are carried forward to the field-test stage.

ODE Content and Assessment Panels exist for each of the content areas for which statewide tests are given: English/Language Arts (this panel reviews Writing and reading/literature assessment items), mathematics, science, social sciences, and English language proficiency.

Most members of these panels are classroom teachers, with some representation from higher education, district curriculum and assessment personnel, and related businesses. Criteria for panel selection include the following:

- Knowledge of Oregon's content standards and expertise in the subject area and its eligible content
- Teaching experience at the grade level or benchmark to which the individual will be assigned
- Geographical location to ensure that all regions of Oregon are represented
- Gender and ethnic diversity to ensure that the panel represents the diversity of Oregon's student population

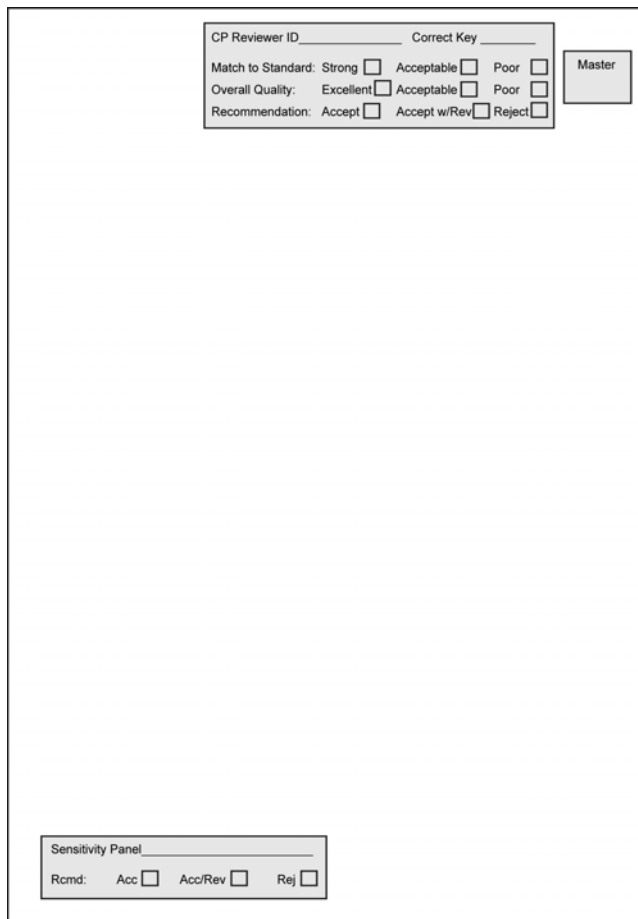
Current item writers are not allowed to serve on item review committees. However, in some cases, content and assessment panel experts may be utilized as item writing facilitators.

Items are accepted, rejected, or modified by the Content and Assessment Panel to make sure they represent the constructs embodied in grade-specific content standards and test specifications. In addition to judgments of content relevance, the panels appraise the technical quality of items, looking for items that are free from such flaws as (a) inappropriate readability level, (b) ambiguity, (c) incorrectly keyed answers and distracters, (d) unclear instructions, and (e) factual inaccuracy. The panels for each content area use the following review process:

1. Three content and assessment panel members review each item independently and complete an Item Review Form (IRF) (Figure 2) using a pre-assigned reviewer ID.
2. Then, the three content and assessment panel members review the item collectively, and item reviewers make a recommendation for each item on the IRF to either (a) accept the item as written, (b) accept the item with revisions, or (c) reject the item (sometimes an alternate question is offered that entails a simple revision).
3. When all three reviewers agree that an item should be accepted or rejected, no further discussion is needed. If one or more of the reviewers feel that an item should be revised, then they attempt to reach a consensus and produce a “master copy” of their recommendation. The same is true if one or two of the reviewers reject an item that another reviewer finds acceptable with or without revisions.

In most cases, recommendations are followed and revisions are made, or items are eliminated. The ODE assessment specialist can override the recommendation, but this occurs rarely and only for compelling reasons.

