



Oregon Department of Education

2009–2010

Technical Report

Oregon's Statewide Assessment System

Annual Report

Volume 1



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, sex, marital status, religion, national origin, age, or handicap 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
503-947-5600
<http://www.ode.state.or.us/>

Susan Castillo
State Superintendent of Public Instruction

Doug Kosty
Assistant Superintendent

Tony Alpert
Director, Assessment and Accountability

Kathleen Vanderwall
Manager, Test Development

Stephen Slater
Manager, Psychometrics and Validity

Melinda Bessner
Manager, Communications

Derek Brown
Manager, Assessment of Essential Skills

Ken Hermens
Language Arts Assessment Specialist

Leslie Phillips
Science and Social Sciences Assessment Specialist

Jim Leigh
Mathematics Assessment Specialist

Cindy Barrick
Research Analyst

Tom Tinkler
Psychometrics Specialist

Saleem Ahmad
Research Analyst

Sheila Somerville
Electronic Publishing Specialist

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: 2009-10 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

Volume 9: ELPA Test Development

Volume 10: ELPA Validity

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

CONTENTS

1. OVERVIEW	1
1.1 NO CHILD LEFT BEHIND	1
1.2 OREGON’S STATEWIDE ASSESSMENT SYSTEM.....	2
1.3 HISTORICAL CONTEXT OF OREGON’S STATEWIDE ASSESSMENT SYSTEM	3
1.4 PURPOSE OF THE STATE ASSESSMENT SYSTEM.....	4
1.5 MONITORING AND IMPROVING THE ONGOING QUALITY OF THE OREGON STATE ASSESSMENT SYSTEM (OSAS).....	6
1.5.1 Solicit External Review from Panels of Experts	6
1.5.2 Solicit Stakeholder Involvement	6
1.5.3 Technical and Research Studies	8
1.5.4 Active Research Agenda.....	8
1.6 RECENT CHANGES TO THE TEST	9
1.7 STUDENT POPULATION AND PARTICIPATION	9
1.7.1 Student Demographics	10
1.7.2 Student Inclusion	11
1.7.3 Participation Formulas.....	14
2. TESTS AND TESTING OPTIONS.....	15
2.1 ACADEMIC CONTENT AND ACHIEVEMENT STANDARDS	17
2.2 THE KNOWLEDGE AND SKILLS ASSESSMENT.....	17
2.2.1 Knowledge and Skills Item Types	17
2.2.2 Knowledge and Skills Assessment Subject Tests	17
2.3 PAPER-AND-PENCIL KNOWLEDGE AND SKILLS TESTS	23
2.4 SIDE-BY-SIDE (RUSSIAN AND SPANISH) ASSESSMENTS FOR ENGLISH LANGUAGE LEARNERS.....	24
2.5 AMERICAN SIGN LANGUAGE AND BRAILLE ASSESSMENTS	25
2.6 PERFORMANCE ASSESSMENTS	25
2.8 ALTERNATE (EXTENDED) ASSESSMENTS FOR STUDENTS WITH DISABILITIES (SWD)	25
3. TEST ADMINISTRATION.....	27
3.1 TESTING WINDOWS	27
3.2 TESTING SESSIONS	27
3.3 OPERATIONAL PROCEDURES.....	27
3.3.1 Standardized Testing Procedures.....	27
3.3.2 Test Security	28
3.3.3 Confidentiality of Student Records	29
3.4 TESTING ACCOMMODATIONS AND MODIFICATIONS	29
3.4.1 Accommodations	30
3.4.2 Modifications.....	30

3.4.3 Testing Accommodation and Modification Review Process.....	30
4. MAINTENANCE OF THE ITEM BANK.....	34
4.1 FIELD TESTING	34
4.1.1 Field-Test Sample Selection	35
4.1.2 Field-Test Analysis.....	36
4.1.3 Linking New Items and Tests to the RIT Scale	39
4.2 ASSIGNMENT OF CALIBRATED ITEMS TO ITEM POOLS	39
4.3 ITEM RELEASE AND RETIREMENT.....	40
5. FORM AND TEST POOL DEVELOPMENT.....	40
5.1 FORM DEVELOPMENT, PAPER-AND-PENCIL TESTS	40
5.2 ITEM POOL CONSTRUCTION.....	41
5.2.1 Summary of Item Selection Algorithms	41
5.3 TEST INFORMATION CURVES	44
6.0 SCORING	47
6.1 OAKS ONLINE SCORING	47
6.2 PAPER-AND-PENCIL TEST SCORING.....	ERROR! BOOKMARK NOT DEFINED.
6.3 WRITING PERFORMANCE ASSESSMENTS SCORING	47
6.3.1 Writing Scoring Process.....	48
6.3.2 Scorer Reliability.....	49

LIST OF TABLES

TABLE 1. DATES OF ADOPTION FOR ACADEMIC CONTENT STANDARDS	5
TABLE 2. STUDENT POPULATION DEMOGRAPHICS, 2009-10	10
TABLE 3. PARTICIPATION RATES, 2009-10	11
TABLE 4. SUMMARY OF TESTS AND TESTING OPTIONS.....	16
TABLE 5. TARGETED PERCENT OF ITEMS ON EACH TEST BY SRC FOR READING/LITERATURE	20
TABLE 6. PERCENT OF ITEMS ON EACH TEST FORM ASSESSING EACH SRC FOR MATH	21
TABLE 7. PERCENT OF ITEMS ON EACH TEST FORM ASSESSING EACH SRC FOR SCIENCE.....	22
TABLE 8. PERCENT OF ITEMS ON EACH TEST FORM ASSESSING EACH SRC FOR SOCIAL SCIENCES	23
TABLE 9. 2009-10 TESTING WINDOW FOR GRADES 3–8 AND HIGH SCHOOL	27
TABLE 10. NUMBER OF ITEMS AND FT ITEM INSERTION POINTS	<u>35</u>
TABLE 11. OAKS ONLINE TESTING ENGINE RULES	42

LIST OF FIGURES

FIGURE 1. OVERVIEW OF THE ASSESSMENT SYSTEM3

LIST OF APPENDICES

APPENDIX A NCLB CRITICAL ELEMENTS	1
APPENDIX B QUICK LINKS TO ODE ASSESSMENT DOCUMENTS.....	1
APPENDIX C TECHNICAL COMMITTEE MEMBERS, 2010-11	1
APPENDIX D RESEARCH STUDIES ON THE OREGON STATEWIDE ASSESSMENT SYSTEM	1

1. OVERVIEW

The present document is part of a series of technical reports designed to provide information about the technical and procedural characteristics of Oregon's Statewide Assessment System (OSAS). OSAS was created by the Office of Assessment in the Oregon Department of Education (ODE), with considerable participation and involvement from Oregon educators.

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

This annual report (Volume 1) describes assessment-related activities undertaken during the specified academic year. Volume 8, which also is updated every year, describes the results of the alternate assessment administered to students with disabilities. The Department updates Volumes 2 through 6 as new information becomes available or as new procedures are implemented.

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

This annual report describes the background, students tested, administration, and statistical properties of Oregon's tests.

1.1 No Child Left Behind

The federal No Child Left Behind Act of 2001 (NCLB)¹ mandated new requirements for state accountability for school improvement, professional development of teachers, assessment of student performance, and reporting student and school information to parents and communities.

NCLB requires the annual determination of whether schools, districts, and states have made adequate annual yearly progress (AYP) toward the goal of having all students meet rigorous state academic standards by the 2013–2014 school year. Each year the performance of all students in the school and district, as well as demographic subgroups such as special education and race/ethnicity, is measured against annual performance targets.

The law requires each state to submit a plan to the U.S. Department of Education, explaining how adequate yearly progress would be determined in that state, how assessments and reporting would be completed, and how additional requirements would be met. The Oregon plan was first approved on May 29, 2003, and last amended in May 2010, and is available at

http://www.ode.state.or.us/initiatives/nclb/pdfs/approvedaypb_current.pdf.

Additionally, NCLB established requirements for the state standards and assessment systems. The requirements are described in seven sections (Academic Content Standards, Academic Achievement Standards, Statewide Assessment System, Technical Quality, Alignment, Inclusion, and Assessment

¹ <http://www.ed.gov/nclb/landing.jhtml?src=pb>

Reports), each consisting of several critical elements. States compile and submit evidence addressing each of the critical elements for review by the U.S. Department of Education peer reviewers—national experts knowledgeable in the fields of standards and assessment.

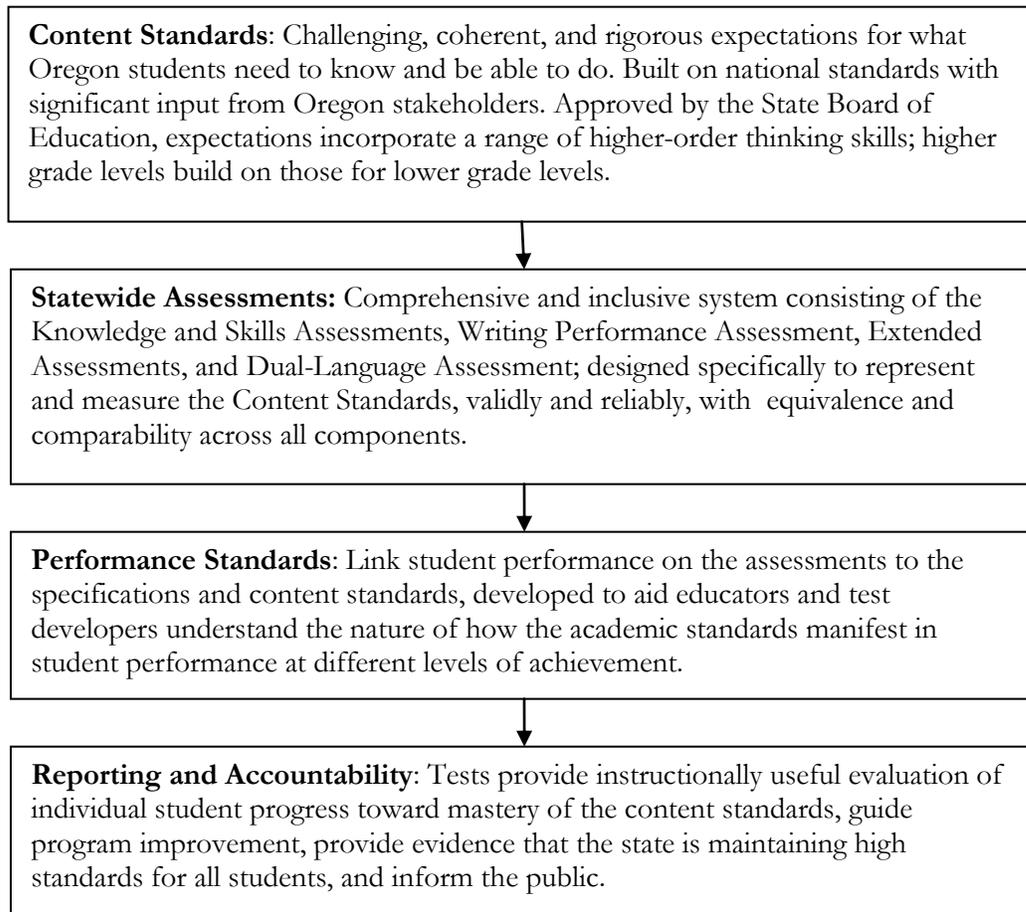
Oregon’s technical reports organize evidence addressing the critical elements to demonstrate the technical quality of the program. The table in Appendix A identifies the required evidence and locates each in Oregon’s eight volumes of technical reports.

1.2 Oregon’s Statewide Assessment System

The Oregon Assessment of Knowledge and Skills (OAKS) consists of the Knowledge and Skills Assessments (KSAs), measuring student performance in Mathematics, Reading/Literature, Science, and Social Science via multiple-choice tests aligned to grade-level content standards, and the Writing Performance Assessment measuring student performance in writing via open-ended essays. The KSAs are administered via OAKS Online, a computer-adaptive testing system. Students unable to take standard administrations of OAKS Online have other options—side-by-side tests in Spanish and Russian for English Language Learners, and Extended Assessments for students with IEP plans. The tests in Mathematics and Reading/Literature are used for NCLB accountability.

The major components of the system are described in Figure 1.

Figure 1.
Overview of Oregon’s Assessment System



1.3 Historical Context of Oregon’s Statewide Assessment System

Well before NCLB, Oregon’s Statewide Assessment System (OSAS) was monitoring student progress and was among the most innovative assessment programs in the country. OSAS is developed and maintained by the Office of Assessment in the Department of Education and thus has historically been responsive to changing state and national educational and assessment challenges and innovations.

In 1989 the Oregon state legislature funded a statewide assessment system based on a series of Oregon Administrative Rules (OAR),² passed in the early 1970s by the State Board of Education. These rules required districts to establish and measure student competencies. In 1991, Oregon administered the first census assessment and passed the Oregon Education Act for the 21st Century (1991, amended 1995),³ requiring that the state assessment system be used in determining student

² http://arcweb.sos.state.or.us/rules/OARS_500/OAR_581/581_tofc.html

³ <http://www.leg.state.or.us/ors/329.html>

progress toward the Certificate of Initial Mastery (CIM). The state tests annually assessed the *Essential Learning Skills* (ELS) at grades 3, 5, 8, and 11 and assessed the *Common Curriculum Goals* (CCG) on a six-year rotating cycle aligned with the textbook and instructional materials adoption cycle. Since 1996, Oregon has been focused on developing a comprehensive assessment system that includes all students.

In 1999 Oregon's legislature required the issuance of school report cards, rating and reporting schools on student performance (performance on the state assessment), student behavior, and school characteristics.⁴ The report cards were used to provide information to parents, identify schools in need of correction, and improve schools.

A year later, in December 2001, Congress passed the No Child Left Behind Act (NCLB), adding accountability in terms of Annual Yearly Progress (AYP) to Oregon's system. To determine AYP, NCLB required states to assess students in reading and mathematics annually in grades 3 through 8 and at least once in high school, and to assess students in science in elementary, middle, and high school (Oregon tests in science at grades 5, 8, and 10).

1.4 Purpose of the State Assessment System

Oregon's Statewide Assessment System provides instructionally useful information to educators about individual student progress toward Oregon's high school-level certification of mastery of the knowledge and skills described by the content standards. The content standards are benchmarked against national standards and were designed with stakeholder involvement to be rigorous, coherent, and demanding. In 1996 a panel of national experts reviewed Oregon's content and performance standards; they concluded that the standards were rigorous and powerful tools for holding students accountable for their learning. Since 1997, *Education Week* has consistently awarded Oregon high marks for its standards and assessment system.

A major effort by the Oregon Department of Education resulted in the establishment of content standards that serve as the goal structure for the state assessments. See *Volume 3: Standard Setting* for a detailed description of the content standards development process.

Table 1 provides the dates of most recent adoption and anticipated revision by the State Board of Education for the content standards in each content area.

⁴ SB 1329

Table 1.
Dates of Adoption for Academic Content Standards

Content Area	Most Recent Adoption Date	Anticipated Revision Date
English Language Arts	K–3, June 2002 4–8 and HS, January 2003	2012
Mathematics	December 2007, June 2009	2015
Science	February 2009	2016
Social Sciences	April 2001	2010
The Arts	October 2004	2011
Second Languages	March 2002	2005
Physical Education	September 2001	2013
Health Education	February 2005	2013
English Language Proficiency	June 2004	2013

Oregon’s Academic Content Standards are available on the Web site via the state’s Searchable Standards Tool that allows you to locate, view, and export standards by subject, grade level (benchmark), and/or strand (subtopic or Score Reporting Category, SRC).

<http://www.ode.state.or.us/teachlearn/real/standards/>

The OSAS is designed to measure the grade-specific content described in the standards and is intended to serve the following goals:

- Provide instructionally useful data as part of an evaluation of individual student progress toward mastery of the academic content standards
- Guide instructional program, school, district and state level improvement
- Ensure that the state is progressing toward the state and federal goals of high standards for all
- Inform the public

All tests are developed to be representative and valid measures of the knowledge required by Oregon’s academic content standards; to facilitate accessibility for all students, the tests are designed according to the principles of universal design.

Expectations for teaching and learning are organized into the following Curriculum Goals, Grade-level Standards, and Foundations:

1. Common Curriculum Goals (CCG) that describe the knowledge and skills expected of all students as a result of their educational experience (OAR 581-022-0102)
2. Grade-level Standards that describe what students should know and be able to do at grades 3 through high school
3. K–2 Grade-level Foundations that describe a suggested curriculum organization to better help students prepare to meet the third grade standards

1.5 Monitoring and Improving the Ongoing Quality of the Oregon State Assessment System (OSAS)

To maintain the ongoing quality of the Statewide Assessment System, Oregon makes continual efforts to monitor and improve the development, administration, and technical adequacy of the assessment system.

Maintaining transparency of the system is important for stakeholder evaluation. A directory of links to important assessment information available on the ODE Web site is provided in Appendix B, and additional links are cited through this report.

1.5.1 External Review from Panels of Experts

In 2006–07, ODE reconvened a TAC consisting of national experts to advise the state on best practices in assessment.

Test development, administration, and analysis are guided by a panel of national and local experts in areas of large-scale state assessment. These panels of experts include the District Advisory Committees (DACs), already in place to provide local education agency (LEA) guidance to the ODE; a Statewide Assessment Advisory Committee (SAAC), consisting of local experts knowledgeable in Oregon’s assessments, provides additional guidance.

In 2006–07, ODE reinstated a panel of national experts to provide additional technical guidance to the ODE, DACs, and SAAC. The Technical Advisory Committee (TAC), composed of national experts in state assessment, computer-based testing, student growth, accountability, special education, and limited English proficiency (LEP), meets regularly to advise ODE on development and improvement of OSAS. The DAC and the SAAC meet quarterly, and the TAC convenes twice each year. Panels discuss current issues and future directions of the assessment program and provide advice and guidance to ODE concerning national best practices. Appendix C identifies the Technical Advisory Committee members.

1.5.2 Stakeholder Involvement

Oregon is unique in the involvement educators have had in the development and maintenance of the OSAS. Unlike most states, Oregon has not relied extensively on testing contractors to develop, administer, score, and report on the assessments. Consequently, Oregon has a long history of extensive stakeholder involvement. Input from education stakeholders is one of the three primary components of Oregon's standards review and revision process.

Education stakeholders include teachers, administrators, higher education (both in departments of education and content area departments), industry, business, parents, and other citizens.

The State has a Content and Assessment Panel for each of the eight legislatively mandated subject areas. Thirty-two educators comprise each panel in English language arts, mathematics and science. These educators represent approximately 27 K-12 teachers and each geographic region of the State. In addition, there are representatives from higher education, school districts, and education service districts. Beginning in 2005, the panels also included representatives with expertise in special education, limited English proficiency, and career pathways.

The panels meet about six times (approximately 8 to 12 total days) during the review and revision process for in-depth analysis of the standards revision and impact on districts, teaching, and learning. ODE staff facilitate these meetings. The panels are led through a series of specific exercises with documented feedback. The goal of these exercises is to ensure: coherency, vertical alignment, and horizontal alignment. The panels also serve on a continuing basis in assessment item review, in curriculum and instruction work. At least two (sometimes up to twelve) drafts of the standards are distributed statewide for educator input. Types of distribution include: content area professional organizations' membership e-mail lists, professional organizations' newsletters and conferences, statewide councils (for example the Oregon Math Education Council and the Oregon Science Education Council - each with broad membership from K-12, community college, private and public universities, professional organizations, and pre-college programs). Oregon's Superintendent Pipeline, and posting on the ODE web. At least once per review and revision cycle a detailed survey is distributed. The survey requests input on clarity of the standards, and grade-level appropriateness and typically results in substantive and specific suggestions that are incorporated as long as they are not in conflict with the national standards or research evidence.

Oregon's 21st Century Advisory Board - with representatives from teaching, administration, higher education, business and industry, and parents - reviewed the standards and provided input. Public input is obtained from the business and industry community, parents, and other citizens. About six public input sessions, per subject, are held across the state, these are held in libraries or in schools where site councils are meeting. Although general solicitations for standards input are sent to businesses and industries; thorough input requires one-on-one contact. Connections are also made with the State parents organization, the Oregon Education Association, Confederation of Oregon State Administrators, and the Oregon School Boards Association.

During development of the content and achievement standards, regional public meetings solicit and incorporate community input after drafts are presented for public review. When establishing performance standards, ODE includes parents and members of the public and business communities on Standard Setting panels and holds multiple public input forums, generally about six for each subject, for both educators and non-educators. When content standards are under development or revision, up to 12 drafts per subject are distributed for stakeholder input.

Additionally, Oregon educators provide input into the quality, content, and alignment between test items and the academic content standards, via participation in Content and Assessment Panels. Teachers contribute invaluable by writing items specifically to measure the objectives and specifications of the content standards. The many different contributors to the item writing and review process contribute to the validity of the assessments. Educators score the state Writing tests and Extended Assessments. Oregon educators are also involved in Bias and Review committees and in committees confirming the accuracy and appropriateness of item translations.

Item and test development committee composition, participation, and responsibilities are described in greater detail in *Volume 2, Test Development* and in the *Test Specifications and Blueprints, 2009-2010*.

Finally, ODE maintains regular communication with District Test Coordinators to obtain feedback on the quality of the assessment system through regional meetings every fall and spring.

1.5.3 Technical and Research Studies

The Standards for Educational and Psychological Testing (AERA/APA/NCME, 1999⁵) describes the characteristics of high-quality assessments and provides a comprehensive process state testing programs can follow to ensure that the results of the system are appropriate, credible, and technically sound. The standards serve as a guide to monitoring and improving the assessment system.

ODE developed the OSAS to be innovative and grounded in research and best practices and frequently contracts studies and evaluations of the system by independent measurement and assessment experts. Results of these studies provide evidence in support of the technical adequacy of the system and identify new directions and technical approaches that continually improve the system.

Studies are reviewed by ODE psychometricians and assessment staff and are typically submitted to technical advisory panels as well. Some of these studies and evaluations are referenced in Appendix D. Additionally; the Department regularly conducts internal studies and evaluations of the system and participates in collaborative investigations with universities and national experts in assessment.

Recent studies are summarized and discussed in *Volume 4: Reliability and Validity*.

1.5.4 Active Research Agenda

Maintaining technical quality is a critical element identified by NCLB, and ODE continues to maintain an active and rigorous research agenda that includes studies to address peer review requests and meet the NCLB requirements. Studies designed, implemented, or underway in 2005–08 include the following:

- A comparability study providing additional evidence supporting the equivalence of the paper-and-pencil and OAKS Online tests and demonstrating that the tests are restricted to grade-level content and fully represent the test blueprints at the achievement levels
- A comparability study documenting the comparability of OAKS Online tests
- Comparison of classification consistency of OAKS Online and paper-and-pencil forms
- A study to provide evidence of strand content consistency across the paper-and-pencil and OAKS Online forms
- Additional reliability and validity providing supplementary evidence of the Alternate Assessments
- A comparability study supporting the comparability of the dual-language tests

⁵ American Educational Research Association, American Psychological Association, The National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington DC: American Psychological Association.

- Alignment studies for Reading/Literature, Science and Mathematics to evaluate the relationship between Oregon’s assessments to its academic content standards
- An alignment study for the Alternate Assessments to provide evidence of alignment with the academic content standards

1.6 Recent Changes to the Tests

Historically, Oregon’s Statewide Assessment System has been responsive, innovative, and continually evolving in response to the changing national educational context and best practices. ODE continues to revise and improve the system to respond to the outcomes of evaluation studies and to address the needs of students and teachers. Additionally, a number of recent changes were implemented to address guidance provided by the U.S. Department of Education (USED) peer review process.

Recent (2006-2010) changes to the system include:

- OAKS Online became mandatory for the 2007-08 school year for Mathematics, Reading, and Science multiple-choice tests for grades 3–8 and high school. Paper-and-pencil tests are now available only to students who need a paper-and-pencil accommodation to demonstrate mastery of the content standards, although district superintendents may request a one-year waiver for individual schools through ODE.
- Oregon updated its performance levels for all grades in the content areas of Mathematics, Reading/Literature, and Science. Revised performance standards were adopted by the state board in March 2007 and were applied to all tests administered during the 2006–2007 school year and later.
- In 2006–07, the state no longer includes juried assessments, targeted down assessments, or modified assessments in calculations of participation or proficiency used for AYP.
- CLRAS, Oregon’s test of life skills, was discontinued in 2006–07. This test no longer is included in calculating AYP participation or performance.
- In 2006–07, the Oregon Department of Education worked with the University of Oregon to expand the extended assessments so that there are three grade bands of assessments aligned to elementary, middle, and high school content standards respectively.
- Oregon revised the K-8 Mathematics content standards in December 2007 for an anticipated implementation in 2010-11 to better integrate process expectations and problem solving throughout the standards.

1.7 Student population and participation

All public school students in grades 3–8 and 10 in Oregon schools participate in the state assessment system. Student participation and demographics are described below.

1.7.1 Student Demographics

Oregon's students are described below by gender, ethnicity, Limited English Proficiency (LEP) status, Individualized Education Program (IEP) status, and economically disadvantaged status.

Table 2.
Student Population Demographics, 2009-10

Number of Students eligible for testing and enrolled at tested Grades on May 1, 2009	Knowledge and Skills Assessments*			Writing Performance Assessment
	Math	Reading/Literature	Science	
Total	299821	299839	129077	128086
Economically disadvantaged	155344	155357	63795	64299
Students with disabilities (IEPs)	47138	47146	19159	19539
Students with Limited English Proficiency	38340	38333	13752	14918
American Indian/Alaskan Native	5769	5785	2511	2508
Asian/Pacific Islander	13692	13678	5802	5904
Black/African American	8229	8220	3642	3527
Hispanic and Latino	58753	58813	23724	24350
White/Non-Hispanic	201846	201597	88509	87183
More than one race	8737	8686	3519	3617
Male	153788	153794	66687	65676
Female	146033	146045	62390	62410

The table below describes the participation rates for Oregon’s tests.