Figure 2.
Sample Oregon Content and Assessment Panel Item Review Form



The form is enclosed in a large rectangular border. At the top left, there are two input fields: "CP Reviewer ID" and "Correct Key". Below these are three rows of checkboxes:

Match to Standard:	Strong <input type="checkbox"/>	Acceptable <input type="checkbox"/>	Poor <input type="checkbox"/>
Overall Quality:	Excellent <input type="checkbox"/>	Acceptable <input type="checkbox"/>	Poor <input type="checkbox"/>
Recommendation:	Accept <input type="checkbox"/>	Accept w/Rev <input type="checkbox"/>	Reject <input type="checkbox"/>

To the right of these checkboxes is a "Master" button. At the bottom left, there is a "Sensitivity Panel" with three checkboxes: "Rcmd: Acc Acc/Rev Rej

The Content and Assessment panels perform specific checks on items to confirm that:

- the SRC and subcategory match.
- the key is correct.
- alternate valid interpretations making the distracters correct do not exist.
- the item is grade-level appropriate in content and reading levels.
- the item is of overall high quality (wording and grammar, graphic quality, curricular importance, etc).
- the identified level of difficulty (i.e., easy, medium, hard) is correct.
- Reading/literature passages are appropriate in content and reading levels. Science and social sciences stimuli align to appropriate content and reading skills.
- the level of cognitive complexity (i.e., recall, skill/concept, or strategic thinking) is appropriate to the item and correctly identified.
- the degree to which the item is accessible to students using Braille is correctly identified.

Following review by the content panel, and according to panel feedback, ODE assessment specialists edit and revise items in ITS in preparation for review by the Sensitivity Panel. The Sensitivity Panel typically convenes day-long in-person meetings at least once a year, supplemented by remote item review sessions. However, for 2012-2013, the full panel has been suspended due to limited budget and reduced item development. Instead, ODE will remotely convene a subset of the full panel to review items for field testing in 2012-2013. The panel reviews items from all grade levels and content areas for bias, controversial content, and overly emotional issues.

In general, the sensitivity panel ensures that items:

- Present racial, ethnic, and cultural groups in a positive light.
- Do not contain controversial, offensive, or potentially upsetting content.
- Avoid content familiar only to specific groups of students because of race or ethnicity, class, or geographic location.
- Aid in the elimination of stereotypes.
- Avoid words or phrases that have multiple meanings.

Following the Sensitivity Panel review and according to panel feedback, ODE assessment specialists edit and revise items in ITS.

6.3 Expert Review

Next, ODE assessment specialists submit the new items for review by experts that have experience in the roles of item writer and content and assessment panel member. Expert reviewers add an additional quality control check for the online assessments. Experts have received extensive professional development in ITS to review items in a web-preview format providing the exact rendering provided to students in the online assessments. Experts review each item and confirm that:

- the key is correct.
- alternate valid interpretations making the distracters correct do not exist.
- the item is grade-level appropriate in content and reading levels.
- the item is of overall high quality (wording and grammar, graphic quality, curricular importance, etc).

Following the expert review, the ODE assessment specialist follows the expert reviewer's recommendations and makes revisions or eliminates items. The ODE assessment specialist can override the recommendation, but this occurs rarely and only for compelling reasons.

6.4 Text-to-Speech Preparation

Oregon's assessment accommodations allow for individual students to have the Mathematics, Science, and Social Sciences items and response choices read aloud to them. OAKS Online Mathematics, Science, and Social Sciences support text-to-speech functionality to provide the read-aloud accommodation through the computer in English. Spanish computer-based read-aloud is also supported for OAKS Online Mathematics. For students using the text-to-speech functionality of OAKS Online, all text that appears in test stimuli, items, or response choices will be read verbatim. ODE has worked closely with both members of the assessment content panels and with expert

reviewers to develop business rules that ensure that non-text components (e.g., tables, diagrams, graphs, or flowcharts) are read in a manner that does not compromise the learning expectations, construct, grade-level standard, or measured outcome of the assessment. Prior to field testing, the text-to-speech output for each item is reviewed by expert reviewers to ensure that the validity of the items is maintained.

6.5 Field Testing

Once the items have been reviewed by the content and assessment panel, the sensitivity panel, and an expert reviewer, all Mathematics, Reading/Literature, Science, and Social Sciences test items are field tested in OAKS Online. Field test items identified by the ODE assessment specialists are embedded in the operational tests by content area. As students take the operational tests, they also respond to approximately 1-9 field test items embedded in the test.

ODE then receives data files of the student responses, which ODE analyzes to determine whether the field test items are behaving as expected. The ODE assessment specialists eliminate those items which the data analysis indicate performed weakly. ODE assessment staff calibrate the difficulty level for those items that performed successfully in preparation for using the item operationally. Section 9 below provides additional information on the embedding of field test items in operational tests.

6.6 Translation of items to Spanish

Concurrent with the field testing of items in English, all new Mathematics, Science, and Social Sciences test items are translated into Spanish. All required grade-level and benchmark-level statewide tests for Mathematics and Science are offered in both English and English/Spanish tests, using stacked English/Spanish items. English/Spanish tests are also available for Social Sciences.

Following translation by ODE's translation vendor, the translated items are reviewed by ODE's Spanish Translation Review and Resolution Panel to ensure that each item accurately conveys the intent of the English text.