Table 3.
Participation Rates, 2009-10

Number and Percent of Eligible Students Tested	Knowledge and Skills Assessments*			Writing Performance Assessment
	Math	Reading/Literature	Science	
Total	297630 99.3%	297351 99.2%	126268 97.8%	123109 96.1%
Economically Disadvantaged	154364 99.4%	154116 99.2%	62295 97.8%	61396 95.5%
Students with Disabilities (IEPs)	46339 98.3%	46351 98.3%	18312 95.6%	18150 92.9%
Students with Limited English Proficiency	38113 99.4%	37653 98.2%	13374 97.3%	14035 94.1%
American Indian/Alaskan Native	5684 98.57%	5714 98.8%	2417 96.3%	2372 94.6%
Asian/Pacific Islander	13605 99.4%	13435 98.2%	5692 98.1%	5656 95.8%
Black/African American	8119 98.7%	8072 98.2%	3495 96.0%	3246 92.0%
Hispanic and Latino	58350 99.3%	58243 99.0%	23092 97.3%	23159 95.1%
White/Non-Hispanic	200428 99.3%	200226 99.3%	86785 98.1%	84234 96.6%
More than one race	8692 99.5%	8643 99.5%	3467 98.5%	3502 96.8%
Male	152509 99.2%	152295 99.0%	65077 97.6%	62622 95.3%
Female	145121 99.4%	145046 99.3%	61191 98.1%	60487 96.9%

1.7.2 Student Inclusion

Since 1996, Oregon’s assessment system has increasingly included all students. Oregon offers a single comprehensive assessment system that includes every child and avoids the creation of separate standards and assessments for subgroups of students. By design, Oregon includes English language learners (students with limited English proficiency), and students with disabilities. The system provides a variety of options to students for participation, and, with these options, more students than ever before are taking tests and receiving feedback on their performance.

Oregon offers multiple testing options including paper-and-pencil tests for those who cannot easily access OAKS Online or who have difficulty with computer testing, accommodations to the testing environment evaluated and approved by a panel of experts, Braille versions, and dual-language

(Russian/English and Spanish/English) forms.⁶ Additionally, Oregon offers Extended Assessments to students with the most significant cognitive disabilities).

Testing may optionally include registered home-schooled students and privately schooled students. Required testing includes students who are exempted from compulsory school attendance or who are attending public schools or programs part time, if the student is claimed by a district financially and the student received instruction provided by the district in the state content standards during the school year. In addition, home-schooled students taking courses in any of the content areas covered by the assessments must take a state test in their grade-level in that area.

Foreign exchange students are not required to participate in state tests if they are not Oregon residents. Parents may request exemptions from testing for religious beliefs or student disabilities; however, exemptions count as nonparticipation in AYP reporting.

1.7.2.1 Guidelines for Including Limited English Proficiency students

Limited English Proficient (LEP) students are those whose native language is other than English and students who are enrolling in a U.S. school for the first time after August 15, 2009. Specifically, the term “Limited English Proficient” means an individual

- A. who is age 3 through 21;
- B. who is enrolled or preparing to enroll in an elementary school or secondary school;
- C. (i) who was not born in the United States or whose native language is a language other than English;
(ii) (I) who is a Native American or Alaska Native, or a native resident of the outlying areas; and
(II) who comes from an environment where a language other than English has had a significant impact on the individual’s level of English language proficiency; or (iii) who is migratory, whose native language is a language other than English, and who comes from an environment where a language other than English is dominant; and
- D. whose difficulties in speaking, reading, writing, or understanding the English language may be sufficient to deny the individual —
(i) the ability to meet the State’s proficient level of achievement on State assessments (described in section 1111(b)(3) of the No Child Left Behind Act);
(ii) the ability to successfully achieve in classrooms where the language of instruction is English; or
(iii) the opportunity to participate fully in society.

Non-English Proficient students during their first year of enrollment in U.S. schools are not required to take the state assessments in Reading and Writing, although they must take a test of English Language Proficiency. Oregon’s test of English Language Proficiency is the English Language Proficiency Assessment (ELPA). “First-year” students taking ELPA during their first year of enrollment in U.S. schools are counted as participants in Adequate Yearly Progress reports in Reading or Writing.

⁶ In spring 2010, Oregon conducted a pilot of a native language Spanish reading assessment for third grade. Use of the results for accountability purposes is pending approval by the U. S. Department of Education.

First-year LEP students enrolled in the assessed grades are required to take state assessments in mathematics and science. The scores of first-year LEP students are included in participation calculations but are not included in determining the percentage of students meeting the standards in AYP determinations or report card ratings.

The Limited English Proficient group includes transitioning students in AYP performance calculations. Transitioning students are LEP students who demonstrate fluency in English on a formal English language proficiency assessment and are on monitoring status until they no longer need instructional services and methods provided by the district's LEP program; they can remain on monitoring status for up to two years. After a maximum of five years, monitoring ends, and the students are not reported as part of the LEP subgroup.

LEP students may take the test under standard administration with or without accommodations.

- Students may take Side-by-Side English/Spanish, English/Russian versions where provided.
- Students may take the test under modified conditions.

Teachers and instructional teams who know individual student strengths and weaknesses decide, in consultation with parents or guardians, which accommodations, if any, are necessary for LEP students.

Each student is considered individually for each assessment on the basis of what is in the best interest of the student, not on participation in a particular program or identification as an English Language Learner.

For more information on test administration to Limited English Proficient students, go to the Department's Web site at

<http://www.ode.state.or.us/search/page/?=1225>.

1.7.2.2 Guidelines for Including Students with Disabilities

Students with disabilities access the assessment system through one of the following means:

- The standard tests
- Accommodations or modifications to the standard tests
- The state's Alternate Assessment System, developed for the students with the most significant cognitive disabilities who cannot demonstrate mastery of the state content standards without a reduction in depth, breadth or complexity.

Most students with disabilities take regular benchmark assessments with or without accommodations; however, the Extended Assessment System includes Reading, Math, Writing, and Science tests and allows access to the testing system for students who meet most or all of the following criteria:

1. The student has a current Individualized Education Program (IEP);

2. The student's psycho-educational evaluation and educational performance data support the following:
 - a. The student exhibits significant cognitive impairments and adaptive skills to the extent he or she requires extensive modifications; and
 - b. The student cannot participate in Oregon's General Statewide Assessment even with accommodations.
3. For students in High School, the students' IEP team determines that he or she is unable to complete the graduation standards necessary to earn a regular high school diploma, even with extended learning opportunities and/or significant modifications.

The IEP team documents eligibility for all students who meet all three criteria. The student is then eligible for participation in Oregon's Alternate Assessment for the grade level at which the student is enrolled. Students may take one, two, three, or all four of the Extended Assessments, depending on the objectives identified in individual IEPs.

1.7.2.3 Guidelines for Identifying and Including Migrant or Mobile Students

Oregon has a system of unique identification codes that is able to track individual students over time as they move within the state. This system allows for the inclusion of students in the assessments despite mobility. Testing via OAKS Online has the additional benefit of allowing new schools and teachers access to current student assessment data, regardless of where a student was enrolled when tested.

Migrant students are those who

- are younger than 22 and have not graduated from high school or earned a GED;
- who are migrant agricultural workers or migrant fishers (or have a parent, spouse, or guardian who is a migrant agricultural worker or a migrant fisher); and
- have moved from one school to another in order to obtain (or to accompany a parent, spouse, or guardian to obtain) temporary or seasonal employment in agriculture or fishing;

and are included in the state's testing system.

1.7.3 Participation Formulas

For AYP, the participation rate formula is as follows: tests scores from students enrolled on the first school day in May / expected number of test scores from students enrolled on the first school day in May minus students ineligible for testing.

For the State Report Card, the participation rate formula is as follows: tests scores from students enrolled on the first school day in May / expected number of test scores from students enrolled on the first school day in May minus students ineligible for testing and students exempted from state tests by parents.

The participation rate formula is the same as the AYP participation rate formula for assessment group reports.

Students who are enrolled on the first school day in May and during the school's test window but who do not participate in the state assessment are counted as nonparticipants. For students who are enrolled on the first school day in May but were not enrolled during the school's test window (or arrived so late in the school's test window that a test could not be completed), the district may select an Administration Code to "excuse" the nonparticipation. Oregon releases annual participation summary reports describing students who were not tested.

The sections below describe the options available for students to access the OSAS.

2. TESTS AND TESTING OPTIONS

Students in grades 3–8 and 10 are tested in Reading/Literature and Mathematics, and students in grades 5, 8, and 10 are tested in Science and Social Science, using Oregon's Knowledge and Skills tests. Students in grades 4, 7, and 10 are tested in Writing, using the Performance Assessment. All are state-developed, criterion-referenced tests designed to align to the content standards and measure what students should know and be able to do in each subject and at each grade level.

Reading/Literature, Mathematics, Social Science, and Science Knowledge and Skills Assessments (KSAs) are available via the web-based Oregon Assessment of Knowledge and Skills (OAKS) Online system. In 2005–06, OAKS Online became the primary method of assessment for the Oregon State Assessment Program with over 90% of the knowledge and skills tests administered online, including approximately 40,000–45,000 students per grade. OAKS Online presents questions based on each student's performance, using only items from that particular grade level, to assess student performance.

Students using OAKS Online have the option to take the standard knowledge and skills grade-level tests up to three times each academic year from October through May. When students test multiple times, their highest test score is reported.

Oregon offers a comprehensive array of options for special populations, including paper-and-pencil tests for those who cannot easily access OAKS Online or who have difficulty with computer testing; accommodations to the testing environment evaluated and approved by a panel of experts; Braille and large print versions; and dual language (Russian/English and Spanish/English) forms. Additionally, Oregon offers Extended Assessments to students with Individual Education Programs (IEPs). The following testing options are available:

- Knowledge and Skills Assessments:
 - OAKS Online
 - Paper-and-pencil form
 - Side-by-Side Assessments (Russian and Spanish, other languages may be provided by districts as testing modifications)
- Extended Assessments (for students with significant cognitive disabilities)

- Writing Performance Assessment

The testing options provided to Oregon’s students are summarized in Table 4 below.

Table 4.
Summary of Tests and Testing Options

Subject Tests	Testing format		Language(s) tested	Grade levels tested
	OAKS Online	Paper and pencil		
Reading/Literature KSA	✓	✓	English Spanish ⁷	3–8, 10 3
Mathematics KSA	✓ ✓	✓ ✓ ✓	English Spanish/English Russian/English	3–8, 10
Science KSA	✓ ✓	✓ ✓ ✓	English Spanish/English Russian/English	5, 8, 10
Social Sciences KSA	✓ ✓		English Spanish/English	5, 8, 10
Writing PA	✓	✓ ✓	English Spanish	4, 7, 10
English Language Proficiency	Web-based		English	K–12
Extended Assessment, Reading		✓	English	3–8, 10
Extended Assessment, Math		✓	English	3–8, 10
Extended Assessment, Writing		✓	English	4, 7, 10
Extended Assessment, Science		✓	English	5, 8, 10

KSA = Knowledge and Skills tests, PA = Performance Assessment.

⁷ Use in accountability pending approval by U. S. Department of Education

2.1 Academic Content and Achievement Standards

All of the state tests are designed to measure the grade-level expectations for what students should know and be able to do as described in Oregon’s Academic Content Standards. The standards for each subject are available on the Departments Web site at

<http://www.ode.state.or.us/teachlearn/real/standards/>.

The development, review, and adoption processes of the content standards are described in *Volume 3: Standard Setting*, along with Academic Achievement (Performance Level) Descriptors and Academic Achievement Standards (cut scores) for each subject and grade.

2.2 The Knowledge and Skills Assessment

Written by ODE specifically to measure the content standards at each grade level, the KSAs measure progress toward and obtainment of the academic content standards in Math,

OAKS Online is designed to administer equivalent tests representing the content standards and SRC across tests and item pools so that students can meet the standards regardless of which items are taken.

Reading/Literature, Science, and Social Science. The tests are administered via OAKS Online (formerly TESA), an electronic, web-based adaptive assessment system that delivers test items in a progressive format, dynamically selecting items for presentation based on a student’s success at answering previous questions. Unlike some adaptive testing programs, OAKS Online adheres to strict content specifications, measuring performance only on items that are aligned with grade-level content standards.

OAKS Online, implemented in 2001, became the mandatory state testing method in 2006 for the KSA. The online administration reduces school and district test administration burden, reduces the material production and handling required for paper-and-pencil tests, allows for flexible test scheduling and multiple testing opportunities, and provides results immediately after a student completes a test. Schools with insufficient technology to access OAKS Online and students who have difficulty with computer testing may request a paper-and-pencil test. In 2008-09, 999 tests were administered via paper and pencil in subjects where OAKS Online was available (0.07% of all tests). In 2009-10, only 639 multiple-choice knowledge and skills tests were administered on paper forms (0.04% of the 1.49 million tests administered).

2.2.1 Knowledge and Skills Item Types

Knowledge and Skills tests are multiple-choice tests, consisting of 40-50 operational items and 5-6 field-test items. Students select answers from four alternatives; each multiple-choice item is scored as right or wrong, and correct answers earn one point. Blocks of field-test items are included on operational tests but are not included in computing student scores.

2.2.2 Knowledge and Skills Assessment Subject Tests

Knowledge and Skills tests are administered in Reading/Literature, Math, Science, and Social Science. Tests for each subject are described below.