The following linguistic guidelines are used by ODE's translation vendor and Spanish-speaking experts:

- Students are expected to have subject knowledge and use proper terminology/vocabulary for that subject. In other words, what is expected from English-speaking students is also expected from Spanish-speaking students.
- ODE uses formal Spanish (usted, not tú) for test items and includes proper verb conjugation.
- ODE strives to use Global Spanish language that will be interpreted and understood by all Spanish speakers from anywhere in the world. Global Spanish language includes words used worldwide by most Spanish speakers.

After the ODE Spanish reviewers complete a review of the newly translated items, extensive research is conducted by a small group of reviewers on any word that has not met group consensus. Every attempt is made to choose the most correct translation based upon grade level and cultural relevance. A variety of resources are used for selecting the proper translated words

including: dictionaries from Mexico, South America, and Spain (e.g. Diccionario Hispanoamericano de Dudas, Diccionario de Matemáticas), and ODE’s list of translated terms for science at <http://www.ode.state.or.us/search/page/?id=517> and for mathematics at <http://www.ode.state.or.us/search/page/?id=500>.

In addition, ODE offers a Native Language Spanish Reading/Literature assessment for grade 3 students who would benefit from taking such an assessment. Consistent with the federal law, only the native language test scores from LEP students who first enrolled in a U.S. School within the last 5 years will be incorporated in the AYP designations. Test scores from students who don’t meet these criteria will be invalidated and will not count toward participation or performance calculations. In addition to LEP students, some districts with Spanish language immersion programs have chosen to administer the Grade 3 Spanish Reading/Literature to its 3rd grade students as one of their three Reading/Literature testing opportunities as part of their program evaluation (even though these test scores will be invalidated for state use as described above. Ten authentic Spanish passages were selected by a search committee as initial selections and items were written in Spanish that align with Oregon’s grade 3 reading standards. These items were also translated into English for initial field testing. In addition, approximately 50 passages from the operational pool of English reading selections and their accompanying items were translated into Spanish and reviewed by another group of bilingual educators. This effort resulted in an opportunity for grade 3 students to use one of their three OAKS Online Reading/Literature testing opportunities to take an adaptive Spanish Reading/Literature assessment.

6.7 Text-to-Braille Preparation

OAKS Online Mathematics, Reading/Literature, Science, and Social Sciences are accessible to students with visual impairments through a Braille interface of OAKS Online. For students using the Braille interface, all items are presented in Braille through either a refreshable Braille display or through a Braille embosser; items with non-text, tactile components (e.g., tables, diagrams, graphs, or flowcharts) are embossed for tactile presentation. In addition, the Braille interface of OAKS Online provides text-to-speech functionality for OAKS Mathematics, Science, and Social Sciences. Volume 5: Test Administration provides additional information about how OAKS Online is administered to students through the Braille interface.

ODE has worked closely with teachers of the visually impaired, certified Brailleists, members of the assessment content panels, and expert reviewers to develop business rules that ensure that items are transcribed into Braille in a manner that complies with the Braille code published by the Braille Authority of North America and that does not compromise the learning expectations, construct, grade-level standard, or measured outcome of the assessment. Prior to field testing, the text-to-speech output for each item is reviewed by expert reviewers to ensure that the validity of the items is maintained.

6.8 Additional Expert Review of Items

On an annual basis, ODE assessment specialists review items from the field test pool for inclusion within the operational test. This level of review acts as an additional quality control for the online assessments. Furthermore, whenever ODE transitions to a different test delivery system, ODE submits all of its Reading/Literature, Mathematics, Science, and Social Sciences items for an another

level of expert review to ensure that all items appear consistently from year to year when presented to students. Finally, all English/Spanish items are reviewed by language experts.

7. ITEM USE AND RELEASE

Approximately every three years, ODE releases one sample test for each content area, grade-level, and benchmark-level comprised of items used on previous test forms. These items are no longer secure and are taken out of the pool of eligible test items.

Released items are provided in the form of sample tests. Sample tests for Reading/Literature, Mathematics, Science, and Social Sciences are available on ODE's Website at <http://www.ode.state.or.us/search/page/?id=1222>.

Sample Writing prompts are also available at <http://www.ode.state.or.us/teachlearn/subjects/elarts/writing/assessment/usingsampleprompts.pdf>

To familiarize students with the online test format, online practice tests for Reading/Literature, Mathematics, Science, Social Sciences, Online Writing, and the ELPA are available at <http://oakspt.tds.airast.org/student>. These practice tests include only a limited number of test items and do not generate a score; their purpose is to provide students experience with the OAKS Online testing environment.