2.2.2.1 Reading/Literature

Reading is available to grades 3–8 and 10 and consists of 50 multiple-choice questions (40 for grade 3 and 45 for grade 4). The tests include passages with multiple associated questions; passages are selected by Oregon teachers and are developed appropriately to each grade in terms of length, subject appropriateness, difficulty, and quality.

Reading passages are selected to reflect three purposes:

- Reading for literary experience (passages include literary texts, such as poetry, fiction, drama)
- Reading to gain information, (passages include informative text, including articles, biographies, autobiographies)
- Reading to perform a task (passages include practical selections, instructions, reference material)

Reading tests are generally administered only in English. Students in 3rd grade who are LEP may demonstrate their mastery of Reading in Spanish by taking the Aprenda, published by Harcourt Assessment.

The content of the Reading/Literature test specifications reflects the skill expectations that were outlined in the content standards most recently adopted by the State Board of Education in June 2002 (for K–3) and January 2003 (4–8 and High School). These standards were developed, in part, to represent the content identified by the National Standards for the Language Arts⁸ and A Compendium of Standards and Benchmarks for K–12 Education⁹ and to correlate to the skills assessed on the reading portion of the National Assessment of Educational Progress (NAEP).¹⁰ As a result, Oregon’s Score Reporting Categories (SRCs) for Reading use similar terminology in their definitions. Based on the standards, the Reading/Literature tests are designed to assess literacy skills overall in the following six SRCs:

- **Vocabulary** (formerly *word meaning*). 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 may be found through careful reading of the entire text. At varying grade levels, students may also be asked to use context clues to determine the meanings of words with multiple meanings or phrases, such as idioms and figurative expressions.
- **Read to perform a task** (formerly *locating information*). 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

⁸ National Council of Teachers in English, www.ncte.org

⁹ McREL and ASCD, www.mcrel.org/standards-benchmarks

¹⁰ <http://nces.ed.gov/nationsreportcard/>

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** (formerly *literal comprehension*). Students show a general understanding by accurately responding to questions that cover 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 to identify details or events that were critical to the development of the plot.
- **Develop an interpretation** (formerly *inferential comprehension*). 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** (formerly *evaluative comprehension*). 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 questions about instances in which the author has relied on facts or opinions, 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** (formerly *literary elements and devices*). 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 5 shows the Score Reporting Categories for each of the grade levels and the range of test questions on the test at each grade level that assesses that category.

Table 5.
Targeted Number of Items on Each Test by 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: informative text	Examine content and structure: literary text
	SRC 1	SRC 2	SRC 3	SRC 4	SRC 5	SRC 6
Grade 3	10-13	5-7	10-13	10-13	N/A	N/A
Grade 4	9-13	5-7	9-13	9-13	4-6	N/A
Grade 5	9-12	5-8	9-12	9-12	5-8	5-8
Grade 6	8-12	5-7	8-12	8-12	6-9	6-9
Grade 7	8-12	5-7	8-12	8-12	6-9	6-9
Grade 8	8-12	5-7	8-16	8-12	7-9	7-9
Grade 10	8-12	5-7	7-10	8-12	7-10	7-10

OAKS Online Reading tests consist of 50 items (40 for grade 3 and 45 for grade 4).

2.2.2.2 Mathematics

Mathematics tests are available for grades 3–8 and 10 and consist of 40 multiple-choice items. Oregon’s revised standards directly align with the *Principles and Standards for School Mathematics*, published by the National Council of Teachers of Mathematics (NCTM) in 2000.¹¹ Mathematics content standards were most recently adopted by the State Board of Education in April 2002. Based on the standards, the tests are designed to assess mathematics skills overall and in the following five SRCs:

- **Calculation and estimation:** Compute with whole numbers, fractions, decimals, and integers using paper and pencil, calculators, and computers. Use estimation to solve problems and check the accuracy of solutions. Apply number theories, mathematical rules, and algorithms to solve problems.
- **Measurement:** Determine appropriate units, tools, and techniques to measure with direct (e.g., U.S. customary, metric) and indirect (e.g., formulas, estimates) methods.
- **Statistics and probability:** Determine the probability that an event will occur. Use measures of central tendency and variability, such as mean, median, and mode. Use statistics to summarize data, draw inferences, and make predictions.
- **Algebraic relationships:** Use mathematical expressions and algebraic operations to solve equations. Represent patterns and mathematical relationships using symbols, graphs, numbers, and words.

¹¹ www.nctm.org

- **Geometry:** Recognize, represent, and use geometric figures and their properties. Use given assumptions to determine properties of geometric figures, solve problems, and prove or justify relationships between figures.

Table 6 shows the SRCs for each of the grade levels and the range of test questions on the test at each grade level that assess that category.

Table 6.
Targeted Number of Items on Each Test by SRC for Mathematics

Math	Score Reporting Category (Strand)				
	Calculations and estimations	Measurement	Statistics and probability	Algebraic relationships	Geometry
	SRC 1	SRC 2	SRC 3	SRC 4	SRC 5
Grade 3	9-12	7-9	5-8	7-9	7-9
Grade 4	7-9	7-9	7-9	7-9	7-9
Grade 5	7-9	7-9	7-9	7-9	7-9
Grade 6	5-8	7-9	7-9	9-12	7-9
Grade 7	5-8	5-8	7-9	11-13	7-9
Grade 8	5-8	5-8	7-9	11-14	7-9
High School	3-6	3-6	7-10	13-16	9-12

OAKS Online Math tests consist of 40 items.

The Math tests rely heavily on graphic representations of core concepts and often have multiple items associated with a single stimulus. Math tests are administered on paper-and-pencil in English, Spanish and Russian. Online tests are delivered in English and Spanish.

2.2.2.3 Science

Science tests are available for grades 5, 8, and 10 and consist of 40 multiple-choice items. As with Math, the Science tests rely heavily on graphic representations of core concepts and often have multiple items associated with a single stimulus. Science tests are administered on paper-and-pencil in English, Spanish and Russian. Online tests are delivered in English and Spanish.

The science specifications reflect the skill expectations outlined in the science content standards, which were most recently adopted in April 2001 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 to align with the National Science Standards¹² and the Benchmarks for Science Literacy.¹³ Oregon's Science tests are designed to assess literacy skills overall and in three SRCs:

¹² National Research Council, www.nas.edu

¹³ American Association for the Advancement of Science, www.project2016.org

- **Physical science:** Understand structures and properties of matter and changes that occur in the physical world.
- **Life science:** Students understand the structure, functions, and interactions of living organisms and the environment, including organisms, heredity, diversity, and interdependence.
- **Earth and space science:** Physical properties of the earth, how those properties change, and Earth’s relationship to other celestial bodies, including the dynamic Earth, Earth in space, and the universe.

Table 7 shows the SRCs for each of the grade levels and the range of test questions on the test at each grade level that assess that category.

Table 7.
Targeted Number of Items on Each Test by SRC for Science

Science	Score Reporting Category (strand)		
	Physical science	Life science	Earth and space science
	SRC 1	SRC 2	SRC 3
Grade 3	N/A	N/A	N/A
Grade 4	N/A	N/A	N/A
Grade 5	12-14	12-14	12-14
Grade 6	N/A	N/A	N/A
Grade 7	N/A	N/A	N/A
Grade 8	12-14	12-14	12-14
High School	12-14	12-14	12-14

Science assessment is administered at grades 5, 8, and High School only. An additional SRC, Scientific Inquiry, is assessed through classroom work samples. Science tests consist of 40 items.

2.2.2.4 Social Sciences

Social Science tests are not required by the state, but are necessary for students seeking a Social Science Area Endorsement in Oregon. Social Science tests are available for grades 5, 8, and 10 and consist of 50 multiple-choice items. The content of the Social Science tests reflects the skill expectations that were outlined in the content standards and were adopted by the State Board of Education in April 2001. The tests are available in English and Spanish only; they rely heavily on graphic representations of core concepts and often have multiple items associated with a single stimulus. They measure knowledge across six strands, or Score Reporting Categories.

- **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.
- **Civics and Government:** Understand and apply knowledge about governmental and political systems and the rights and responsibilities of citizens.
- **Geography:** Understand and use geographic skills and concepts to interpret contemporary and historical issues.
- **Economics:** Understand economic concepts and principles and how available resources are allocated in a market economy.
- **Historical Skills:** Design and implement strategies to analyze issues, explain perspectives, and resolve issues using the social sciences.

Table 8.
Targeted Number of Items on Each Test by SRC for Social Sciences

Social Sciences	Score Reporting Category (Strand)					
	U.S. History	World History	Civics and Government	Economics	Geography	Historical Skills
	SRC 1	SRC 2	SRC 3	SRC 4	SRC 5	SRC 7
Grade 3	N/A	N/A	N/A	N/A	N/A	N/A
Grade 4	N/A	N/A	N/A	N/A	N/A	N/A
Grade 5	6-9	N/A	11-14	11-14	11-14	4-6
Grade 6	N/A	N/A	N/A	N/A	N/A	N/A
Grade 7	N/A	N/A	N/A	N/A	N/A	N/A
Grade 8	6-9	6-9	9-11	9-11	9-11	4-6
High Schl	6-9	6-9	9-11	9-11	9-11	4-6

Social Science assessments are administered at grades 5, 8, and High School only. Tests consist of 50 items.

2.3 Paper-and-Pencil Knowledge and Skills Tests

When statewide OAKS Online testing became mandatory in 2006–07, approximately 100 schools requested waivers to continue paper-and-pencil tests. In 2009-10, only 639 paper-and-pencil tests were given (0.07% of 1.43 million knowledge and skills tests). Paper and pencil test forms are made available to individual students who may need them as an accommodation (e.g., large print version). The availability of two accommodations in OAKS Online, variable screen magnification and the ability to print reading passages, beginning in 2008-09, was a major factor in reducing the necessity for paper-and-pencil tests.

2.4 Side-by-Side (Russian and Spanish) Assessments for English Language Learners

Side-by-side assessments allow students to take the KSAs or Performance Assessments in Spanish and English or Russian and English. A Spanish/English Side-by-Side Writing Performance Assessment is also available. Responses in languages other than Spanish or Russian are modifications and are not included in federal participation or performance calculations. Tests are constructed so that each item is presented in English on the left side and the second-language translation on the right. Designed to minimize English language barriers while testing English Language Learners (ELL: includes Limited English Proficient [LEP], Non-English Proficient [NEP], and English as a Second Language [ESL] students) on Oregon's content standards, the side-by-side tests allow these students the opportunity to demonstrate mastery of academic skills while learning English.

Teachers and instructional teams determine whether side-by-side tests provide the best opportunity for students to demonstrate mastery of the content standards. Side-by-Side tests are translations of the paper-and-pencil form administered each year. Current assessment information for English Language Learners (ELL) can be found at the assessment ELL Web site: <http://www.ode.state.or.us/search/page/?=1225>.

In 2009-10, Side-by-Side assessments were available for the following:

- Mathematics Knowledge and Skills Test in Spanish/English via OAKS Online and paper and pencil
- Mathematics Knowledge and Skills Test in Russian/English via paper and pencil
- Reading/Literature (in Grade 3 only) Knowledge and Skills Test in Spanish via paper and pencil (Modified administration if student is not eligible)
- Science Knowledge and Skills Test in Spanish/English (via OAKS Online and paper and pencil) and Russian/English (paper and pencil only)
- Social Science Knowledge and Skills Test in Spanish/English via OAKS Online
- Writing Test in Spanish/English (Modified administration if response is in Spanish and student is not eligible)

Translations are made by a state contractor employing expert translators knowledgeable in the subject area of the test. Tests are iteratively reviewed and revised, using forward translation. Once the translators are finished, drafts are submitted to ODE for review. ODE convenes a panel of 20 to 25 dual-language teachers from around the state to review the translation, line by line and item by item, for accuracy and consistency. The translation company revises the translated tests according to the panel's input and recommendations and then sends the tests to print.

Translated items are examined for differential item functioning (DIF) after test administration to monitor unintended effects of presenting the item in the dual-language format. For details of the translation DIF studies, see “Comparability of English and Spanish/English Science Tests”.¹⁴

2.5 American Sign Language and Braille Assessments

Braille, large-print, and American Sign Language (ASL) Knowledge and Skills tests are available in the same form as the paper-and-pencil test. Signed tests are provided by districts, but because of the inability to ensure consistency and quality control, signed tests are considered modified administrations

2.6 Performance Assessments

The Writing Performance Assessment is an open-ended test administered at grades 4, 7, and 10. Students write an essay in response to one of three writing prompts. The assessment is not timed, but students generally complete the assessment within three class periods (approximately 120 minutes). The following practice is recommended:

Day 1: Complete student information; prewrite and begin the rough draft

Day 2: Continue draft and begin revising and editing

Day 3: Revise and edit; copy into writing folder

The traits scored include the following:

Ideas and content: Present clear, complete, and well-developed ideas

Organization: Effectively introduce, structure, and conclude the ideas presented and the writing as a whole

Sentence fluency: Write smooth and easy-to-read sentences

Conventions: Proofread; correct spelling, punctuation, and grammatical errors

Voice: Integrate honest thoughts, feelings, and personality into the writing

Word choice: Use accurate and interesting words to make the writing easy to understand

Voice and word choice are not tied to the state performance standard but are valuable components of writing and are scored to provide students with feedback. Writing Assessments are not machine scored but evaluated by committees of trained raters according to grade-level standards.

2.8 Alternate (Extended) Assessments for Students with Disabilities (SWD)

The Alternate Assessments comprise tasks designed to measure basic academic skills that are preparatory to meeting Oregon’s Benchmark 1 Content Standards.

Oregon Department of

Students in Oregon with disabilities who are unable to take the Knowledge and Skills tests, even with accommodations, take the Extended Assessments. Consistent with NCLB legislation requiring that all students in Oregon schools be tested, the Alternate Assessments are developed for assessing the progress of students with profound cognitive disabilities. There are four Extended Assessments, one each for Reading, Writing, Mathematics, and Science. The Extended

Assessments measure emerging skills and performance as a prelude to the Benchmark 1 assessment.

Two separate volumes describe the Alternate Assessments system, *Volume 7, Alternate Assessment, Program Description* provides a detailed description of the assessments and *Volume 8, Alternate Assessment, 2007-08 Statistical Summary* provides participation and results summaries. These are available on the Department's Web site at <http://www.ode.state.or.us/search/page/?id=787>. A summary of the Extended Assessment Program is provided below.

The Extended Assessments evaluate whether students with significant disabilities are learning to read, write, and manipulate numbers. Each assessment consists of between 10 and 22 tasks, and each task is measured by between one and eight items varying in difficulty. Many of the tasks involve the use of flashcards, from which students identify or order information. Other tasks require students to copy information (words, numbers or letters) or to produce text (write their name, write a story) or solve math problems.

3. Test Administration

3.1 Testing Windows

The OAKS Online testing window is mid-October through mid-May of each year. Typically, districts and schools determine when students test, although with OAKS Online, testing windows can be determined for individual students to best accommodate readiness.

Table 9.
2009-10 Testing Window for Grades 3–8 and High School

Assessment	Testing Window
OAKS Online Reading/Literature and Math KSA	10/7–5/19
OAKS Online Science and Social Science KSA	10/7–5/19
Writing Performance Assessments	1/11–2/26, 4/12–4/30
Paper-and-pencil KSA –Reading/Literature, Math and Science	3/8–4/9
Extended Assessments – Reading, Math and Science	2/18–4/28
Extended Assessments – Writing	2/18–4/28
English Language Proficiency Assessment (ELPA)	1/21–5/19

OAKS = Oregon Assessment of Knowledge and Skills, KSA = Knowledge and Skills Assessment

3.2 Testing sessions

Although OAKS Online testing is untimed, most students complete the math test in a single 50- to 60-minute session, reading and science in two 50-minute sessions, and social sciences in two 40- to 50-minute sessions. Writing is usually administered over three sessions across three days. Typically, districts and schools determine the number and timing of testing sessions although, with the OAKS Online, testing windows may be determined to best accommodate individual student needs and classroom schedules, and the number of sessions and time of test may vary. Generally, students are allowed to continue the test as long as they continue to progress.

3.3 Operational Procedures

Test Administration Manuals provide detailed information about the testing environment, procedures, security, and instructions to students. This information is described in detail in *Volume 5: Test Administration* available at <http://www.ode.state.or.us/search/page/?id=787> and in test Administration Manuals available on the ODE Web site at <http://www.ode.state.or.us/search/page/?=486>.

3.3.1 Standardized Testing Procedures

Test administrators follow testing procedures as outlined in the Administration Manual available for each test. Test administrators are required to review the Administration Manual prior to the

beginning of testing and verify that they have the correct quantity of material, ensure that the testing setting is prepared for testing (e.g., removing certain classroom posters, arranging desks, etc.) and establishing makeup procedures for any students who are absent on the day(s) of testing.

The manuals contain administration instructions and scripts in English, Spanish, and Russian for administrators to follow, ensuring standardized administration conditions. Teachers follow the instructions and read the boxed scripts verbatim to students. The manuals also include instructions for properly collecting and storing student test materials between sessions to ensure security and provide a troubleshooting guide for the OAKS Online administration.

3.3.2 Test Security

Oregon's test materials are secure documents and NO PART of the actual test materials may be reproduced in any way prior to, during, or after testing, nor may any copies of the materials be kept in the school building or district office after the test administration is completed (although copies may be made of the student response portion of the state writing assessment). Descriptions of individual test items MUST NOT be shared or released.

The security of tests and student information for OAKS Online was maintained by the following:

- Administering tests over secure Internet browsers
- Requiring user passwords to access the system
- Limiting users to appropriate levels within the system (proctors have access to the test delivery system, teachers have access to their students' scores, etc.)
- Authenticating students logging-in to a testing session
- Utilizing security codes/test keys, changed daily to access testing sessions

ODE's testing contractors and application hosts are required to offer additional security by the following:

- Maintaining data hosting in a physically secured, remote environment with limited key-card access only to senior, authorized personnel
- Thoroughly testing system security and formally auditing security
- Training all staff in security procedures
- Encrypting administrative system functions
- Requiring secure passwords for administrative and user log-in to the system

OAKS Online test administration reduces the burden on schools of having to handle test booklets and eliminates the storage/security problems inherent in large shipments of secure test materials.

Prior to beginning preparations for either the printed or computer-based assessments, District Test Coordinators, School Test Coordinators (often, but not necessarily principals), and Test Administrators are required to read and sign the Assurance of Test Security acknowledging that all

materials will be kept secure and no inappropriate assistance will be provided to any student. Signed forms are given to the School Test Coordinators. A blank copy of the Assurance of Test Security is included in the Administration Manual, and School Coordinators have additional copies. In addition, District Test Coordinators, School Test Coordinators, Test Administrators, and principals are encouraged to review the Test Security PowerPoint and attend a test administration training session.

Beginning in 2006-07 ODE released a test administrator handbook and training to increase test security.

3.3.3 Confidentiality of Student Records

Confidentiality of student records will be protected in compliance with the Family Educational Rights and Privacy Act, 34 CFR & 99 et. Seq., and Oregon Administrative Rules relating to student records. The relevant Oregon Administrative Rules addressing the rights of parents and students and the confidentiality of student records include the following:

- 581-021-0230: The Rights of Parents
- 581-021-0270: Rights of Inspection and Review of Education Records
- 581-021-0330: Prior Consent to Disclose Information
- 581-021-0400: Recordkeeping Requirements
- 581-021-0250: An Educational Agency or Institution's Policy Regarding Student Education Records

All are available at

http://arcweb.sos.state.or.us/rules/OARS_500/OAR_581/581_tofc.html.

3.4 Testing Accommodations and Modifications

Accommodations
Panel members
research and
review all testing
accommodations
and modifications

Oregon's assessments system is designed to be flexible and inclusive of all students. To measure the achievement of students with diverse learning needs relative to state standards, the following testing adaptations are provided.

Accommodations are changes in how a test is presented to the student or how the student responds that do not change the content and/or performance standards of what is being measured by the test. The changes are made in

order to provide a student with equal access to learning and equal opportunity to demonstrate what is known. All students are eligible to use accommodations. The use of accommodations is considered standard administration.

Modifications are alterations in test presentation or in the response format or substance that substantially change the content and/or performance standards of what is being measured by the test. Students' scores will count as not meeting the performance standard. Beginning in 2006-07 students only testing with modifications will be counted as nonparticipants.

3.4.1 Accommodations

Common accommodations include the following:

- Braille
- Large print
- Extended time
- Multiple testing sessions
- Frequent breaks
- Reading or signing directions
- Providing written versions of oral directions
- Dual-language tests

3.4.2 Modifications

Common modifications include the following:

- Limiting the number of multiple-choice answer options
- Thesaurus or dictionary
- Translations other than what the state provides (Spanish/Russian)
- Rewording questions to simplify the item
- Reading aloud or signing the reading/literature passages, items, and distractors to the student
- Providing synonyms or definitions upon request

ODE does not identify all possible modifications. Rather, it supplies districts with exemplars of changes to the administration of the test that results in a score that is not comparable with the scores of other students.

3.4.3 Testing Accommodation and Modification Review Process

Oregon continually assesses the needs of its students and addresses those needs as they arise. Accommodations are recommended, evaluated, and made available on an ongoing basis by ODE through a formal review process, involving an Accommodations Panel. The accommodations panel is comprised of 8-10 educators from around the state who are well versed in how to assess students with disabilities and to provide access to the assessment for all students without changing the nature of the construct. The Accommodations Panel uses a combination of policy, judgments, and research to ensure that accommodations are fair and provide valid scores that enable all students to demonstrate what they know and can do.

Teachers and administrators are encouraged to nominate accommodations or modifications appropriate for meeting their students' needs for review and approval through the following process:

1. The request to add an adaptation to the accommodations table in the OSAS test Administration Manual may be brought forward by any representative of the school district (e.g., teacher, school person involved in the testing process, or administrator). A parent wanting such consideration must go through the local school team to submit a request
 - a. For students with disabilities on IEPs, the decision is made by the IEP team, and one person is selected to advance the request.
 - b. For students with disabilities on 504 plans, the decision is made by the 504 team, and one person is selected to advance the request.
 - c. For all other students, the request should be advanced through some coordinated effort at the school level.
2. A form on which to present the request is available on the assessment web page or from the Office of Assessment and Evaluation (OAE).
3. The request is sent to ODE, OAE attention: Accommodations Panel. Within three working days an acknowledgment is sent to the requestor indicating that the request has been received and, if necessary, that further information is needed.
4. Upon receipt of all necessary information regarding the request, staff from ODE, OAE complete research on the item and summarize the issues relative to adding the item as an accommodation to the test Administration Manual, including the accuracy and validity of inferences made from scores based on the requested accommodation.
 - a. This information is prepared for the Accommodations Panel.
 - b. A copy of the information is also sent to the requestor with a request to offer any further information to the OAE if deemed appropriate or necessary by the requestor within three working days.
 - c. If the item requested for consideration is on the current list of accommodations in the test Administration Manual, the requestor is notified that this is the case, and no further action is necessary.
5. The summarized information and research are provided to the Accommodations Panel for review and consideration. The Accommodations Panel uses preset criteria to determine its recommendation.
6. If the Panel feels there is insufficient information to make a decision, staff from ODE, OAE gather further information as requested by the Panel.
7. If the Panel has enough information to review the request against the preset criteria, the Panel documents the findings and makes a recommendation.

8. The recommendation from the Panel is forwarded immediately to the Associate Superintendent for Assessment and Evaluation.
9. The Associate Superintendent for Assessment and Evaluation makes a decision about the requested item through a process that includes the State Board of Education.
10. As soon as the decision is made, the requestor, the Accommodations Panel, and the field are informed of the status of the requested item.
11. Staff from ODE, OAE develop a compendium of decisions and inform the districts of the change (if any) and place the new accommodation (if appropriate) in the next test Administration Manual.

Current descriptions of accommodations and modifications are provided in annual test Administration Manuals. For a complete list of current accommodations and modifications, please see the Department's Web site at <http://www.ode.state.or.us/search/page/?=487>.

3.5 Student Testing

OAKS Online is the online component of Oregon’s assessment system. Students access the tests using a special secure browser that locks down the student’s machine, preventing printing, access to the Internet, copying, and pasting, and screenshots, thus ensuring the security of the testing environment.

3.5.1 Student Testing Tools

During testing, students have access to many tools that enhance accessibility and help students with the testing experience. Students are able to:

- Navigate back in tests to change answers to previously answered items (except after pausing the test 20 minutes or more)
- Mark answers they are unsure of and return to review them later. Marked items are displayed in a drop-down menu to facilitate navigation.
- Strikethrough incorrect responses to remove them from consideration
- Highlight text on the computer screen
- Pause tests and return to complete them in another session
- Access a formula page for Math tests, a Periodic Table for Science tests, and a Guide to Revision for Online Writing tests
- Comment on test items. Comments are reported directly to ODE and are associated with individual items.
- Use “Go to” navigation to review and return to previously seen items
- Access an online help system, available to students at any time during the test, that provides assistance to students using the available testing tools
- Print paper copies of reading passages and items
- Enlarge the display size of online text and graphics
- Select a blue or yellow background color
- Access a basic calculator for grade 3 and 4 Math and grade 5 Math and Science; a scientific calculator for grades 6 and 7 Math; and scientific/graphing calculator for grade 8 and high school Math and Science.

Answers to all items can be changed, although students will be unable to review items following any test break of 20 minutes or more, even if they are marked for review when the break begins.¹⁵ Test items cannot be skipped, and students must finish presented items (including all items associated with a stimulus) before moving on to the next item. To allow students to move through the test at different speeds within a specified time period, tests may be paused in the middle of item sets.

Student tests automatically close after 45 minutes of inactivity and expire 45 days after test start. Additionally, online sample tests in Math, Reading, Science, and Social Sciences are available that ensure that students have the opportunity to see representative content and to practice using the tools.

¹⁵ The exception to the 20 minute rule is made for the online writing assessment. Students may revise their essay after a pause exceeding 20 minutes.

3.5.2 Test Administrator Tools

Test Administrators approve students to test and then monitor students testing throughout the session via the Test Administration component of the system. This shows the progress of each student currently testing and allows TAs to monitor student progress and be alerted to any student who may need technical assistance, has paused the test for a break, or may need additional time to complete the test.