8. TEST DEVELOPMENT

OAKS Online is the required test format for Reading/Literature, Mathematics, and Science, unless a student has an IEP that specifies the student must take the Extended Assessment. The optional Social Sciences assessment is also available through OAKS Online. OAKS Online is delivered to students through the secure, adaptive online testing system. While the Writing Performance Assessment is still primarily administered using a paper-based format, students have the option to test online instead.

Testing opportunities by content area and grade are described in greater detail in Volume 5: Test Administration.

8.1 Using Field-Test and Operational Test Data to Choose Items

ODE considers the statistical properties of existing and previously administered operational items when developing OAKS Online item pools.

ODE psychometricians examine item performance, including the percentage of students who answer each item correctly (the p value) and the correlation between each item and the total test score (the point-biserial correlation). Point-biserials for the correct answer should be 0.20 or higher, and for all incorrect options, point-biserials should be 0.0 or less. P values less than 0.25 indicate that students selected the correct response less often than would be expected by chance. If an item also has a high point-biserial, then the item discriminates and only the most successful students get it correct. If the point-biserial is low, then the item does not discriminate and is not included on subsequent tests.

Finally, field-tested and operational items undergo a differential item functioning (DIF) analysis, and items exhibiting significant DIF are referred back to content specialists for additional review.

8.2 OAKS Online Item Pools

OAKS Online is an adaptive assessment, which means that the items presented to each student vary in difficulty based on the student's cumulative performance on the test. Therefore, the state creates a grade-level item pool rather than a single pre-made test for each grade level and content area. The computer selects questions based on the answer a student gives to a test item, which in turn helps determine the difficulty of the next item that the computer will select.

ODE develops a separate item pool by grade for each content area: Reading/Literature, Mathematics, Science, and Social Sciences. Within each item pool, all content area SRCs are represented by a range of item difficulty levels that are approximately equivalent across SRCs. Each item pool contains enough items per SRC to ensure that as the test delivery system adaptively selects items to present to each student, students see a representative selection of items across SRC. This ensures the maintenance of the SRC weighting for each student's test.

To ensure that the OAKS Online item pools remain stable over time, each pool contains a percentage of items, typically approximately 80%-90%, that have been previously used operationally and are psychometrically sound.

8.3 OAKS Online Adaptive Item Selection Algorithm

Because OAKS Online is an adaptive assessment, the accuracy of a student's responses to one item determines the difficulty of the next item that the student will see. Thus, each student is presented with a set of items that most accurately aligns with his or her ability level.

Constraints used to ensure representative item selection: OAKS Online's adaptive test delivery system gives ODE the ability to develop an adaptive test that meets the requirements of a traditional test blueprint. Test blueprints and associated general test specifications typically include the following guidelines:

1. Length of the test
2. Content areas to be covered and acceptable range of items within each content area
3. Acceptable range of item difficulty for the specified grade level
4. Items that cannot appear on the same test
5. Number and location of field-test items, if applicable

In much the same way that these rules are used to build linear test forms, the same rules can be used by OAKS Online. Table 7 summarizes the test development rules and describes how the OAKS Online adaptive test engine implements the rules.

Table 7.
OAKS Online Testing Engine Rules

Rule	Implementation
Test length = X	The testing engine stops the examination when the student submits his or her answer to the X th item.
Items	The testing engine will deliver only the items specified as available for the adaptive delivery for that subject, grade, and adaptive test. The items included in the delivery rules file limit the difficulty range and grade-level coverage. If an item is in the item bank but is not in the adaptive test delivery specification file, then the item will not be delivered.
Field-test items	The test developer specifies the items for field testing. Field-test items do not contribute to the final proficiency score. The OAKS Online system ensures that the correct number of field-test items are delivered and that they are delivered within the specified location range for each test. (The test developer can specify item slot locations or random delivery.) A field-test item cannot be in the first or last few items on the adaptive test.
Item statistics	The engine uses the statistical data provided for each operational item by the test developer. These data are used by the engine to deliver the most appropriate subsequent item based on the previous response string.
Item content codes	The engine uses this classification data to monitor the content coverage. The content codes are used directly with the constraint groups, and the engine keeps track of each item that was delivered and keeps score against the blueprint ensuring that all rules are followed and that tests adequately cover the content as defined by the blueprint.
Item sets/groups	The item group information tells the OAKS Online engine to deliver a set of items together on the examination form. This is helpful, for example, for sets of items related to the same stimulus. Item sets are not limited to operational items. Field-test items can also be delivered in sets.