After tests are completed, scores are transferred to the OAKS Online Reporting System where teachers and administrators can monitor and analyze student test scores and quickly apply testing data to student instruction.

Test Administrators are provided with detailed User Guides for both the Test Administration and Student Testing components of OAKS Online and are supported by local (Educational Service Districts) and vendor (American Institutes for Research) helpdesks.

4. MAINTENANCE OF THE ITEM BANK

Oregon develops and tests new items annually. As items are used and become over-exposed or outdated, they are replaced with new items. Field tests provide ODE with psychometric information about newly developed items (item difficulty, discrimination, whether the items perform as expected for all groups of students, etc.). This information about each item is used to make judgments about which items are placed on future operational tests.

4.1 Field Testing

ODE uses embedded field tests to maintain a sufficient number of items in the item bank. For Mathematics, Reading/Literature, Science and Social Science, each operational test includes 5 to 6 field-test items. Table XX below shows by grade and subject, the number of field test items included during the 2009-10 administrations.

ODE also field tests to maintain a sufficient number of writing prompts for the multiple operational test forms. Approximately 15 writing prompts are field tested each fall for each grade.

	Math	Reading	Science	Social Studies
Grade 3	381	108		
Grade 4	395	118		
Grade 5	311	119	204	60
Grade 6	616	118		
Grade 7	428	117		
Grade 8	478	122	203	60
Grade 10	516	139	218	57

4.1.1 Field-Test Item Insertion

Field-test items are inserted in each student’s test randomly, within a specified range of locations on each test. The insertion points for each test are described below.

Table 10. Number of Items and FT Item Insertion Points

	Operational	Field Test	Total test length w/o FT	Total test length w/FT	FT insertion point range
Math	40 items	5	40	45	>5 & <35
Reading grade 3	40	5-7	40	45-47	>5 & <30
Reading grade 4	45	6-8	45	51-53	>5 & <34
Reading grade 5	50	6-8	50	56-58	>5 & <38
Reading grade 6	50	6-8	50	56-58	>5 & <38
Reading grade 7	50	6-8	50	56-58	>5 & <38
Reading grade 8	50	6-7	50	56-57	>5 & <38
Reading grade 10	50	6-8	50	56-58	>5 & <38
Science	40	6	40	46	>5 & <35
Social Science	50	6	50	56	>5 & <38

4.1.1 Field-Test Sample Selection

All operational multiple-choice tests contain embedded field-test items. Field-test Mathematics, Reading, Science and Social Science items are presented randomly at intervals throughout the test in the OAKS Online test. In all tests, field-test items are not included at the beginning or the end of the test. In reading tests, six field-test items are administered with 50 previously-calibrated items (5 and 40 for third grade; 6 and 45 for fourth grade). In mathematics, the ratio is 5 field-test items to 40 operational items that have previously been calibrated on Oregon’s mathematics scale.

Given the increased reliance on OAKS Online, field testing on paper-and-pencil has been discontinued. To complete the field test, items are assigned to item pools after completing the review cycle.¹⁶ Items that will fill anticipated gaps in future item pools are selected first. If there is additional space available for more field test items, they are selected based on anticipated level of quality. The embedded field test items are administered randomly within a grade, independent of student proficiency.

Performance assessment writing prompts are field tested in the fall in districts participating in the voluntary field testing. To be included in the field test, schools must be willing to administer the writing assessments in the fall according to standard administration requirements. Students in the next higher grade level are included in the study (i.e. students in 5th grade write to the 4th grade prompts). Schools are compensated for their participation on a per-student basis. The writing papers are scored by Oregon teachers employed by the state consistent with the process used in the operational test.

¹⁶ See Test Specifications and Blueprints for a description of the item review process.

Sample sizes for field-tested items range from 500 to 2000, depending on the item pool and the number of students tested using items from each item pool.

4.1.2 Field-Test Analysis

ODE content specialists and psychometricians use field-test data to examine new 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 (point-biserial, or “Pbis”). Items with 500 or more responses were included in the analysis.

The following criteria were used to determine which items moved from field-test to operational status and which were flagged for additional review:

- Point-biserials for the correct answer must be 0.20 or higher.
- P values (percent of students selecting each option)
 - For item keys, p -values less than 0.25 indicate that students selected the correct response less often than would be expected by chance, and these items are flagged. If these items also have high point-biserials, then the item discriminates and only most successful students get them correct. If the point-biserial is low, then the item does not discriminate and is flagged for review. P -values for item keys greater than .95 are too easy and are flagged for review.
 - Distractors are flagged if the p -value is higher than the key. Items may be unexpectedly difficult if a concept is not included in the curriculum.
- Correlation with total test score ranges from -1 to $+1$. Keys are flagged if < 0.20 ; distractors are flagged if > 0.05 . For the item key, $+$ is good, $-$ is bad (unacceptable). For distractors, $-$ is good, $+$ is bad (unacceptable). Large positive numbers indicate that high performing students are more likely to answer the item correctly than low performing students.
- Item Quality – Model Fit (Infit and Outfit). Items are flagged if Infit or Outfit statistics < 0.70 or > 1.30 . Outfit can indicate:
 - A difficult item that low performing students are getting correct
 - An easy item that high performing students are getting incorrect
- Infit can indicate:
 - Difficult items that are not correctly answered by high performing students
 - Easy items that are not correctly answered by low performing students

Finally, field-tested items undergo a differential item functioning (DIF) analysis, and items exhibiting significant DIF are referred back to content specialists for additional review. Oregon uses the Mantel-Haenszel χ^2 test for DIF. Groups include Female/Male, African American/White,

Hispanic/White, SPED/non-SPED, LEP/non-LEP, and Economically Disadvantaged/non-Economically Disadvantaged.

Fairness Statistics identify items that behave differently across groups. DIF analysis requires at least 400 responses per group and resulted in the following classifications:

- A = no statistical evidence of DIF
- B = potential statistical evidence of mild DIF
- C = potential statistical evidence of severe DIF

We use “+” and “-” to denote significant DIF. A “+” item favors the focal group (listed first in the list above), while a “-” item favors the referent group. Bias cannot be determined by the DIF analysis alone; in some cases there may be real differences between groups due to instructional differences. Operational items also are examined for DIF, particularly for focal groups with small numbers of students.

Although students in Oregon may take each test up to three times, items are not repeated across attempts, so field-test items from each attempt are included in analysis. The high school level test can be taken by students in grades 8 through 12; all student data from high school tests were included in the analyses regardless of grade.

The scaling of multiple-choice assessments is accomplished using the Rasch item response theory (IRT) model. The estimation of item difficulty parameters is performed in a concurrent calibration separately for paper-and-pencil and OAKS Online tests. The item difficulties from a concurrent calibration with the operational items are converted to the Rasch Unit, or RIT Scale, on which scores are reported. A summary of IRT, the Rasch model, and the RIT Scale is provided below.

4.1.2.1 Item Response Theory

Item response theory (Hambleton & Swaminathan, 1985; Lord & Novick, 1968) gets its name because the data elements are individual responses to test questions (items). In classical or true score theory (Allen & Yen, 1979), the central data elements are responses of entire populations to whole tests. IRT methods start with a matrix of individual examinee responses to individual test items as input. Estimates of examinee proficiency and item parameters, including item difficulty, are derived from this input. Because we can't directly measure mental qualities such as reading or math proficiency, we call them latent traits. The scale for the latent (unknown) trait is inferred from manifest (known) responses.

The primary advantage of item response theory over classical or true score testing theory is this stability of scale. Tests based on classical theory interpret scores in terms of rank, percentile, or distance from the mean. Item difficulties are expressed in terms of the percentage of students answering the item correctly. These are functions of the population tested. Under classical testing theory, if the population changes, estimates of proficiency and difficulty also change. Because items have no identity independent of the population or the test in which they are embedded, they cannot be combined with new items to form tests with known characteristics. Through application of IRT,

the existence of an independent scale and the ability to separate item characteristics from population characteristics are what makes adaptive testing possible.

Rasch models decompose an observed variable into a “true” and an orthogonal “measurement error” component. Suppose there are a total of J items available, with a difficulty parameter b_j for each item j . Let the i th person’s response to the j th item be z_{ij} . Furthermore, assume that θ_i represents examinee i ’s true proficiency on the latent trait, and $f(\theta)$ is the distribution associated with the latent trait; it is usually further assumed that $f(\theta)$ has finite moments. If item responses were continuous variables, they could be described by the linear model, $y_{ij} = \theta_i + e_{ij}$ where y_{ij} is person i ’s unobserved response to item j , and e_{ij} is the individual and item-specific measurement error in the response.

If item responses are binary measures, instead of y_{ij} , let $z_{ij} = \begin{cases} 1 & \text{if } y_{ij} > b_j \\ 0 & \text{otherwise} \end{cases}$, where b_j is a threshold along the true-score dimension. With this, the relationship between proficiency and item response can be stated as a probability,

$$\begin{aligned} p(z_{ij} = 1 | \theta_i) &= p(\theta_i + e_{ij} > b_j) \\ &= p(-e_{ij} < \theta_i - b_j). \end{aligned}$$

This relationship forms the basis of most models from item response theory (IRT), of which the Rasch model is a special case.

4.1.2.2 The Rasch Model

The simplest IRT model estimates only one item parameter, difficulty. One-parameter models are also called Rasch models after Danish mathematician Georg Rasch who developed it separately from the other IRT models (Rasch, 1960,¹⁷ Wright & Stone 1979¹⁸).

In the standard Rasch model, the distribution of the measurement error takes a logistic form,

$$p(z_{ij} = 1 | \theta_i) = \frac{1}{1 + e^{-(\theta_i - b_j)}}.$$

The Rasch model yields a monotonic, one-to-one correspondence between maximum likelihood estimates of scale scores for individual examinees and the number of raw “points” the examinee scores on a test. In other words, the Rasch model dictates that the probability of success on an item is dependent upon the difficulty of the item and the proficiency of examinees and assumes that items vary only in difficulty.

¹⁷ Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Copenhagen: Denmark's Paedagogiske Institut.

¹⁸ Wright, B.D., & Stone, M.H.. (1979). Best test Design. Chicago: MESA.

IRT parameter estimation procedures do not provide finite estimates of student proficiency for either a perfect score or a zero score without a certain form of approximation procedure. The approximated proficiency for a perfect score is set at the proficiency that corresponds to 0.3 less than a perfect score, whereas 0.3 is added to a zero score to estimate a corresponding proficiency measure.

Currently, the KSA fixed-form multiple-choice tests (paper-and-pencil and Side-by-Side tests) are pre-calibrated and use known parameters generated from the Rasch model. The estimation of item difficulty parameters was performed by BIGSTEPS (Linacre & Wright, 1998),¹⁹ and the fixed form administered each year is constructed of items with known difficulty estimates. Item difficulties obtained from previous concurrent calibrations are used to generate a raw score to RIT scale score conversion table.

Computer adaptive testing (CAT) must maintain the rigor of accepted equating and scaling procedures in use for fixed forms, although methods for achieving this differ somewhat from those for paper-and-pencil tests. In an adaptive testing session, all items used for computing the score have established parameters. Field-test items are embedded in operational tests, and responses for the new items accumulate until there are enough responses over an appropriate proficiency range to estimate item parameters. At this point the item's performance is evaluated, and if it is statistically sound, it becomes part of the item bank.

4.1.2.3 The RIT Scale

Rasch calibration includes standardization of the item difficulties (mean logit set to zero) and a bias correction (Wright & Stone, 1979). All item parameters can be converted back to logits by subtracting 200 and dividing by 10.

4.1.3 Linking New Items and Tests to the RIT Scale

Fixed forms (paper-and-pencil tests) use a non-equivalent groups anchor design to link student scores back to the scale. OAKS Online tests use items with known parameters that have demonstrated consistent difficulty on at least two operational assessments. Because these item parameters are already linked to the scale, scores derived from student responses to these items are also linked. New items are calibrated to the scale, using operational items in the same item bank as the anchors.

4.2 Assignment of Calibrated Items to Item Pools

Field tests provide ODE with psychometric information about newly developed items (item difficulty, discrimination, whether the items perform as expected for all groups of students, etc.). This information is used to determine which items are administered on future operational tests.

Embedded multiple-choice field-test items are flagged for possible exclusion if they exhibit any of the following:

¹⁹ www.winsteps.com/bigsteps.htm.

- Low test-item correlation (point-biserial less than .2)
- If the percent of students correctly answering the item (p-value) is less than chance (25%) or too low or too high for the targeted grade (RIT difficulty) level
- Extreme INFIT, OUTFIT, or DIF statistics
- Significant departure from ideal item-characteristic curves

For the Writing test, the score distributions of field-tested Writing task prompts are compared to previously used operational prompts. Prompts producing similar score distributions are retained for operational use.

Following review of the field-test results, items that are well-fitting and psychometrically sound are combined with previously administered items and assembled into equivalent test forms and item pools. This process is described in Section 5 below.

4.3 Item Release and Retirement

Approximately every three years, ODE releases one sample test for each content area and grade level, comprising 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 practice tests. Practice tests for Reading/Literature, Math, Social Science, and Science are available on ODE's Web site at

<http://www.ode.state.or.us/search/page/?=1222>.

Sample Writing prompts are also available at <http://www.ode.state.or.us/search/page/?id=443>.

5. FORM AND TEST POOL DEVELOPMENT

5.1 Form Development, Paper-and-Pencil Tests

Since 2005–06 when testing via OAKS Online became mandatory, a single paper-and-pencil form is created for each grade and subject. Paper-and-pencil forms are available only to students needing them to demonstrate their mastery of the content standards. Every couple of years, new forms are developed for paper-and-pencil administration.

ODE staff assign items to test forms according to each subject and grade test specifications. In addition to balancing content and graphics, ODE considers the statistical properties of field-tested and previously administered operational items in forms development, including the following:

- The range of item difficulty, including a minimum, maximum, and average difficulty across all items on the form
- The weighting or distribution of items across SRC as prescribed by the test specifications

- The balance of item graphics types (charts, spinners, box-and-whisker plots, etc.). Most items in some content areas contain graphics. For example, mathematics items frequently contain charts, spinners, box-and-whisker plots, line or parabolic graphics, thermometers, clocks, stem-and-leaf plots, geometric shapes, and dice. ODE ensures a balance of these types of graphics across forms (i.e., to see that a test does not include probability items that all use a spinner or math problems that all involve telling time and use clocks).

The OAKS item selection algorithm ensures that students do not receive the same item more than once across all three available testing opportunities.

New forms are compared to previous years and are created to be equivalent across year.

5.2 Item Pool Construction, OAKS Online

For OAKS Online, each grade has a single pool of items. Each grade level item pool is developed following the procedures for paper-and-pencil forms construction. Items in each pool are specifically selected to represent the content and range of difficulty described by the grade-level content standards, for students at all performance levels.

5.2.1 Item Selection Algorithms

For OAKS Online, item pool construction and item selection rules ensure that each student receives a test representing an adequate sample of the domain with appropriate difficulty. The algorithm maximizes the information for each student and allows for certain constraints to be set, ensuring that items selected represent the required content distribution..

The test delivery system ensures that students are not exposed to the same items in subsequent tests, should they participate in multiple testing opportunities. The starting point for each test is selected to be near the average ability of students at that grade.

In OAKS Online, the accuracy of the student responses to items determines the next block of items and passage that the student will see. Thus, each student is presented with a set of items that most accurately aligns with his or her proficiency level based on grade level content. The adaptive delivery system gives test developers the ability to develop tests that meet 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
6. Other constraints to control test composition (e.g., number of “male” names, number of “female” names, number of situational items, types of graphics, etc.)

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 11 summarizes the test development rules and the specifications for each and describes how OAKS Online (the engine) implements the rules.

Table 11.
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.

Each grade level item pool is designed to support up to three testing opportunities. 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 stand (SRC) and coverage of the specified breadth and depth.

The OAKS Online item pools each contain approximately 500 items per grade and subject, a sufficient number to ensure that students administered more or less difficult tests are provided items representing the breadth and depth identified in the test specifications and content standards. Because the test adapts to each student’s performance, while maintaining accurate representation of the required grade level knowledge and skills in content breadth and depth, the OAKS Online results provide precise estimates of each student’s true achievement level across the range of proficiency.

OAKS Online tests are administered under constraints to ensure accurate representation of grade-level content standards and the specified distribution of items in each Score Reporting Category

Because OAKS Online is adaptive in nature, a unique test form is constructed for each student participating in the test. Conceptually, it is then possible that there are as many unique test forms as there are students taking the test, although this does not occur in practice.

5.2.2 OAKS Online Adaptive Algorithm Overview

Objectives

The OAKS Online adaptive algorithm controls how tests start, how items are selected, and how tests are ended. The algorithm selects items to administer on each student’s test to meet three goals: (1) that tests are administered according to the test specifications; (2) that tests best assess examinees’ proficiency in each content strand; and (3) that tests are administered with items targeted to a student’s ability and are not too easy or too difficult.

Starting a Test

The algorithm starts each test with an item of average difficulty for the specific subject and grade . Subsequent items are selected for administration by the algorithm based on student responses.

Item Selection

After the initial item is administered, the algorithm identifies the best item to administer using the following criteria:

1. Match to the blueprint. The algorithm first selects items to maximize fit to the test blueprint. Blueprints specify a range of items to be administered in each strand for each test, and during item selection, the algorithm “rewards” strands that have not yet reached the minimum number of items and then “rewards” strands that have not yet reached the maximum number of items. Item sets (groups of items associated to a graphic or passage) are selected based on an average fit (weighting) based on the items comprising the set.
2. Increased precision. After identifying eligible items that meet the blueprint, the algorithm selects items that maximize the precision with which proficiency is assessed for each strand

by selecting the best fitting item from the available items within the targeted strand. For item sets, this step uses the average difficulty of items for each strand within the set.

3. Match to student ability. From the items identified in step 2, the algorithm selects the item or item set that best matches the student's estimated ability. Higher performance (answering items correctly) is followed by more difficult items, and lower performance (answering items incorrectly) is followed by less difficult items.

The algorithm allows previously answered items to be changed; it does not allow items to be skipped. Item selection requires iteratively updating the estimate of the overall and strand ability estimates after each item is answered. When a previously answered item is changed, the proficiency estimate is adjusted to account for the changed responses when the next new item is selected.

While the update of the ability estimates is performed at each iteration, the overall and strand scores are recalculated using all data at the end of the test for the final score.

Items administered on previous tests are blocked from the item selection algorithm, ensuring that the student is not exposed to any item more than once. In all but rare cases relating to item pool composition, the algorithm prevents the selection of an item that the student has seen before.

Ending the Test

The algorithm stops administering items when the specified test length is met. If the end of the test is reached on an item set, the item set is truncated.

Scoring the Test

After each response is submitted, the algorithm recalculates a score. As more answers are provided, the estimate becomes more precise and the difficulty of the items selected for administration more closely aligns to the student's ability level. When the test is completed, the algorithm scores the overall test and each content strand.

After the last item has been answered on the test, the score is displayed in the session component of the Test Administrator screen, cutscores are applied, and the student's scores, overall and strand, are transferred to the online reporting system.

5.3 Test Information Curves and Conditional Standard Error of Measurement

The usefulness of proficiency estimates depends upon the standard error of measurement. Smaller standard errors imply more precise and interpretable results than larger standard errors. Unlike classical test theory in which the standard error of measurement is assumed to be the same for all respondents, IRT assumes different standard errors for different proficiency estimates.

Computer adaptive testing goes one step further than what is typically possible with paper and pencil testing. With paper and pencil tests, often referred to as *fixed forms*, all students see the same test items. Hence, the items may not be well-suited for very low or for very high performing

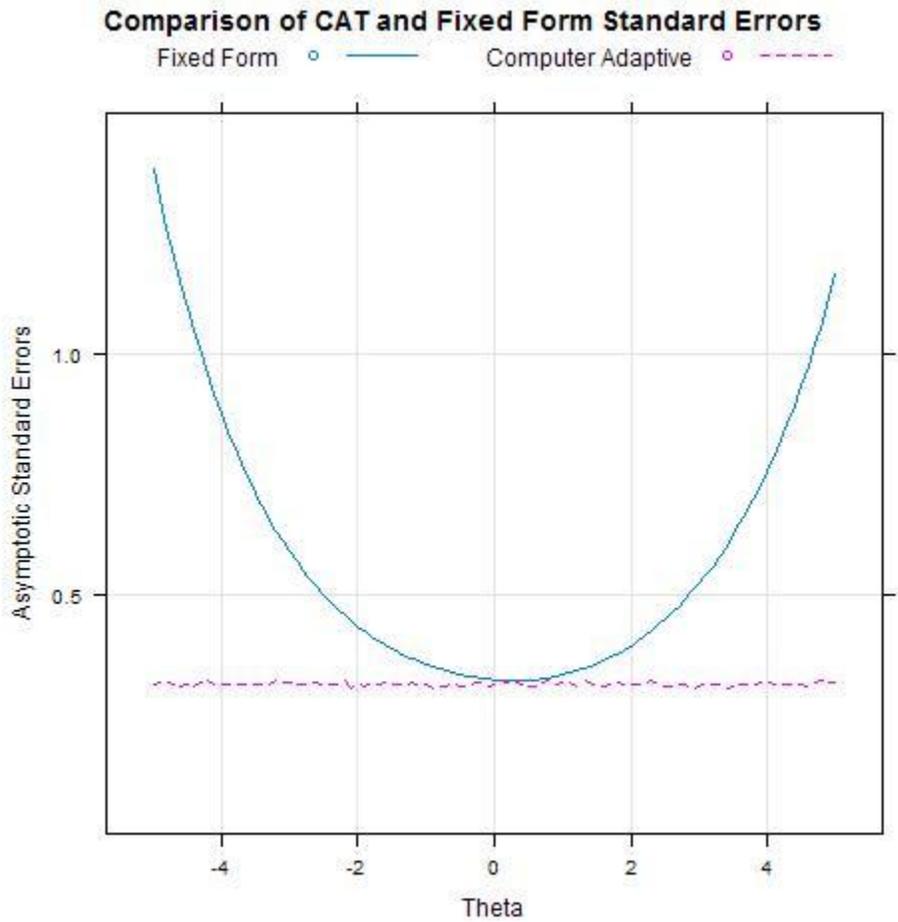
students. The consequence of this mismatch between a student's true ability level and the items they are administered is a lack of measurement precision for some students.

In contrast, computer adaptive test aim to administer each student test items that are best aligned with their true level of ability. As a result, scores for students from an adaptive test are measured with better precision at (virtually) all points along the ability scale.

The following figure illustrates this scenario using a hypothetical 50 item test. The standard error curve for the fixed form, which is the inverse of the test information function at each level of ability, forms a "U" shape, thus denoting that standard errors at low and high levels of ability are higher than at other points in the middle of the ability scale. In other words, a fixed form provides decent measurement properties for some students, but not for all students.

The standard errors from a computer adaptive test, also consisting of 50 items, are relatively flat at all points along the ability scale. This denotes that the measurement properties of the test are the same for all students, not only those in the middle of the score distribution. In addition, standard errors from a CAT test tend to be lower than what is typically observed in a fixed form assessment given that items are intentionally matched with a student's level of ability.

This example illustrates a best-case scenario when a sufficient number of items exist in the pool for each of the reporting categories, such that items can be drawn to match test blueprints as well as a student's level of ability.



Conditional standard errors of measurement at each achievement standard are shown in the following tables for the 2009-10 mathematics, reading and science assessments.

Conditional Standard Error of Measurement by Achievement Standard Oregon Mathematics Assessment, 2009-10						
	Nearly Meets		Meets		Exceeds	
Grade	Standard	Conditional SEM	Standard	Conditional SEM	Standard	Conditional SEM
3	201	3.35	205	3.37	217	3.61
4	208	3.36	212	3.36	225	3.54
5	214	3.35	218	3.34	229	3.45
6	216	3.38	221	3.34	232	3.43
7	221	3.38	226	3.33	238	3.42
8	225	3.35	230	3.32	241	3.43
10	231	3.37	236	3.33	246	3.40

Conditional Standard Error of Measurement by Achievement Standard Oregon Reading Assessment, 2009-10						
Grade	Nearly Meets		Meets		Exceeds	
	Standard	Conditional SEM	Standard	Conditional SEM	Standard	Conditional SEM
3						
4						
5						
6						
7						
8						
10						

Conditional Standard Error of Measurement by Achievement Standard Oregon Science Assessment, 2009-10						
Grade	Nearly Meets		Meets		Exceeds	
	Standard	Conditional SEM	Standard	Conditional SEM	Standard	Conditional SEM
5						
8						
10						

6.0 SCORING

Tests administered via OAKS Online are scored immediately. Paper-and-pencil tests are returned by schools to District Test Coordinators, who, in turn, return tests to the state’s paper-and-pencil scoring contractor.

6.1 OAKS Online Scoring

The multiple-choice items are scored online, using a scoring algorithm based on Rasch methods, as the student is taking the test. Responses are maintained in a student score file and are compared against the answer key to produce a raw score. The raw score is converted to a scale score called a Rasch Unit or RIT score and is based on the score most likely to have led to the observed responses—taking into account the difficulty of the questions.

ODE verifies the item, test, and pool specifications and independently replicates student scores prior to and during each operational testing window.

6.2 Writing Performance Assessments Scoring

Students must respond to a writing prompt by writing an essay for the writing test; each essay is scored and is awarded points from a range of possible points, based on scoring rubrics. Students respond to the test on answer sheets (for paper-and-pencil administration). Teachers are trained to score the tests and double score all essays in scoring sessions held following the testing window.

6.2.1 Writing Scoring Process

About 1,000 Oregon classroom teachers gather for six days at 16 or more sites around the state to score the state Writing Performance Assessments.

Tests are scored by two readers, usually classroom teachers but sometimes other individuals with language arts backgrounds. Readers undergo a rigorous training and evaluation process to ensure that they have the expertise to score papers using the analytic traits and that they score consistently with readers at the various scoring sites throughout the state.

Writing essays are scored using a six-point scale. Oregon uses an analytic trait scoring guide to assess student writing. The traits (Score Reporting Categories, SRC) scored are Ideas and Content, Organization, Sentence Fluency, Conventions, Voice, and Word Choice. Although the scoring guide includes specific descriptors for each trait, or SRC, and score point, the score point can also be described in a more global perspective:

Exemplary (6 points). Writing at this level is both exceptional and memorable. It is often characterized by distinctive and unusually sophisticated thought processes with rich details and outstanding craftsmanship.