The adaptive test pools are designed to be equivalent in content and difficulty so that students can meet the standards regardless of which items the test engine presents to them. Items selected for each student depend on the student’s performance on previously selected items. Higher performance is followed by more difficult items, and lower performance is followed by less difficult items until test length constraints are met. Item selection is limited to items written for the specified grade level and is constrained to represent the test specifications, ensuring the appropriate representation of each SRC and coverage of the specified breadth and depth.

Each item pool contains approximately 1,000 items by grade and content area, a sufficient number of items to ensure students are presented with a test representing the breadth and depth identified in the test specifications and content standards, regardless of the item difficulty. Because the test adapts to each student’s performance while maintaining accurate representation of content breadth and depth, OAKS Online results provide precise estimates of each student’s true achievement level across the range of proficiency.

ODE and contractor staff work together to define and verify all rules used in OAKS Online.

8.4 OAKS Online Test Stopping Rules and Target Score Precision

Adaptive tests can be stopped based on the number of items presented to a student or on the desired precision of the student's estimated achievement level. The test engine will stop the assessment when the student submits his or her answer to the Xth item, as specified by the test specifications for each content area and grade level.

9. EMBEDDED FIELD TESTS

Each year, ODE embeds field-test items in the OAKS Online item pools by grade and content area. As described in Section 2.1, ODE will continue field testing new machine-scored graphic response items for Mathematics and Science. This new item type asks students to plot their answers on a grid and allows students to demonstrate their knowledge and skills in a more complex fashion than permitted by multiple choice items. These items are scored using a rubric and may be weighted to be worth multiple points.

Because students only respond to one prompt when taking the Writing Performance Assessment, ODE field tests new prompts for the Writing Performance Assessment in stand-alone field tests by grade rather than embedding the prompts in the operational test. These stand-alone field tests are described in greater detail in Volume 1: Annual Technical Report (available at <http://www.ode.state.or.us/search/page/?id=787>).

10. ALIGNMENT OF STANDARDS AND ASSESSMENTS

Oregon manages the alignment of content standards and assessments by carefully controlling every step of the test development process—from the creation of test specifications to item writing, content and sensitivity review, field testing, review of item performance, and test development.

The test specifications that ODE creates to describe how the content standards will be assessed by the OAKS and the Writing Performance Assessment at each grade level (3–8 and 11) and direct item writing and test construction play a major role in the alignment of standards and assessments. Items are written to specific standards and are reviewed for alignment to the standards through peer review, panel review, and ODE review. Each level of review assesses the alignment of the test item to the content standard that the item was written to measure. Only items that pass all reviews become operational.

Oregon maintains a consistent and coherent system of standards and assessments by closely coupling:

- development of grade-level content standards;
- item writing to match item specifications based on the content standards;
- Content and Assessment Panel judgments of fidelity to item specifications;
- development of test specifications;
- test assembly to ensure the appropriate coverage across SRCs; and
- score reporting according to the original content standards frameworks.

APPENDIX

Technical Advisory Committee Membership and District Advisory Committee

Name	Affiliation	Expertise
TAC Members		
Stanley Rabinowitz	WestEd	Assessment Policy and Content
Randy Bennett	ETS	Assessment of Complex Performance
Joseph Stevens	University of Oregon	Longitudinal Growth Measurement
Wayne Neuburger	ODE (Retired)	Large Scale Assessment
Gerald Tindal	University of Oregon	Assessment of Students with Disabilities
Evelyn Brzezinski	Portland Public Schools	Assessment Design and Interpretation
Tom Haladyna	Arizona State University (Retired)	Test Development and Validation
DAC Members		
Brian Bain	North Clackamas SD	
Todd Bloomquist	Medford Schools 549C	Director of Curriculum and Assessment
Karen Brown-Smith	Umatilla-Morrow ESD	Data Specialist
Catherine Carlson	Salem-Keizer Public Schools	
Lori Cullen	Clackamas ESD	School Improvement Consultant
Tim Drilling	Gresham-Barlow SD	Director, Student Achievement
Derek Edens	David Douglas SD	District Test Coordinator
David Marshall	Milton Freewater USD	Media Specialist/Technology Teacher
Kim Maurer	Santiam Canyon SD	Director of Curriculum/Federal Programs/Assessment
Kathi Robinson	Hillsboro SD	
Leigh Santy	Corvallis SD	ELL Coordinator
Tami Schild	NW Regional ESD	Assessment Coordinator
Don Staples	Newberg SD	Math Teacher/District Test Coordinator
Bill Stewart	Gladstone SD	Director of Curriculum and Assessment
Joe Suggs	Portland Public Schools	
Kelvin Webster	Multnomah ESD	Associate Director, Dept. of Instruction