Strong (5 points). Writing at this level exceeds the standard. It is thorough, complex, and consistently portrays exceptional control of content and skills.

Proficient (4 points). Writing at this level meets the standard. It is solid work that has more strengths than weaknesses. The writing demonstrates mastery of skills and reflects care and commitment.

Developing (3 points). Writing at this level shows basic, although sometimes inconsistent, mastery and application of content and skills. It shows some strengths but tends to have more weaknesses overall.

Emerging (2 points). Writing at this level is often superficial, fragmented, or incomplete. It may show a partial mastery of content and skills, but it needs considerable development before reflecting the proficient level of performance.

Beginning (1 point). Writing at this level is minimal. It typically lacks understanding and use of appropriate skills and strategies. Writing at this level may contain major errors.

In addition to the composite score, the State Board set minimum individual scores allowable to meet the standard. Students at grades 4, 7, and High School must receive a score of 3 or more on each required trait (Ideas and Content, Organization, Sentence Fluency, and Conventions). Scores of a 2 or 1 in any one trait from either rater would result in a paper not meeting the standard, regardless of the composite score.

Special scoring codes are applicable to Writing, including codes for essays that are blank, too long or short to score, written in a language other than English, illegible, profane or graphically violent, off-topic, or that show signs of plagiarism. All are not scored.

Student papers that do not address the topic are coded as “off-topic.” This means that the student receives scores on the paper, but data for that student are not included in group reports and the scores do not count toward meeting a standard. Student papers at grade 4 are not coded as off-topic.

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 receive feedback and consider these traits as valuable components of writing along with the required traits.

Teachers indicate the scores on each SRC for each essay on scannable documents that are sent to EDS to be scanned and scored. Composite scores are computed as follows:

The final scores for the two raters are summed across the SRC of Ideas and Content, Organization, and Sentence Fluency. For Benchmark 1 only, this is the final score; for benchmarks 2, 3, and High School, the scores for Conventions are doubled and added to the first sum.

Results are typically available two months after the testing window.

6.2.2 Scorer Reliability

Two teachers independently read and score each student paper using the state scoring guide. If scores given by the first two teachers differ by more than one point, a third expert teacher reads and scores the paper. The third teacher’s score replaces the first two scores. Additionally, papers are reviewed if one rater scored a single trait as “emerging” and both raters otherwise scored the rest of the traits as “Proficient” or higher.

During operational scoring, refresher papers are rated at specified times twice per day, and scorers use these papers to adjust their scoring if necessary. Site scoring directors monitor inter-rater reliability continuously in real time by comparing the scoring of the refresher papers.

Scoring sites utilize a software package to generate reliability reports. The first is a *Score Distribution Summary and Rater Discrepancy with Score Given by the Third Rater* report, which lists the percent frequency and mean discrepancy of each score for each rater. The second is a *Rater Discrepancy Summary* report, which lists the percents in the categories of Agreement, Leniency, and Severity for each rater in the six domains.

APPENDIX A
NCLB CRITICAL ELEMENTS

Location of Evidence Provided in Technical Reports Supporting the Critical Elements of Oregon’s Statewide Assessment System

Sections and Critical Element		Evidence provided in	
		Volume	Section(s) in each volume
Section 1	Content Standards		
1.1	State adoption for Reading, Language Arts, and Math Content Standards	1 3	
1.2	State adoption of Science Content Standards	1 3	
1.3	Challenging, coherent and rigorous Content Standards	1 3	
1.4	Stakeholder involvement in development of Content Standards	1 3	
Section 2	Academic Achievement Standards		
2.1	State adoption for Reading, Language Arts, Math and Alternate Assessment Achievement Standards	1 3	2.1.2.1, 2.1.1.2
2.2	State adoption of Science descriptors and cut scores	1 3	2.1.2.3
2.3	At least 3 levels, competencies, and cut scores	3 6	
2.4	Application to all students	1	2.0, 1.5, 1.7
2.5	Alternate achievement and content alignment	7	
2.6	Stakeholder involvement in development of Achievement Standards	1 3	1.5
Section 3	Statewide Assessment System		
3.1	Current assessment system	1	2.0
3.2	Local assessments	1	2.0
3.3	Matrix design		
3.4	Coherency system across grades and subjects	1 4	1.2, 1.4, 2.0
3.5	Comparability	4	
3.6	Higher order thinking skills	2 4	
3.7	Implementation of alternate assessments		
Section 4	Technical Quality		
4.1	Validity	4	
4.2	Reliability	4	
4.3	Fair and Accessible	1	1.5, 1.7.2., 3.4
4.4	Consistent Interpretation	6	
4.5	Clear procedures and quality monitoring	1	1.5, 3.0, 4.0, 5.0, 6.0

Sections and Critical Element		Evidence provided in	
		Volume	Section(s) in each volume
		2 5 7	
4.6	Evaluation of accommodations	1	3.4.1, 3.4.3
Section 5	Alignment		
5.1	System alignment	2 4	
5.2	Comprehensive and complex alignment	4	
5.3	Content and procedural representation	4	
5.4	Emphasis	4	
5.5	Test specifications and development	2	
5.6	Express results as achievement standards	6	
5.7	Maintain and improve alignment	3 4	
Section 6	Inclusion		
6.1	Include all students in tested grades	1	1.7
6.2	Including SWD	1 7	1.7.2.2
6.3	Including LEP students	1	1.7.2.1.
6.4	Including migrant students	1	1.7.2.3.
Section 7	Assessment reports	6	
7.1	Interpretation support	6	
7.2	Participation reports	1 6	1.7, App B
7.3	Instructionally useful individual reports	6	
7.4	Record security and confidentiality	1 6	3.3.2, 3.3.3
7.5	Instructionally useful strand level reports	6	

APPENDIX B

QUICK LINKS TO ODE ASSESSMENT DOCUMENTS

Academic Content standards

<http://www.ode.state.or.us/teachlearn/real/standards/>

Test Administration Materials

<http://www.ode.state.or.us/search/page/?id=625>

Test Manuals

<http://www.ode.state.or.us/search/page/?id=486>

Information about Extended Assessments

<http://www.ode.state.or.us/search/results/?id=178>

Accommodations and Modifications

<http://www.ode.state.or.us/search/page/?id=487>

Sample tests

<http://www.ode.state.or.us/search/page/?=1222>

Other sites

- ELL –
<http://www.ode.state.or.us/search/page/?=1225>
- National Assessment of Educational Progress (NAEP) –
<http://www.ode.state.or.us/search/results/?id=163>
- Technology Enhanced Student Assessment (TESA) System –
<http://www.ode.state.or.us/search/page/?=391>

Overview of Oregon’s Assessment System and General Overview of Assessments by Subject Area

- Mathematics
<http://www.ode.state.or.us/search/results/?id=239>
- Science
<http://www.ode.state.or.us/search/results/?id=240>
- Social Sciences
<http://www.ode.state.or.us/search/results/?id=241>
- Reading
<http://www.ode.state.or.us/search/results/?id=236>

- Writing
<http://www.ode.state.or.us/search/results/?id=238>

APPENDIX C TECHNICAL COMMITTEE MEMBERS, 2010-11

Technical Advisory Committee (TAC) District Advisory Committee (DAC), and Statewide Assessment Committee (SAC) Membership

Name	Affiliation	Expertise
TAC Members		
Stanley Rabinowitz	WestEd	Policy and Content
Randy Bennett	ETS	Psychometrics
Joe Stevens	U of O	Longitudinal Growth
Wayne Neuburger	ODE-Retired	Large Scale Assessment
Gerald Tindal	U of O	Special Education
Tom Haladyna	Arizona State University – West	Test Development and Validation
DAC Members		
Brian Bain	Tigard-Tualatin SD	District Assessment Coordinator
Todd Bloomquist	Medford Schools 549C	Director of Curriculum and Assessment
Michael Boyles	Harney County SD (Burns-Hines SD)	Curriculum and Assessment Director
Karen Brown Smith	Umatilla Morrow ESD	Data Specialist
Evelyn Brzezinski	Portland Public Schools	Director, Research, Evaluation and Assessment
Catherine Carlson	Salem-Keizer Public Schools	District Assessment Coordinator
Dee Carlson	Beaverton SD	Research Specialist
Elaine Drakulich	North Clackamas SD	Asst Superintendent of Admin Serv
Tim Drilling	Gresham-Barlow SD	Director, Student Achievement
Derek Edens	Gresham-Barlow SD	District Assessment Coordinator
Jim Harrington	Hillsboro SD/NWRESD	CIO
David Hicks	Jefferson County SD 509-J	District Technology Trainer
David Marshall	Milton Freewater USD	Media Specialist/Technology Teacher
Kim Maurer	Santiam Canyon SD	Director of Curriculum/Federal Programs/Assessment
John O'Neill, Jr.	Forest Grove SD & PAC-9	Director of Student Achievement
Kathi Robinson	Hillsboro SD	Executive Director
Tami Schild	NW Regional ESD	Assessment Coordinator
Doug Smith	Klamath County SD	Director of Curriculum/Instruction
Don Staples	Newberg SD	Math Teacher/District Test Coordinator
Bill Stewart	Gladstone SD	Director of Curriculum and Assessment
Joe Suggs	Portland Public Schools	District Assessment Coordinator
Kelvin Webster	Multnomah ESD	Associate Director, Dept. of Instruction
SAC Members		
Doug Kosty	ODE	Large Scale Assessment
Tony Alpert	ODE	Large Scale Assessment
Steve Slater	ODE	Psychometrics

APPENDIX D

RESEARCH STUDIES ON THE OREGON STATEWIDE ASSESSMENT SYSTEM

- Almond, P. J., Tindal, G., Slater, S., and Tinkler, T. (2005). Statewide reading and mathematics achievement for Oregon’s students with disabilities – then and now. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.
- Oregon Department of Education (1992). California Achievement Test Equating study. Report submitted to U. S. Department of Education.
- Choi, S. W. and Tinkler, T. Evaluating comparability of paper and pencil and computer based assessment in a K-12 setting (2002). Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.
- Oregon Department of Education (2005). Differential Item Functioning Study of Students with Learning Disabilities
- Duran, S.W., Brown, C., and McCall, M. (2002). Assessment of English language learners in the Oregon Statewide Assessment System: National and state perspectives. In Gerald Tindal and Thomas Haladyna (Eds.), *Large-Scale Assessment Programs for All Students* (pp. 371-394). Mahwah, NJ; Lawrence Erlbaum Associates, Publishers.
- Choi, S. W. and McCall, M. (2002). Linking Bilingual Mathematics Assessments: A Monolingual IRT Approach. In Gerald Tindal and Thomas Haladyna (Eds.), *Large-Scale Assessment Programs for All Students* (pp. 317-338). Mahwah, NJ; Lawrence Erlbaum Associates, Publishers.>
- Haladyna, T. (2005). Validity Measuring Mathematics Problem Solving Ability in a Statewide Assessment Program.
- Kingsbury, G.C. and Hauser, C. (2004). Computerized adaptive testing and No Child Left Behind, Northwest Evaluation Association. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Sanford, E. (2004). Oregon Lexile Measure Equating Study . MetaMetrics, Inc.
- McCall, M.S. (2005). Technical review of the Oregon Department of Education assessment system. Northwest Evaluation Association.
- McLaughlin, D. (2000). NAEP State Reading Assessment Linkage Study: Report to Oregon. The American Institutes for Research.
- Measure Up: A report on Education Standards and Assessments (2000), conducted by ACHIEVE, Inc.
- Measures of Academic Progress (2004), Northwest Evaluation Association.
- Oregon Department of Education, (2003). The first year: Student performance on the 10th grade benchmark standards and subsequent performance in the first year of college. Oregon Office of Community Colleges and Workforce Development, The Oregon University System
- Oregon Department of Education (2005). Oregon Statewide Assessment System. Standard errors of measurement, reading and mathematics knowledge and skills tests.

- Relationships between student’s performance on Oregon’s 10th grade Certificate of Initial Mastery (CIM) Assessments and Selected Oregon Business Pre-Employment Tests (1999). The Oregon Business Council, The Oregon Department of Education, and The Oregon University System.
- Surveys of Enacted Curriculum report.
- Thomas Haladyna (2001). Evaluation of the Oregon Statewide Assessment Program. at the Arizona State University.
- Tindal, G. (2005). Evaluation of extended science assessment: Volume II – 2005 Analysis *Behavioral Research and Teaching*, University of Oregon.
- Tindal, G. (2006). Alignment of alternate assessments using the Webb system. In *Aligning Assessment to Guide the Learning of All Students: Six Reports on the Development, Refinement, and Dissemination of the Web Alignment Tool* .
- Tindal, G. (2006). Models for Understanding Task Comparability in Accommodated testing.
- Tindal, G., Glasgow, A., Vanloo, D., Chow, E., and Gall, J, (2002). 2001 Oregon extended assessment reliability summary behavioral research and teaching. University of Oregon.
- Yovanoff, P and Tindal, G. (in press). Scaling early reading alternate assessments with statewide measures. *Exceptional Children*.

